Los Alamos
NATIONAL LABORATORY
EST.1943

LA-UR-15-27742
Approved for public release; distribution is unlimited.

Title: HIV Sequence Compendium 2015

Author(s): Foley, Brian Thomas
Leitner, Thomas Kenneth
Korber, Bette Tina Marie
Apetrei, Cristian
Hahn, Beatrice
Mizrachi, Ilene
Mullins, James
Rambaut, Andrew
Wolinsky, Steven

Intended for: Report

Issued: 2015-10-05

Disclaimer:
Los Alamos National Laboratory, an affirmative action/equal opportunity employer,is operated by the Los Alamos National Security, LLC for the National NuclearSecurity Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Departmentof Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

HIV Sequence Compendium 2015

Editors

Brian Foley
Los Alamos National Laboratory

Cristian Apetrei
University of Pittsburgh

Ilene Mizrachi
National Center for Biotechnology Information

Andrew Rambaut
University of Edinburgh

Bette Korber
Los Alamos National Laboratory

Thomas Leitner
Los Alamos National Laboratory

Beatrice Hahn
University of Pennsylvania

James Mullins
University of Washington

Steven Wolinsky
Northwestern University

Contracting Officer's Representative
Anjali Singh
Division of AIDS
National Institute of Allergy and Infectious Diseases

Los Alamos HIV Sequence Database and Analysis Staff
Werner Abfalterer, Mira Dimitrijevic, Shihai Feng, Will Fischer, Peter Hraber, Jennifer Macke, James J. Szinger, Hyejin Yoon

This publication is funded by the Division of AIDS, National Institute of Allergy and Infectious Diseases, through interagency agreement AAI 12007-0000-01000 "HIV/SIV Database and Analysis Unit" with the U.S. Department of Energy.

Published by
Theoretical Biology and Biophysics
Group T-6, Mail Stop K710
Los Alamos National Laboratory
Los Alamos, New Mexico 87545 U.S.A.

LA-UR-15-xxxxx

http://www.hiv.lanl.gov/

HIV Sequence Compendium 2015

Published by
Theoretical Biology and Biophysics
Group T-6, Mail Stop K710
Los Alamos National Laboratory
Los Alamos, New Mexico 87545 U.S.A.

LA-UR-15-xxxxx
Approved for public release; distribution is unlimited.

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Los Alamos National Security, LLC, for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396.

This report was prepared as an account of work sponsored by an agency of the U.S. Government. Neither Los Alamos National Security, LLC, the U.S. Government nor any agency thereof, nor any of their employees make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by Los Alamos National Security, LLC, the U.S. Government, or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of Los Alamos National Security, LLC, the U.S. Government, or any agency thereof.

Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

This report was prepared as an account of work sponsored by NIH/NIAID/DAIDS under contract number AAI 12007-0000-01000 "HIV/SIV Database and Analysis Unit".

Contents

Contents **iii**

I Preface **1**
I-1 Introduction 1
I-2 Acknowledgements 1
I-3 Citing the compendium or database 1
I-4 About the PDF 1
I-5 Genome maps 2
I-6 HIV/SIV proteins 3
I-7 Landmarks of the genome 4
I-8 Amino acid codes 6
I-9 Nucleic acid codes 6

II HIV-1/SIVcpz Complete Genomes **7**
II-1 Introduction 7
II-2 Annotated features 8
II-3 Sequences 11
II-4 Alignments 18

III HIV-2/SIV Complete Genomes **153**
III-1 Introduction 153
III-2 Annotated features 154
III-3 Sequences 156
III-4 Alignments 158

IV PLV Complete Genomes **223**
IV-1 Introduction 223
IV-2 Sequences 224
IV-3 Alignments 226

V HIV-1/SIVcpz Proteins **307**
V-1 Introduction 307
V-2 Annotated features 308
V-3 Sequences 310
V-4 Alignments 318

VI HIV-2/SIV Proteins **371**
VI-1 Introduction 371
VI-2 Annotated features 372
VI-3 Sequences 373
VI-4 Alignments 378

VII PLV Proteins **405**
VII-1 Introduction 405
VII-2 Sequences 406
VII-3 Alignments 413

I

Preface

I-1 Introduction

This compendium is an annual printed summary of the data contained in the HIV sequence database. We try to present a judicious selection of the data in such a way that it is of maximum utility to HIV researchers. Each of the alignments attempts to display the genetic variability within the different species, groups and subtypes of the virus.

This compendium contains sequences published before January 1, 2015. Hence, though it is published in 2015 and called the 2015 Compendium, its contents correspond to the 2014 curated alignments on our website.

The number of sequences in the HIV database is still increasing. In total, at the end of 2014, there were 624,121 sequences in the HIV Sequence Database, an increase of 7% since the previous year. This is the first year that the number of new sequences added to the database has decreased compared to the previous year.

The number of near complete genomes (>7000 nucleotides) increased to 5834 by end of 2014. However, as in previous years, the compendium alignments contain only a fraction of these. A more complete version of all alignments is available on our website, `http://www.hiv.lanl.gov/content/sequence/NEWALIGN/align.html`

As always, we are open to complaints and suggestions for improvement. Inquiries and comments regarding the compendium should be addressed to seq-info@lanl.gov.

I-2 Acknowledgements

The HIV Sequence Database and Analysis Project is funded by the Vaccine and Prevention Research Program of the AIDS Division of the National Institute of Allergy and Infectious Diseases (Anjali Singh, Contracting Officer's Representative) through interagency agreement AAI 12007-0000-01000 "HIV/SIV Database and Analysis Unit" with the U.S. Department of Energy.

I-3 Citing the compendium or database

The LANL HIV Sequence Database may be cited in the same manner as this compendium:

HIV Sequence Compendium 2015. Brian Foley, Thomas Leitner, Cristian Apetrei, Beatrice Hahn, Ilene Mizrachi, James Mullins, Andrew Rambaut, Steven Wolinsky, and Bette Korber editors. 2015. Publisher: Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, New Mexico. LA-UR-15-xxxxx.

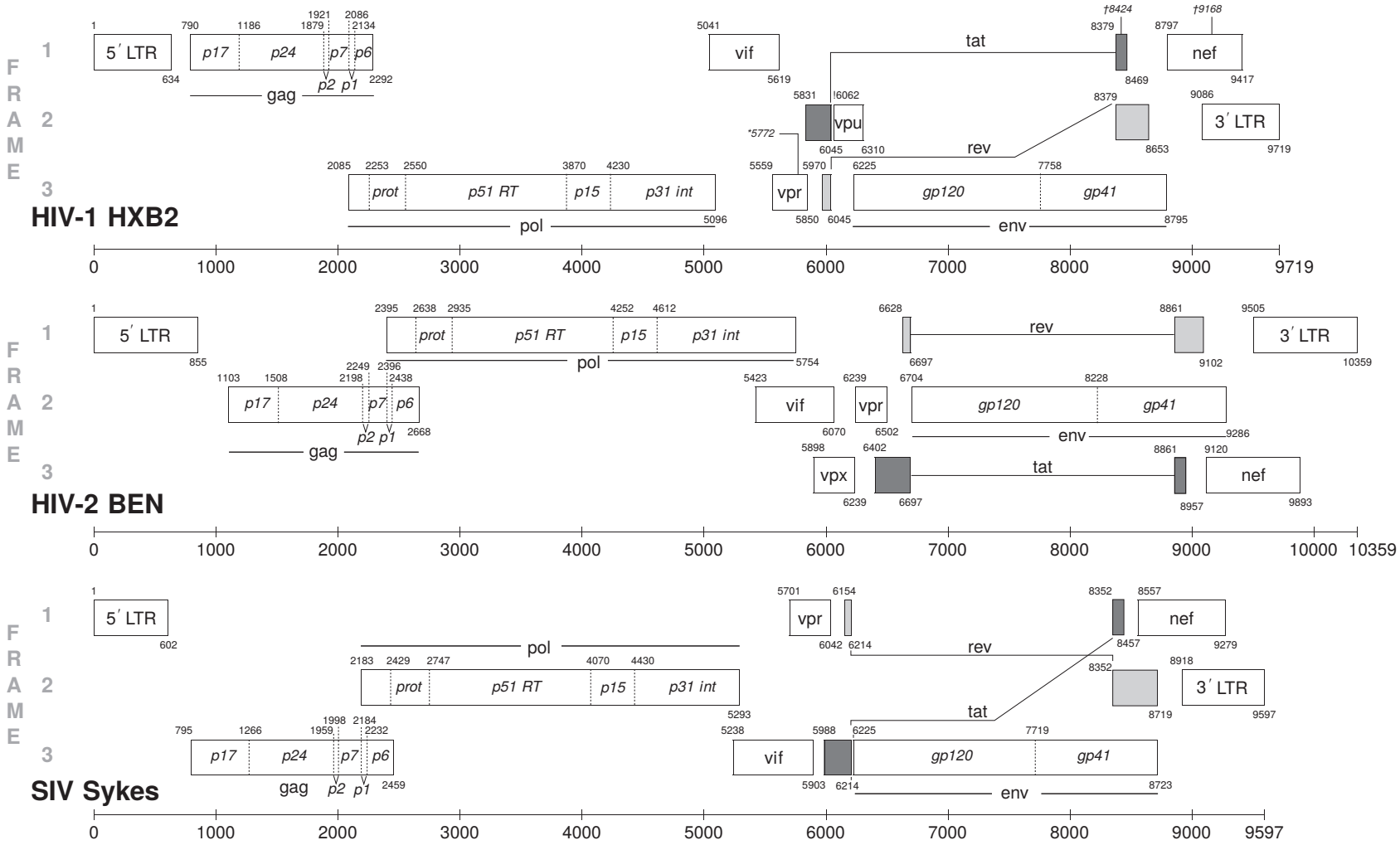
I-4 About the PDF

The complete *HIV Sequence Compendium 2015* is available in Adobe Portable Document Format (PDF) from our website, `http://www.hiv.lanl.gov/`. The 2015 compendium is the first to be available only as PDF; print copies will not be available.

The PDF version is hypertext enabled and features 'clickable' table-of-contents, indexes, references and links to external web sites.

This volume is typeset using LaTeX.

I-5 Genome maps



Landmarks of the HIV-1, HIV-2, and SIV genomes. Open reading frames are shown as rectangles. The gene start, indicated by the small number in the upper left corner of each rectangle normally records the position of the `a` in the `atg` start codon for that gene, while the number in the lower right records the last position of the stop codon. For *pol*, the start is taken to be the first `t` in the sequence `ttttttag`, which forms part of the stem loop that potentiates ribosomal slippage on the RNA and a resulting −1 frameshift and the translation of the Gag-Pol polyprotein. The *tat* and *rev* spliced exons are shown as shaded rectangles. In HXB2, *5772 marks the position of a frameshift in the *vpr* gene caused by an "extra" `t` relative to most other subtype B viruses; !6062 indicates a defective `acg` start codon in *vpu*; †8424 and †9168 mark premature stop codons in *tat* and *nef*. See Korber *et al.*, Numbering Positions in HIV Relative to HXB2CG, in *Human Retroviruses and AIDS*, 1998, p. 102. Available from `http://www.hiv.lanl.gov/content/sequence/HIV/REVIEWS/HXB2.html`

Draft of 30 September 2015, not for public release
preface.tex r8876–10546

I-6 HIV/SIV proteins

Name	Size	Function	Localization
Gag			
MA	p17	membrane anchoring; env interaction; nuclear transport of viral core (myristylated protein)	virion
CA	p24	core capsid	virion
NC	p7	nucleocapsid, binds RNA	virion
	p6	binds Vpr	virion
Pol			
Protease (PR)	p15	Gag/Pol cleavage and maturation	virion
Reverse Transcriptase (RT)	p66, p51	reverse transcription, RNAse H activity	virion
RNase H	p15		virion
Integrase (IN)	p31	DNA provirus integration	virion
Env	gp120/gp41	external viral glycoproteins bind to CD4 and secondary receptors	plasma membrane, virion envelope
Tat	p16/p14	viral transcriptional transactivator	primarily in nucleolus/nucleus
Rev	p19	RNA transport, stability and utilization factor (phosphoprotein)	primarily in nuleolus/nucleus shuttling between nucleolus and cytoplasm
Vif	p23	promotes virion maturation and infectivity	cytoplasm (cytosol, membranes), virion
Vpr	p10-15	promotes nuclear localization of preintegration complex, inhibits cell division, arrests infected cells at G2/M	virion nucleus (nuclear membrane?)
Vpu	p16	promotes extracellular release of viral particles; degrades CD4 in the ER; (phosphoprotein only in HIV-1 and SIVcpz)	integral membrane protein
Nef	p27-p25	CD4 and class I downregulation (myristylated protein)	plasma membrane, cytoplasm, (virion?)
Vpx	p12-16	Vpr homolog present in HIV-2 and some SIVs, absent in HIV-1	virion (nucleus?)
Tev	p28	tripartite tat-env-rev protein (also named Tnv)	primarily in nucleolus/nucleus

I-7 Landmarks of the genome

HIV genomic structural elements

LTR Long terminal repeat, the DNA sequence flanking the genome of integrated proviruses. It contains important regulatory regions, especially those for transcription initiation and polyadenylation.

TAR Target sequence for viral transactivation, the binding site for Tat protein and for cellular proteins; consists of approximately the first 45 nucleotides of the viral mRNAs in HIV-1 (or the first 100 nucleotides in HIV-2 and SIV.) TAR RNA forms a hairpin stem-loop structure with a side bulge; the bulge is necessary for Tat binding and function.

RRE Rev responsive element, an RNA element encoded within the env region of HIV-1. It consists of approximately 200 nucleotides (positions 7327 to 7530 from the start of transcription in HIV-1, spanning the border of gp120 and gp41). The RRE is necessary for Rev function; it contains a high affinity site for Rev; in all, approximately seven binding sites for Rev exist within the RRE RNA. Other lentiviruses (HIV-2, SIV, visna, CAEV) have similar RRE elements in similar locations within env, while HTLVs have an analogous RNA element (RXRE) serving the same purpose within their LTR; RRE is the binding site for Rev protein, while RXRE is the binding site for Rex protein. RRE (and RXRE) form complex secondary structures, necessary for specific protein binding.

PE Psi elements, a set of 4 stem-loop structures preceding and overlapping the Gag start codon which are the sites recognized by the cysteine histidine box, a conserved motif with the canonical sequence CysX2CysX4HisX4Cys, present in the Gag p7 NC protein. The Psi Elements are present in unspliced genomic transcripts but absent from spliced viral mRNAs.

SLIP A TTTTTT slippery site, followed by a stem-loop structure, is responsible for regulating the -1 ribosomal frameshift out of the Gag reading frame into the Pol reading frame.

CRS Cis-acting repressive sequences postulated to inhibit structural protein expression in the absence of Rev. One such site was mapped within the pol region of HIV-1. The exact function has not been defined; splice sites have been postulated to act as CRS sequences.

INS Inhibitory/Instability RNA sequences found within the structural genes of HIV-1 and of other complex retroviruses. Multiple INS elements exist within the genome and can act independently; one of the best characterized elements spans nucleotides 414 to 631 in the gag region of HIV-1. The INS elements have been defined by functional assays as elements that inhibit expression posttranscriptionally. Mutation of the RNA elements was shown to lead to INS inactivation and up regulation of gene expression.

Genes and gene products

GAG The genomic region encoding the capsid proteins (group specific antigens). The precursor is the p55 myristylated protein, which is processed to p17 (MAtrix), p24 (CApsid), p7 (NucleoCapsid), and p6 proteins, by the viral protease. Gag associates with the plasma membrane where the virus assembly takes place. The 55 kDa Gag precursor is called assemblin to indicate its role in viral assembly.

POL The genomic region encoding the viral enzymes protease, reverse transcriptase, RNAse, and integrase. These enzymes are produced as a Gag-Pol precursor polyprotein, which is processed by the viral protease; the Gag-Pol precursor is produced by ribosome frameshifting near the 3′end of gag.

ENV Viral glycoproteins produced as a precursor (gp160) which is processed to give a noncovalent complex of the external glycoprotein gp120 and the transmembrane glycoprotein gp41. The mature gp120-gp41 proteins are bound by non-covalent interactions and are associated as a trimer on the cell surface. A substantial amount of gp120 can be found released in the medium. gp120 contains the binding site for the CD4 receptor, and the seven transmembrane domain chemokine receptors that serve as co-receptors for HIV-1.

TAT Transactivator of HIV gene expression. One of two essential viral regulatory factors (Tat and Rev) for HIV gene expression. Two forms are known, Tat-1 exon (minor form) of 72 amino acids and Tat-2 exon (major form) of 86 amino acids. Low levels of both proteins are found in persistently infected cells. Tat has been localized primarily in the nucleolus/nucleus by immunofluorescence. It acts by binding to the TAR RNA element and activating transcription initiation and elongation from the LTR promoter, preventing the 5′LTR AATAAA polyadenylation signal from causing premature termination of transcription and polyadenylation. It is the first eukaryotic transcription factor known to interact with RNA rather than DNA and may have similarities with prokaryotic anti-termination factors. Extracellular Tat can be found and can be taken up by cells in culture.

REV The second necessary regulatory factor for HIV expression. A 19 kDa phosphoprotein, localized primarily in the nucleolus/nucleus, Rev acts by binding to RRE and promoting the nuclear export, stabilization and utilization of the unspliced viral mRNAs containing RRE. Rev is considered the most functionally conserved regulatory protein of lentiviruses. Rev cycles rapidly between the nucleus and the cytoplasm.

VIF Viral infectivity factor, a basic protein of typically 23 kDa. Promotes the infectivity but not the production of viral particles. In the absence of Vif the produced viral particles are defective, while the cell-to-cell transmission of virus is not affected significantly. Found in almost all lentiviruses, Vif is a cytoplasmic protein, existing in both a soluble cytosolic form and a membrane-associated form. The latter form of Vif is a peripheral membrane protein that is tightly associated with the cytoplasmic side of cellular membranes. In 2003, it was discovered that Vif prevents the action of the cellular APOBEC-3G protein which deaminates DNA:RNA heteroduplexes in the cytoplasm.

VPR Vpr (viral protein R) is a 96-amino acid (14 kDa) protein, which is incorporated into the virion. It interacts with the p6

Gag part of the Pr55 Gag precursor. Vpr detected in the cell is localized to the nucleus. Proposed functions for Vpr include the targeting the nuclear import of preintegration complexes, cell growth arrest, transactivation of cellular genes, and induction of cellular differentiation. In HIV-2, SIV-SMM, SIV-RCM, SIV-MND-2 and SIV-DRL the Vpx gene is apparently the result of a Vpr gene duplication event, possibly by recombination.

VPU Vpu (viral protein U) is unique to HIV-1, SIVcpz (the closest SIV relative of HIV-1), SIV-GSN, SIV-MUS, SIV-MON and SIV-DEN. There is no similar gene in HIV-2, SIV-SMM or other SIVs. Vpu is a 16 kDa (81-amino acid) type I integral membrane protein with at least two different biological functions: (a) degradation of CD4 in the endoplasmic reticulum, and (b) enhancement of virion release from the plasma membrane of HIV-1-infected cells. Env and Vpu are expressed from a bicistronic mRNA. Vpu probably possesses an N-terminal hydrophobic membrane anchor and a hydrophilic moiety. It is phosphorylated by casein kinase II at positions Ser52 and Ser56. Vpu is involved in Env maturation and is not found in the virion. Vpu has been found to increase susceptibility of HIV-1 infected cells to Fas killing.

NEF A multifunctional 27-kDa myristylated protein produced by an ORF located at the 3′end of the primate lentiviruses. Other forms of Nef are known, including nonmyristylated variants. Nef is predominantly cytoplasmic and associated with the plasma membrane via the myristyl residue linked to the conserved second amino acid (Gly). Nef has also been identified in the nucleus and found associated with the cytoskeleton in some experiments. One of the first HIV proteins to be produced in infected cells, it is the most immunogenic of the accessory proteins. The nef genes of HIV and SIV are dispensable *in vitro*, but are essential for efficient viral spread and disease progression *in vivo*. Nef is necessary for the maintenance of high virus loads and for the development of AIDS in macaques, and viruses with defective Nef have been detected in some HIV-1 infected long term survivors. Nef downregulates CD4, the primary viral receptor, and MHC class I molecules, and these functions map to different parts of the protein. Nef interacts with components of host cell signal transduction and clathrin-dependent protein sorting pathways. It increases viral infectivity. Nef contains PxxP motifs that bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of HIV but not for the downregulation of CD4.

VPX A virion protein of 12 kDa found in HIV-2, SIV-SMM, SIV-RCM, SIV-MND-2 and SIV-DRL and not in HIV-1 or other SIVs. This accessory gene is a homolog of HIV-1 vpr, and viruses with Vpx carry both vpr and vpx. Vpx function in relation to Vpr is not fully elucidated; both are incorporated into virions at levels comparable to Gag proteins through interactions with Gag p6. Vpx is necessary for efficient replication of SIV-SMM in PBMCs. Progression to AIDS and death in SIV-infected animals can occur in the absence of Vpr or Vpx. Double mutant virus lacking both vpr and vpx was attenuated, whereas the single mutants were not, suggesting a redundancy in the function of Vpr and Vpx related to virus pathogenicity.

Structural proteins/viral enzymes The products of *gag*, *pol*, and *env* genes, which are essential components of the retroviral particle.

Regulatory proteins Tat and Rev proteins of HIV/SIV and Tax and Rex proteins of HTLVs. They modulate transcriptional and posttranscriptional steps of virus gene expression and are essential for virus propagation.

Accessory or auxiliary proteins Additional virion and non-virion-associated proteins produced by HIV/SIV retroviruses: Vif, Vpr, Vpu, Vpx, Nef. Although the accessory proteins are in general not necessary for viral propagation in tissue culture, they have been conserved in the different isolates; this conservation and experimental observations suggest that their role *in vivo* is very important. Their functional importance continues to be elucidated.

Complex retroviruses Retroviruses regulating their expression via viral factors and expressing additional proteins (regulatory and accessory) essential for their life cycle.

I-8 Amino acid codes

A	Alanine
B	Aspartic Acid or Asparagine
C	Cysteine
D	Aspartic Acid
E	Glutamic Acid
F	Phenylalanine
G	Glycine
H	Histidine
I	Isoleucine
K	Lysine
L	Leucine
M	Methionine
N	Asparagine
P	Proline
Q	Glutamine
R	Arginine
S	Serine
T	Threonine
V	Valine
W	Tryptophan
X	unknown or “other” amino acid
Y	Tyrosine
Z	Glutamic Acid or Glutamine
.	gap
-	identity
*	stop codon
#	incomplete codon

I-9 Nucleic acid codes

A	Adenine
C	Cytosine
G	Guanine
T	Thymine
U	Uracil
M	A or C
R	A or G
W	A or T
S	C or G
Y	C or T
K	G or T
V	A or C or G
H	A or C or T
D	A or G or T
B	C or G or T
N	unknown
.	gap
-	identity

II

HIV-1/SIVcpz Complete Genomes

Contents

II-1	Introduction	**7**
II-2	Annotated features	**8**
II-3	Sequences	**11**
II-4	Alignments	**18**

II-1 Introduction

While the "web-alignment" this year has grown to 2635 full-length genomes, the compendium alignment herein has strict page limitations. The selection of sequences for this year's compendium was slightly different than in past years. We began the selection from the filtered web alignments, thus making the selection slightly more stringent for overall sequence quality. Sequences for each subtype were selected to maximize geographic and phylogenetic diversity, with up to 10 sequences chosen for each pure subtype, CRF01, CRF02, and unclassified (U). When possible, recent sequences were prioritized over older sequences. For all other CRFs, only the reference sequence was included. Representatives were included from groups N, O, and P, and also from the HIV-1-related primate lentiviruses CPZ and GOR. Unique recombinants were all omitted.

The HXB2 sequence (accession K03455) is the master sequence in this alignment. This is also the genome coordinate standard used throughout the HIV Database. The alignment was generated by MAFFT v7.245b (E-INS-i with gap open penalty 2.0) accessed at `http://mafft.cbrc.jp/alignment/server/`. The alignment was subsequently codon-aligned using GeneCutter, followed by a few manual edits to fix obvious misalignments. The alignment presented cannot be considered an "optimal alignment" to any single criterion; it is a compromise between optimal alignment, readability, and codon alignment. Most gaps have been introduced in multiples of 3 bases to maintain open reading frames when the alignment is translated.

Also part of this nucleotide alignment is a translation to protein sequence based on the HXB2 sequence; the HIV genome has many overlapping coding regions, and all are shown. For more complete annotation of functional domains, see the protein sequence alignments in Chapter V.

II-2 Annotated features

Features of HIV-1 annotated in the alignment that follows.

Feature	Location	Page
5' LTR U3 start	1	18
TCF-1 alpha	315-329	20
NF-κ-B-II	350-359	22
NF-κ-B-I	364-373	22
Sp1-III	375-386	22
Sp1-II	388-397	22
Sp1-I	398-408	22
TATA Box	427-431	24
TAR element start	453	24
5' LTR U3 end	455	24
+1 mRNA start site	456	24
5' LTR R repeat begin	456	24
TAR element end	513	24
Poly-A signal	527-532	24
5' LTR R repeat end	551	24
5' LTR U5 start	552	24
Extensive secondary structure	568-605	24
5' LTR U5 end	633	26
Lys tRNA primer binding site	634-653	26
Packaging loops begin	681	26
Packaging loops end	789	28
Gag and Gag-Pol start	790	28
Gag p17 Matrix end	1185	34
Gag p24 Capsid start	1186	34
Gag p24 Capsid end	1881	42
Gag p2 start	1882	42
Gag p2 end	1920	44
Gag-Pol fusion TF protein start	1921	44
Gag p7 nucleocapsid start	1921	44
Gag p7 nucleocapsid end	2085	46
Gag-Pol -1 ribosomal slip site	2085	46
Pol start	2085	46
Gag p1 start	2086	46
Gag p1 end	2133	48
Gag p6 start	2134	48
Gag-Pol TF end	2252	50
Pol protease start	2253	50
Gag p6 end	2292	50
Gag end	2292	50
Pol Protease end	2549	54
Pol p66 and p51 RT start	2550	54
p51 end and p66 RT continue	3869	70
Pol p15 RNAse H start	3870	70
Pol p66 RT, Pol p15 Rnase H end	4229	74
Pol p31 Integrase start	4230	74
Vif start	5041	84
Pol, Gag-Pol, and p31 integrase end	5096	84
Vpr start	5559	90

Feature	Location	Page
Vif end	5619	90
frameshift insert in HXB2	5772	92
Vpr premature end (HXB2 only)	5795	92
Tat exon 1 start	5831	94
Vpr end	5850	94
Rev exon 1 start	5970	94
Tat Rev exon 1 end	6045	96
intron start	6046	96
Vpu start (ACG in HXB2)	6062	96
Vpu transmembrane domain start	6062	96
Vpu transmembrane domain end	6143	98
Env start	6225	100
Vpu end	6310	100
Env signal peptide end	6314	100
Env gp120 start	6315	100
V1 loop start	6615	104
V1 loop end	6692	106
V2 loop start	6696	106
V2 loop end	6812	108
V3 loop start	7110	112
V3 loop end	7217	114
V4 loop start	7377	116
V4 loop end	7478	118
V5 start	7602	120
V5 end	7634	122
Rev Responsive Element (RRE) region	7710	122
Env gp120 end	7757	122
Env gp41 start	7758	122
RRE end	8061	126
Env gp41 transmembrane domain	8277-8336	130
Tat Rev intron end	8378	130
Tat Rev exon 2 start	8379	130
Tat premature stop in HXB2	8424	132
Tat end	8469	132
Rev end (TAA) in some lineages	8605	134
Rev end	8653	134
Env gp41, gp160 end	8795	136
Nef start	8797	136
3' LTR U3 start	9086	142
Nef premature end in HXB2	9168	142
TCF-1 alpha binding	9400-9414	146
Nef end	9417	146
NF-κ-B-II	9435-9444	148
NF-κ-B-I	9449-9458	148
Sp1-III	9462-9471	148
Sp1-II	9473-9482	148
Sp1-I	9483-9493	148
TATA box	9512-9516	148
TAR element start	9538	148
3' LTR U3 end	9540	148
3' LTR repeat start	9541	148
TAR element end	9599	150
Poly-A signal	9612-9617	150

Feature	Location	Page
3’ LTR R repeat end	9636	150
3’ LTR U5 start	9637	150
3’ LTR U5 end	9719	150

II-3 Sequences

Sequences included in the HIV-1/SIVcpz complete genome alignment.

Name	Accession	Country	Author	Reference
B.FR.83.HXB2	K03455	France	Wong-Staal, F.	*Nature* **313**(6000):277-284 (1985)
A1.AU.03.PS1044_Day0	DQ676872	Australia	Li, B.	*J Virol* **81**(1):193-201 (2007)
A1.CH.03.HIV_CH_BID_V3538	JQ403028	Switzerland	Henn, M.R.	*PLoS Pathog* **8**(3):E1002529 (2012)
A1.CY.08.CY236	JF683783	Cyprus	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
A1.ES.05.X1608_8	FJ670519	Spain	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
A1.KE.11.DEMA11KE001	KF716475	Kenya	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A1.RU.11.11RU6950	JX500694	Russia	Baryshev, P.B.	*ARHR* **30**(6); 592-7 (2014)
A1.RW.11.DEMA111RW002	KF716472	Rwanda	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A1.SN.01.DDI579	AY521629	Senegal	Meloni, S.T.	*J Virol* **78**(22):12438-12445 (2004)
A1.UG.11.DEMA110UG001	KF859745	Uganda	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A1.ZA.04.04ZASK162B1	DQ396400	S. Africa	Rousseau, C.M.	*J Virol Methods* **136**(1-2):118-125 (2006)
A2.CM.01.01CM_1445MV	GU201516	Cameroon	Carr, J.K.	*Retrovirology* **2010** Apr 28;7:39 doi: 101186/1742-4690-7-39
A2.CY.94.94CY017_41	AF286237	Cyprus	Gao, F.	*ARHR* **17**(8):675-688 (2001)
B.BR.10.10BR_MG029	KJ849799	Brazil	Pessoa, R.	*Transfusion* **55**(5); 980-90 (2015)
B.CA.07.502_1191_03	JF320424	Canada	Rolland, M.	*Nat Med* **17**(3); 366-71 (2011)
B.CH.08.M2_0803101_NFLG8	KC797225	Switzerland	Castro, E.	*AIDS* **28**(12); 1840-4 (2014)
B.CN.10.DEMB10CN002	JX140658	China	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.ES.10.DEMB10ES002	KC473842	Spain	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.FR.11.DEMB11FR001	KF716496	France	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.GB.05.MM45d213_GN1	HM586212	United Kingdom	Turnbull, E.L.	*J Immunol* **182**(11):7131-7145 (2009)
B.HK.06.HK003	FJ460500	Hong Kong	Tsui, S.K.W.	*ARHR* **26**(1); 117-22 (2010)
B.HT.05.05HT_129389	EU839602	Haiti	Nadai, Y.	*PLoS ONE* **4**(3):E4814 (2009)
B.JP.12.DEMB12JP001	KF716498	Japan	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19

Name	Accession	Country	Author	Reference
B.KR.07.HP_18_07JHS10_3909	KJ140263	S. Korea	Kim, B.-R.	*Haemophilia* **21**(1); e1-11 (2015)
B.PE.07.502_0525_wg5	JF320191	Peru	Rolland, M.	*Nat Med* **17**(3); 366-71 (2011)
B.RU.11.11RU21n	JX500708	Russia	Baryshev, P.B.	Unpublished
B.TH.08.MERLBDTRC10	JN860769	Thailand	Rutvisuttinunt, W.	*ARHR* **28**(12); 1703-11 (2012)
B.US.11.ES38	JN397362	United States	Buckheit, R.W.3.	*Nat Commun* **3**, 716 (2012)
C.AR.01.ARG4006	AY563170	Argentina	Carrion, G.	*ARHR* **20**(9):1022-1025 (2004)
C.BR.07.DEMC07BR003	JX140663	Brazil	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
C.BW.00.00BW5031_1	AF443115	Botswana	Novitsky, V.	*J Virol* **76**(11):5435-5451 (2002)
C.CN.10.YNFL19	KC870038	China	Wei, H.	Unpublished
C.CY.09.CY260	JF683803	Cyprus	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
C.ES.08.X2363_2	EU786681	Spain	Fernandez-Garcia, A.	*ARHR* **25**(1); 93-102 (2009)
C.ET.02.02ET_288	AY713417	Ethiopia	Brown, B.K.	*J Virol* **79**(10):6089-6101 (2005)
C.IN.09.T125_2139	KC156210	India	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.KE.00.KER2010	AF457054	Kenya	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
C.MW.09.703010256_CH256.w96	KC156214	Malawi	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.TZ.08.707010457_CH457.w8	KC156220	Tanzania	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.YE.02.02YE511	AY795906	Yemen	Saad, M.D.	*ARHR* **21**(7):644-648 (2005)
C.ZA.08.705010534_CH534.w12	KC156221	S. Africa	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.ZA.10.DEMC10ZA001	JX140669	S. Africa	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
C.ZM.11.DEMC11ZM006	KF716467	Zambia	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.CM.10.DEMD10CM009	JX140670	Cameroon	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.CY.06.CY163	FJ388945	Cyprus	Kousiappa, I.	*ARHR* **25**(8); 727-40 (2009)
D.KE.11.DEMD11KE003	KF716476	Kenya	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.KR.04.04KBH8	DQ054367	S. Korea	Cho, Y.-K.	*ARHR* **29**(4); 738-43 (2013)
D.SN.90.SE365	AB485648	Senegal	Takekawa, N.	Unpublished
D.TZ.01.A280	AY253311	Tanzania	Arroyo, M.A.	*ARHR* **20**(8):895-901 (2004)
D.UG.10.DEMD10UG004	KF716479	Uganda	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19

Name	Accession	Country	Author	Reference
D.UG.11.DEMD11UG003	KF716480	Uganda	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.YE.02.02YE516	AY795907	Yemen	Saad, M.D.	*ARHR* **21**(7):644-648 (2005)
D.ZA.90.R1	EF633445	S. Africa	Jacobs, G.B.	*ARHR* **23**(12):1575-8 (2007)
F1.AO.06.AO_06_ANG32	FJ900266	Angola	Guimaraes, M.L.	*Retrovirology* **6**, 39 (2009)
F1.AR.02.ARE933	DQ189088	Argentina	Aulicino, P.C.	*ARHR* **21**(2):158-164 (2005)
F1.BR.10.10BR_RJ015	KJ849791	Brazil	Pessoa, R.	*Transfusion* **55**(5); 980-90 (2015)
F1.CY.08.CY222	JF683771	Cyprus	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
F1.ES.02.ES_X845_4	FJ670516	Spain	Fernandez-Garcia, A.	*ARHR* **25**(11):1187-1191 (2009)
F1.RO.96.BCI_R07	AB485658	Romania	Takekawa, N.	Unpublished
F1.RU.08.D88_845	GQ290462	Russia	Fernandez-Garcia, A.	*ARHR* **25**(11):1187-1191 (2009)
F2.CM.02.02CM_0016BBY	AY371158	Cameroon	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
F2.CM.10.DEMF210CM001	JX140672	Cameroon	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
F2.CM.10.DEMF210CM007	JX140673	Cameroon	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.BE.96.DRCBL	AF084936	Belgium	Oelrichs, R.B.	*ARHR* **15**(6):585-589 (1999)
G.CM.10.DEMG10CM008	JX140676	Cameroon	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.CN.08.GX_2084_08	JN106043	China	Liu, W.	*Zhonghua Liu Xing Bing Xue Za Zhi* **34**(1); 53-6 (2013)
G.CU.99.Cu74	AY586547	Cuba	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
G.ES.09.X2634_2	GU362882	Spain	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
G.GH.03.03GH175G	AB287004	Ghana	Takekawa, N.	Unpublished
G.KE.09.DEMG09KE001	KF716477	Kenya	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.NG.09.09NG_SC62	JN248593	Nigeria	Charurat, M.	*J Infect Dis* **205**(8); 1239-47 (2012)
G.ZA.01.TV546	KJ948662	S. Africa	Wilkinson, E.	*ARHR* **31**(4); 412-20 (2015)
H.BE.93.VI991	AF190127	Belgium	Janssens, W.	*AIDS* **14**(11):1533-1543 (2000)
H.CF.90.056	AF005496	C.A.R.	Gao, F.	*J Virol* **72**(7):5680-5698 (1998)
H.GB.00.00GBAC4001	FJ711703	United Kingdom	Holzmayer, V.	*ARHR* **25**(7):721-726 (2009)
J.CM.04.04CMU11421	GU237072	Cameroon	Yamaguchi, J.	*ARHR* **26**(6); 693-7 (2010)
J.SE.93.SE9280_7887	AF082394	Sweden	Laukkanen, T.	*ARHR* **15**(3):293-297 (1999)
J.SE.94.SE9173_7022	AF082395	Sweden	Laukkanen, T.	*ARHR* **15**(3):293-297 (1999)
K.CD.97.97ZR_EQTB11	AJ249235	D.R.C.	Triques, K.	*ARHR* **16**(2):139-151 (2000)
K.CM.96.96CM_MP535	AJ249239	Cameroon	Triques, K.	*ARHR* **16**(2):139-151 (2000)
U.CA.01.TV749	HM215251	Canada	Quesnel-Vallieres, M.	*Emerging Infect Dis* **17**(2):271-274 (2011)
U.CA.99.TV721	HM215249	Canada	Quesnel-Vallieres, M.	*Emerging Infect Dis* **17**(2):271-274 (2011)

Name	Accession	Country	Author	Reference
U.CD.83.83CD003_Z3	AF286236	D.R.C.	Gao, F.	*ARHR* **17**(12):1217-1222 (2001)
U.CD.90.90CD121E12	AF457101	D.R.C.	Mokili, J.L.	*ARHR* **18**(11):817-823 (2002)
U.CY.05.CY090	FJ388921	Cyprus	Kousiappa, I.	*ARHR* **25**(8); 727-40 (2009)
U.CY.08.CY223	JF683772	Cyprus	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
U.ES.10.DEURF10DZ001	JX140679	Spain	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
U.GR.99.99GR303	AY046058	Greece	Paraskevis, D.	*J Gen Virol* **82**(Pt 10):2509-2514 (2001)
U.NL.95.U_NL_95_H10986_D1	EF029066	Netherlands	van der Hoek, L.	*ARHR* **23**(3):466-470 (2007)
01_AE.AF.07.569M	GQ477441	Afghanistan	Sanders-Buell, E.	*ARHR* **26**(5):605-608 (2010)
01_AE.CF.90.90CF11697	AF197340	C.A.R.	Anderson, J.P.	*J Virol* **74**(22):10752-10765 (2000)
01_AE.CN.10.YNFL03	KC870029	China	Wei, H.	Unpublished
01_AE.HK.04.HK001	DQ234790	Hong Kong	Tsui, S.K.W.	Unpublished
01_AE.IR.10.10IR.THR48F	AB703616	Iran	Jahanbakhsh, F.	*ARHR* **29**(1); 198-203 (2013)
01_AE.JP.x.JRC77AE	AB565504	Japan	Umeki-Sakamoto, Y.	Unpublished
01_AE.TH.09.AA090a_WG11	JX447727	Thailand	Rolland, M.	*Nature* **490**(7420); 417-20 (2012)
01_AE.TH.90.CM240	U54771	Thailand	Carr, J.K.	*J Virol* **70**(9):5935-5943 (1996)
01_AE.US.05.306163_FL	JX863920	United States	Heipertz, R.A. Jr.	*ARHR* **29**(10):1310-1320 (2013)
01_AE.VN.98.98VNND15	FJ185235	Viet Nam	Liao, H.	*Virology* **391**(1):51-56 (2009)
02_AG.CM.10.DE00210CM013	KF859739	Cameroon	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
02_AG.CY.09.CY256	JF683799	Cyprus	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
02_AG.ES.06.P1423	EU884501	Spain	Fernandez-Garcia, A.	*ARHR* **25**(1); 93-102 (2009)
02_AG.GW.05.CC_0048	FJ694792	Guinea-Bissau	Vinner, L.	*APMIS* **119**(8); 487-97 (2011)
02_AG.KR.12.12MHR9	KF561435	S. Korea	Cho, Y.-K.	Unpublished
02_AG.LR.x.POC44951	AB485636	Liberia	Takekawa, N.	Unpublished
02_AG.NG.09.09NG_SC61	JN248592	Nigeria	Charurat, M.	*J Infect Dis* **205**(8); 1239-47 (2012)
02_AG.NG.x.IBNG	L39106	Nigeria	Howard, T.M.	*ARHR* **10**(12):1755-1757 (1994)
02_AG.SE.94.SE7812	AF107770	Sweden	Carr, J.K.	*AIDS* **13**(14); 1819-26 (1999)
02_AG.SN.98.98SE_MP1211	AJ251056	Senegal	Toure-Kane, C.	*ARHR* **16**(6):603-609 (2000)
03_AB.RU.97.KAL153_2	AF193276	Russia	Liitsola, K.	*ARHR* **16**(11):1047-1053 (2000)
04_cpx.CY.94.94CY032_3	AF049337	Cyprus	Gao, F.	*J Virol* **72**(12):10234-10241 (1998)
05_DF.BE.x.VI1310	AF193253	Belgium	Laukkanen, T.	*Virology* **269**(1):95-104 (2000)
06_cpx.AU.96.BFP90	AF064699	Australia	Oelrichs, R.B.	*ARHR* **14**(16):1495-1500 (1998)
07_BC.CN.98.98CN009	AF286230	China	Rodenburg, C.M.	*ARHR* **17**(2):161-168 (2001)
08_BC.CN.97.97CNGX_6F	AY008715	China	Piyasirisilp, S.	*J Virol* **74**(23):11286-11295 (2000)
09_cpx.GH.96.96GH2911	AY093605	Ghana	McCutchan, F.E.	*ARHR* **20**(8):819-826 (2004)
10_CD.TZ.96.96TZ_BF061	AF289548	Tanzania	Koulinska, I.N.	*ARHR* **17**(5):423-431 (2001)
11_cpx.CM.95.95CM_1816	AF492624	Cameroon	Wilbe, K.	*ARHR* **18**(12):849-56 (2002)
12_BF.AR.99.ARMA159	AF385936	Argentina	Carr, J.K.	*AIDS* **15**(15):F41-F47 (2001)
13_cpx.CM.96.96CM_1849	AF460972	Cameroon	Wilbe, K.	*ARHR* **18**(12):849-56 (2002)
14_BG.ES.05.X1870	FJ670522	Spain	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
15_01B.TH.99.99TH_MU2079	AF516184	Thailand	Viputtijul, K.	*ARHR* **18**(16):1235-1237 (2002)
16_A2D.KR.97.97KR004	AF286239	S. Korea	Gao, F.	*ARHR* **17**(8):675-688 (2001)
17_BF.AR.99.ARMA038	AY037281	Argentina	Carr, J.K.	*AIDS* **15**(15):F41-F47 (2001)
18_cpx.CU.99.CU76	AY586540	Cuba	Thomson, M.M.	*AIDS* **19**(11):1155-63 (2005)

Name	Accession	Country	Author	Reference
19_cpx.CU.99.CU7	AY894994	Cuba	Casado, G.	*JAIDS* **40**(5):532-537 (2005)
20_BG.CU.99.Cu103	AY586545	Cuba	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
21_A2D.KE.99.KER2003	AF457051	Kenya	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
22_01A1.CM.01.01CM_0001BBY	AY371159	Cameroon	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
23_BG.CU.03.CB118	AY900571	Cuba	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
24_BG.ES.08.X2456_2	FJ670526	Spain	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
25_cpx.CM.02.1918LE	AY371169	Cameroon	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
26_AU.CD.02.02CD_MBTB047	FM877782	D.R.C.	Vidal, N.	*ARHR* **25**(8):823-832 (2009)
27_cpx.FR.04.04CD_FR_KZS	AM851091	France	Vidal, N.	*ARHR* **24**(2):315-321 (2008)
28_BF.BR.99.BREPM12609	DQ085873	Brazil	Sa Filho, D.J.	*ARHR* **22**(1):1-13 (2006)
29_BF.BR.01.BREPM16704	DQ085876	Brazil	Sa Filho, D.J.	*ARHR* **22**(1):1-13 (2006)
31_BC.BR.04.04BR142	AY727527	Brazil	Sanabani, S.	*ARHR* **22**(2):171-176 (2006)
32_06A1.EE.01.EE0369	AY535660	Estonia	Adojaan, M.	*JAIDS* **39**(5):598-605 (2005)
33_01B.ID.07.JKT189_C	AB547463	Indonesia	SahBandar, I.N.	*ARHR* **27**(1); 97-102 (2011)
34_01B.TH.99.OUR2478P	EF165541	Thailand	Tovanabutra, S.	*ARHR* **23**(6):829-833 (2007)
35_AD.AF.07.169H	GQ477446	Afghanistan	Sanders-Buell, E.	*ARHR* **26**(5):605-608 (2010)
36_cpx.CM.00.00CMNYU830	EF087994	Cameroon	Powell, R.L.	*ARHR* **23**(8):1008-1019 (2007)
37_cpx.CM.00.00CMNYU926	EF116594	Cameroon	Powell, R.L.	*ARHR* **23**(7):923-933 (2007)
38_BF1.UY.03.UY03_3389	FJ213783	Uruguay	Ruchansky, D.	*ARHR* **25**(3); 351-6 (2009)
39_BF.BR.04.04BRRJ179	EU735535	Brazil	Guimaraes, M.L.	*AIDS* **22**(3):433-435 (2008)
40_BF.BR.05.05BRRJ055	EU735537	Brazil	Guimaraes, M.L.	*AIDS* **22**(3):433-435 (2008)
42_BF.LU.06.luBF_18_06	EU170139	Luxembourg	Struck, D.	*ARHR* **31**(5); 554-8 (2015)
43_02G.SA.03.J11223	EU697904	Saudi Arabia	Badreddine, S.	*ARHR* **23**(5):667-674 (2007)
44_BF.CL.00.CH80	FJ358521	Chile	Delgado, E.	*ARHR* **26**(7); 821-6 (2010)
45_cpx.FR.04.04FR_AUK	EU448295	France	Frange, P.	*Retrovirology* **2008** Aug 1;5:69 doi: 101186/1742-4690-5-69
46_BF.BR.07.07BR_FPS625	HM026456	Brazil	Sanabani, S.S.	*Virol J* **2010** Apr 16;7:74 doi: 101186/1743-422X-7-74
47_BF.ES.08.P1942	GQ372987	Spain	Fernandez-Garcia, A.	*ARHR* **26**(7); 827-32 (2010)
48_01B.MY.07.07MYKT021	GQ175883	Malaysia	Li, Y.	*JAIDS* **54**(2):129-136 (2010)
49_cpx.GM.03.N26677	HQ385479	Gambia	de Silva, T.I.	*Retrovirology* **7**(1):82 (2010)
50_A1D.GB.10.12792	JN417240	United Kingdom	Foster, G.M.	*PLoS One* **9**(1); e83337 (2014)
51_01B.SG.11.11SG_HM021	JN029801	Singapore	Ng, O.T.	*ARHR* **28**(5); 527-30 (2012)
52_01B.MY.03.03MYKL018_1	DQ366664	Malaysia	Tee, K.K.	*JAIDS* **43**(5):523-529 (2006)
53_01B.MY.11.11FIR164	JX390610	Malaysia	Chow, W.Z.	*J Virol* **86**(20):11398-11399 (2012)
54_01B.MY.09.09MYSB023	JX390976	Malaysia	Ng, K.T.	*J Virol* **86**(20):11405-11406 (2012)
55_01B.CN.10.HNCS102056	JX574661	China	Han, X.	*Genome Announc* **1**(1):E00050-12 (2013)
56_cpx.FR.10.URF5_patient_A	JN882655	France	Leoz, M.	*AIDS* **25**(11):1371-1377 (2011)
57_BC.CN.09.09YNLX19sg	KC899008	China	Han, X.	*PLoS ONE* **8**(5):E65337 (2013)
58_01B.MY.09.09MYPR37	KC522031	Malaysia	Chow, W.Z.	*PLoS ONE* **9**(1):E85250 (2014)
59_01B.CN.09.09LNA423	JX960635	China	An, M.	*J Virol* **86**(22); 12402-6 (2012)
60_BC.IT.11.BAV499	KC899079	Italy	Simonetti, F.R.	*Infect Genet Evol* **2014** Apr;23:176-81 doi: 101016/jmeegid201402007 Epub 2014 Mar 3
61_BC.CN.10.JL100010	KC990124	China	Li, X.	*Genome Announc* **2013** Jun 27;1(3) pii: e00326-13 doi: 101128/genomeA00326-13
62_BC.CN.10.YNFL13	KC870034	China	Wei, H.	*ARHR* **30**(4):380-383 (2014)
63_02A1.RU.10.10RU6637	JN230353	Russia	Baryshev, P.B.	*Arch Virol* **157**(12); 2335-41 (2012)

Name	Accession	Country	Author	Reference
64_BC.CN.09.YNFL31	KC870042	China	Hsi, J.	*ARHR* **30**(4); 389-93 (2014)
65_cpx.CN.10.YNFL01	KC870027	China	Feng, Y.	*ARHR* **30**(6); 598-602 (2014)
67_01B.CN.11.ANHUI_HF115	KC183779	China	Wu, J.	*PLoS ONE* **8**(1):E54322 (2013)
68_01B.CN.11.ANHUI_WH73	KC183782	China	Wu, J.	*PLoS ONE* **8**(1):E54322 (2013)
70_BF1.BR.10.10BR_PE004	KJ849758	Brazil	Pessoa, R.	*PLoS One* **9**(11); e112674 (2014)
71_BF1.BR.10.10BR_PE008	KJ849759	Brazil	Pessoa, R.	*PLoS One* **9**(11); e112674 (2014)
72_BF1.BR.10.10BR_MG002	KJ671534	Brazil	Pessoa, R.	*Genome Announc* **2**(3):e00386-14 (2014)
O.BE.87.ANT70	L20587	Belgium	Vanden Haesevelde, M.	*J Virol* **68**(3):1586-1596 (1994)
O.CM.98.98CMA104	AY169802	Cameroon	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.98.98CMABB141	AY169807	Cameroon	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.98.98CMABB212	AY169804	Cameroon	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.98.98CMU5337	AY169808	Cameroon	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.99.99CMU4122	AY169815	Cameroon	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.FR.92.VAU	AF407418	France	Vartanian, J.P.	*J Gen Virol* **83**(Pt 4):801-805 (2002)
O.GA.11.11Gab6352	JX245015	Gabon	Liegeois, F.	*ARHR* **29**(7); 1085-90 (2013)
O.SN.99.99SE_MP1299	AJ302646	Senegal	Toure-Kane, C.	*ARHR* **17**(12):1211-1216 (2001)
O.US.10.LTNP	JN571034	United States	Buckheit, R.W.3.	*ARHR* **30**(6); 511-513 (2014)
N.CM.02.DJO0131	AY532635	Cameroon	Bodelle, P.	*ARHR* **20**(8):902-908 (2004)
N.CM.02.SJGddd	GQ324959	Cameroon	Vallari, A.	*ARHR* **26**(1):109-115 (2010)
N.CM.04.04CM_1015_04	DQ017382	Cameroon	Yamaguchi, J.	*ARHR* **22**(1):83-92 (2006)
N.CM.06.U14296	GQ324962	Cameroon	Vallari, A.	*ARHR* **26**(1):109-115 (2010)
N.CM.06.U14842	GQ324958	Cameroon	Vallari, A.	*ARHR* **26**(1):109-115 (2010)
N.FR.11.N1_FR_2011	JN572926	France	Delaugerre, C.	*Lancet* **378**(9806); 1894 (2011)
P.CM.06.U14788	HQ179987	Cameroon	Vallari, A.	*J Virol* **85**(3); 1403-7 (2011)
P.FR.09.RBF168	GU111555	France	Plantier, J.-C.	*Nat Med* **15**(8); 871-2 (2009)
CPZ.CD.06.BF1167	JQ866001	D.R.C.	Li, Y.	*J Virol* **86**(19):10776-10791 (2012)
CPZ.CM.05.SIVcpzMT145	DQ373066	Cameroon	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.GA.88.GAB1	X52154	Gabon	Huet, T.	*Nature* **345**(6273):356-359 (1990)
CPZ.TZ.06.SIVcpzTAN13	JQ768416	Tanzania	Takehisa, J.	Unpublished
CPZ.US.85.US_Marilyn	AF103818	United States	Gao, F.	*Nature* **397**(6718):436-441 (1999)
GOR.CM.07.SIVgor2139_287	FJ424866	Cameroon	Takehisa, J.	*J Virol* **83**(4):1635-1648 (2009)
GOR.CM.07.SIVgorCP2135con	FJ424863	Cameroon	Takehisa, J.	*J Virol* **83**(4):1635-1648 (2009)

5' LTR U3 start

```
B.FR.83.HXB2                   TGGAAGGGCTAATTCACTCCCAACGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGAGAAGTTA 170
A1.AU.03.PS1044_Day0           .......................................................................................................................................................................... 0
A1.CH.03.HIV_CH_BID_V3538      .......................................................................................................................................................................... 0
A1.CY.08.CY236                 .......................................................................................................................................................................... 0
A1.ES.05.X1608_8               .......................................................................................................................................................................... 0
A1.KE.11.DEMA11KE001           .......................................................................................................................................................................... 0
A1.RU.11.11RU6950              .......................................................................................................................................................................... 0
A1.RW.11.DEMA111RW002          .......................................................................................................................................................................... 0
A1.SN.01.DDI579                .......................................................................................................................................................................... 0
A1.UG.11.DEMA110UG001          .......................................................................................................................................................................... 0
A1.ZA.04.04ZASK162B1           .......................................................................................................................................................................... 0
A2.CM.01.01CM_1445MV           .......................................................................................................................................................................... 0
A2.CY.94.94CY017_41            .......................................................................................................................................................................... 0

B.BR.10.10BR_MG029             .......................................................................................................................................................................... 0
B.CA.07.502_1191_03            .......................................................................................................................................................................... 0
B.CH.08.M2_0803101_NFLG8       .......................................................................................................................................................................... 0
B.CN.10.DEMB10CN002            .......................................................................................................................................................................... 0
B.ES.10.DEMB10ES002            .......................................................................................................................................................................... 0
B.FR.11.DEMB11FR001            .......................................................................................................................................................................... 0
B.GB.05.MM45d213_GN1           .......................................................................................................................................................................... 0
B.HK.06.HK003                  .......................................................................................................................................................................... 0
B.HT.05.05HT_129389            .......................................................................................................................................................................... 0
B.JP.12.DEMB12JP001            .......................................................................................................................................................................... 0
B.KR.07.HP_18_07JHS10_3909     .......................................................................................................................................................................... 0
B.PE.07.502_0525_wg5           .......................................................................................................................................................................... 0
B.RU.11.11RU21n                .......................................................................................................................................................................... 0
B.TH.08.MERLBDTRC10            .......................................................................................................................................................................... 0
B.US.11.ES38                   .............................................................................----G---A---------------------CC---------C-------------------T------------------T-----T--AG-- 93

C.AR.01.ARG4006                .......................................................................................................................................................................... 0
C.BR.07.DEMC07BR003            .......................................................................................................................................................................... 0
C.BW.00.00BW5031_1             .......................................................................................................................................................................... 0
C.CN.10.YNFL19                 .......................................................................................................................................................................... 0
C.CY.09.CY260                  .......................................................................................................................................................................... 0
C.ES.08.X2363_2                .......................................................................................................................................................................... 0
C.ET.02.02ET_288               .......................................................................................................................................................................... 0
C.IN.09.T125_2139              --------T-----T----TA-GAA---G-----G---------T-----G----T-----G-------------------G---T-----------G--A-----AG-------C--------T-----------T-T-------A----------C---A--G-AG-- 170
C.KE.00.KER2010                .......................................................................................................................................................................... 0
C.MW.09.703010256_CH256.w96    --------T-----T-----A-GAA---G-----C---------T-----G-T--TA------------T---T-------G---A-----------G--A------G------------------------------T------------------C---AG-G-AG-- 170
C.TZ.08.707010457_CH457.w8     --------T-----T-----A-GAA---G-----C---------T-A---G----T-----------T-------------G---A--------------A------G-------------------G--------------------G--------T---A--G-AG-- 170
C.YE.02.02YE511                .......................................................................................................................................................................... 0
C.ZA.08.705010534_CH534.w12    --------T-----T-----AGGAA---G-----G---------T-----G--C-TA------------T-----------G---A--------------A------G------------------------------T---------G-----C--C---AG-G-AG-G 170
C.ZA.10.DEMC10ZA001            .......................................................................................................................................................................... 0
C.ZM.11.DEMC11ZM006            .......................................................................................................................................................................... 0

D.CM.10.DEMD10CM009            .......................................................................................................................................................................... 0
D.CY.06.CY163                  .......................................................................................................................................................................... 0
D.KE.11.DEMD11KE003            .......................................................................................................................................................................... 0
D.KR.04.04KBH8                 ...........................................------------------------------------C-G---T--------------------------------------------T-------T-G----------------AG-----G--G-- 127
D.SN.90.SE365                  --------------TGG---A-GAA---------G-----------T---G------------------T-----------G---T----------------------CT--------C---TG--------------T-C--T-------------T------G--G-- 170
D.TZ.01.A280                   .......................................................................................................................................................................... 0
D.UG.10.DEMD10UG004            .......................................................................................................................................................................... 0
D.UG.11.DEMD11UG003            .......................................................................................................................................................................... 0
D.YE.02.02YE516                .......................................................................................................................................................................... 0
D.ZA.90.R1                     .......................................................................................................................................................................... 0

F1.AO.06.AO_06_ANG32           .......................................................................................................................................................................... 0
F1.AR.02.ARE933                .......................................................................................................................................................................... 0
F1.BR.10.10BR_RJ015            .......................................................................................................................................................................... 0
F1.CY.08.CY222                 .......................................................................................................................................................................... 0
F1.ES.02.ES_X845_4             .......................................................................................................................................................................... 0
F1.RO.96.BCI_R07               ----T---------T-------GAA-----G---G---------------G-T---A------------------------G--------------------------CT-----C--G-------------------T------------------T------G--G-- 170
F1.RU.08.D88_845               .......................................................................................................................................................................... 0
F2.CM.02.02CM_0016BBY          .......................................................................................................................................................................... 0
F2.CM.10.DEMF210CM001          .......................................................................................................................................................................... 0
F2.CM.10.DEMF210CM007          .......................................................................................................................................................................... 0

G.BE.96.DRCBL                  .....................................A-----C--A---G.---TA-T--CA----A-T------A----G--------------------------CT----TC---------C----G-------T---A--G------A-G--T---TCAG--G-- 132
G.CM.10.DEMG10CM008            .......................................................................................................................................................................... 0
G.CN.08.GX_2084_08             .......................................................................................................................................................................... 0
G.CU.99.Cu74                   .......................................................................................................................................................................... 0
G.ES.09.X2634_2                .......................................................................................................................................................................... 0
G.GH.03.03GH175G               ----T---T-----TGG---A-GAA---------G---------------G----TA----------A-T------A----G--------------------------CT---CTC--------------G-------T---A---------A-A--T----CAG--G-- 170
G.KE.09.DEMG09KE001            .......................................................................................................................................................................... 0
G.NG.09.09NG_SC62              .......................................................................................................................................................................... 0
G.ZA.01.TV546                  .......................................................................................................................................................................... 0

H.BE.93.VI991                  .......................................................................................................................................................................... 0
H.CF.90.056                    .......................................................................................................................................................................... 0
H.GB.00.00GBAC4001             .......................................................................................................................................................................... 0

J.CM.04.04CMU11421             .......................................................................................................................................................................... 0
J.SE.93.SE9280_7887            .......................................................................................................................................................................... 0
J.SE.94.SE9173_7022            .......................................................................................................................................................................... 0

K.CD.97.97ZR_EQTB11            .......................................................................................................................................................................... 0
K.CM.96.96CM_MP535             .......................................................................................................................................................................... 0

U.CA.01.TV749                  --------T-----T-------GAA---------C---------------G-----A------------------------G------------------------A--------C-----A--A--------------------------------T------G-AG-- 170
U.CA.99.TV721                  --------T-----T-------G-A---G--G--C---------T-----G-----A----G-------T-----------G------------------------A--------C-----A--A-------------T---------C--------T-----TG-AG-- 170
U.CD.83.83CD003_Z3             .......................................................................................................................................................................... 0
U.CD.90.90CD121E12             .......................................................................................................................................................................... 0
U.CY.05.CY090                  .......................................................................................................................................................................... 0
U.CY.08.CY223                  .......................................................................................................................................................................... 0
U.ES.10.DEURF10DZ001           .......................................................................................................................................................................... 0
U.GR.99.99GR303                .......................................................................................................................................................................... 0
U.NL.95.U_NL_95_H10986_D1      .........................................................................................................----------------------------------------------------T---AGAG-AG-- 65

01_AE.AF.07.569M               .......................................................................................................................................................................... 0
01_AE.CF.90.90CF11697          ...............................................................----------------------------------------------------C------TGT--------------------------------T---A-AG-AG-- 107
01_AE.CN.10.YNFL03             .......................................................................................................................................................................... 0
01_AE.HK.04.HK001              .......................................................................................................................................................................... 0
01_AE.IR.10.10IR.THR48F        .......................................................................................................................................................................... 0
01_AE.JP.x.JRC77AE             ----T---------T----TA-GAA---------G---------T-A---G----TA-T----------T-----------G---A----------------------C------C------TGT-------------T------------------C---ACAG--G-- 170
01_AE.TH.09.AA090a_WG11        .......................................................................................................................................................................... 0
01_AE.TH.90.CM240              .......................................................................................................................................................................... 0
01_AE.US.05.306163_FL          .......................................................................................................................................................................... 0
01_AE.VN.98.98VNND15           .......................................................................................................................................................................... 0
```

```
                                 5’ LTR U3 start
B.FR.83.HXB2                     TGGAAGGGCTAATTCACTCCCAACGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGAGAAGTTA 170
02_AG.CM.10.DE00210CM013         .......................................................................................................................................................................... 0
02_AG.CY.09.CY256                .......................................................................................................................................................................... 0
02_AG.ES.06.P1423                .......................................................................................................................................................................... 0
02_AG.GW.05.CC_0048              .......................................................................................................................................................................... 0
02_AG.KR.12.12MHR9               .......................................................................................................................................................................... 0
02_AG.LR.x.POC44951              ----T---T-----T------GAA----------------------G----T----------A-T-----A---G-----------------------CT---TC-------------G------T---A--------A-G--T----CAG--G-- 170
02_AG.NG.09.09NG_SC61            .......................................................................................................................................................................... 0
02_AG.NG.x.IBNG                  .......................................................................................................................................................................... 0
02_AG.SE.94.SE7812               .......................................................................................................................................................................... 0
02_AG.SN.98.98SE_MP1211          .......................................................................................................................................................................... 0
03_AB.RU.97.KAL153_2             .......................................................................................................................................................................... 0
04_cpx.CY.94.94CY032_3           .......................................................................................................................................................................... 0
05_DF.BE.x.VI1310                .......................................................................................................................................................................... 0
06_cpx.AU.96.BFP90               ----T---------T-----A-GAA------G--G--------------G----T----------T-T----------G---A-------------------CC------C---T----T----------------------------T------G--G-- 170
07_BC.CN.98.98CN009              .......................................................................................................................................................................... 0
08_BC.CN.97.97CNGX_6F            .......................................................................................................................................................................... 0
09_cpx.GH.96.96GH2911            .......................................................................................................................................................................... 0
10_CD.TZ.96.96TZ_BF061           .......................................................................................................................................................................... 0
11_cpx.CM.95.95CM_1816           .......................................................................................................................................................................... 0
12_BF.AR.99.ARMA159              --------T-----T-T---A-GAA---------G---------------------------T----------G-----------------------G------TG-------------G------T-----------------C----T-G--G-- 170
13_cpx.CM.96.96CM_1849           .......................................................................................................................................................................... 0
14_BG.ES.05.X1870                .......................................................................................................................................................................... 0
15_01B.TH.99.99TH_MU2079         .......................................................................................................................................................................... 0
16_A2D.KR.97.97KR004             .......................................................................................................................................................................... 0
17_BF.AR.99.ARMA038              .......................................................................................................................................................................... 0
18_cpx.CU.99.CU76                .......................................................................................................................................................................... 0
19_cpx.CU.99.CU7                 .......................................................................................................................................................................... 0
20_BG.CU.99.Cu103                .......................................................................................................................................................................... 0
21_A2D.KE.99.KER2003             .......................................................................................................................................................................... 0
22_01A1.CM.01.01CM_0001BBY       .......................................................................................................................................................................... 0
23_BG.CU.03.CB118                .......................................................................................................................................................................... 0
24_BG.ES.08.X2456_2              .......................................................................................................................................................................... 0
25_cpx.CM.02.1918LE              .......................................................................................................................................................................... 0
26_AU.CD.02.02CD_MBTB047         --------------T------G-A----A----C--------C-----G----TA-T---------T--------C-G-----T-------------------T-----C-----A--A-----G------T----------------T----CTG--G-- 170
27_cpx.FR.04.04CD_FR_KZS         ----T---------TGG---A-GAA---------A---------T-----G----TA----------T----------C-G---A-------------A------GAG----T---C--------A----------T--------------A--C--CA--G--G-- 170
28_BF.BR.99.BREPM12609           .......................................................................................................................................................................... 0
29_BF.BR.01.BREPM16704           .......................................................................................................................................................................... 0
31_BC.BR.04.04BR142              .......................................................................................................................................................................... 0
32_06A1.EE.01.EE0369             .......................................................................................................................................................................... 0
33_01B.ID.07.JKT189_C            .......................................................................................................................................................................... 0
34_01B.TH.99.OUR2478P            .......................................................................................................................................................................... 0
35_AD.AF.07.169H                 .......................................................................................................................................................................... 0
36_cpx.CM.00.00CMNYU830          .......................................................................................................................................................................... 0
37_cpx.CM.00.00CMNYU926          .......................................................................................................................................................................... 0
38_BF1.UY.03.UY03_3389           .......................................................................................................................................................................... 0
39_BF.BR.04.04BRRJ179            .......................................................................................................................................................................... 0
40_BF.BR.05.05BRRJ055            .......................................................................................................................................................................... 0
42_BF.LU.06.luBF_18_06           .......................................................................................................................................................................... 0
43_02G.SA.03.J11223              .......................................................................................................................................................................... 0
44_BF.CL.00.CH80                 .......................................................................................................................................................................... 0
45_cpx.FR.04.04FR_AUK            --------T--G--T-------GAA---------C----------A---G---------------------------G-----T-------------------T--------------------------TT---T---------G--T---A--G-AG-- 170
46_BF.BR.07.07BR_FPS625          .......................................................................................................................................................................... 0
47_BF.ES.08.P1942                .......................................................................................................................................................................... 0
48_01B.MY.07.07MYKT021           .......................................................................................................................................................................... 0
49_cpx.GM.03.N26677              .......................................................................................................................................................................... 0
50_A1D.GB.10.12792               .......................................................................................................................................................................... 0
51_01B.SG.11.11SG_HM021          .......................................................................................................................................................................... 0
52_01B.MY.03.03MYKL018_1         .......................................................................................................................................................................... 0
53_01B.MY.11.11FIR164            .......................................................................................................................................................................... 0
54_01B.MY.09.09MYSB023           .......................................................................................................................................................................... 0
55_01B.CN.10.HNCS102056          .......................................................................................................................................................................... 0
56_cpx.FR.10.URF5_patient_A      .......................................................................................................................................................................... 0
57_BC.CN.09.09YNLX19sg           .......................................................................................................................................................................... 0
58_01B.MY.09.09MYPR37            .......................................................................................................................................................................... 0
59_01B.CN.09.09LNA423            .......................................................................................................................................................................... 0
60_BC.IT.11.BAV499               --------------T-----A---A-----G------A--GT--------G-T---A---------T--T----------G---A----T------G---------------------TG-------------T--------G-----G---A--A--...... 164
61_BC.CN.10.JL100010             .......................................................................................................................................................................... 0
62_BC.CN.10.YNFL13               .......................................................................................................................................................................... 0
63_02A1.RU.10.10RU6637           .......................................................................................................................................................................... 0
64_BC.CN.09.YNFL31               .......................................................................................................................................................................... 0
65_cpx.CN.10.YNFL01              .......................................................................................................................................................................... 0
67_01B.CN.11.ANHUI_HF115         .......................................................................................................................................................................... 0
68_01B.CN.11.ANHUI_WH73          .......................................................................................................................................................................... 0
70_BF1.BR.10.10BR_PE004          .......................................................................................................................................................................... 0
71_BF1.BR.10.10BR_PE008          .........................................................................................-GCTA-T--AAA--GGCG--CT---CAGT-GCC-CCCCTC--CT-TT--CGTGCG--C--C---A--........ 67
72_BF1.BR.10.10BR_MG002          .......................................................................................................................................................................... 0
O.BE.87.ANT70                    --------T-----T-------TAA----GC---A-----G-----T---G-G--TA----T--G--A-T-----------G------------------A-----A-C---G-TC--------A----------TG-TT--A-----------GTCAGA---AG--GC- 170
O.CM.98.98CMA104                 .......................................................................................................................................................................... 0
O.CM.98.98CMABB141               .......................................................................................................................................................................... 0
O.CM.98.98CMABB212               .......................................................................................................................................................................... 0
O.CM.98.98CMU5337                .......................................................................................................................................................................... 0
O.CM.99.99CMU4122                .......................................................................................................................................................................... 0
O.FR.92.VAU                      .......................................................................................................................................................................... 0
O.GA.11.11Gab6352                .......................................................................................................................................................................... 0
O.SN.99.99SE_MP1299              ----T---T-----T-------TAA----GC---G-----G-----T---G-A--T-----T--G--A-T-----------G----GG------------A------CCA----TC--------A-----G----TA-TT---T----------GTCAGA--CTG--GC- 170
O.US.10.LTNP                     ........................................................................-..T------T-----------A-----A-CT----TC--------A---------CTA-TT--------------GACAGAG---G--GC- 90
N.CM.02.DJO0131                  .......................................................................................................................................................................... 0
N.CM.02.SJGddd                   .......................................................................................................................................................................... 0
N.CM.04.04CM_1015_04             .......................................................................................................................................................................... 0
N.CM.06.U14296                   .......................................................................................................................................................................... 0
N.CM.06.U14842                   .......................................................................................................................................................................... 0
N.FR.11.N1_FR_2011               .......................................................................................................................................................................... 0
P.CM.06.U14788                   .......................................................................................................................................................................... 0
P.FR.09.RBF168                   --------T-----T-T--A-CTG-G---GC---A--A--A-----C---G----T-----T--G--G-T---------C-G----C-T-----T----------AGAA----TC-------------G---CTG-TT--A-----------CTCTGAG-TAG-AG-T 170
CPZ.CD.06.BF1167                 -----------TA-TGG--T-C-AAG---GC--CA--A--A---ACC-----G-TTA----T-----AATA----------G------T-----------------AG-T--------T---TGTAGA--T----TA-TT---T---------CA--C---......... 161
CPZ.CM.05.SIVcpzMT145            .......................................................................................................................................................................... 0
CPZ.GA.88.GAB1                   --------T--G--T-----AGGA----------G--------C--C---G----T-------------T---------C-G-------------A----A-----A-CA----TC-----CTGT-------------T---A-----G--CC-GACAGAG---C--G-- 170
CPZ.TZ.06.SIVcpzTAN13            .......................................................................................................................................................................... 0
CPZ.US.85.US_Marilyn             --------T--G--T-T---AG-A----------G-----A-----C---G-T-----T--------AAT---------C-G------T-----C--------T--AG-A---------C--A--A-A---GC------TT---T----C--TC-CACAGA----G-AG-- 170
GOR.CM.07.SIVgor2139_287         .......................................................................................................................................................................... 0
GOR.CM.07.SIVgorCP2135con        .......................................................................................................................................................................... 0
```

TCF-1 alpha

```
B.FR.83.HXB2                GAAGAA...GCCAACAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGAATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCACATGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGA..... 332
A1.AU.03.PS1044_Day0        ..........................................................................................................................................................  0
A1.CH.03.HIV_CH_BID_V3538   ..........................................................................................................................................................  0
A1.CY.08.CY236              ..........................................................................................................................................................  0
A1.ES.05.X1608_8            ..........................................................................................................................................................  0
A1.KE.11.DEMA11KE001        ..........................................................................................................................................................  0
A1.RU.11.11RU6950           ..........................................................................................................................................................  0
A1.RW.11.DEMA111RW002       ..........................................................................................................................................................  0
A1.SN.01.DDI579             ..........................................................................................................................................................  0
A1.UG.11.DEMA110UG001       ..........................................................................................................................................................  0
A1.ZA.04.04ZASK162B1        ..........................................................................................................................................................  0
A2.CM.01.01CM_1445MV        ..........................................................................................................................................................  0
A2.CY.94.94CY017_41         ..........................................................................................................................................................  0
B.BR.10.10BR_MG029          ..........................................................................................................................................................  0
B.CA.07.502_1191_03         ..........................................................................................................................................................  0
B.CH.08.M2_0803101_NFLG8    ..........................................................................................................................................................  0
B.CN.10.DEMB10CN002         ..........................................................................................................................................................  0
B.ES.10.DEMB10ES002         ..........................................................................................................................................................  0
B.FR.11.DEMB11FR001         ..........................................................................................................................................................  0
B.GB.05.MM45d213_GN1        ..........................................................................................................................................................  0
B.HK.06.HK003               ..........................................................................................................................................................  0
B.HT.05.05HT_129389         ..........................................................................................................................................................  0
B.JP.12.DEMB12JP001         ..........................................................................................................................................................  0
B.KR.07.HP_18_07JHS10_3909  ..........................................................................................................................................................  0
B.PE.07.502_0525_wg5        ..........................................................................................................................................................  0
B.RU.11.11RU21ln            ..........................................................................................................................................................  0
B.TH.08.MERLBDTRC10         ..........................................................................................................................................................  0
B.US.11.ES38                ------..--------C----------A-T--C----------A-----A-----G--------A-----A----------TA---A---------A--------A-C---G-A-------------------A----------..... 255
C.AR.01.ARG4006             ..........................................................................................................................................................  0
C.BR.07.DEMC07BR003         ..........................................................................................................................................................  0
C.BW.00.00BW5031_1          ..........................................................................................................................................................  0
C.CN.10.YNFL19              ..........................................................................................................................................................  0
C.CY.09.CY260               ..........................................................................................................................................................  0
C.ES.08.X2363_2             ..........................................................................................................................................................  0
C.ET.02.02ET_288            ..........................................................................................................................................................  0
C.IN.09.T125_2139           -----G.--------G----------A-T-----C----------T---A----------G--TGAAC-C-------A---AT-----A--------T-AA-----CGCAGA---C-------T---AA---------TT-A---AG--------..... 332
C.KE.00.KER2010             ..........................................................................................................................................................  0
C.MW.09.703010256_CH256.w96 ----------G------A---A-T-TC--C--------A-----AA----------TGATC-C-------A---A-----A--------T-AA-----CACA-A---------C----A---------TT-A---AG--------..... 332
C.TZ.08.707010457_CH457.w8  -----G...AA-CGG----------A---T---C-------A----------G--TGAAC-C-------A---A-----A--------TATG-----CGCAGA----------C---AAA---------T-A---AG--------..... 332
C.YE.02.02YE511             ..........................................................................................................................................................  0
C.ZA.08.705010534_CH534.w12 -----G.--------TG------A---A-T-T---C-G-----A-----A----------G--TGAA--C-------C---AT-----A-------TTA--------CACAGA----------C---A-A---------T-A---AG--------..... 332
C.ZA.10.DEMC10ZA001         ..........................................................................................................................................................  0
C.ZM.11.DEMC11ZM006         ..........................................................................................................................................................  0
D.CM.10.DEMD10CM009         ..........................................................................................................................................................  0
D.CY.06.CY163               ..........................................................................................................................................................  0
D.KE.11.DEMD11KE003         ..........................................................................................................................................................  0
D.KR.04.04KBH8              -----G.-------TG----------A-T--C----------A-----A--------A--G----------C-----C--A------A---A-----AAA---------G-A--T-A------AT-------------TT-A---AG--------..... 289
D.SN.90.SE365               -----G..-----CTG---A---AG---A-T------------C--T---A--------A--G---GA------C--------T------A---A-----AAA---------G-A----A----A--A-------------T--A---AG--------..... 332
D.TZ.01.A280                ..........................................................................................................................................................  0
D.UG.10.DEMD10UG004         ..........................................................................................................................................................  0
D.UG.11.DEMD11UG003         ..........................................................................................................................................................  0
D.YE.02.02YE516             ..........................................................................................................................................................  0
D.ZA.90.R1                  ..........................................................................................................................................................  0
F1.AO.06.AO_06_ANG32        ..........................................................................................................................................................  0
F1.AR.02.ARE933             ..........................................................................................................................................................  0
F1.BR.10.10BR_RJ015         ..........................................................................................................................................................  0
F1.CY.08.CY222              ..........................................................................................................................................................  0
F1.ES.02.ES_X845_4          ..........................................................................................................................................................  0
F1.RO.96.BCI_R07            ---A-G.---T--TG----------T-A-T----AC-------CA------AA----------A--TGAA--C-------AC-GA-----CA-------------G------A-A--TGAA---A----AGA---------T--A-C-AG---AAGAC..... 332
F1.RU.08.D88_845            ..........................................................................................................................................................  0
F2.CM.02.02CM_0016BBY       ..........................................................................................................................................................  0
F2.CM.10.DEMF210CM001       ..........................................................................................................................................................  0
F2.CM.10.DEMF210CM007       ..........................................................................................................................................................  0
G.BE.96.DRCBL               --G---.-------T------------A---TC-A--------CA-CT---A----------G---GAA--C---------C-G-T-----AA--------TA--------CGGAGA-----A----------------------A---AG--------..... 294
G.CM.10.DEMG10CM008         ..........................................................................................................................................................  0
G.CN.08.GX_2084_08          ..........................................................................................................................................................  0
G.CU.99.Cu74                ..........................................................................................................................................................  0
G.ES.09.X2634_2             ..........................................................................................................................................................  0
G.GH.03.03GH175G            --G--G.-------T----------C--A---TC-C--------CA-CT---A--------A--G--TGAA--C-A-------C-G-TA----A--------TA-------CGGAGA--TC-T----------------------A----------..... 332
G.KE.09.DEMG09KE001         ..........................................................................................................................................................  0
G.NG.09.09NG_SC62           ..........................................................................................................................................................  0
G.ZA.01.TV546               ..........................................................................................................................................................  0
H.BE.93.VI991               ..........................................................................................................................................................  0
H.CF.90.056                 ..........................................................................................................................................................  0
H.GB.00.00GBAC4001          ..........................................................................................................................................................  0
J.CM.04.04CMU11421          ..........................................................................................................................................................  0
J.SE.93.SE9280_7887         ..........................................................................................................................................................  0
J.SE.94.SE9173_7022         ..........................................................................................................................................................  0
K.CD.97.97ZR_EQTB11        ..........................................................................................................................................................  0
K.CM.96.96CM_MP535          ..........................................................................................................................................................  0
U.CA.01.TV749               ---------.---T-CTG-R--R-------A------A--------A--T---AA-----GCA--G---T-T--------------AGA----A--------TAAA--G---C-GA-D-----A---AA--------------C-T--A---AG--------..... 332
U.CA.99.TV721               ----------.-T-CTG-----------A------A--------A-AT---AA-----G-A-------T----A---------A-A----A--------T-AA--G---CGGAGA-----A---A-------T-------C-T--A---AG--------..... 332
U.CD.83.83CD003_Z3          ..........................................................................................................................................................  0
U.CD.90.90CD121E12          ..........................................................................................................................................................  0
U.CY.05.CY090               ..........................................................................................................................................................  0
U.CY.08.CY223               ..........................................................................................................................................................  0
U.ES.10.DEURF10DZ001        ..........................................................................................................................................................  0
U.GR.99.99GR303             ..........................................................................................................................................................  0
U.NL.95.U_NL_95_H10986_D1   -----G.-------TG------G----A-T--C----------CA------A----------TG-A--C-----------AGA---GA---------TC------CGAGTA---CAA------T---A-----------A---AG--------..... 227
01_AE.AF.07.569M            ..........................................................................................................................................................  0
01_AE.CF.90.90CF11697       --G------A--------AT--A-G--A-T--C----------CA------AA------G-A-----TGAA--A--------C-GAT-----A--------TTC-------CGAAGA-----A----------G---A------A---AG--------..... 269
01_AE.CN.10.YNFL03          ..........................................................................................................................................................  0
01_AE.HK.04.HK001           ..........................................................................................................................................................  0
01_AE.IR.10.10IR.THR48F     ..........................................................................................................................................................  0
01_AE.JP.x.JRC77AE          --G--G...A----------AG---A-T----------CA------A--------A-----TGAA--C--------C-GAT-----A--------TGC-------CGAA-A-----A-------A-------A------AT--AG--------..... 332
01_AE.TH.09.AA090a_WG11     ..........................................................................................................................................................  0
01_AE.TH.90.CM240           ..........................................................................................................................................................  0
01_AE.US.05.306163_FL       ..........................................................................................................................................................  0
01_AE.VN.98.98VNND15        ..........................................................................................................................................................  0
```

TCF-1 alpha

```
B.FR.83.HXB2                GAAGAA...GCCAACAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGAATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCACATGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGA..... 332
02_AG.CM.10.DE00210CM013    ..........................................................................................................................................................  0
02_AG.CY.09.CY256           ..........................................................................................................................................................  0
02_AG.ES.06.P1423           ..........................................................................................................................................................  0
02_AG.GW.05.CC_0048         ..........................................................................................................................................................  0
02_AG.KR.12.12MHR9          ..........................................................................................................................................................  0
02_AG.LR.x.POC44951         --G---..-----T-----------A---T--A--------CA-CT-T-A--------A--G---A-C--C---C-----C-G-TC-----A--------A----G---C-CA-A----AA--T-A---------------T--A---AG-------..... 332
02_AG.NG.09.09NG_SC61       ..........................................................................................................................................................  0
02_AG.NG.x.IBNG             ..........................................................................................................................................................  0
02_AG.SE.94.SE7812          ..........................................................................................................................................................  0
02_AG.SN.98.98SE_MP1211     ..........................................................................................................................................................  0
03_AB.RU.97.KAL153_2        ..........................................................................................................................................................  0
04_cpx.CY.94.94CY032_3      ..........................................................................................................................................................  0
05_DF.BE.x.VI1310           ..........................................................................................................................................................  0
06_cpx.AU.96.BFP90          --------..CTT-CT-----------A-T--C---------CA-AT---A------GCA--G--TGAA--A----------A-----A---------TC------AGGAGA-----A--------AAA---------TT-A---AG-------CAAGG 337
07_BC.CN.98.98CN009         ..........................................................................................................................................................  0
08_BC.CN.97.97CNGX_6F       ..........................................................................................................................................................  0
09_cpx.GH.96.96GH2911       ..........................................................................................................................................................  0
10_CD.TZ.96.96TZ_BF061      ..........................................................................................................................................................  0
11_cpx.CM.95.95CM_1816      ..........................................................................................................................................................  0
12_BF.AR.99.ARMA159         ---A-G..-------TG----------A-T-----C-------CA------AA----------A--TGAA--C--------AC-GATA----A----------G--GC-AAGA---T-A---A-----AAA-----A---GG-AT--AG----AGAC..... 332
13_cpx.CM.96.96CM_1849      ..........................................................................................................................................................  0
14_BG.ES.05.X1870           ..........................................................................................................................................................  0
15_01B.TH.99.99TH_MU2079    ..........................................................................................................................................................  0
16_A2D.KR.97.97KR004        ..........................................................................................................................................................  0
17_BF.AR.99.ARMA038         ..........................................................................................................................................................  0
18_cpx.CU.99.CU76           ..........................................................................................................................................................  0
19_cpx.CU.99.CU7            ..........................................................................................................................................................  0
20_BG.CU.99.Cu103           ..........................................................................................................................................................  0
21_A2D.KE.99.KER2003        ..........................................................................................................................................................  0
22_01A1.CM.01.01CM_0001BBY  ..........................................................................................................................................................  0
23_BG.CU.03.CB118           ..........................................................................................................................................................  0
24_BG.ES.08.X2456_2         ..........................................................................................................................................................  0
25_cpx.CM.02.1918LE         ..........................................................................................................................................................  0
26_AU.CD.02.02CD_MBTB047    ----------T--T--G-----A----A----C-AC-G--TG--A-------A----------G--TAGT--AG--------C-CAT-----A--------T-AA------CGGA-G----CA-----C-----------------A---AG-------..... 332
27_cpx.FR.04.04CD_FR_KZS     -----G..-------TG------A----A----C--C----T---A----------------G---GAA--A----------AT-----A--------T---------C-AAGA--TG-A---------AAA---------T--A---AG-------..... 332
28_BF.BR.99.BREPM12609       ..........................................................................................................................................................  0
29_BF.BR.01.BREPM16704       ..........................................................................................................................................................  0
31_BC.BR.04.04BR142          ..........................................................................................................................................................  0
32_06A1.EE.01.EE0369         ..........................................................................................................................................................  0
33_01B.ID.07.JKT189_C        ..........................................................................................................................................................  0
34_01B.TH.99.OUR2478P        ..........................................................................................................................................................  0
35_AD.AF.07.169H             ..........................................................................................................................................................  0
36_cpx.CM.00.00CMNYU830      ..........................................................................................................................................................  0
37_cpx.CM.00.00CMNYU926      ..........................................................................................................................................................  0
38_BF1.UY.03.UY03_3389       ..........................................................................................................................................................  0
39_BF.BR.04.04BRRJ179        ..........................................................................................................................................................  0
40_BF.BR.05.05BRRJ055        ..........................................................................................................................................................  0
42_BF.LU.06.luBF_18_06       ..........................................................................................................................................................  0
43_02G.SA.03.JJ1223          ..........................................................................................................................................................  0
44_BF.CL.00.CH80             ..........................................................................................................................................................  0
45_cpx.FR.04.04FR_AUK        -----G..-----CTG-----------TT-T----------A-A----AA------A--G---A-A--A--------C--ATA---A--------T-A------CAAAGA---T----A----AAA---------T--A---AG-------..... 332
46_BF.BR.07.07BR_FPS625      ..........................................................................................................................................................  0
47_BF.ES.08.P1942            ..........................................................................................................................................................  0
48_01B.MY.07.07MYKT021       ..........................................................................................................................................................  0
49_cpx.GM.03.N26677          ..........................................................................................................................................................  0
50_A1D.GB.10.12792           ..........................................................................................................................................................  0
51_01B.SG.11.11SG_HM021      ..........................................................................................................................................................  0
52_01B.MY.03.03MYKL018_1     ..........................................................................................................................................................  0
53_01B.MY.11.11FIR164        ..........................................................................................................................................................  0
54_01B.MY.09.09MYSB023       ..........................................................................................................................................................  0
55_01B.CN.10.HNCS102056      ..........................................................................................................................................................  0
56_cpx.FR.10.URF5_patient_A  ..........................................................................................................................................................  0
57_BC.CN.09.09YNLX19sg       ..........................................................................................................................................................  0
58_01B.MY.09.09MYPR37        ..........................................................................................................................................................  0
59_01B.CN.09.09LNA423        ..........................................................................................................................................................  0
60_BC.IT.11.BAV499           --------A--TTG---A----G--G-T-------------A--CAA-AA---A--G-A--A-----------------AT----A--------A------G--C-----------------AA----------------AG-------..... 320
61_BC.CN.10.JL100010         ..........................................................................................................................................................  0
62_BC.CN.10.YNFL13           ..........................................................................................................................................................  0
63_02A1.RU.10.10RU6637       ..........................................................................................................................................................  0
64_BC.CN.09.YNFL31           ..........................................................................................................................................................  0
65_cpx.CN.10.YNFL01          ..........................................................................................................................................................  0
67_01B.CN.11.ANHUI_HF115     ..........................................................................................................................................................  0
68_01B.CN.11.ANHUI_WH73      ..........................................................................................................................................................  0
70_BF1.BR.10.10BR_PE004      ..........................................................................................................................................................  0
71_BF1.BR.10.10BR_PE008      ..........................................................................................................................-----TC----G--GA---AG--C-TCTGGT-CTAC..... 103
72_BF1.BR.10.10BR_MG002      ..........................................................................................................................................................  0
O.BE.87.ANT70                ---AG-CTA-GAG-T-C-T-T----GGG-T-ATC-CC-G--T--A-CAT-TGCC------T-T--A--TA-AC-T-A----A-AC-GAT-----A------T--ATCT----GCAACACC--TG-T--TATGATAACT--C--A---CT----C----GGA--A-AAAC. 339
O.CM.98.98CMA104             ..........................................................................................................................................................  0
O.CM.98.98CMABB141           ..........................................................................................................................................................  0
O.CM.98.98CMABB212           ..........................................................................................................................................................  0
O.CM.98.98CMU5337            ..........................................................................................................................................................  0
O.CM.99.99CMU4122            ..........................................................................................................................................................  0
O.FR.92.VAU                  ..........................................................................................................................................................  0
O.GA.11.11Gab6352            ..........................................................................................................................................................  0
O.SN.99.99SE_MP1299          --G---CTA-GA-----GT-T----GGG-T---C-CC-G--T--A-CTT--AAC-----CT-T--A---AACC-CG-GC--A-AC-GA-A---CA------T--ATCA----GCAGCACC--TG-T--TATG-TAACCA-C--A---CT---T--C--GGA--A-AAC.. 338
O.US.10.LTNP                 ---AG-CTA-GA--T-C-T-T----GGG-T-ATC-CC-G--T--A-CAT-TGCC-----CTCT-----T-AAC-TG-G---C-AC-GAT-----A------T--ATCT----GCAACACC--TG-T--T-AGATAACT--C--A---CT---C----GGA--A-AAGC. 259
N.CM.02.DJO0131              ..........................................................................................................................................................  0
N.CM.02.SJGddd               ..........................................................................................................................................................  0
N.CM.04.04CM_1015_04         ..........................................................................................................................................................  0
N.CM.06.U14296               ..........................................................................................................................................................  0
N.CM.06.U14842               ..........................................................................................................................................................  0
N.FR.11.N1_FR_2011           ..........................................................................................................................................................  0
P.CM.06.U14788               ..........................................................................................................................................................  0
P.FR.09.RBF168               --G---ATG-GAG-TG-GCAG-----AG-T-AGC--C----T--A-CCT--ACTT----GGATTCA--T-ATC-T--G--GA-C---AT-----A---------ATCA--G-G-GAAAGA--TG-T---TT-C-AAA---C-----ACTG--T-CT--AGA--A-..... 335
CPZ.CD.06.BF1167             .........C-AG-AG-G-AT--A-GG-ATCTAC-CA-G-------CTT--TCCT------GA--------TA-TG-----C----G-TA----A-------TCTGAA--G--TAGAAGA-----A---A-----AGA---------------AT-AG-ACTGCTTCC 322
CPZ.CM.05.SIVcpzMT145        ..........................................................................................................................................................  0
CPZ.GA.88.GAB1               ---C-..------TG-------T----A-T--C----G--T--CA-AT-T-A------G-----A--TGA---C-A---G--C--G-TC---C-C---------A-G--G--C--AAGA--T--T---A----A-AA-----------A---AG----ACTTTAGAC 337
CPZ.TZ.06.SIVcpzTAN13        ..........................................................................................................................................................  0
CPZ.US.85.US_Marilyn         --GCGG.-------T---------C--A--TAC--C-G-----CA--T---A----------A--TGAAC-TG-C-----C---ATC---CA--------CTGAA--G--TCGGAGA----GA--TAA----------A--------CG------A--GCTGTA 337
GOR.CM.07.SIVgor2139_287     ..........................................................................................................................................................  0
GOR.CM.07.SIVgorCP2135con    ..........................................................................................................................................................  0
```


Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

NF-κ-B-II
NF-κ-B-I
Sp1-III
Sp1-II
Sp1-I
B.FR.83.HXB2 ...CATCGAGCTTGCT...ACAAGGGACTTT...CCGCTGGGGACTTTC...CAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGA 419
A1.AU.03.PS1044_Day0
A1.CH.03.HIV_CH_BID_V3538
A1.CY.08.CY236
A1.ES.05.X1608_8
A1.KE.11.DEMA11KE001
A1.RU.11.11RU6950
A1.RW.11.DEMA111RW002
A1.SN.01.DDI579
A1.UG.11.DEMA110UG001
A1.ZA.04.04ZASK162B1
A2.CM.01.01CM_1445MV
A2.CY.94.94CY017_41
B.BR.10.10BR_MG029
B.CA.07.502_1191_03
B.CH.08.M2_0803101_NFLG8
B.CN.10.DEMB10CN002
B.ES.10.DEMB10ES002
B.FR.11.DEMB11FR001
B.GB.05.MM45d213_GN1
B.HK.06.HK003
B.HT.05.05HT_129389
B.JP.12.DEMB12JP001
B.KR.07.HP_18_07JHS10_3909 ..T-- 81
B.PE.07.502_0525_wg5
B.RU.11.11RU21n
B.TH.08.MERLBDTRC10
B.US.11.ES38 ..CT-----T-- 343
C.AR.01.ARG4006
C.BR.07.DEMC07BR003
C.BW.00.00BW5031_1
C.CN.10.YNFL19
C.CY.09.CY260
C.ES.08.X2363_2
C.ET.02.02ET_288
C.IN.09.T125_2139 ..CA--AGGGA--TTCC 419
C.KE.00.KER2010
C.MW.09.703010256_CH256.w96 ..CAA-AGGGA--TTCC 420
C.TZ.08.707010457_CH457.w8 ..CA--AGGGA--TTCC 419
C.YE.02.02YE511
C.ZA.08.705010534_CH534.w12 ..CA--AGGGA--TTC 420
C.ZA.10.DEMC10ZA001
C.ZM.11.DEMC11ZM006
D.CM.10.DEMD10CM009
D.CY.06.CY163
D.KE.11.DEMD11KE003
D.KR.04.04KBH8 ..CT---T--T-- 376
D.SN.90.SE365 ..CA--AG----- 418
D.TZ.01.A280
D.UG.10.DEMD10UG004
D.UG.11.DEMD11UG003
D.YE.02.02YE516
D.ZA.90.R1
F1.AO.06.AO_06_ANG32
F1.AR.02.ARE933
F1.BR.10.10BR_RJ015
F1.CY.08.CY222
F1.ES.02.ES_X845_4
F1.RO.96.BCI_R07 ......CTCTGA--CA---A-----GACATAGAAGATTAGAGATTGCTGACATAGAAGATTCTA 462
F1.RU.08.D88_845
F2.CM.02.02CM_0016BBY
F2.CM.10.DEMF210CM001
F2.CM.10.DEMF210CM007
G.BE.96.DRCBL ..CA--AG-----GAC 384
G.CM.10.DEMG10CM008
G.CN.08.GX_2084_08
G.CU.99.Cu74 13
G.ES.09.X2634_2
G.GH.03.03GH175G ..CA--AGG--T-GACTGC..TGACACAGAAGTTGCTG 442
G.KE.09.DEMG09KE001
G.NG.09.09NG_SC62
G.ZA.01.TV546
H.BE.93.VI991
H.CF.90.056
H.GB.00.00GBAC4001
J.CM.04.04CMU11421
J.SE.93.SE9280_7887
J.SE.94.SE9173_7022
K.CD.97.97ZR_EQTB11
K.CM.96.96CM_MP535
U.CA.01.TV749 ..CA--AG----- 419
U.CA.99.TV721 ..CA--AG----- 419
U.CD.83.83CD003_Z3
U.CD.90.90CD121E12
U.CY.05.CY090
U.CY.08.CY223
U.ES.10.DEURF10DZ001
U.GR.99.99GR303
U.NL.95.U_NL_95_H10986_D1 ..T-AA--AG--T-- 314
01_AE.AF.07.569M
01_AE.CF.90.90CF11697 ..GA--AG--T-- 356
01_AE.CN.10.YNFL03
01_AE.HK.04.HK001
01_AE.IR.10.10IR.THR48F
01_AE.JP.x.JRC77AE ..AA--AG--T-- 418
01_AE.TH.09.AA090a_WG11
01_AE.TH.90.CM240
01_AE.US.05.306163_FL
01_AE.VN.98.98VNND15

NF-κ-B-II
NF-κ-B-I
Sp1-III
Sp1-II
Sp1-I

```
                                                                    +1 mRNA start site
                                                                  TAR element start
                                   TATA Box           5' LTR U3 end 5' LTR R repeat begin                         TAR element end                Poly-A signal   5' LTR R repeat end 5' LTR U5 start   Extensive secondary structure
B.FR.83.HXB2                 TCCTGCATATAAGCAGCTGCTTTTTGCC.TGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTA....GGGAACCCACTGCTT.AAGCCTCAATAAAGCTTGCCTTGAGTGCTT.CAAGTAGTGTGTGCCCGTCTG.TTGTGT.GACT 580
A1.AU.03.PS1044_Day0         .......................................................................................................................................................................... 0
A1.CH.03.HIV_CH_BID_V3538    .......................................................................................................................................................................... 0
A1.CY.08.CY236               .......................................................................................................................................................................... 0
A1.ES.05.X1608_8             ..................................................................................................................................-CTTC..A-GT-----------------.----A-.---- 36
A1.KE.11.DEMA11KE001         .......................................................................................................................................................................... 0
A1.RU.11.11RU6950            .....................................---------------------------------------------G---....---------------.------------------------------.T--------------------.----A-.---- 125
A1.RW.11.DEMA111RW002        .......................................................................................................................................................................... 0
A1.SN.01.DDI579              .......................................................................................................................................................................... 0
A1.UG.11.DEMA110UG001        .......................................................................................................................................................................... 0
A1.ZA.04.04ZASK162B1         ..........................................................................................................................................................----.------.-G-- 14
A2.CM.01.01CM_1445MV         .......................................................................................................................................................................... 0
A2.CY.94.94CY017_41          .......................................................................................................................................................................... 0
B.BR.10.10BR_MG029           ...........................................................................................--C-------A---................................G--------------------.------.---- 45
B.CA.07.502_1191_03          .......................................................................................................................................................................... 0
B.CH.08.M2_0803101_NFLG8     ..................................................................................................................................-C-GA-.-C-------------------.------.---- 37
B.CN.10.DEMB10CN002          .......................................................................................................................................................................... 0
B.ES.10.DEMB10ES002          .......................................................................................................................................................................... 0
B.FR.11.DEMB11FR001          .......................................................................................................................................................................... 0
B.GB.05.MM45d213_GN1         .......................................................................................................................................................................... 0
B.HK.06.HK003                .......................................................................................................................................................................... 0
B.HT.05.05HT_129389          .......................................................................................................................................................................... 0
B.JP.12.DEMB12JP001          .......................................................................................................................................................................... 0
B.KR.07.HP_18_07JHS10_3909   -G-------A------------------.----------------------------------------------------G----....---------------.------------------------------.T--------------------.------.---- 242
B.PE.07.502_0525_wg5         ...............................................................................................................................................---------------.------.---- 25
B.RU.11.11RU21n              .....................................--------------------------------------------G----....---------------.-----------------------------C.T--------------------.------.---- 125
B.TH.08.MERLBDTRC10          .......................................................................................................................................................................... 0
B.US.11.ES38                 -G--------------------------.-----------------------------------------------------G---....---------------.------------------------------.---------------------.----T-.---- 504
C.AR.01.ARG4006              .......................................................................................................................................................................... 0
C.BR.07.DEMC07BR003          .......................................................................................................................................................................... 0
C.BW.00.00BW5031_1           .......................................................................................................................................................................... 0
C.CN.10.YNFL19               .......................................................................................................................................................................... 0
C.CY.09.CY260                .......................................................................................................................................................................... 0
C.ES.08.X2363_2              ...........................................................................................................................................................---.------.---- 13
C.ET.02.02ET_288             .......................................................................................................................................................................... 0
C.IN.09.T125_2139            -G-------------------G------.----------------A-G-G------------------------------C-T---....---------------.-----------------------------C.TG--C----------------.------.---- 580
C.KE.00.KER2010              .......................................................................................................................................................................... 0
C.MW.09.703010256_CH256.w96  -G----------------------C---.----------------A-G---------------------------------GC---....---------------.-----------------------------C.T--------------------.------.---- 581
C.TZ.08.707010457_CH457.w8   -G----------------------C---.----------------A-G----------------------------------T--G....---------------.-----------------------------C.T--------------------.------.---- 580
C.YE.02.02YE511              .......................................................................................................................................................................... 0
C.ZA.08.705010534_CH534.w12  -G----------------------C---.----------------A-G-----------------C----------------T---....---------------.------------------------------.G--------------------.------.---- 581
C.ZA.10.DEMC10ZA001          .......................................................................................................................................................................... 0
C.ZM.11.DEMC11ZM006          .......................................................................................................................................................................... 0
D.CM.10.DEMD10CM009          .......................................................................................................................................................................... 0
D.CY.06.CY163                .......................................................................................................................................................................... 0
D.KE.11.DEMD11KE003          .......................................................................................................................................................................... 0
D.KR.04.04KBH8               -G---------------------C----.-----------------------------------------------------G---....---------------.----------------------A--C--C-.----------A----------.------.A--- 537
D.SN.90.SE365                -G---------------------C----.-----G-----------------------T--------A--------------G---....---------------.------------------------------.----------------A----.------.---- 579
D.TZ.01.A280                 .......................................................................................................................................................................... 0
D.UG.10.DEMD10UG004          .......................................................................................................................................................................... 0
D.UG.11.DEMD11UG003          .......................................................................................................................................................................... 0
D.YE.02.02YE516              .......................................................................................................................................................................... 0
D.ZA.90.R1                   .......................................................................................................................................................................... 0
F1.AO.06.AO_06_ANG32         .......................................................................................................................................................................... 0
F1.AR.02.ARE933              .......................................................................................................................................................................... 0
F1.BR.10.10BR_RJ015          ................................................................................................................---C.-GA-GG--ACAC-C-A---.G---Y----------------.------.---- 54
F1.CY.08.CY222               .......................................................................................................................................................................... 0
F1.ES.02.ES_X845_4           ..................................................................................................................................TGCT-C.A-G---AG-G-----------.------.---- 37
F1.RO.96.BCI_R07             -G----------------------C--T.----------------A------------T-----------------------G---....---------------.------------------------------.---------------------.----T-.---- 623
F1.RU.08.D88_845             ..................................................................................................................................T-T---.---------------------.------.---- 37
F2.CM.02.02CM_0016BBY        .......................................................................................................................................................................... 0
F2.CM.10.DEMF210CM001        .......................................................................................................................................................................... 0
F2.CM.10.DEMF210CM007        .......................................................................................................................................................................... 0
G.BE.96.DRCBL                AG---------------C----C-C---.----------------T------------T-----------------------G-AG....---------------AG-----------------------------.---------------------.------.---- 546
G.CM.10.DEMG10CM008          .......................................................................................................................................................................... 0
G.CN.08.GX_2084_08           .......................................................................................................................................................................... 0
G.CU.99.Cu74                 -G---------------C----C-C---.----------------T-C----------T-------------------A---G-AG....---------------.------------------------------.-----G---------------.------.---- 174
G.ES.09.X2634_2              .........................................................................................................................----.--C-AGTA.........---------------.------.---- 37
G.GH.03.03GH175G             AG---------------C----C-C--T.--------------TCT-C----------------------------------GTAG....A--------------.------------------------------.---------------------.------.---- 603
G.KE.09.DEMG09KE001          .......................................................................................................................................................................... 0
G.NG.09.09NG_SC62            .......................................................................................................................................................................... 0
G.ZA.01.TV546                .......................................................................................................................................................................... 0
H.BE.93.VI991                .......................................................................................................................................................................... 0
H.CF.90.056                  .......................................................................................................................................................................... 0
H.GB.00.00GBAC4001           ......................................................................------------G---....---------------.----------------------------C-.T--------------------.------.---- 92
J.CM.04.04CMU11421           ..................................................................----------------G---....---------------.------------------------------.---------------------.------.---- 96
J.SE.93.SE9280_7887          .......................................................................................................................................................................... 0
J.SE.94.SE9173_7022          .......................................................................................................................................................................... 0
K.CD.97.97ZR_EQTB11          .......................................................................................................................................................................... 0
K.CM.96.96CM_MP535           .......................................................................................................................................................................... 0
U.CA.01.TV749                -G---------------C-----C----.----------------T---G-------.TA----------------------G-AG....---------------.------------------------------.---------------------.------.---- 579
U.CA.99.TV721                -G---------------C-----C----.----------------T-C---------.T-------------------A---G-AG....---------------.------------------------------.---------------------.------.---- 579
U.CD.83.83CD003_Z3           ......................................................................................TA..---------------.------------------------------.A--------------------.------.---- 78
U.CD.90.90CD121E12           .......................................................................................................................................................................... 0
U.CY.05.CY090                .......................................................................................................................................................................... 0
U.CY.08.CY223                .......................................................................................................................................................................... 0
U.ES.10.DEURF10DZ001         .......................................................................................................................................................................... 0
U.GR.99.99GR303              .......................................................................................................................................................................... 0
U.NL.95.U_NL_95_H10986_D1    -G---------------C------C---.----------------A------------------------------------G---....---------------A------------------------------.A----G---------------.------.---- 476
01_AE.AF.07.569M             .......................................................................................................................................................................... 0
01_AE.CF.90.90CF11697        -G-------A-------C------C--T.----------------T-A--------G.TC-----C----------------G-A-....---------------A------------------------------.A----G---------------.-GA-AG.---- 517
01_AE.CN.10.YNFL03           .......................................................................................................................................................................... 0
01_AE.HK.04.HK001            .......................................................................................................................................................................... 0
01_AE.IR.10.10IR.THR48F      .......................................................................................................................................................................... 0
01_AE.JP.x.JRC77AE           -G-------A-------C------C--T.----------------T-G--------G.TC-----C----------------G-A-....---------------G------------------------------.A----G---------------.-GT-AG.---- 579
01_AE.TH.09.AA090a_WG11      .......................................................................................................................................................................... 0
01_AE.TH.90.CM240            ...................................----------T----------G.TC-----C----------------G-A-....---------------A------------------------------.A----G---------------.-GT-AG.---- 127
01_AE.US.05.306163_FL        .......................................................................................................................................................................... 0
01_AE.VN.98.98VNND15         .......................................................................................................................................................................... 0
```

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

Extensive secondary structure | 5' LTR U5 end | Lys tRNA primer binding site | Packaging loops begin

```
B.FR.83.HXB2                 CTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGG.AAAATCTCTAGCAGTGGCGCC.........................CGAACAGGG.ACC.TGA.AAGCGAAAG.....................GGAAACCAGAGG...AGCTCTCTCGACGCA.GGA 696
A1.AU.03.PS1044_Day0         .......................................................................................................................................................................... 0
A1.CH.03.HIV_CH_BID_V3538    .......................................................................................................................................................................... 0
A1.CY.08.CY236               .......................................................................................................................................................................... 0
A1.ES.05.X1608_8             -------------------------AC-GG--A-C---C-..---------------------.........................---------.--T.C--.---------........................TT------A...--T------------.--- 148
A1.KE.11.DEMA11KE001         ...............................................................................................--.--T.---.---------........................TG------A...--T------------.--- 44
A1.RU.11.11RU6950            -------------------------AC-C---G-G----A..---------------------.........................---------.--T.---.---------........................TT------A...--T------------.--- 237
A1.RW.11.DEMA111RW002        .........................................................................................--------.--T.C--.---------........................TG------A...--T------------.--- 50
A1.SN.01.DDI579              .......................................................................................................................................................................... 0
A1.UG.11.DEMA110UG001        ..............................................................................AAAGCGAAAGTT--T----.--T.---.---------........................TT------A...--T------------.--- 61
A1.ZA.04.04ZASK162B1         -------------------------AC-C---A-T-A--A..G--------------------.........................---------.--T.C--.---------........................TT------A...--T------------.--- 126
A2.CM.01.01CM_1445MV         .......................................................................................................................................................................... 0
A2.CY.94.94CY017_41          ...........................................................................TTGAAAAGCGAAAGT-------.--T.C--.---------........................TT-------...--T------------.--- 64
B.BR.10.10BR_MG029           ------------------------A----------------.---------------------.........................---------.---.C--.---------.....................-A----------...---------------.--- 161
B.CA.07.502_1191_03          .------------------------A---------------.---------------------.........................---------.---.C--.---------.....................TAG---------...--A------------.--- 115
B.CH.08.M2_0803101_NFLG8     -------------------------.---------------.---------------------.........................---------.---C---.---------.....................TAGG-------A...--T------------.--- 153
B.CN.10.DEMB10CN002          .......................................................--------.........................---------.---.---.---------.....................TAG---------...---------------.--- 63
B.ES.10.DEMB10ES002          ...................................................--T---------.........................---------.---.A--.---------.....................TA----------...---------------.--- 66
B.FR.11.DEMB11FR001          .......................................................................................................................................................................... 0
B.GB.05.MM45d213_GN1         ..............................................-----------------.........................---------.--T.---A---T-----.....................A----------A...---------------.--- 72
B.HK.06.HK003                ..............................................................-.........................---------.---.---.---------......................AG---------...---------------.--- 54
B.HT.05.05HT_129389          .......................................................................................................................................................................... 0
B.JP.12.DEMB12JP001          ................................................--CCT----------.........................---------.---.---.---T-----.....................T--G-------A...---------------.--- 69
B.KR.07.HP_18_07JHS10_3909   -----------------------------------------.---------------------.........................---------.--T.---.---------.....................TAG---------...---------------.--- 358
B.PE.07.502_0525_wg5         ----------------------------AA-----A-----.---------------------.........................---------.--T.---.---T-----.....................TA---------A...--------T------.--- 141
B.RU.11.11RU21n              ------------------------T----------------.---------------------.........................---------.--T.---.---T-----.....................TA--C-AGAG-A......G----T------.--- 238
B.TH.08.MERLBDTRC10          .......................................................................................................................................................................... 0
B.US.11.ES38                 ------------------------T--CG-----------A.---------------------.........................---------.---T-A-.---------.....................A-----------...---------------.--- 621
C.AR.01.ARG4006              .......................................................................................................................................................................... 0
C.BR.07.DEMC07BR003          .............................................................--.........................---------.--T.C--.---------.....................TA-G-------A...--A------------.--- 56
C.BW.00.00BW5031_1           .............................................TCTCTAGCAG--------.........................---------.--T.---.---------.....................TA-G--------...--A------------.--- 72
C.CN.10.YNFL19               .....................................................T---------.........................---------.--T.---.---------.....................TA-G--------...--A------------.--- 64
C.CY.09.CY260                .......................................................................................................................................................................... 0
C.ES.08.X2363_2              ------G----------------T-AA---T-CTCT-----.---------------------.........................---------.--G.C--.---------.....................TAG--------A...--A------------.--- 129
C.ET.02.02ET_288             .......................................................................................................................................................................... 0
C.IN.09.T125_2139            -------------------------T.---T-GT-------.---------------------.........................---------.--T.---.---------.....................TA-G-------A...--A------------.--- 695
C.KE.00.KER2010              .......................................................................................................................................................................... 0
C.MW.09.703010256_CH256.w96  -------------------------A.---T-GTT-C----.---------------------.........................---------.--T.---.---------.....................TA-G--------...--A------------.--- 696
C.TZ.08.707010457_CH457.w8   -------------------------AA--GT-GT-------.---------------------.........................---------.---.C--.---------.....................TA-G-------A...--A------------.--- 696
C.YE.02.02YE511              .......................................................................................................................................................................... 0
C.ZA.08.705010534_CH534.w12  ----------------------------G-T-GT-------.---------------------.........................---------.---.C--.T--------.....................T-----------...--A------------.--- 697
C.ZA.10.DEMC10ZA001          ...........................................................----.........................---------.--T.---.---------.....................TA-G--------...--A------------.--- 58
C.ZM.11.DEMC11ZM006          ............................................................................................-----.--G.C--.---------.....................T---G-------...--A------------.--- 50
D.CM.10.DEMD10CM009          ......................................................GT.------.........................---------.--T.---.---------.....................TAG---------...---------------.--- 62
D.CY.06.CY163                .......................................................................................................................................................................... 0
D.KE.11.DEMD11KE003          ..............................................................................................---.--T.---.---------.....................TAG--------A...---------------.--- 48
D.KR.04.04KBH8               -------------------------TCG--T--------AA.---------C-----------.........................---------.--T.---.---------.....................TAG---------...-------C-------.--- 653
D.SN.90.SE365                -----------------------------------------.---------------------.........................---------.---.C--.---------.....................TAG---------...---------------.--- 695
D.TZ.01.A280                 .......................................................................................................................................................................... 0
D.UG.10.DEMD10UG004          ................................................................................................................---.....................TAG---------...--A------------.--- 33
D.UG.11.DEMD11UG003          ............................................................................................-----.--T.---.---------.....................TAG---------...---------------.--- 50
D.YE.02.02YE516              .......................................................................................................................................................................... 0
D.ZA.90.R1                   .........................................................................................................TCG--TGG-A.....................C--GTTA--T-A...G-AGTCTCTCG----.--- 40
F1.AO.06.AO_06_ANG32         .......................................................................................................................................................................... 0
F1.AR.02.ARE933              ................................................................................................................................................................TT.---.--- 8
F1.BR.10.10BR_RJ015          -------------------------A.--------------.---------------------.........................---------.--G.C--.---------.....................TAG--------C...--A------------.--- 169
F1.CY.08.CY222               .......................................................................................................................................................................... 0
F1.ES.02.ES_X845_4           -------------------------A----------A---T.---------------------.........................---------.--G.C--A--------A.....................TA---------A...--A------------.--- 154
F1.RO.96.BCI_R07             -------------------------A---------------.---------------------.........................---------.--T.---.---------.....................TAG--------A...--A------------.--- 739
F1.RU.08.D88_845             -------------------------A---------------.---------------------.........................---------.--G.C--.---------.....................TAG---------...--A------------.--- 153
F2.CM.02.02CM_0016BBY        .......................................................................................................................................................................... 0
F2.CM.10.DEMF210CM001        .......................................................................................................................................................................... 0
F2.CM.10.DEMF210CM007        ......................................................GT.------.........................---------.---.---.---------.....................TAG---------...--A------------.--- 62
G.BE.96.DRCBL                -------------------------AC-C---AT-----A..---------------------.........................---------.--T.---.--A------TTAATAGGGACTCGAAAGCGAAAGTT------A...--T------------.--- 682
G.CM.10.DEMG10CM008          ........................................................-------.........................---------.--T.---.---T-----TTAATAGGGACTCGAAAGCGAAAGTT------A...--T------------.--- 82
G.CN.08.GX_2084_08           .......................................................................................................................................................................... 0
G.CU.99.Cu74                 -------------------------AC-C---A-G----A..---------------------.........................---------.--T.---.---------TTAACAGGGACTCGAAAGCGAAAGTT------A...--T------------.--- 310
G.ES.09.X2634_2              -------------------------AC-C----G-----A..---------------------.........................---------.--T.---.---------TTAACAGGGACTCGAAAGCGAAAGTT------A...--T------------.--- 173
G.GH.03.03GH175G             -------------------------AC-C---A------A..---------------------.........................---------.---.C--.---T-----TTAATAGGGACTCGAAAGCGAAAGTT------A...--T------------.--- 739
G.KE.09.DEMG09KE001          .........................................................------.........................---------.--T.---.---------TTAACAGGGACTCGAAAGCGAAAGTT------A...--A------------.--- 81
G.NG.09.09NG_SC62            .......................................................................................................................................................................... 0
G.ZA.01.TV546                .......................................................................................................................................................................... 0
H.BE.93.VI991                ............................................-------------------.........................---------.--T.---.---------.....................CA---------A...---------------.--- 74
H.CF.90.056                  ....................................................................................................T.---.T--T-----.....................TA---------A...--A------------.--- 43
H.GB.00.00GBAC4001           ------W------------------A--Y------------.---------------------.........................---------.--T.---.---------.....................TA---------A...--T------------.--- 208
J.CM.04.04CMU11421           -------------------------.---------------.---------------------.........................---------.---T---.---T-----.....................TAGG-------A...--T------------.--- 212
J.SE.93.SE9280_7887          .........................................................................................................................................................T------------.--- 16
J.SE.94.SE9173_7022          .........................................................................................................................................................T------------.--- 16
K.CD.97.97ZR_EQTB11          .......................................................................................................................................................................... 0
K.CM.96.96CM_MP535           .......................................................................................................................................................................... 0
U.CA.01.TV749                -------------------------AC-C---A-------A.---------------------.........................---------.---.C--.---------TTAATAGGGACTCGAAAACGAAAGTT------A...--T------------.--- 716
U.CA.99.TV721                -------------------------AC-C---A-------A.---------------------.........................---------.---.---.---------TTAATAGGGACTCGAAAACGAAAGTT------A...--T------------.--- 716
U.CD.83.83CD003_Z3           ------T------------------AC-----GT-------.---------------------.........................---------.---.G--.---------.....................TA---------A...--T------------.--- 194
U.CD.90.90CD121E12           .......................................................................................................................................................................... 0
U.CY.05.CY090                .......................................................................................................................................................................... 0
U.CY.08.CY223                .......................................................................................................................................................................... 0
U.ES.10.DEURF10DZ001         .......................................................................................................................................................................... 0
U.GR.99.99GR303              .......................................................................................................................................................................... 0
U.NL.95.U_NL_95_H10986_D1    -------------------------A.--------------.---------------------.........................---------.--T.---.---------.....................TAGG-------A...--T------------.--- 591
01_AE.AF.07.569M             .......................................................................................................................................................................... 0
01_AE.CF.90.90CF11697        -------------------------AC-C---A-T-A--A..---------------------CGAACAGGGACTTGAAAGCGAAAG.TT--T----.--T.C--.---------........................TT------A...--T------------.--- 653
01_AE.CN.10.YNFL03           .....................................................AT--------CGAACAGGGACTTGAAAGTGAAAG.TT--T----.--T.C--.---------........................TT------A...--T------------.--- 85
01_AE.HK.04.HK001            ...............................................................CGAACAGGGACTTGAAAGCGAAAG.TT--T----.--T.C--.---------........................TT------A...--T--------GCGCA--- 76
01_AE.IR.10.10IR.THR48F      .......................................................................................................................................................................... 0
01_AE.JP.x.JRC77AE           -------------------------AC-C---A-T-A--A..---T-----------------CGAACAGGGACTTGAAAACGAAAG.TT--T----.--T.C--.---------........................TT------A...--T------------.--- 715
01_AE.TH.09.AA090a_WG11      .------------------------AC-C---A-T-A--A..---------------------CGAACAGGGACTTGAAAGCGAAAG.TT--T----.--T.C--.---------........................TT------A...--T------------.--- 135
01_AE.TH.90.CM240            -------------------------AC-C---A-T-A--A..----------C----------CGAACAGGGCACTCGAAAGCGAAAGTT--T----.--T.C--.---------........................TT------A...--T------------.--- 264
01_AE.US.05.306163_FL        .......................................................................................................................................................................... 0
01_AE.VN.98.98VNND15         .......................................................................................................................................................................... 0
```

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

Packaging loops end Gag and Gag-Pol start

```
B.FR.83.HXB2                CTCGGCTTGCTG..AAGCGC..GCACGGCAAGAGGCGAGGGGCGGC...GAC....TGGTGAGTACGCCAAAAA........TTTTGACTAGCGGAGGCTAGA.........................AGG.....AGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCG 817
Gag                                                                                                                                                                           _M__G__A__R__A__S__V__L__S__
A1.AU.03.PS1044_Day0        ..............................................................................................................................................------------------A--------- 28
A1.CH.03.HIV_CH_BID_V3538   ..............................................................................................................................................------------------A--------- 28
A1.CY.08.CY236              ..............................................................................................................................................--------------------------T- 28
A1.ES.05.X1608_8            ------------..-G-T--..A---A------------.A-----...---....------------.........CAT---------------------.........................---.....--------------------------------T- 265
A1.KE.11.DEMA111KE001       ------------..-G-T--..A-G-A------------.A-----...---....---------------........T..-----------A---------.........................---.....---------------------------------- 162
A1.RU.11.11RU6950           ------------..-G-T--..A---A------------.A-----...---....-------------T---G.....AT.---------------------.........................---.....---------------------------------- 359
A1.RW.11.DEMA111RW002       ------------..-G-T--..A---A------------.A-----...---....-------------T---------T..---------------------.........................---.....---------------------------------- 171
A1.SN.01.DDI579             ..................................................................................................................................................------------------------ 24
A1.UG.11.DEMA110UG001       ------------..---T--..A-G-A------------.A-----...-G-....-------------TT--------ATT---------------------.........................---.....--------------------------------T- 184
A1.ZA.04.04ZASK162B1        ------------..-G-T--..A---A------------.A-----...---....-------------GGG--.....ATT---------------------.........................---.....---------------------------------- 249
A2.CM.01.01CM_1445MV        ....................................................................................................................................................---------------------- 22
A2.CY.94.94CY017_41         ------------..-G-T--..A-----------------------...---....-------------T--T-.....T..-----------A---------.........................---.....------------------------A--------- 186
B.BR.10.10BR_MG029          ------------..------..------------------------...---....-------------G---T........------------A--------.........................---.....---------------------------------- 282
B.CA.07.502_1191_03         ------------..------..--G---------------------...---....-----------------T........---------------------.........................---...AG---------------------------------- 238
B.CH.08.M2_0803101_NFLG8    ------------..C-----..----A-------------------...---....------------------.....TTT---------------------.........................---.....---------------------------------- 277
B.CN.10.DEMB10CN002         ------------..------..----A-------------------...---....C----------------.........C--------------------.........................---.....---------------------------------- 183
B.ES.10.DEMB10ES002         ------------..------..------------------------...---....-----------------.........---------------------.........................---.....---------------------------------- 185
B.FR.11.DEMB11FR001         .................................................................-CAA---T.........---------------------.........................---.....---------------------------------- 66
B.GB.05.MM45d213_GN1        ------------..------..------------------------...---....C-----------------......T.---------------------.........................---.....---------------------------------- 194
B.HK.06.HK003               ------------..------..--G---------------------...---....C----------------T........---------------------.........................---.....---------------------------------- 175
B.HT.05.05HT_129389         ....................................................................................................................................................---------------------- 22
B.JP.12.DEMB12JP001         -----------A..------..--G---------------------...---....C------------G---....AATT.---------------------.........................---.....-----------------------G---A---AT- 195
B.KR.07.HP_18_07JHS10_3909  ------------..------..----A-------------------...---....------------------.....TT.---------------------.........................---.....-----------------------G---------- 481
B.PE.07.502_0525_wg5        ------------..------..-----------------A------...-G-....------------------...ATTTT---------------------.........................---.....------------------------A--------- 267
B.RU.11.11RU21n             ------------..---T--..------------------------...---....------------------.....TT.---------------------.........................---.....------------------------A--------- 361
B.TH.08.MERLBDTRC10         .............................................................................................................................................----------------------------- 29
B.US.11.ES38                -----------A..------..------------------A-----...---....--------------T-..........---------------------.........................---.....-----------------------G---------- 741
C.AR.01.ARG4006             ....................................................................................................................................................------------A-------A- 22
C.BR.07.DEMC07BR003         ------------..---T--..A----------------.A-----...-G-....---------------TTTTTT.....A--------------------.........................---.....------------------------A-------A- 179
C.BW.00.00BW5031_1          ------------..---T--..A-T--------------.A-----...-G-....---------------......TTTTTA--------------------.........................---.....------------------------A-----C-A- 194
C.CN.10.YNFL19              ------------..---T--..A-T--------------.A-----...---....---------------......TTTT.A--------------------.........................---.....------------------------A-------A- 185
C.CY.09.CY260               ..............................................................................................................................................------------------A-------A- 28
C.ES.08.X2363_2             ------------..-G-T--..A----------------.A-G---...---....---------------......ATTTTAA-------------------.........................---.....------------------------A-------A- 251
C.ET.02.02ET_288            ....................................................................................................................................................------------------C-A- 22
C.IN.09.T125_2139           ------------A.---T--..A-T--------------.A-----...---....---------------......TTTT.A--------------------.........................---.....------------------------A-------A- 817
C.KE.00.KER2010             ....................................................................................................................................................------------A-------A- 22
C.MW.09.703010256_CH256.w96 ------------..---T--..A-T--------------.A-----...---....---------------......TTTTA---------------------.........................---.....------------------------A-------A- 818
C.TZ.08.707010457_CH457.w8  ------------..---T--..A-T--------------.A-----...-G-....---------------......TTTTTA--------------------.........................---.....-----------------T------A-----C-A- 818
C.YE.02.02YE511             ....................................................................................................................................................------------A-------A- 22
C.ZA.08.705010534_CH534.w12 ------------..---T--..A----------------.A-----...---....---------------......ATTTTA--------------------.........................---.....------------------------A-------A- 819
C.ZA.10.DEMC10ZA001         ------------..---T--..A-T---------------------...---....---------------......ATTTTA--------------------.........................---.....------------------------A-------A- 181
C.ZM.11.DEMC11ZM006         ------------..---T--..A-T--------------.A-----...---....---------------......TTTTTA--------------------.........................---.....------------------------A-------A- 172
D.CM.10.DEMD10CM009         ------------...-----..--------------------TA--...---....------------T----------T..-----------A---------.........................---.....---------------------------------- 183
D.CY.06.CY163               ..............................................................................................................................................--------------------------G- 28
D.KE.11.DEMD11KE003         ------------..------..---------------------A--...--A....C-----------TG--..........A--------------------.........................---.....---------------------------------- 167
D.KR.04.04KBH8              --T---------..------..----A---------------TA--...---....------------T----------TT.-----------A---------.........................---.....G--A----------A--------T---------- 776
D.SN.90.SE365               ------------..------..---------------------A--...---....------------TG---T........---------------------.........................---.....---------------------------------- 816
D.TZ.01.A280                ....................................................................................................................................................---------------------- 22
D.UG.10.DEMD10UG004         ------------..------..----A----------------A--...--A....C-----------T----T.....T..---------------------.........................---.....------------------------------C--- 155
D.UG.11.DEMD11UG003         ------------..------..---------------------A--...-GA....C-----------T----T.....T..--G------------------.........................---.....---------------------------------- 172
D.YE.02.02YE516             ....................................................................................................................................................---------------------- 22
D.ZA.90.R1                  ------------..------..--G------------------A--...---....------------T----T.....AA.-------------------A-.........................---.....--------------------------T------- 163
F1.AO.06.AO_06_ANG32        ..............................................................................................................................................---------------------------- 28
F1.AR.02.ARE933             ------------..---T--..A----------------.A-----...---....----------------..........-----------A---------.........................---.....---------------------------------- 126
F1.BR.10.10BR_RJ015         ------------..---T--..A----------------.A-----...---....--------------T.........T.-----------A---------.........................---.....------------------------A--------- 287
F1.CY.08.CY222              ..............................................................................................................................................---------------------------- 28
F1.ES.02.ES_X845_4          ------------..---T--..A----------------.--GA--GGC---....----------------........T.---------------------.........................---.....---------------------------------- 275
F1.RO.96.BCI_R07            ------------..---T--..A----------------.A-----...---....----------------........T.---------------------.........................---.....----.----------------------------- 856
F1.RU.08.D88_845            ------------..-G-T--..A---A------------.A-----...-G-....----------------.......TTT---------------------.........................---.....---------------------------------- 273
F2.CM.02.02CM_0016BBY       ....................................................................................................................................................---------------------- 22
F2.CM.10.DEMF210CM001       ............................................................------C---T--T........---------------------.........................--A.....G-----------------------A--------- 72
F2.CM.10.DEMF210CM007       ------------..---T--..A-G-A------------AA-----...---....-----------------T........-----------A---------.........................---.....---------------------------------- 182
G.BE.96.DRCBL               ------------..-G-T--..A---A------------.A-----...---....-----------------T........---------------------.........................---.....---------------------------------- 802
G.CM.10.DEMG10CM008         -----------A..---T--..A---A------------.A-----...---....-----------------.........-----------A---------.........................---.....--------------------------------T- 201
G.CN.08.GX_2084_08          .....................................................................................--------C.--------.........................---.....---------------------------------- 54
G.CU.99.Cu74                ------------..-G-T--..A---A------------.A-----...---....----------------CT........---------------------.........................---.....--------------------------G------- 430
G.ES.09.X2634_2             ------------A.-G-T--..A-G-A------------.A-----...-G-....--------------T--.........---------------------.........................---.....---------------------------------- 294
G.GH.03.03GH175G            ------------..---T--..A---A------------.A-----...---....----------------TC.....C..---------------------.........................---.....---------------------------------- 860
G.KE.09.DEMG09KE001         ------------..-G-T--..A---A------------.A-----...---....--------------T--T........---------------------.........................---.....------------------------A-----C--- 201
G.NG.09.09NG_SC62           ..................................................................................................................................................--------------------TC-- 24
G.ZA.01.TV546               ....................................................................................................................................................---------------------- 22
H.BE.93.VI991               ------------..---T--..A---A-------------------GT.---....-------------G----TTTTTA..---------------------.........................---.....---------------------------------- 203
H.CF.90.056                 ------------..---T--..A---A------------.A-----...---....--------------T......TTTG.---------------------.........................---.....---------------------------------- 164
H.GB.00.00GBAC4001          ------------..---T--..A-----------------------GC.---....-------------G----ATTTTTGT---------------------.........................---.....---------------------------------- 338
J.CM.04.04CMU11421          ------------..---T--..A---A------------.A-----...---....--------------TTTT.....GTT---------------------.........................---.....---------------------------------- 335
J.SE.93.SE9280_7887         ------------..---T--..A-----------------------...---....------------------.....TAT---------------------.........................---.....------------------------A-------T- 140
J.SE.94.SE9173_7022         ------------..---T--..A---------------A-------...---....------------------...TATT.---------------------.........................---.....------------------------A-------T- 141
K.CD.97.97ZR_EQTB11         ..............................................................................................................................................--------------T------------- 28
K.CM.96.96CM_MP535          ..............................................................................................................................................---------------------------- 28
U.CA.01.TV749               ------------..-G-T--..A---A------------.A-----...---....-------------T---------T..---------------------.........................---.....---------------------------------- 837
U.CA.99.TV721               ------------..-G-T--..A---A------------A-CG-CGAAC.......-------------T--G-.....TT.---------------------.........................---.....---------------------------------- 839
U.CD.83.83CD003_Z3          ------------..---T--..A-----------------------...---....-------------G---T........---------------------.........................---.....---------------------------------- 315
U.CD.90.90CD121E12          .....................................................................................................................................................-A----------------- 19
U.CY.05.CY090               ..............................................................................................................................................---------------------------- 28
U.CY.08.CY223               ..............................................................................................................................................------------------A----G--T- 28
U.ES.10.DEURF10DZ001        ..........--..-G-T--..A---A------------.A-----...-G-....---------------........T..---------------------.........................---.....---------------------------C------ 108
U.GR.99.99GR303             .........................-A------------AC-G-T-GGC---....------------------.....ATT---------------------.........................---.....--------------------------G-----T- 106
U.NL.95.U_NL_95_H10986_D1   ------------..---T--..A-T--------------.A-----...---....-------------G---......AAT---------------------.........................---.....------------------------------C-T- 714
01_AE.AF.07.569M            .............................................................................................................................................---------------------------T- 29
01_AE.CF.90.90CF11697       ------------..-G-T--..T---A------------.A-----...---....---------------........ATT--C------------------.........................---.....--------------------------------T- 772
01_AE.CN.10.YNFL03          ------------..-G-T--..A---A------------.A-----...---....---------------........ATT---------------------.........................---.....------------------------A-------T- 204
01_AE.HK.04.HK001           ------------..-G-T--..A---A------------.A-----...---....---------------........A..---------------------.........................---.....------------------------A------CA- 194
01_AE.IR.10.10IR.THR48F     ........................................................................................................................................--------------------------------T- 34
01_AE.JP.x.JRC77AE          ------------..-G-T--..A---A------------.A-----...---....---------------........ATT---------------------.........................---.....--------------------------------T- 835
01_AE.TH.09.AA090a_WG11     ------------..-G-T--..A---A------------.A-----...---....---------------...........-----------A---------.........................---.....-------------------------------CT- 252
01_AE.TH.90.CM240           ------------..-G-T--..A---A------------.A-----...---....---------------........A..---------------------.........................---.....--------------------------------T- 382
01_AE.US.05.306163_FL       .............................................................................................................................................---------------------------T- 29
01_AE.VN.98.98VNND15        ..........................................----...---....---------------........A..---------------------.........................---.....--------------------------------T- 81
```

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

B.FR.83.HXB2 GGGGAGAATTAGATCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCA 987
Gag G G E L D R W E K I R L R P G G K K K Y K L K H I V W A S R E L E R F A V N P G L L E T S E G C R Q I L G Q L Q P 198
A1.AU.03.PS1044_Day0 A1.CH.03.HIV_CH_BID_V3538 A1.CY.08.CY236 A1.ES.05.X1608_8 A1.KE.11.DEMA11KE001 A1.RU.11.11RU6950 A1.RW.11.DEMA111RW002 A1.SN.01.DDI579 A1.UG.11.DEMA110UG001 A1.ZA.04.04ZASK162B1 A2.CM.01.01CM_1445MV A2.CY.94.94CY017_41 B.BR.10.10BR_MG029 B.CA.07.502_1191_03 B.CH.08.M2_0803101_NFLG8 B.CN.10.DEMB10CN002 B.ES.10.DEMB10ES002 B.FR.11.DEMB11FR001 B.GB.05.MM45d213_GN1 B.HK.06.HK003 B.HT.05.05HT_129389 B.JP.12.DEMB12JP001 B.KR.07.HP_18_07JHS10_3909 B.PE.07.502_0525_wg5 B.RU.11.11RU2ln B.TH.08.MERLBDTRC10 B.US.11.ES38 C.AR.01.ARG4006 C.BR.07.DEMC07BR003 C.BW.00.00BW5031_1 C.CN.10.YNFL19 C.CY.09.CY260 C.ES.08.X2363_2 C.ET.02.02ET_288 C.IN.09.T125_2139 C.KE.00.KER2010 C.MW.09.703010256_CH256.w96 C.TZ.08.707010457_CH457.w8 C.YE.02.02YE511 C.ZA.08.705010534_CH534.w12 C.ZA.10.DEMC10ZA001 C.ZM.11.DEMC11ZM006 D.CM.10.DEMD10CM009 D.CY.06.CY163 D.KE.11.DEMD11KE003 D.KR.04.04KBH8 D.SN.90.SE365 D.TZ.01.A280 D.UG.10.DEMD10UG004 D.UG.11.DEMD11UG003 D.YE.02.02YE516 D.ZA.90.R1 F1.AO.06.AO_06_ANG32 F1.AR.02.ARE933 F1.BR.10.10BR_RJ015 F1.CY.08.CY222 F1.ES.02.ES_X845_4 F1.RO.96.BCI_R07 F1.RU.08.D88_845 F2.CM.02.02CM_0016BBY F2.CM.10.DEMF210CM001 F2.CM.10.DEMF210CM007 G.BE.96.DRCBL G.CM.10.DEMG10CM008 G.CN.08.GX_2084_08 G.CU.99.Cu74 G.ES.09.X2634_2 G.GH.03.03GH175G G.KE.09.DEMG09KE001 G.NG.09.09NG_SC62 G.ZA.01.TV546 H.BE.93.VI991 H.CF.90.056 H.GB.00.00GBAC4001 J.CM.04.04CMU11421 J.SE.93.SE9280_7887 J.SE.94.SE9173_7022 K.CD.97.97ZR_EQTB11 K.CM.96.96CM_MP535 U.CA.01.TV749 U.CA.99.TV721 U.CD.83.83CD003_Z3 U.CD.90.90CD121E12 U.CY.05.CY090 U.CY.08.CY223 U.ES.10.DEURF10DZ001 U.GR.99.99GR303 U.NL.95.U_NL_95_H10986_D1 01_AE.AF.07.569M 01_AE.CF.90.90CF11697 01_AE.CN.10.YNFL03 01_AE.HK.04.HK001 01_AE.IR.10.10IR.THR48F 01_AE.JP.x.JRC77AE 01_AE.TH.09.AA090a_WG11 01_AE.TH.90.CM240 01_AE.US.05.306163_FL 01_AE.VN.98.98VNND15

```
B.FR.83.HXB2                 GGGGAGAATTAGATCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCA  987
Gag                          G__G__E__L__D__R__W__E__K__I__R__L__R__P__G__G__K__K__K__Y__K__L__K__H__I__V__W__A__S__R__E__L__E__R__F__A__V__N__P__G__L__L__E__T__S__E__G__C__R__Q__I__L__G__Q__L__Q__P_
02_AG.CM.10.DE00210CM013     -----A---------A------G-----------C-----------------------G-C----G---C-----------------------G--A---------C----C---------T---------G-------A---CA-------A---A---AT-----T--  354
02_AG.CY.09.CY256            -----A--------GC------------------------------C------G-C-------C-----------------------G--GA-G------C----C--C-----T---------G-------A--CCA----C--A-A-A----T-----T--  198
02_AG.ES.06.P1423            -----A--------GC------G-----------M-----------------------G-C-------T--A---------------T-G---A---------C----C---------T------T--G-------A---CA-------A---A---AT-----T--  440
02_AG.GW.05.CC_0048          -----A--------GC------G-G---------------------------------G----------T----------------A---T-G---A---------C----C---------T------T--G-------A---CAG---C--A---A----T-----T--  343
02_AG.KR.12.12MHR9           -----A--------GCG-----G-----------------------------------G-C-------T-----------------------G--A---------C----C--C-----T---------G-------A---CAG---C-GA-A-A----A----GT--  649
02_AG.LR.x.POC44951          -----A--------GC------G-G----A---------G------------------G-C-------C-------------------T-G---A---------A----C---------T---------G-------G---CA-------A---AG---T-----T--  997
02_AG.NG.09.09NG_SC61        -----A--------GC------G----------G-------A--------C------G-C-------C-----------------------G--A---------C----C---------T---------G-------A---CA---GC--A---A----T-----T--  194
02_AG.NG.x.IBNG              -----A--------GC------G-----------------------------------G-C-------C-----------------------G--A---------C----C---------T---------G---G--A---CA----C--A---A----T-----T--  527
02_AG.SE.94.SE7812           -----A--------GC------G-----------------------------------G-C-------C----------------A---T-G---A---------C----C-------------------G-------A---CA----C--A---A---AT-----T--  374
02_AG.SN.98.98SE_MP1211      -----A--------GC------G-G---------------------------------G-C-------C----------------A---T-G---A---------C----C---------T---------G-------A---CA-------A-A-A----T-G---T--  199
03_AB.RU.97.KAL153_2         -----A--------GC-------------------------------G----------G-A-----CC-----------------------G--A--------GC----C--A---T----------------------A---CA-----------A----T--------  219
04_cpx.CY.94.94CY032_3       -----A--------GC------G-GG--------------------------------G-C-G------C----------------A---T-G---A---------C----C---------T---------G-------A---CA----T--A---A----T-----T--  353
05_DF.BE.x.VI1310            -----A--------GC------------------C-------A---------G---G-C-------T------------------------------T---C---------------T----G---------------A-------A-AA-C---T--------  363
06_cpx.AU.96.BFP90           -----A--------GA------------------C------------------------A-G-----CC-----------------------G--A---------A----C---------T---------G-------A---CA---G---A-A-AG---T-----T-- 1015
07_BC.CN.98.98CN009          -----A--------AA------------A-------------------GC-C----TGC-------CC-----------------------G--A---T---C---------------T----G----------------------A--AA-------------  334
08_BC.CN.97.97CNGX_6F        -----A--------AA------------A-----------------A---C-C----TGC-------CC-----------------------G--A---T---C----C---------T----G----------------AG------A-AAA-------------  194
09_cpx.GH.96.96GH2911        -----A--------GC------G--------------A--------------------G-A-G--G---C-------------------T-G---A--------------G-----A---C---------A---G---G---CA-------T---A----T----G---  198
10_CD.TZ.96.96TZ_BF061       -----A--------GA---------------C--C---------------C------C-------T------------------------C----T---A---------------T----------------------A-------A-A---------------  366
11_cpx.CM.95.95CM_1816       -C---A--------AA--------------------------GA---C-----GGC-G-----C-----------------------G--A---------C----C---A---T---------G-------A---CA-------A---T---AA--------  363
12_BF.AR.99.ARMA159          --------------A-----------------G-----------------------G---------------------------------------------------------------A--------------------C--A----TA-A--------------- 1007
13_cpx.CM.96.96CM_1849       -----A--------AA----------------G-----------------G---C-G-A-G-----C------------------A--G--GA---T---C----C---------T---------A-------T---CA----------A-----T-G------  365
14_BG.ES.05.X1870            -----A--------TCC---------------G-----------------------G-C-------T-----------------------G--A---T---C-C--C---A---T---------G-------G---CA-------A--------T-G------  463
15_01B.TH.99.99TH_MU2079     -----A--------GC------------------C------A---------G-----C-------C----------------A---T-G---A---------A---C--A---T------T--G-------A---CA---------A-A----T---GT--  395
16_A2D.KR.97.97KR004         -----A--------GC------------------C------A---------------G-C-------C------------------------C-T---A-------------T----G--------------------A-------A-A--------G-----  354
17_BF.AR.99.ARMA038          --------------A-------------------C------------G---------G--------------------------------T----C----------A---------A--------G---C-----C--A-G--A-A-----------------  192
18_cpx.CU.99.CU76            -----A--------AA-------------------------A---------------G-C-------T-G-------------------G--GA---T---C-C--C-------------------G-------A---CAG-----A-A-A----T-----T--  304
19_cpx.CU.99.CU7             -----A--------GC-----------------------T--C-----------C-----G-A-G-----T-------------------T-G---A-------A---C--A---T---A---T------T--G-------A---CAG-----A-A-A----T-----T--  213
20_BG.CU.99.Cu103            -----A--------GA------G---------G-----------------------G-A-G-----C-C-----------------------G--A---T---C---C-------T---------GT------T---CA-------A-AAAC---T--------  462
21_A2D.KE.99.KER2003         -----A--------GCT-----------------C------A--------C---C-G-C-------T-----------------------------T---A-------------T----G-------------------A-------A-A--G-----------  192
22_01A1.CM.01.01CM_0001BBY   -----A--------GC------G--------------A--------------C-----G-A-G-----T----------------A---T---CA--------C--G-C-------T---------G-------A---CA-------A-A-A----T--------  192
23_BG.CU.03.CB118            -----A--------GC------G---------G-----------------------G-A-G-----C-C-----------------------G--CA---T---C---C-------T---------G-------T---CA-------A--A-C---T-G------  436
24_BG.ES.08.X2456_2          -----A--------GC------G------A--------------------------G-A-G-----C-C-----------------------G--CA---T---C-C--C-------TC--------G---G---T---CA-------T--AAC--A--G------  447
25_cpx.CM.02.1918LE          -----AG-------GC------G---------G---G-----CG----G---G-----------T----------------A-----G--A---------C----C--A---C---A----C------ATGG------A---CA----------------AT-----T--  192
26_AU.CD.02.02CD_MBTB047     -A---C--------A-------G------A--C---A------------------G-C-------T-----------------------G--AA---TT--A----C---------T---------G-------A---CA----------A-A----T----G--- 1007
27_cpx.FR.04.04CD_FR_KZS     -----A--------GC-------------------------------G---------GGA-G-----C-----------------------G--A---------A----C---------C------G--G---T------TT-G-C------A----------G---  990
28_BF.BR.99.BREPM12609       ----T-----G---A-----------------------------------------G-------------------------------------------T---------------------C---------------------------A-A------------G---  356
29_BF.BR.01.BREPM16704       --------------A-------------------------------T-----------C-------------------------------G-------------------------------A----------------------------ATC-------------  358
31_BC.BR.04.04BR142          -C-A-A--------ACT-------G----A--------------------C------TGA-G-----CT----C-----------------G--A---T--C--G-C-------T-------------G--------A-------A-ACA-------------  444
32_06A1.EE.01.EE0369         -----A--------GA------G------------------------G-------------G-----C---------------------G--A--------GC----C---------T---------G-G-----A---CAG--G---G---A----G--------  637
33_01B.ID.07.JKT189_C        -----A--------GC------------------C--------------------GG--G-----T----------------A---T-G---A---------A----C---------T---------G-------A--CCA---G---A-A-AC---T----GT--  300
34_01B.TH.99.OUR2478P        -----A--------A-------------------C------A---------GT---C-GA-G-----C------------------T----A---------C----C---------T---------G-------A---CA-------A-A-A----T----GT--  192
35_AD.AF.07.169H             -A---A--------AC------G----------------------G-----------G-C-G--G---T-----------------------G--A---T---C----C--A---T---------A-------A--CCA-------A---A----T--------  199
36_cpx.CM.00.00CMNYU830      -----A--------GC------G-----------------------------------G-A-------T----------------A---T----A-----------G-C-------T------T--G---G---A---CA-------A-A-A------------  192
37_cpx.CM.00.00CMNYU926      -----A--------GC-------------------------------C-----G-A-G-----T-----------------------G--A---------C----C---------T---------G-------A--CCAG-----A-A-A----T-----T--  192
38_BF1.UY.03.UY03_3389       -----A--------GC----------------------------T-------------------C--C-----------------------------T----------------A-----T--------------C--A-----A---C---T---------  413
39_BF.BR.04.04BRRJ179        -C--C---------A---------------------------------------G-C-----------------------------------------T-----C-----------------------------------------A-A-C-------------  461
40_BF.BR.05.05BRRJ055        --------------AA---------------A-----A-------G---------C-----------------------AT-----------------------A---T------------A---T----G---G------------------------T----C---  470
42_BF.LU.06.luBF_18_06       --------------A-------------------------------------------------------------------------G-----------A--------------A---------------G----C---------A--------------------  532
43_02G.SA.03.J11223          -----A--------GC------G---------G--R-------------------G-A-G-----Y-----------------------G--A---T--A---C-----T--T------T--G-------A---CA-------A---A----Y-M--RT--  521
44_BF.CL.00.CH80             -----A--------GC---------------A-----A----------------G-A-G--G---C--A---------------------T-----G---------TC--------C-----------C--A-----A-A-----A---------  426
45_cpx.FR.04.04FR_AUK        ----GA--------ATG------------------------------------C-G-----CT-----------------------G--CA-------C----C-------------------G-------A---CAG-----A-A-A----T---------  979
46_BF.BR.07.07BR_FPS625      -----A--------GC--------------------G------------------G-A-G-----C-GA----------------------T--C--G--T------TC--------C------------C-GA-----A-A-C---A------T--  440
47_BF.ES.08.P1942            ----------G---A-------------------------A---------C--------G---------------------------------------TC--------------------------------------------------------------  432
48_01B.MY.07.07MYKT021       -----A--------GC------------------C-----------------------GGC-G-----T----------------A---T----A---------C----C---------T---------G-------A---CA----C--A-A-A----T----GT--  198
49_cpx.GM.03.N26677          -----A--------AA------G------------------------C-----G-C-G-----C-----------------------G--A-------------C---------T---------G-------A---CA-------A-A-A----T--------  422
50_A1D.GB.10.12792           -C---A-------CAC------G--G-------------------GA-----G-C-G-----T-----------------------G--A---T--TC---C---A---T---------G-G-----A---CT-------A---AC---A--------  403
51_01B.SG.11.11SG_HM021      -----A--------AA--------G------------------G-----G-------------------------------------------------C--C-----------------------------C----G------------------  194
52_01B.MY.03.03MYKL018_1     -----A--------GC------------------C---------------------C-GA-G-----T----------------A---T-R---A---------C----C---------K---------G-------A---CA-------A-A-A----T----RT--  304
53_01B.MY.11.11FIR164        -----A--------GC--------G------C--C---------------------GSA---------------------------A---T-G---A---------C----C---------T---------G-G-----A---CA-------A-A-A----A----G---  333
54_01B.MY.09.09MYSB023       -----C----G---A---------G--------------------------------G-C-------------G--------------A---------A---------------------------T--------------A----------A--------------  386
55_01B.CN.10.HNCS102056      -----A--------GCT-----------C-----C-------------------GGC--------C----------------A---T-G--GA-------C----C---------A---------A-------A---CA----C-GA-A-A----T----GT--  303
56_cpx.FR.10.URF5_patient_A  -----A--------GC------G--------------A-------T-------G---G-C---G---C-----------------------G--TA-------C--G-C-------T---------T-G-----A---CAG---C--A-A-A----T--------  338
57_BC.CN.09.09YNLX19sg       -----A--------AA------------A------------------CGC---TGC-------CT-----------------------G--A---T---C----C---AG--T----G---G--------------A----T--A--AA-------------  198
58_01B.MY.09.09MYPR37        -----A--------GC------------------C-----------------------GA-G----------------------A---T----A---T---C----C---------T---------G-------A---CA-------A-A-A----T----GT--  341
59_01B.CN.09.09LNA423        -----A--------GCG-----------------C-----------------------GGA-------T----------------A--AT----A-------------G-C--C----T--------AA------A---CA-------A-A-A---AT----G---  198
60_BC.IT.11.BAV499           -C-A-A--------ACT-------G----AA-----A-----------C-T---TGC-G----CC---C----------------T-G---A---T---C--G-C-------T----G------G-------A-------A--AAC---T--------  974
61_BC.CN.10.JL100010         -----A-G------AA------------A--------------------C-T---TGC-------CC-----------------------G--A---T---C----C---------T----G---------------------A-AAA-------------  435
62_BC.CN.10.YNFL13           -----A--------AA------------A---G-----------------C-T---TGC-------CT-----------------------G--A---T---C----C---------T-------------C------A---C--A--AA-------------  355
63_02A1.RU.10.10RU6637       -----A--------GC-------------------------------------GGC-------C-----------------------G--A---------C----C--------TC--------G-------A--CCA---C--A---A----T-----T--  543
64_BC.CN.09.YNFL31           -----A--------AA------------A--------------------C-T---TGC-------CC-----------------------G--A---T---C----C---------T----------------------A-------A--AA---T--------  356
65_cpx.CN.10.YNFL01          -----A--------GC--------G---------C-----------------------GGA-G-----T----------------A---T-G--GA-------A----C---------T---------G-------A---CA-------A-A-A----T----G---  376
67_01B.CN.11.ANHUI_HF115     -----A--------GC------------------C---------------C-------A-G-----T-------------------T----A--------C----C-------------------G-------A---CA----------A-A----T----GT--  342
68_01B.CN.11.ANHUI_WH73      -----A--------GC------------------C---------------------GGA-G-----T-------------------T----A---------C----C----------------------------A---CA----------A-A----T----GT--  333
70_BF1.BR.10.10BR_PE004      ----------G-----------------C---A---------------C------G---------G--------------------A------------------------C----------A-------------------------------------A-C--------------  434
71_BF1.BR.10.10BR_PE008      -----A--------AA------G----------------------G------------------G---------------------------------------------------------A----G------------------------------------  657
72_BF1.BR.10.10BR_MG002      ----------G---A-------G-G---------------A--G--T--------------------------------A-----------T-------------------A-------C--G----------------A-A-----------G---  495
O.BE.87.ANT70                -AA-CA----G---GC-------C-----A------A------ATCT--A--G------G-C------CT----------------A----G--A--------TG--------AG--AC-------TG----G--TAA-GAGA-GC-G--ACAG---T--G-G--- 1038
O.CM.98.98CMA104             --A-CA----G---GC-------C-----AA-----A------AT-T--A--G------G-C-------T-----------------------G--A------TTG--------AG--AC-------TG-------TAA-GAGA--T-GT-ACA----T--G-G---  470
O.CM.98.98CMABB141           -AA-CA----G---ACG------C-----A--C---A------AT-T--A--G------GGC-------T--------------A----G--A-------TG--------AG--AC-------TG----G---AA-GAG---C-GT-ACAG--AT--G-G---  469
O.CM.98.98CMABB212           --A-TA----G---GC-------------A------A------ATCT--A--G------GGC-------T--A------------AA----G--A-----A-TG----------T--AC-----TATG-------GAA-GAG---C-GT--CAG---T--G-G---  474
O.CM.98.98CMU5337            --A-TA----G---GC-------C-T----AT----A------AT-T--A--G------G-C------CT----------------A----G--A--------TG--------AG--AC-------TG---G---AA-GA-A-GC-GT-ACA----T--G-----  471
O.CM.99.99CMU4122            --A-CA----G---GC-------C-----A------A------ATCT--A--G------G-C-------T-----------------------G--A--------TG--------AG--AC-------TG-C--G---AACGAGA--C-GT-ACAG---T--G-G---  469
O.FR.92.VAU                  --ACTA----G---GC-------C-----A------A------ATCT--A--G-------GC-------T--A-------------G---G--CA---T---TG--------AG--AC-------TG-------TAA-GA----C-GT-ACA----T--G-G---  547
O.GA.11.11Gab6352            --A-TA----G---TC-------C-----A------A------AA-T--A--G------G-C-------T-----------------------G--A--------TG--------AG--AC-------TG-------TAA-GAGA--C-GT-ACA----T--G-G---  198
O.SN.99.99SE_MP1299          --A-CA----G---GC-------C-----A------A------AT-T--A--G----C-G-C-------T-----------------------G--CA--C----TG--------AG--AC-------TG----G---AA-GAGA--C-GT-ACAG---T--G-G--- 1040
O.US.10.LTNP                 -AA-CCG---G---GC-------C-----A------A------ATCT--A--G------G-C-------T----T-----------------G--A--------TG--G-----AG--A--------TG----G--TAA-GAGA-GC-GT-ACAG---T--G-G---  959

N.CM.02.DJO0131              -----A---------A-------GC----TAT--G--A-------------------C-G-A-G-----T-----------------------G--A------TTG---C--A--T--CA-G--C---G-GAGC------GCC--GT---AAAC--AT--G-----  457
N.CM.02.SJGddd               -----A---------A-------TC----TAT--G--A-------------------C-G-C-G-----T-----------------------G--CA------TTG---C--A--T--CA-G--C---G-G-T-------GCC--GT--A-AAAC--AT--G-----  456
N.CM.04.04CM_1015_04         -----A--------AA-------GC----TAT--G--A-------------------C---A-G-----T-----------------------G--A------TTG---C--A--T--CA-G--C---G-GC-C-------AC--GT---AAAC--AT--G-----  459
N.CM.06.U14296               -----A---------A-------GC----TAT--G--A-------------------C-G-A-G-----T-----------------------G--A------TTG---C--A--T--CA-G--C---G-GA-C------GCC--GT----AAAY---T--G-----  457
N.CM.06.U14842               -----A----R----A------GGG----TAT--G--A--------------G--C-G-A-G-----T-----------------------G--A------TTG---C--A--T--CA-G--C---G-G-CC------GCC--GT-R--AAAT--AT--G-----  458
N.FR.11.N1_FR_2011           -----A---------A-------GC----TAC--G--A-------------------C-G-A-G-----T-----------------------G--A------TTG---C--A--T--CA-G--C---G-G-------KGTC--GT---AAAC--AT--G-----  352

P.CM.06.U14788               ----CCG---G---GC--------G----A--C-T--A----A--T-------------GC-------G----------------A----G--A---T---TG---------G--TA-G------G-G-------CGA-A--T--T-A-AG---T--G-----  473
P.FR.09.RBF168               ----CCG---G---GC--------G----A--C-T-A-----AA-T-----------TGC------------------------A----G--A---T---TG--------AG--TA-G------G-----------GA-A--T--T-A-AG---T--G----- 1031

CPZ.CD.06.BF1167             -A-AGA-G--------------TCGG-AA--C-TC-C-----A--C-----------CTTA---G---C----G-----CG---AT--AT--C-G--------TT-G-----C------A-G------G----T--T--C---A-G--CA-TATG--A--CTC----  996
CPZ.CM.05.SIVcpzMT145        -----AG-------GCT-------------A--C----A-----------------C-TGA-G-----T----------------A----G--A------TTG--C----T--AA-G---ATG-------T---GAGA-GT--T--AA-------G-----  556
CPZ.GA.88.GAB1               -----A----------C---------G----C-T----C------GA---G---TGA-G-----T-----------------------G--A-------TG-G-C--C--G--AA-G----GTAAG----A----CTA--T-GT-ACA---AT--G-G--- 1024
CPZ.TZ.06.SIVcpzTAN13        -A-A-A-GC-T---ACT-------GC---A--C-T--A--C--A--C--------G--CTT-A-------C---------CG-A--C-----C-----C----GA-G-----C--T--TA-G----ACAA----------CT-GA----CA-CCTC--------G--C  547
CPZ.US.85.US_Marilyn         -----CG-------GCT--------G------C-T--A--T---------------G---TGA-G-----T----T-----------------G--A---T--TTG---C--A--T--AA-G------G----C------CTC---T--T-AAA----T--G----- 1029

GOR.CM.07.SIVgor2139_287     -----CG---G---GC-------CG----A--C-T-A-----AA-T--------G---TGC-------G-----------------------G--A---T---TG--------AG--TA-G------G-C--G------GA-A--T--T-A--G--A---G--T--  548
GOR.CM.07.SIVgorCP2135con    -----CG---G---GCG------CG-------C-T-A-----AA-T--------G---TGC-------G------T-------------G--A-G--T---TG--------AG--TA-G------G-Y--G------GA-A--T--T-A--G--A---G-----  547
```

```
B.FR.83.HXB2                TCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATA.........GAGGAAGAGCAAAACAAAAGT......................... 1122
Gag                         _S__L__Q__T__G__S__E__E__L__R__S__L__Y__N__T__V__A__T__L__Y__C__V__H__Q__R__I__E__I__K__D__T__K__E__A__L__D__K__I__.__.__.__E__E__E__Q__N__K__S__.__.__.__.__.__.__.__.__.
A1.AU.03.PS1044_Day0        G-T--YA--------W--------A--A--------T-------------------------A------------TG----------------Y---T--A---..................------ATAA----T--G--C.........................  333
A1.CH.03.HIV_CH_BID_V3538   G-T--CA----------------TA--A----------------------------------A-----------G-G-G--G----------------------T--A---.................-----GATA-----T--G--C.........................  333
A1.CY.08.CY236              G-T--CA-T------------------A--------T-------------------------A------------TG-----------------CC---T--A---.................------ATAA----T--G--C.........................  333
A1.ES.05.X1608_8            G-T---A--------A--------------------T----------------Y--------A----------A-TG------------------C---T--A---.................------TTA-----T--G--C.........................  570
A1.KE.11.DEMA11KE001        G-T--CA-C------A----------C--G--------------------------------A--------A----TG------------------C---T--A---.................-----GATA-----T--G--C.........................  467
A1.RU.11.11RU6950           A-T---A----------------T---A----------------------T-----------A-----C------------------------------T--A---.................------ATA-----T--G--C.........................  664
A1.RW.11.DEMA111RW002       G-T--C--------GA--------G---A----------------A--------------C--A-----------CTG---G--------A-----------T--A---.................-----GATA-----G-----C.........................  476
A1.SN.01.DDI579             G-T--CA--------A-------G---------C--------------------------C---------C-----TG-----------------------T--A---.................------ATA-----T--G-AC.........................  329
A1.UG.11.DEMA110UG001       G-T--CGG-------A------G-----A------G-T----------------------C--A-----CA------G------------------C---T--A---.................------ATA-----G--G--C.........................  489
A1.ZA.04.04ZASK162B1        G-T--CA--------A--------A---A---C-----------------------------A------------TG-----------------------T--A---.................------ATA-----T--G--C.........................  554
A2.CM.01.01CM_1445MV        G-T--C-CAGT----A------------A----C----------A----GT------C-A--A-------AA-----G------------------CC---T--AT--.................-----G------------TAC.........................  327
A2.CY.94.94CY017_41         G-T--C--A------A------------A--------------------T-GT------C--G--A--------G----TG------------------C---T--A---.................---------A------..........................  485
B.BR.10.10BR_MG029          G-T------------A--------A---A--------------------C------------A----G-A-----T----G------------------C-----A---G--.................---------------------.........................  587
B.CA.07.502_1191_03         --T-----------------------------C----TC-----------GT----------A---------------------------------------A-------.................--------A----G-------.........................  543
B.CH.08.M2_0803101_NFLG8    --T---A--------------------AG--------C------------------------A-------A-----TG----G--------------------T-------.................------------------T--.........................  582
B.CN.10.DEMB10CN002         ---------------------------A------------------------------------------------TG---G------------------C---T-------.................---------------------.........................  488
B.ES.10.DEMB10ES002         G-T------------A---------------------C--C-------------T-------A---G-G--------G----------------------A---G--.................-----------------G-A.........................  490
B.FR.11.DEMB11FR001         ---------GT----------------------------------A----GT-------------------A------G----------------------A-------.................---------------------.........................  371
B.GB.05.MM45d213_GN1        G------------------------------------------------------A--------------------TG------------------C-----A-------.................--------A--G-----GG--.........................  499
B.HK.06.HK003               G-------------------------A-----------------------GT---T------A-----G------T-----------------------T-------.................--------A------------.........................  480
B.HT.05.05HT_129389         G------------------------------------C------------------------A--C----A------G------------------C---T-G---.................-----G-----------T--.........................  327
B.JP.12.DEMB12JP001         G-T--------------------------------TC---GT------GT---T------A--------------A-------------------G---------.................-------------G-------.........................  500
B.KR.07.HP_18_07JHS10_3909  ---------------------------A--------T-------------------------A----------------------------------------------.................---------------------.........................  786
B.PE.07.502_0525_wg5        ---------A-----G---------A---A------------------------------A-------A------G---G-----------------C----T--AG--.................---------------G-----.........................  572
B.RU.11.11RU21n             ---------------------------A--------T-------------GT-------------------------TG-----------------------T--A---.................---------------------.........................  666
B.TH.08.MERLBDTRC10         --------------G------------A--------T--------G---------------------------------------------A-----C------G---.................--------A----T-------.........................  334
B.US.11.ES38                G-T--------------------G----AG------T-------------------------------G------TG-----------------------A---.................---------------------......................... 1045
C.AR.01.ARG4006             G-T--------------A-G------T-----------------T-----T----------A--GC--AA---------CG--------A-----C--------A---..........A----------G-A------.........................  327
C.BR.07.DEMC07BR003         G-T------------A--A-G------GT-------C----------T--GTT----------A--G---A-------G--CG--------------A--------A---..........A-------A------------.........................  484
C.BW.00.00BW5031_1          G-T---A--------GA----G----------------C-----G-----T----------A--AG--AC-----G--C-----------------C---------G--.................--------A--------CT--.........................  499
C.CN.10.YNFL19              G-T-----A------A--------------------TC--C-----------T----------A--G--G------AG--CG--------A----C---------G--.................-----G--A----------T-.........................  490
C.CY.09.CY260               --T---A--------A----G---------------T--C------------T----------A--GC-G-------G--CG-----------------C----------.................--------A------------.........................  333
C.ES.08.X2363_2             G-T---A--------A----G-------A--------C--C------------T----------A--G---AA-----G--CG-----------------C----------.................--------A--------GC-.........................  556
C.ET.02.02ET_288            G-T---A-A------A----G---------------T--C------------T----------A--G--G------T---CG-----------------C-------G----.................--------A--------GTG.........................  327
C.IN.09.T125_2139           G-T------------A----G-----G---------T--C------------T----------A--G--G------AG--CG--------A----C----------.................-----G--A--------GT--......................... 1122
C.KE.00.KER2010             G-T------------A----G----------------C--C------------T----------A--A-T-A-----AG--CG-----------------C----------.................--------A------------.........................  327
C.MW.09.703010256_CH256.w96 G-T--C---------A----G-----C-A----C------C------------T----------A--A---AT----TG-GCG-----------G--C-------G----.................-----G--A---------T--......................... 1123
C.TZ.08.707010457_CH457.w8  G-T------------A----G----------------C--C--------GTT----------A--A----------AG--CG----------------------G----.................--------A-----------C......................... 1123
C.YE.02.02YE511             G-T------------A---CG-------A---------------C-----G-----T--T--C-----A--A--GAC---CAG-G-----------------C--------A--.................--------A------------.........................  327
C.ZA.08.705010534_CH534.w12 G-T--C---------A----G---------------T--C------------T----------A--G-GGA-----TG--CG--------A-----CC------G----.................--------A------------......................... 1124
C.ZA.10.DEMC10ZA001         G-T------------A----G-------A-------TC--C------------T----------A--G--G-----TG--CG-----------------C----------.................--------A------------.........................  486
C.ZM.11.DEMC11ZM006         G-T--C---------A----G-------------------C------------T----------A--GC--------TG--CGG----------------C----------.................--------A------G-----.........................  477
D.CM.10.DEMD10CM009         G--A--A-------------------G-T-------C----C-----------T----------A--A---A----TG---G-----A--A------------A------.................--------A-----G--G-T-.........................  488
D.CY.06.CY163               G-T---------A-T---------T----------G------G----C--A-------A-----------A--GGC-G-A-----------G--T--A------------A------.................--------A--------C---.........................  333
D.KE.11.DEMD11KE003         G-TG----------------G-------A--------T------------------GG----A--G--G----A--G-----------------C---------------.................--------A----C-------.........................  472
D.KR.04.04KBH8              --TA-A--------------A-----A--------T-------------------------A--G---A--A-TG---G-------------------------A---.................-----G--A------------......................... 1081
D.SN.90.SE365               G-TA--------------------------------------A------------------A--G--A------G-----------------G---A----G.........................--------A-G-----G---......................... 1121
D.TZ.01.A280                G-TA-----------------------A--------T-------------------------AG--A------G-------------------------A--AT--.................------------C-------.........................  317
D.UG.10.DEMD10UG004         --TA--A-------------G-------A------------------------A---------A--GC-G----A--G--C---------------------T------.................--------A----C-------.........................  460
D.UG.11.DEMD11UG003         G-TA----------------G---A---A----------------------C----A--G-------A-AG------T--------------------------------.................--------A----C-------.........................  477
D.YE.02.02YE516             G-TT-GA-------T---------------------T--------------A---------A--G--------G-G---G----------------------T------.................--------A----------T-.........................  327
D.ZA.90.R1                  --TA----------------------------------------------------------A--G-------TG----------C-------------------------.................--------A------------.........................  468
F1.AO.06.AO_06_ANG32        ----------------------G----A-----G-----------T-GT--------A--A---------GA-A-----GT---------------------T---C--.................---A----A--G--T------.........................  333
F1.AR.02.ARE933             ---A------------------G----A--G-------------------GT--------C---A-------------AG---G---------------------G---C--.................--------C-----T------.........................  431
F1.BR.10.10BR_RJ015         ----------------------G-----------T-------A----GT------C-T---A-------A-G----G----G-----A-------------G--AC--.................--------A------------.........................  592
F1.CY.08.CY222              -----------------G----G-----------G-------A---T-GT-------A---A-----G-A------G----G---------------------A---C--.................----C---A----C---G---.........................  333
F1.ES.02.ES_X845_4          G----GA---------------G----A-----G-T---TTG-----GT--------A---A-------A-GA-A------G----------C---------T---C--.................--------A------------.........................  580
F1.RO.96.BCI_R07            ----------------------G-----------G-------A-----GT------------A--------G-----G---G---------------------G---C--.................--------A---CG-------......................... 1161
F1.RU.08.D88_845            -----C-C---------G--G--GA----------------------GT--------A---A------------CCTG-------T-----A-----------G---T--.................---------------------.........................  578
F2.CM.02.02CM_0016BBY       ----------------------GA---A------------C--------GT--------A---A--------A----C-A----G---------------------T---C--..........C--------A---G-----TA-.........................  327
F2.CM.10.DEMF210CM001       ----------------------G----A------------------G------T-GTT--------A---A-----------T------------------------T---C-G..........C--------A------------.........................  377
F2.CM.10.DEMF210CM007       ----------------------G----AG-------T----G------T-GTT--------A---A--------A-----AG----------------------------T---C--..........C--------A---G--------.........................  487
G.BE.96.DRCBL               G-T--C--A------A----G--GA---A----C---T------------------------A--------A------G---G---------A--G---C----GG-AG-G..........--AA-GATA-----G--G---......................... 1107
G.CM.10.DEMG10CM008         --T--C---------A----G-------A--------------------------T----------A----G-----------G--------------A------C---TG-A---..........--AA-GATA---C-G--G---.........................  506
G.CN.08.GX_2084_08          G-T--C---------A----G---A---A--------T----G-------------------A--------A------G------------C-A------A--A-GG-AG-G..........--AA--ATAAR-C-G------.........................  359
G.CU.99.Cu74                G-T--C--A----C-A----G--G----AG--------------------------------A--------A------G-------------A------C---GG-AG-G..........--AA-G-CA-----G------.........................  735
G.ES.09.X2634_2             G-T--C---------A----G--GA---A-------------------------------A--------A----G-------G-------------A------C---TG-AG--..........--AA-GATA-----G---G--.........................  599
G.GH.03.03GH175G            G-T--C---------A----G--G------------TC-------------------C----A-------A-----AG--------------A------C---G--AG-G..........--AA--CTA-----G--T---......................... 1165
G.KE.09.DEMG09KE001         --TA-C--A------A----G--GG---A--------------------------GG-----A----GC-A-------G---G--------A------C---GG-AG--..........--AA-G-CA-----G------.........................  506
G.NG.09.09NG_SC62           ---A-C---------A-------GA---A--------T-------------------------A--------A------G--------------A------C---GG-AG-G..........--AA--ATA-----G------.........................  329
G.ZA.01.TV546               --TY-C-------C-A----R--G----------------------------------A--A----------G--------------A------C---GG-AG-G..........--AA-GATA-----G--C---.........................  327
H.BE.93.VI991               G-T---A--------A-------C---CA--------------A-----GT---------C--A-------A-----TG-G-------------------GG------.................------ATA--G--T--G-AC.........................  508
H.CF.90.056                 G-TA--A--------A------------A--------T---CT------GT--------C--A----G--AA-----TG-G-------------G-------T------.................------ATA-------------.........................  469
H.GB.00.00GBAC4001          G-T------------A--------A---A---G----T---CT------GT---------------A----G-AA-----G--C----T----GT--------A-T---G--..........A----GATA--G-----G-AC.........................  643
J.CM.04.04CMU11421          --TTA---A-----GA--------A---AG-----------C--------------------C--A---G-GGRS-----G----------------R-----------------.................-----CTA-----G--G-AC.........................  640
J.SE.93.SE9280_7887         G-T--C---------A--C-----A---A-------G-----------------------------C--A----G---------A-------------T-----------G------.................------ATT--------G-AC.........................  445
J.SE.94.SE9173_7022         G-T------------A--------A---A---------------------------------C--A----G-----------------------------------------------.................------ATT----------AC.........................  446
K.CD.97.97ZR_EQTB11         ---------A-----C--G-----G---------C---T------------------------G----------T---C--G---GG-------------C--------AC--.................--------A-------G--C-.........................  333
K.CM.96.96CM_MP535          ---A----A--------------GA---A----C-----------A-----GT-------T---A-------A------G--------------------C--------AC--.................--------A------------.........................  333
U.CA.01.TV749               G-T--C---------A------------A--------------------------------A------G-------G-----------------------------GA---.................--------A-----G---T-C......................... 1142
U.CA.99.TV721               G-T--C---------A------------A--------T-------------GT--------------A----------------G---------------------------T-GA---.................--------A-----T---T-C......................... 1144
U.CD.83.83CD003_Z3          G-T-----------------G-------A----A-------CTT-----------------C--A-----------------G-------------A--------T-------.................-------TA---------G-C.........................  620
U.CD.90.90CD121E12          G-TT----------------G-------A----C-----------------------------C--A-------------------------------------A----------A-------.................-------CA---------G--.........................  324
U.CY.05.CY090               G--A--A-C-----------G--G----AG--TC------------------------------A----GG-------G---G-----------------C--------AT--.................--------A------------.........................  333
U.CY.08.CY223               A-T---AG--------------------AG--C--G-TC------A------T-------------A---------------G-G-----T--------------------T--A---.................------CTAA----T--G--C.........................  333
U.ES.10.DEURF10DZ001        A-T--CA---------------------AG-------T-----------------T-------A------------A-AG--GG----------------C-----T--A---.................------AT-A--C-A--G--C.........................  413
U.GR.99.99GR303             G-T--CA---------------------A-------------------------------C--A-----------TG------T-------------------T--A---.................------CTA-----A-G----.........................  411
U.NL.95.U_NL_95_H10986_D1   G-T------------A-------------------------------T------------------A--TGT-A----------------------------C------------.................-----GATT----------AC......................... 1019
01_AE.AF.07.569M            A-T--CA---------------------A---C---------------T--------GG--C--A--C-----------G------------------------T--A---.................--------A-----G--G--C.........................  334
01_AE.CF.90.90CF11697       AAT--CA------------------A--A--------T-----CA-----------GG-----A-----------TG----------------------------T--A---.................-------TA-----T--G-AC......................... 1077
01_AE.CN.10.YNFL03          A-T--CA---------------G-----A---C-G------------T--------GG--C----C--------C--G---T--------------------------T--AT--.................-------T------A-GG--C.........................  509
01_AE.HK.04.HK001           A-T--CA---------------------A--------T--------------------GG--C--A--C-----------G-----------------------T--AT--.................------ATA-----T--G-AC.........................  499
01_AE.IR.10.10IR.THR48F     A-T--CA-------G-------------A--------T--------------------GG--C--A--C-----------G-----------------------T---C--.................-------CA----CT--G--C.........................  339
01_AE.JP.x.JRC77AE          A-T--CA---------------------A----C---T---G---G-T-GT----GG--C-----C-----------G-----------------------T--AT--.................-----G-TA-----G--GGAA......................... 1140
01_AE.TH.09.AA090a_WG11     A-T--CA------------------------------T---G--A-----GT---------C--C--CA-----------------------------C---T--A---.................-------CA-----G--G--C.........................  557
01_AE.TH.90.CM240           A-T--CA---------T-----------A--------T--------------------GG--C--A--C-----------G-----------------------T--A---.................-------TA-----T--G--C.........................  687
01_AE.US.05.306163_FL       A-T--CA------------------------------T---G--A------------GG--C--A--CA-----------G--------------------------AC--.................--.........G-G--C.........................  322
01_AE.VN.98.98VNND15        A-T--CA---------------------A--------T-------A---------------C--A--C-G---------G-----------------------T--A---.................-------TA-----T--G--C.........................  386
```

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546


```
B.FR.83.HXB2                  TCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATA.........GAGGAAGAGCAAAACAAAAGT......................... 1122
Gag                           _S__L__Q__T__G__S__E__E__L__R__S__L__Y__N__T__V__A__T__L__Y__C__V__H__Q__R__I__E__I__K__D__T__K__E__A__L__D__K__I__.__.__._E__E__E__Q__N__K__S__.__.__.__.__.__.__.__.__.
02_AG.CM.10.DE00210CM013      G-T--CGG-------A-----------AG--C----T------A----------T-GG----A-------------C------------------CC----T--AG--.........-------TA-----T--G--CAGGCAAAAGACACAG.......... 504
02_AG.CY.09.CY256             A-T---A-------------------A------G-TC------A----------T-GG--C--A---AGC--A---ACC------------------C----T--A--.........-------TA-----T--G--C......................... 333
02_AG.ES.06.P1423             A-T---AG-------------------A--------------------GT-----C--C--A-------AT-T------G-------------C----T--A--.........--A---ATAA----T--G--C......................... 575
02_AG.GW.05.CC_0048           A-T--CAG-----------------C-A-----------------A---------T-GG----A-------------C---G-------------C----T--AG--.........-------TA-----T--G--CAAACAAAAGGCACAG.......... 493
02_AG.KR.12.12MHR9            G-T--CAG---------------T---A-----------------A---T-----T-GG----A-------------C------------------C----T--AT-G.........-------TA-----T--G--C......................... 784
02_AG.LR.x.POC44951           A-T--CAGT-----------G-------A--------T-------A---------G-GG----A---G---T---C---C-------------C----T--A--.........------AT------T--G--C......................... 1132
02_AG.NG.09.09NG_SC61         A-T---A-------------------A--------T------T---------T-GG----A-------------C---G------------------T--AG--.........--A---AT------G--G--C......................... 329
02_AG.NG.x.IBNG               G-T--CAG-------------------A--------T------A---------T-GG-C----------------C-------------------C----T--A--G.........------ATA-----T--G--C......................... 662
02_AG.SE.94.SE7812            A-T--CAG-------------------A--------T-----CA---------T-GG-C--A-------------C-----------A----C----T--AG--.........-----G-TA-----G--G--C......................... 509
02_AG.SN.98.98SE_MP1211       A-T--CA--------------------A--------T-----------------T-------A----G--A----AG-------------------C----T--AT--.........------ATA---G-G--G--C......................... 334
03_AB.RU.97.KAL153_2          A-T--CA--------------------A------------------------------------A-----C------------------------------------T--A--.........------ATA-----T--G--C......................... 354
04_cpx.CY.94.94CY032_3        A-T--CA-A-------------------------------------------TA--A---------GG--C--A--------A----TG--C-------------------------T--A--.........------ATA----GT--G--C......................... 488
05_DF.BE.x.VI1310             G-TA----------A-----------A--------C------------------------C---G-----------G-G---------------------------A-------.........--------------------AAA....................... 501
06_cpx.AU.96.BFP90            A-T--CA--------------------A--------T-------------------------A-----------A-AG--C---------------CA---T--A--.........------ATA-----G--G--C......................... 1150
07_BC.CN.98.98CN009           G-T-----------A----G---------------TC--C----------T----------A---AC-GA-----TG--CG---------A-----C---------------.........--------A----------T-......................... 469
08_BC.CN.97.97CNGX_6F         G-T-----------A----G---------------TC--C----------T----------A---G--GA-----AG--CG---------A-----C---------------.........--------A-----T----T-......................... 329
09_cpx.GH.96.96GH2911         --T---AG-------------------A-------------------------------------A-------A------G--G---------------G----T--A--.........------ATA---GGT--G--C......................... 333
10_CD.TZ.96.96TZ_BF061        G-TA-------------G---A---A----------------------------------------A---G------A--G--GC----------------------------------.........--------A----C-------......................... 501
11_cpx.CM.95.95CM_1816        G-T--CA-T------A--------------------C---------------------C--A----GCA------G--G---------G-------T--A--.........------ATA-----T--G--C......................... 498
12_BF.AR.99.ARMA159           -----------------------G-----------T-----------GT------------A-------A------G---G--------------------G--C--.........--------A------------......................... 1142
13_cpx.CM.96.96CM_1849        --T--C-------------G--G---A-------------------------------------A--C-------------G-----------A-----C---GG-AG-G.........--AA-G-TA-----T-GC--C......................... 500
14_BG.ES.05.X1870             G-T--C---------A----G--GA-----------T---------------A-----------A----------------------------A-----C---GG-AG-G.........--AA-G-CA--G--G------......................... 598
15_01B.TH.99.99TH_MU2079      A-T--CA--------------------A-----------------------------GG--C--A--AA-C---------G---------------------------T-------.........-------TA-----T--G-AC......................... 530
16_A2D.KR.97.97KR004          --TA--A----------------A---AG----G-T---G------T---------------A-------A-----G--G---------------C---T--AT--.........--------A---------CAC......................... 489
17_BF.AR.99.ARMA038           -----------------------G-----------------A----GT-------A---A-------A-G---TG---G-------------------G---C--.........--------A--G---------......................... 327
18_cpx.CU.99.CU76             --TA-CA-A--------------A---AG----------------------T----------A----GG------TG-------------------C---T--A--.........------ATA-----T--G--C......................... 439
19_cpx.CU.99.CU7              A-T--CA--------------------------T-----------------GG--C--A----GCA------G-------------------------T--AT--.........-------TA--G--T--G--C......................... 348
20_BG.CU.99.Cu103             G-T--CA-A------A----G--G----AG-------T---------------------A-------------G-G-----------A-----C---TG-AG-G.........--AA-GTTA-----G------......................... 597
21_A2D.KE.99.KER2003          --TA-------------------A--------T---------------------A---GC---A----TG------------------------------------.........--------A------------......................... 327
22_01A1.CM.01.01CM_0001BBY    AGT--CA----------------A---A----------------------A-----------------A-----C------TG--------------------------T--A--.........-----GATA-----T--G--C......................... 327
23_BG.CU.03.CB118             G-T--CA-A-----A----G---A---AG------T---G-------------------A-------------G-----------A-----C---TG-AG-G.........--AA-GTTA-----G------......................... 571
24_BG.ES.08.X2456_2           G-T--C--A-----GA----G--GA---AG----------------------------A------A------G---G--------A-----C---TG-AG-G.........--AA-GTTA-----G------......................... 582
25_cpx.CM.02.1918LE           G-T---A--------------------A-----------------------GT-----------A----G--A----TG------------------------T--A--.........-----GATA-----T--G-AC......................... 327
26_AU.CD.02.02CD_MBTB047      G-TG--A----------------G----A--------T-----------GT-----C-----A-----------A-TG--------------------G--T--AG--.........------AT-------............................... 1136
27_cpx.FR.04.04CD_FR_KZS      G-T--C--------------------A-----G-T---CT------------------C--A---------A----TG-----------A-----C---GG-AG-G.........--AA---TA-----G--C--G......................... 1125
28_BF.BR.99.BREPMI2609        -----A------------------------------C-----------GT---------------C----A-------------------------------------------.........-----G---------------......................... 491
29_BF.BR.01.BREPM16704        ---------------------------A-------T-----------GT-----------------------T-----------------------A----------------.........---------------------......................... 493
31_BC.BR.04.04BR142           G-T---A--------A--A-G------GT------------------T-------------A--G--A-------G--G---------------A------A---.........A-------A------------......................... 579
32_06A1.EE.01.EE0369          G-T--CAG------G-------------AG------T---------------------A-----------A-AG--GC---------------C---T--A--.........------ATA-----T--G--C......................... 772
33_01B.ID.07.JKT189_C         A-T--CA-------------------CA-------C-------A---T-------GG----A--C-------------G-------------------------T--A--G.........--A---ACA-----T-----G......................... 435
34_01B.TH.99.OUR2478P         A-T--CA------------G------A--------T------A-------------GG----A--C--G------TG-------------------------T--AT--.........-------TA-----T--G---......................... 327
35_AD.AF.07.169H              G-T--CA----------------------------T--------------------C--A-----CA-----CA----------------------------T--A--.........------ATA-----T-----C......................... 334
36_cpx.CM.00.00CMNYU830       A-T--CA--------------------A--------------------------------A-----C------TG-------------------------T--A--.........---A-GTTA----GT--C--C......................... 327
37_cpx.CM.00.00CMNYU926       A-T--CA--------------------A--------T------------------GG----A--C---C------TG---G-------------------T--AT--.........-------TA-----T................................ 321
38_BF1.UY.03.UY03_3389        -----------------------G---A----------------G---GTT-----C-T--A--------AG---TG---G--T-----------C----A---C--.........--------A-----G------......................... 548
39_BF.BR.04.04BRRJ179         -----------------G-----G-----------------------A-----GT--------A---A-------A-G----G------------------------T---C--.........---------------------......................... 596
40_BF.BR.05.05BRRJ055         ---------------------------A---------------------A----------------------------A--------G------------------------------.........---------------------......................... 605
42_BF.LU.06.luBF_18_06        G--------------------------A--------------------------------------A-------A-----TG---G-------------------------------.........---------------------......................... 667
43_02G.SA.03.J11223           A-T---AG-------------------A--------T------A---M-----T-GG--Y--A---------A-----C--------------------C----T--AG--.........------AT------T--G--C......................... 656
44_BF.CL.00.CH80              -----------------------G-----------G-------------T-GT--------T---A---------G-----G----G-----A--------------G---C--.........--------A------------......................... 561
45_cpx.FR.04.04FR_AUK         AAT--CA----------------A---A-----------------------T-----------A--------------TG-----------A--G---------A---.........------ATA-----T--G--C......................... 1114
46_BF.BR.07.07BR_FPS625       -----------------------G-----------T-----------T-GT--------T---A--------A-G-----G---G-----------------------G---C--.........--------A------------......................... 575
47_BF.ES.08.P1942             G-------------------------------C---T---G---------------------------A---C--TG------------------------T--A---.........---------------------......................... 567
48_01B.MY.07.07MYKT021        A-T--CA-------------------------C---T------A---T-------TG--C-----C--G---------G----G-------A-----C---T--A--T.........------AT------G------......................... 333
49_cpx.GM.03.N26677           G-T--CA-A---------------------------------------------------A---------G-G-------------G--G----T--A--.........------ATA-----G--G--C......................... 557
50_A1D.GB.10.12792            G-T--CAG------A-----------------C---T------------------------C--A-----GC-A-----G-------------A--G--C---T--A--.........-------TA-----T--G--A......................... 538
51_01B.SG.11.11SG_HM021       --T--------------R----------------T--------G--GT------------------------T-----TG---------A------C----T----------.........---------------------......................... 329
52_01B.MY.03.03MYKL018_1      A-T--CA--------------------A---------------C------T--------GG----A--C--G--------TG------------------------T--A--G.........--R----CA-----T--G--C......................... 439
53_01B.MY.11.11FIR164         R-T---A--------------------A------RC----CT------GT----GG----A--C------------G---W--T------------------T--A--G.........-------MA-----G---G--......................... 468
54_01B.MY.09.09MYSB023        G--------------------G-T------------TC-----------GTT-----------A-------A------G-------------------------T--A--.........-------CA-----R--GG--......................... 521
55_01B.CN.10.HNCS102056       A-T--CA----------------------------T------A---T-------GG--C--A--C-----------TG------------------------T--AT--.........-------TA-----G--G--C......................... 438
56_cpx.FR.10.URF5_patient_A   G-T--CA---G-------------------A--------T------A---------T-TG----A----T---A---A-C-------------------C---------G---.........---------------------......................... 473
57_BC.CN.09.09YNLX19sg        G-T-----------A----G-G-T---A--------TC--C-----G----T------------A--GC-GA-----CG--CG---------A-----C------G-------.........---A---A-----------T-......................... 333
58_01B.MY.09.09MYPR37         A-T--CA--------------------A--------T------------------GG--C--A--C--G-----------G-------------------------T-GAC-G.........-------CA-----T---G--......................... 476
59_01B.CN.09.09LNA423         G-T---A------C---------------------T------------------GG--C-----C-----------TG---------------------C----T--A--G.........-------CA---.................................. 324
60_BC.IT.11.BAV499            G-T---------------A-G------GT-------C-----G----T-----T-----------A---G--A--------G--CG---T-----------C--------A---.........A----G--A----GG................................ 1103
61_BC.CN.10.JL100010          G-T-----------A----G----------------T---C-----------T---------------TC-GAA-----TG-GCG---------A-----C----------------.........--------A-------C----......................... 570
62_BC.CN.10.YNFL13            G-T-----------A----G-----G---------T---C---C-------T-----------A---A--G-------TG--C-----------A-----C--G------------.........--------A-----.................................. 484
63_02A1.RU.10.10RU6637        A-T--CA---------------G---A--------T------A----------T-GG--C--A--------------T---CG-----------G--C----T--A--.........-------T------T--G--C......................... 678
64_BC.CN.09.YNFL31            G-T-----------A----G---------------TC--C-----------T-----------A---G--GA------CG--CG---------A--G--C---------G---.........-----G--A----------T-......................... 491
65_cpx.CN.10.YNFL01           A-TG-CA--------------------A--------TC---G-----T-GTT------------A--------------T-----------------C----T--AC--.........--------A-----T--G--C......................... 511
67_01B.CN.11.ANHUI_HF115      A-T--CA--------------------A--------T------A----GT---T-GG----A--C---------C----------------------------T--A--.........-------TA-----T--G--C......................... 477
68_01B.CN.11.ANHUI_WH73       A-T--CA--------------------A--------T---C--------------T-GG----A--C---------------------------------------T--A--.........-------TA--G--A-GG--C......................... 468
70_BF1.BR.10.10BR_PE004       ------------A-----------------------T----------------------------------------TG---------------------------A--A---.........---------------------......................... 569
71_BF1.BR.10.10BR_PE008       G------------------------------C-----------------------------------T-----TG--------------------------------------.........---------------------......................... 792
72_BF1.BR.10.10BR_MG002       -----------------------G---A--------T------A----GT-----------A--------------TG-------------------------G--------.........---------------------......................... 630
O.BE.87.ANT70                 G-T--CA-------G-----CAGC--GCAG---C-C-GG--CG--A---T-GTG---GG-----T--CA-C--ATATA-A--TGG---T--GC--C-G--AA--C-A---T--.........A-----TAATGGGG-GC--GAAG.................... 1176
O.CM.98.98CMA104              G-T--CA-------G------AGC--GCAG---C-C-GG--C---A-----GTG---GG-----T--CA-C--AT-TA-AG-TG-G--T--GC--C-G--AA--C-A---C--.........AGA----AA-GGGGGGC--GAAG.................... 608
O.CM.98.98CMABB141            G-T--CA-------G------AGC--GCAG---C-C-GG--C---A-----GTG---GG-----T--CA-C--AT-TA-AG-TGG---T--GC--C-G--AA--C-A--AT--.........AG----TCATGGGG--C--GAAG.................... 607
O.CM.98.98CMABB212            G-T--AGC------G----GAGCT-G-A---TC-C-GG--CTT-T-----GTA---GG-----T---AGC--A--T--C--TC-G--T--AC--C-G--AG--C-A---G--.........A-A---A-AATGGCA--C--GAAG.................... 612
O.CM.98.98CMU5337             G-T--CA-------G------AGC--GCA---GC-C-GG--C---A-----GTG---GG--C--T--CA-C--ATACA--G-TGG---T--GC--C-G--AA--C-A---T--.........A-----TAATGGGG-GC--GAAG.................... 609
O.CM.99.99CMU4122             G-T--CA-A-----G-----CAGC--GCAG---C-C-GG--C---A---T-GTG---GG-----T--YA-C--ATATAGA--TG----T--AC--C-G--AA--A-A--AT-G.........AG----TAATGGGR-GC-RGAAA.................... 607
O.FR.92.VAU                   G-T--CA-------G------AGC--GTAG---C-C-GG--CG-TA-----GTG---GG-----T---A-C--ATATA-AG-TG----T--AC--C-G--AG--AGA--AT--.........A----G--AATG--G-----AAAG.................... 685
O.GA.11.11Gab6352             G-T--CA-------G-----CAGC--GCAG---C-C-GG--C---A-----GTG---GG-----A--CA-TG-AT-TAGA--T--T-----AC--C-G--AA--C-A--AT--.........A-----TAATG....--C-AAAAG.................... 333
O.SN.99.99SE_MP1299           G-T--CA-------G------AGC--GCAG---C-C-GG--C--TA-----GTG---GG-----T--CA-C--A--TA-AG-TG----T--GC--C-G--AA--C-G--AC--.........A-----TAATGGGG-GC--GAAG.................... 1178
O.US.10.LTNP                  G-T--CA-A-----G------AGC--GCAG---C-C-GG--C---A-----GTG---GG-----T--CA-C--ATATA-A--TGG-----GC--C-G--AA--C-A--AC--.........A-----T-ATGGGG-GC--GAAA.................... 1097

N.CM.02.DJO0131               G-T--CGCA-----G-------G---GC-C--C--------CG-CC-----GTT--T---------C---AGT-AC---C--G--C-CA-T--AC-A--G--G--G---------AAAGAGATAA-A--GA-AA-GG-ACGGCACAAGTCC.................... 607
N.CM.02.SJGddd                G-T--CA-------G-------G---GC-C--T-----------CC-----GTT--T---------C---AGT--A---C--G--C-CA-T--AC-A---------G-----A---.........A-A--GA-A--GG-AC-GCACAAG...................CC 596
N.CM.04.04CM_1015_04          G-T--CA-T-----G-------G---GC-C--C--------Y--CA-----GTT--T---------C---AGT-CA---CC-G--CTCA-T--AC-A--G-----G----------.........A-A--GA-AA-GG-AC-GCACAAGTCT.................. 600
N.CM.06.U14296                G-T--CA-------G-------G---GC-C--C--------C--CC-----GT---T--C-----C---AGT--A---C--G--C--A-T--AC-A---------G---M--R--.........A-AS-GA-A--GG-MC-GCACAAGCCC.................. 598
N.CM.06.U14842                G-T--CA-T-----G-------G---GC-C--T--------CR-CT-----GTT--T---------C---A-T-AA-G-C-----C-TA-T--AC-A---------G---C-C---.........A-A--G-TAA-GG-AC-GCACAAGCCC.................. 599
N.FR.11.N1_FR_2011            G-T--CA-------G-------G---GC-C--T--------C--CT-----GTT--T---------C---AGT--A---TC----C-G--T--AC-A-C------G-----------.........A-A--GA-AA-GG-AC-GCACAAG...................CA 492

P.CM.06.U14788                G-T--CAGA-----C--Y--T-GC--GCAG--TC-T-GG--C--CC---T-GTTY-A-GG-----T--CA----A----------GT-----AC--C-G--CA-TACA--ATGG.........A----G--AATGC-G-----A......................... 608
P.FR.09.RBF168                G-T--CAGA-----T--T--T-GC--GCAG--TC-T-GG--C--CC---T-GTTT-A-GG-----T--CAG---AG-R--T----GT-----AC--C-G--CA-TACA--ATGG.........A----G--AATGC-G-G---A......................... 1166

CPZ.CD.06.BF1167              --AG-AG-A-T---G---CC----A-A-TC--C----T---C--CA-T---GTGT-AC----CA-A---G--G-AGA-A-A--TC-G-----TG-A-C---AG--A-A-TTG-G.........A-A--GA-A-T---A-T--CGAGCAGTGAAACAACAGGAAATGAGGA 1157
CPZ.CM.05.SIVcpzMT145         G-T---A-------G-------G---GC-C--T----T---C--TT-G---GT---T-GG----A---A-G--A-----AG-GG-------AC-AC-G------C-G--AC-G.........A-A-CT-CCGTG---CC-GAA........................ 691
CPZ.GA.88.GAB1                G-T--CA-A-----C-------G---GC-G--C--G-T---C--TC-G---GTA--G-GG--CA-A---AGTGAC--CACTG--G-------AC--A----C----AC--C--.........A--CGGC-T--TGGAG--CAACAGAGC.................. 1165
CPZ.TZ.06.SIVcpzTAN13         --AA-AG-C-TT--C--TCC----A-AGTC--C--G--C--C--T--CTGT--A--A-T------A--CGC-G-AGA-A--G--CGG-----GG-A-----CG-TA-A-TTG-G.........A-A-T-A-A-T--CT-C-GTA........................ 682
CPZ.US.85.US_Marilyn          G-T--AA-A--------G----GC--GC----GC-C-T---C--CT----GGT---T-GG-----C--CAGT---G-GACAG-GG-G--T--GC--C-G--C----TG--AC-G.........A-A--G-TAGTGC-GTTGCAGAAA...................CA 1169

GOR.CM.07.SIVgor2139_287      G-T--CA-A-----G-----T-GCT-GC----TC-G-GG--C--TT-G-T-GTG--G-GG-----C--C-GG--A--T--TG--TC------TC--C----C-----A--AT-T.........A-----TCATGTCA-----AAAG.................... 686
GOR.CM.07.SIVgorCP2135con     G-T--CA-A-----G-----T-GCT-AC----TC-G-GG--C--TT-G-T-GTGT-G-GG-----C--C-GG--A--T--TG--TC---T--TC--C----C-----A--AT-T.........A------CCATGTCA-----AAAG.................... 685
```

Gag p17 Matrix end Gag p24 Capsid start

```
B.FR.83.HXB2                ....AAGAAAAAAGCACAGCAA...............GCAGCAGCTGAC......ACAGGA..........................................CACAGCAAT.........CAGGTCAGCCAAAATTACCCTATAGTGCAGAACATCCAGGGGCAAATGG 1216
Gag                         __.__K__K__K__A__Q__Q__.__.__.__.__.__A__A__A__D__.__.__T__G__.__.__.__.__.__.__.__.__.__.__.__.__.__.__H__S__N__.__.__.__Q__V__S__Q__N__Y__P__I__V__Q__N__I__Q__G__Q__M__
A1.AU.03.PS1044_Day0        ....---C---GGA-T--A--G...............------------......------..........................................A-----GGC.........A----------------------------A--TGCA------------A 427
A1.CH.03.HIV_CH_BID_V3538   ....---C---GGA--------...............---A--A--A--......------..........................................AG-----GC.........AGT-----T--------------------A--TGCA--A---------A 427
A1.CY.08.CY236              ....---C----GA-T--A--G...............-----C------......------.........AACAGCAGCACAGGA..................A------GC.........---------AG------------------A--TGCA--A---------A 442
A1.ES.05.X1608_8            ....---C-G.........................GCA------------......------..........................................A------GC.........A----------R-----------------A--TGCA--A------TG-A 655
A1.KE.11.DEMA11KE001        ....--------GA-------G...............A-----------......------..........................................A--C---GC.........A-T--------------------------A--TGCA--A---------A 561
A1.RU.11.11RU6950           ....---C----GA-C--A--G...............------A---G-......------...................................................................-----T-----------C--------A--TGCA--A---------A 746
A1.RW.11.DEMA111RW002       ....---C----GA--ACA--GGCA............------------......------..........................................A------GC.........A-T--------------------------A--TGCA--A---------- 573
A1.SN.01.DDI579             ....---C----GA-------G...............------A-----......------..........................................A------GA.........--------T--------------------A--TGCA--A---------- 423
A1.UG.11.DEMA110UG001       ....---C----GA----A---...............----A-----G-......------..........................................AC---T...............---------T----------------A--TGCA--A---------- 577
A1.ZA.04.04ZASK162B1        ....---C----GA-------GGCAGCCGCAGCCCAG---------A--......------..........................................A------GC.........--------T--------T-----------A--TGCA--C---------- 663
A2.CM.01.01CM_1445MV        ....---C-G--GA-------G...............----------C-......------..........................................A-T---...............AG---T--G-----T--C--------A--TGCA--A---------- 415
A2.CY.94.94CY017_41         ....---C-G---A-------T...............------------......-----G..........................................A-----...............AG---T--------T--C--------A--TGCA--A---------- 573
B.BR.10.10BR_MG029          ....------------------...............------------......T-----..........................................A------GC.........A-------T--------------------A--TGCA--A---------- 678
B.CA.07.502_1191_03         ....--------G-T-------...............--------CACAGCTGAC------..........................................A------GC.........AC-------------------------------C-T--A---------- 643
B.CH.08.M2_0803101_NFLG8    ....-------GG-A-------...............-----------T......-T-AA-..........................................G---A--GC.........---A-------------T--------A--A---C-T------------- 676
B.CN.10.DEMB10CN002         ....--------G---------...............-A----------......G----G..........................................A------GC...AAGGCCAGC-----------------------A------C-T------------- 588
B.ES.10.DEMB10ES002         ....------G-G-------C-...............-----G--A-C-......------..........................................A------GC.........-----------T----T---------------TC-A------------- 581
B.FR.11.DEMB11FR001         ....------------------...............-----C------......------..........................................A-----GGC.........----C---T---------------------------------------- 465
B.GB.05.MM45d213_GN1        ....--------GA--------...............---A-----A--......------AACCAAGCAGCARCTGCCTCAGGA..................AG-----G-.........--------------------------A---------------------- 617
B.HK.06.HK003               ....-G------G-T-------...............------------......------..........................................A------GC.........--------------------------A--------G------------- 574
B.HT.05.05HT_129389         ....--------G-----A---...............------------......------..........................................A------GC.........--------------------------A--------G------------- 421
B.JP.12.DEMB12JP001         ....------------------...............------------......------..........................................A---A--GC.........---------------------------------C--------------- 594
B.KR.07.HP_18_07JHS10_3909  ....--------G---------...............---A------C-......------..........................................A------GC.........---------------------------------C-T------------- 880
B.PE.07.502_0525_wg5        ....--------G---------...............------A--A--......------..........................................A-G----GC.........--------T-----------------A------C---R------G---- 666
B.RU.11.11RU21n             ....-------GG---------...............------------......------AACAGCAGTCCGGGA...........................A------GC.........-C---T--T-----C-----------A------C-T-----A------- 775
B.TH.08.MERLBDTRC10         ....---C------------C-...............--G---------......------..........................................A---A--GC.........--------------------------A------C-G------------- 428
B.US.11.ES38                ....----------------C-...............---A--------......------..........................................A------GC.........---------------------------------C-T------------- 1139
C.AR.01.ARG4006             ....C--C-----A-------G...............----A---G-CT.....................................................................GA---AGGA......A-------T--------T--------------TC----A---------- 415
C.BR.07.DEMC07BR003         .............AGG-----G...............----A---G-CT.....................................................................GA---AAGA......A-------T--------T--------------TC----A---------- 563
C.BW.00.00BW5031_1          ....C--C-----A-------G...............----A---A-CT.....................................................................G-TGGA.........A--A----T--------T-----------A--TC----A--A------- 584
C.CN.10.YNFL19              ....C--C-----A-----G-G...............---AA--AG-CT.....................................................................GA-G-G.........A-------T-GG-----T-----------A---C-T------------- 575
C.CY.09.CY260               ....C--C-----A--------AAAACACAGCAG...-T--A---G-CT......GGC--G...................................................................A-------T--------T--------------TC-A--A---------- 430
C.ES.08.X2363_2             ....C--C----G---------...............AA-A--CAGC-G........................................................G-TGAT--A.........A--A----T--------T--------A-----TC-A--A---------- 644
C.ET.02.02ET_288            ....C--C-----A-------G...............----G-.............-A-..........................................GCTGA---AGGA......A-------T--------T--------------TC--------------- 415
C.IN.09.T125_2139           ....C--C-----A-------G...............---AA-.............-AG..........................................GCTGA-GGG.........A-------T--------T--------------TC----A---------- 1207
C.KE.00.KER2010             ....C--C-----A-------G...............-G-AA-.............-CG..........................................GCTGATGGA.........A-------T--------T--------A-----TC----A---------- 412
C.MW.09.703010256_CH256.w96 ....C--C-----A-------G...............------.............-CG..........................................GCTGA-GGA.........--------T-G-----TT--------A-----TC-T--A---------- 1208
C.TZ.08.707010457_CH457.w8  ....C--C-----A--------...............--C-A-.............-CG..........................................GCTGA-............A-------T--------T--------------TC-T--A---------- 1205
C.YE.02.02YE511             ....C--C-----A-------G...............---A--.............-CG..........................................G-GGA---A.........A-A--T--T--------T--------------T-----A---------- 412
C.ZA.08.705010534_CH534.w12 ....C--C--GC----.....................................-CG..........................................GCTGA---AGGA......A--A----T--------T--A-----------TC----A---------- 1200
C.ZA.10.DEMC10ZA001         ....C--C-----A-------G...............---AA-.............-CG..........................................ACTGAAGGGAAA......GCA-----T--------T--------------TC----A--A------- 574
C.ZM.11.DEMC11ZM006         ....---C-----A-------G...............----A-.............-CG..........................................GCTGA---GGGG......A-------T--------T-----------A--TC----A---------- 565
D.CM.10.DEMD10CM009         ....-------GG-A-...GC-...............------------......------..........................................A------GC.........---------------------------------T-A------------- 579
D.CY.06.CY163               ....------------------...............------------......------..........................................A------G-.........-----------------T---------------T-A--A---------- 427
D.KE.11.DEMD11KE003         ....--------G---------...............------------......---A--..........................................A------GC.........-----------------T---------------T-A------------- 566
D.KR.04.04KBH8              ....--A---------------...............---A--------......---AA-..........................................A---A--GC.........-----------------T-----------A---C-A------------- 1175
D.SN.90.SE365               ....--------G-T-------...............---A--------......TT----..........................................A---A--GC.........-----------------------------A---C-A------------- 1215
D.TZ.01.A280                ....--------G---------...............---A--------......------..........................................AG-----GC.........-----------------T-----------A---C-A------------- 411
D.UG.10.DEMD10UG004         ....---G------T---A---...............---A--A-----......---A--..........................................A------GC.........--------T--G-----T-----------A---C-A-----A------- 554
D.UG.11.DEMD11UG003         ....--------G---------...............--G---------......---AA-..........................................A------G-.........-----------------T-----------A---C-A--A---------- 568
D.YE.02.02YE516             ....--------G----C----...............------------......------..........................................A------GC.........---------------------------------T-A------------- 421
D.ZA.90.R1                  ....-----G--G---------...............-----G------......------..........................................A------GC.........-----------------T---------------C-A------------- 562
F1.AO.06.AO_06_ANG32        ....C--C-----A--------...............-TGA----A-TG......----C-................................................GC-..............-C---------------------A--A--TT-T-----A------- 418
F1.AR.02.ARE933             ....C--C----GAG-------...............--G---------......-----G.............................................................-----T------------G----A-----TCCT-----A------- 513
F1.BR.10.10BR_RJ015         ....---C----GA--------...............--GA----A-CG......---.............................................GCTGC-GCA.........GG------T--G--------------A-----TC-T------------- 683
F1.CY.08.CY222              ....C--C----GA--------...............--G---------......-A----.............................................................-----T-----------------A--A--TC-T-----A------- 415
F1.ES.02.ES_X845_4          ....C--C-----A--------...............--G-----A-CTGCC...-A----.............................................................-----T-----------------A--A--TC-T-----A------- 665
F1.RO.96.BCI_R07            ....C--C----GA--------...............--G---------......-A--T-.............................................................-----T-----------------A--A--TC-T-----A------- 1243
F1.RU.08.D88_845            ....C--CG---GC----A---...............--------A-CA......----............................................GCAGCAGCAGCTGACCAAGGA-----T--------------------A--TCAT-----A------- 678
F2.CM.02.02CM_0016BBY       ....C--C-----A----A-C-...............-----G-----T......-A---G.............................................................-----T-----------------A-----TC-T------------- 409
F2.CM.10.DEMF210CM001       ....C--C-----A----A---...............------------......-A---G.............................................................-----T-----------------A--A--TC-T--A---------- 459
F2.CM.10.DEMF210CM007       ....C--------A----A--G...............-----G------......-A---G.............................................................-----T-----------------A-----TC-T--A---------- 569
G.BE.96.DRCBL               ....C--C------A-.........................................AAC..........................................AG-----GC.........--A-----T-----------------------TGCA--A---------- 1180
G.CM.10.DEMG10CM008         ....C--C-----A-------G...............-------GG-GT......GA----..........................................A-------C.........--A--T-----------T--------A-----TGCA--A------CCA- 600
G.CN.08.GX_2084_08          ....C--C--GR-------A-G...............------.............-C-..........................................A-T----GC.........--A--T--T--------T-----------A--TGCA--A---------- 444
G.CU.99.Cu74                ....C--C-----A-G------...............------AGG--T......GA----..........................................A---A--GC.........--A-----T--------T--------------TGCA--A------CCA- 829
G.ES.09.X2634_2             ....C--------...-----G...............------ATG--T......GA----..........................................G-T-A--GC.........---AC------------T--------------TGCA--A---------- 690
G.GH.03.03GH175G            ....C--C-----...-----G...............------AAG-CT......GA----..........................................A---A---C.........--A--------------T--------------TGCA--A--A------- 1256
G.KE.09.DEMG09KE001         ....C--C----C-T------G...............-----CAGA---......GG----..........................................A---A--GC.........--A--------------T---G----------TGCA--A---------- 600
G.NG.09.09NG_SC62           ....C--C--G---AA---A-G...............------.......................................................................-A--G-.........--A--T-----------T--------------TGCA--A-----G---- 408
G.ZA.01.TV546               ....C--C-----A-------G...............------ATG-MT......----R-..........................................AR-----GC.........--A--------------T--------------TGCA--A---------- 421
H.BE.93.VI991               ....---C---G-A-------G.........GCCCCA-----------T......-A--A-..........................................A-GGA--GC.........A--A----T--------T--------A-----TGC-------------- 608
H.CF.90.056                 ....C--C-----A--------...............-----------T......-AG-A-..........................................A-AGA---C.........A-------T--------T--------A-----TGCT--A-----G---- 563
H.GB.00.00GBAC4001          ....C--C---C-A-------G...............---A-----RGT......-A--AG..........................................A-GGA--GC.........A--A----T--------T--------A-----TGC---A---------- 737
J.CM.04.04CMU11421          ....---C--C-G------A--...............----A-R-----......-A-A-R..........................................R---R--G-.........--A-----T--------T-----------------G--A------CCA- 734
J.SE.93.SE9280_7887         ....--AC-GC-G------A--...............----A-A-----......-A-AA-..........................................G---A--G-.........--------T--------T--------------TC-G--A------CC-- 539
J.SE.94.SE9173_7022         ....--AC-GC-GA-----A--...............----A-A-----......-A-AA-..........................................G---A--G-.........--------T--------T--------------TC-G--A------CC-- 540
K.CD.97.97ZR_EQTB11         ....C--C----GA--------...............-G-AA-------......-A---G.............................................................--T--T-----------------A-----TC-T------------- 415
K.CM.96.96CM_MP535          ....C--CG---GA----A---...............-A----------......-A---G.............................................................-----T-----------------A-----TC-G------------- 415
U.CA.01.TV749               ................A-A--G...............-------A--C-......------..........................................AGTG--CGC.........G-------T-------T-------A----A--GC----A---------A 1224
U.CA.99.TV721               ................A----G...............---A---G--C-......------..........................................A-TG--CGC.........G-------T---------------A----A--TGC---A---------A 1226
U.CD.83.83CD003_Z3          ....--AC-------T--A--G...............------------......-A----..........................................AG--A--G-.........-----------------T--------A-----T--G--A---------- 714
U.CD.90.90CD121E12          ....---C----------A--G...............------------......-A----..........................................G---A--GC.........A-------T--------T--------A-----T--G--A---------- 418
U.CY.05.CY090               ....---CG---G---------...............-A-A--------......-A-...................................................................G-------T-------T---------A-----TG-T--A---------- 415
U.CY.08.CY223               ....---C-----A--------...............----------CTGAT...------..........................................AGT----G-.........A----------------------------A--TGCA--A---------- 430
U.ES.10.DEURF10DZ001        ....--AC----GA-------G...............----------C-......------..........................................GG-T--.....................----------T------------A--TGCA--A------CC-- 498
U.GR.99.99GR303             ....---C----GA--------...............------------......------..........................................GG--C--GC.........A----------------------------A--TGCA--A---------A 505
U.NL.95.U_NL_95_H10986_D1   ....--------GA--..........................-A-------......------..........................................A-A-C--G-.........--------------------C-----A-G---TC-T------------- 1104
01_AE.AF.07.569M            ....CG-C----GA-------G...............----------G-......------..........................................A------GC.........--A-C------------------------A--TGCA--A---------- 428
01_AE.CF.90.90CF11697       ....---C----GA-------G...............----------G-......------..........................................AG--A--GC.........A-A--------------------------A--TGCA--A---------- 1171
01_AE.CN.10.YNFL03          ....C--C----GA-------G...............------A---G-......------..........................................AG-----GC.........A-A---------------------A-A--A--TGCA--A---------- 603
01_AE.HK.04.HK001           ....C-------GA-------G...............----------G-......------..........................................AG--A--GC.........A-A--------------------------A--TGCA--A--A------A 593
01_AE.IR.10.10IR.THR48F     ....C--C----GA--------...............A---------G-......------..........................................AG-----GC.........A-A--------G-----------------A--TGCA--A---------- 433
01_AE.JP.x.JRC77AE          ....C--C----GA-------G...............----------G-......------..........................................A------GC.........A-A--------------------------A--TGCA--A---------- 1234
01_AE.TH.09.AA090a_WG11     ....C--C----GA-------G...............---------AG-......------..........................................AG-G-T.....................--------------------------A--TGCA--A---------- 645
01_AE.TH.90.CM240           ....C--CG---GA-------G...............----------G-......------..........................................AG-----GC.........A-A--------------------------A--TGCA--A---------- 781
01_AE.US.05.306163_FL       ....C--C----GA-------G...............------------......------..........................................AG-----GC.........A-A------------------G-------A--TGCA--A-----C---- 416
01_AE.VN.98.98VNND15        ....C--C----GA-------G...............----------G-......C-----..........................................AG-----GC.........A-A--------------T-----------A--TGCA--A---------- 480
```

HIV-1 Genomes

```
B.FR.83.HXB2                 TACATCAGGCCATATCACCTAGAACTTTAAATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTGATACCCATGTTTTCAGCATTATCAGAAGGAGCCACCCCACAAGATTTAAACACCATGCTAAACACAGTGGGGGGACATCAAGCAGCCATGCAA 1386
Gag                          V__H__Q__A__I__S__P__R__T__L__N__A__W__V__K__V__V__E__E__K__A__F__S__P__E__V__I__P__M__F__S__A__L__S__E__G__A__T__P__Q__D__L__N__T__M__L__N__T__V__G__G__H__Q__A__A__M__Q_
A1.AU.03.PS1044_Day0         ----Y--AAA-T-G---------------------------------A-------A--------T--T--------------------C--------------------------------------TGTA-----G----T------------C--G-----T------ 597
A1.CH.03.HIV_CH_BID_V3538    ----C---AATT----------------G------------------A-------A--------------------A-----------CA-------------G-----------------------T-TG-----G--T-T------------C--G-----T------ 597
A1.CY.08.CY236               -------AAATT-G--------------------------G------A-------A--------T--------------------------------------------------------------TGTG----------T------------C--G-----T------ 612
A1.ES.05.X1608_8             C---C---AA-T-G--------------------------G--G---A----------------------------------------CA-T--------------------------T--------T-TG----------T---------------G------------ 825
A1.KE.11.DEMA11KE001         --------AA-T-G--------G-----G------------------A-------A------------------A-A-----------C--------------T------------T-T--------T-TG-----G--T-T---------------G-----T------ 731
A1.RU.11.11RU6950            ----C---T----G--------G-----G-----------G------A-------A--------------------A------------A-------------G-------------------------TG-----G----T------------C--G-----T-----G 916
A1.RW.11.DEMA111RW002        --------T--T----------G-----G------------------A----------------------------A-----------C--------------------------------------T-TG-----G----T------------C--G-----T------ 743
A1.SN.01.DDI579              --------T--T----------G-----G-----------G------A-------A--------------------A--------------------------G---A-------------------T-TG---T-G--T-T---------------------T------ 593
A1.UG.11.DEMA110UG001        ----C-----------------G-----G-----------G------A-------A-----------------------------------------G-----T-----A-----------------T-TG-----G--T-T------------C--G-----T------ 747
A1.ZA.04.04ZASK162B1         ----C--AC--C----------G--A--G--------------G--GA-------A----G-----A---------A-----------C--------------G-----------------------T-TG-----G----T---------------G-----T------ 833
A2.CM.01.01CM_1445MV         -G--C-----------G-----G-----G-----------C--------------A--A-----------------A--------------------------------------------------T--T--------------------------------T------ 585
A2.CY.94.94CY017_41          ----C-----T-----------G--G--G-----C-----C--------------A--------------------A-----T------A-------------------------------C-----T--T--------------------------------T------ 743
B.BR.10.10BR_MG029           --T----AT----C--------------------------------------------A-----------------A--------------------------------TC------------------T----------G----------------------T------ 848
B.CA.07.502_1191_03          -----------T-----------------------------------A-------------------------G--A--------------------G-----------------T------------------T----------------------------T------ 813
B.CH.08.M2_0803101_NFLG8     -----------T-------------------------------------------A--------T-----------A---------------------------------------------C-------------------------A--------G------------ 846
B.CN.10.DEMB10CN002          --------C----------------------------------------------------------------G--A---------------------G----------------------------------------------------------------------G 758
B.ES.10.DEMB10ES002          ----C--------G--------------------------------------------------------------A---------------------------------------------C----------------------------------------------- 751
B.FR.11.DEMB11FR001          ----------------------------------------------------------------T-----------A--------------------------------T-----------------T-----------T------------------------------ 635
B.GB.05.MM45d213_GN1         -----------C-------------------------------------------A-----C--------------A-----------------TC-------------------------------------------------------G------------------ 787
B.HK.06.HK003                ----------------------------G-----------------------------------T-----------A---------------------G----------------------------T--------G--------------------------T------ 744
B.HT.05.05HT_129389          -------A--------------------------------------------G-----------T--------G--A------------A-------------------------------------------------T------------------------------ 591
B.JP.12.DEMB12JP001          ----------------------G-----G-----------------------------------------------A-----T--------------------------------------------------------T------------------------------ 764
B.KR.07.HP_18_07JHS10_3909   -------------------------------------------------T--------------T-----------A------------------------------------------------------------------------------------T------ 1050
B.PE.07.502_0525_wg5         -------------------------------------------------------A--------T-----------A--------R---------C--------------------------C----------------------------------------T------ 836
B.RU.11.11RU21n              -----------C-----------------------------------------------------------------------------A-------------------------------C--------------------------A--G--C--G------------ 945
B.TH.08.MERLBDTRC10          --------C-----------------C-------------G-----------------------------------A-----------------Y--------------------------------------------------------------------------- 598
B.US.11.ES38                 -----------T----------------------------------------------------------G-----A------------A-------------------------------------T--------------T--------------------------- 1309
C.AR.01.ARG4006              ----C---C--C-G--------------G--------------G---A----G--A-----------------G--C------------A--------------------------------------------T----T------------------------------ 585
C.BR.07.DEMC07BR003          ----C---C-----G--G----------G-----G--------G--------G--------------------G--A------------A-----C--------------------------------------T----T-----A------------------------ 733
C.BW.00.00BW5031_1           ----C-----T-----------------G------------------A---GG-------------A------GA-A--------A---A---------------------------G----------------T----T-----------------------------G 754
C.CN.10.YNFL19               --------C----------------------------------------G-----A--------T---------------------------------------------------------------------T----T------------------------------ 745
C.CY.09.CY260                ----------------------------G------------------A----G----------G------T--G--AG---G----G--A-----G-------------------CTC----C-----T-A--------T------------------------------ 600
C.ES.08.X2363_2              -C--C-----------C-----------G------------------A----G-----------T--------G--A------------A-----C--------------------------------------T----T--------A--------------------- 814
C.ET.02.02ET_288             ----C---T--C-------------C--G------------------A-------A-----------------GA-A------------A-------G--------------------------------------C--T------------------------------ 585
C.IN.09.T125_2139            ----C---T-------------------G------------------A----G-----------T--------G--A------------A--------------------------------------------T----T-----A------------------------ 1377
C.KE.00.KER2010              ----C------C----------------G------------------A----G--A--------T--------G--A------------A--------------------------------------------T----T-----A-----------------------G 582
C.MW.09.703010256_CH256.w96  ----C---C--T----------------G-----T-----------------G-------G-----A------G--A------------A--------------------------------A-G-----T---T----T------------------------------ 1378
C.TZ.08.707010457_CH457.w8   ----C------C-------G--------G-----C------------A----G--A-----------------G--A-----------CA----C---------------------------------------T----T------------------------------ 1375
C.YE.02.02YE511              ----C--AT-TC----------------G------------------A----G-----------------------A------------A--------------------------------C-----------T----T------------------------------ 582
C.ZA.08.705010534_CH534.w12  ----C---C----------A--G--------------------G--------G-----------------------A--C---------A--------------------------------------------T----T------------------------------ 1370
C.ZA.10.DEMC10ZA001          ----C---T--T-------------C--G------------------A----G--A-----C-----------G--A------------A-------G------------------------------------T----T-----A------------------------ 744
C.ZM.11.DEMC11ZM006          ----C------T----------------G------------------A-------------------------G--A-----------CA-------------------T-----------CC-------A---T----T--------A--------------------G 735
D.CM.10.DEMD10CM009          ----C--A---C-------------C--G--C---------------A----G--C--A--------T--C-----A---------------------------------------GC------G-----T--------------------------------A------ 749
D.CY.06.CY163                ----C--A----------------TC--G-G----------------A----G--A--------------------A-----------------C-----------G-----------C--------------------------------------------------- 597
D.KE.11.DEMD11KE003          ----------------------------G--C---------------A----G-------------A-------A-A---------------------------------------------C----------------TG----------------------------- 736
D.KR.04.04KBH8               ----C---C--C-------------C--G--C----------------------------------AT-----G--A------------A-------------------------------------TT-----------G----------------G------------ 1345
D.SN.90.SE365                ----C--A---C----------G-----G--C--------C------A-----------------------------------------A-----C----------------------G---------------T----------------------------------- 1385
D.TZ.01.A280                 ----C--------------------C--G--C---------------A----------------------------------------------------------------------------------A--------------------------------T------ 581
D.UG.10.DEMD10UG004          ----C-----------G-----------G------------------A----G--A--------------------A-----------------G--------------------------------------------------A-----G-----------T------ 724
D.UG.11.DEMD11UG003          ----C-----------------------G-----C------------A----G-----------------------------------------------------------------------G--------------------------------------------- 738
D.YE.02.02YE516              ----C--A-----------------C--G--C---------------A----G--A--------T-A---------A--------------------------------T-----------C-----T-----------------------------------A------ 591
D.ZA.90.R1                   --------C-------------------G--------------G---A-------A-----C--------------A---------------------------------------------------------------------------------------------G 732
F1.AO.06.AO_06_ANG32         ----------T--------------------------------G--GA----------------T-----------A-----------C-----------------G--------------------------------T------------------------------ 588
F1.AR.02.ARE933              --T-C---T--T----------------------------G--G--G-----------------T-----------A-----------------------------G-----T---------------------T----T------------------------------ 683
F1.BR.10.10BR_RJ015          ----C---T-TC-------------------------------G--GA----------------T-----------A-----------------------G-----G-----T---------------------T----T------------------------------ 853
F1.CY.08.CY222               --------T-T--------------------------------G---A-------A--------T-----------A-----------C---------------------------------C-----------T----T------------------------------ 585
F1.ES.02.ES_X845_4           ----------A--G--------------------------------GA----------------T-----------------------CGT---------------G---------------------G----------------------G------------------ 835
F1.RO.96.BCI_R07             ----------T-----------G--------------------G--GA----------------T-----------A-----------------------------G--------------------T-----------T------------------------------ 1413
F1.RU.08.D88_845             --------T----------------------------------G--GA----G-----------T-----------A--------------------------------------------------T------------------------------------------ 848
F2.CM.02.02CM_0016BBY        ----------T---------------C--------------------A----------------------------C-----------------------------G--------------------------------------A------------------------ 579
F2.CM.10.DEMF210CM001        ----C--------------------------------------G--GA-------------------------G--A------------------------------------------------------------------------------------T------ 629
F2.CM.10.DEMF210CM007        ----C------------------------------------------A----G--A--------------------A-----------------------------G---------------------GT---------T-----------------------T------ 739
G.BE.96.DRCBL                ----C-----------------------G--------------------------A-----C-----T--------A------------A-------------------------------------T-----------T-----------G-----------T------ 1350
G.CM.10.DEMG10CM008          ----C--A---C-------------G--------------G--------------A--------T------------------------A-------------------------------------T-----------------------G-----------T------ 770
G.CN.08.GX_2084_08           ----C---T--T-------C--G------------------------A-------A-----C-----T--------A------------A----------------------------G-----G--T--------------C--------G-----------T------ 614
G.CU.99.Cu74                 ----C-----------------------------------C--------------A-----C--T--T--------A-----------------------------G--------------------T------T----------------G------------------ 999
G.ES.09.X2634_2              -------------------C--------------------------G--------A-----C-----T--------A--------------------------------------G-----------T------T----------A-----G-----------T------ 860
G.GH.03.03GH175G             ----C--AC----------------C-----------------------------A-----C-----T--------A--------------------------G-----------------------T------------------------------------------ 1426
G.KE.09.DEMG09KE001          ----C-----------------------G----------------C---------A-----C--------------A------------GT------------------------------------T-----------------------G-----------T------ 770
G.NG.09.09NG_SC62            ----C--A-------------A-------------------------A----------A--C-----T--------A------------A----------------------------------G--T--------------C--------------------T------ 578
G.ZA.01.TV546                ----C---C-------------G------------------------A----G--A--A--C-----T--------A--------------------------------------------------T-----------------------G-----------T------ 591
H.BE.93.VI991                ----C-----A-----------G--C--------------------------------------T-----------A--------------------------------------------C-----TG----------T------------------------------ 778
H.CF.90.056                  ----C-----------------G--C-----------------------------A--------T-----------A--------------------------------------------C-----TG-T--------T-----------------------------G 733
H.GB.00.00GBAC4001           ----C-----------------G--C--------------------------------------T-----------A--------------------------R-----M-----------CY----T------T-----T----------------G------------ 907
J.CM.04.04CMU11421           ----Y--R---C----------------------------------GR-------R--------Y-----------A-----------------Y--------------------R-----------T-----------T-----R--------C--R-----T------ 904
J.SE.93.SE9280_7887          ----C------C----------------------------------GA-------A--A--------------------------------------------------------G-----------T---------------A-A--------C--------T------ 709
J.SE.94.SE9173_7022          ----C------T----------------------------------GA----------A--C-----------------------------------------------------G---------------------------A-A--------C--------T------ 710
K.CD.97.97ZR_EQTB11          ----C------C----------------------------T------A----------------------------A-----------------------------------T--------------------------------------------------------- 585
K.CM.96.96CM_MP535           ----C------C----------------------------G--G---A----G-----------------------A------------A----------------------T---------C----------------------------------------------- 585
U.CA.01.TV749                C-T-----AG-C--A-------------G-----------G------A-------A---AA------T--------A--------------------------------------------------T--T---A-G--------------------------T------ 1394
U.CA.99.TV721                C-T-C---AG-C--A--------------------------------A-------A---AA------T--------A--------------------------------T-----TC----------T--T---A-G--------A-----------------T------ 1396
U.CD.83.83CD003_Z3           -----------------------------------------------A-------A-----C--------------A---------------------------------------------------------T----------------------------------- 884
U.CD.90.90CD121E12           ----C--AC--------------------------------------A-------A--------------------A---------------------------------------------------------T----------------------------------- 588
U.CY.05.CY090                -----------C-----------------------------------A-------A---T----------------A------------GT---------------------------------G--T--------G--------------G------------------ 585
U.CY.08.CY223                --------T-T--G--------G-----G--------------------------A----G---T--T--------A-----------C------C-------G--G--------G------C----T-TG---T------T------------C--G-----T------ 600
U.ES.10.DEURF10DZ001         -------------T-----A--G--------------------G---A-------A-----------------G--A------------A-------------G-----------------------T-TG-----G----T------------C--G-----T------ 668
U.GR.99.99GR303              --------C-T--G--------G-----G-----------G------A-------A--------T-----------A-----------CA-------------D-----------------------T-TG-----G--T-T-A-T--------C--G-----T------ 675
U.NL.95.U_NL_95_H10986_D1    ----C------T-------------C--------G-----G--------------A-----------T--------A-----------------------------------T-----G--C------------------------------------------------ 1274
01_AE.AF.07.569M             --------C-TT----------------G-----------G--G-----------A----G---T-A---------A---T-------C--------------G-----------------------T-TG--------T-T---A--------C--G-----A------ 598
01_AE.CF.90.90CF11697        --------C--T----------------G--------------------------A----G---T-A---------A-----------C--------------------------------------T-TG----------T---------------G-----A------ 1341
01_AE.CN.10.YNFL03           --------C-TT----------------G-----------G------A------------G---T-A---------A-----------C--------G-----------------------------T-TG--------T-T---------------G-----A------ 773
01_AE.HK.04.HK001            --------T-TT----------------G-----------G--------------A----GC--T-A---------A-----------C--------------G------G-T--------------T-TG--------T-T------------C--G-----A------ 763
01_AE.IR.10.10IR.THR48F      --------T-TC----------G-----G-----------G--------------A----G---T-A---------A-----------C--------------T--------T--------------T-TG--------T-T------------C--G-----G------ 603
01_AE.JP.x.JRC77AE           --------T-TT-----------------------------------A-------A----G---T-A---------A-----------C------C-------G-----------------------T-TG--------T-T------------C--G-----A------ 1404
01_AE.TH.09.AA090a_WG11      --------C-TT----------------G--------------------------A----G---T-AT--------A-----------C--------------------------------------T-TG--------T-T------------C--G-----A------ 815
01_AE.TH.90.CM240            C-------C-TT----------------G-----------G--------------A----G---T-A---------A-----------C--------------G-----------------------T-TG--------T-T------------C--G-----A------ 951
01_AE.US.05.306163_FL        --------C-TT----------------G--------------------------A----G---T-A---------------------C--------------G-----------------------T-TG--------T-T------------C--G-----A------ 586
01_AE.VN.98.98VNND15         --------C-TT----------------G--------------------------A----G---T-A---------A-----------C--------------G-----------------------T-TG--------T-T------------C--G-----A------ 650
```

HIV-1 Genomes

```
B.FR.83.HXB2       ATGTTAAAAGAGACCATCAATGAGGAAGCTGCAGAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAACTACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAAT...AATCCACC 1553
Gag                _M__L__K__E__T__I__N__E__E__A__A__E__W__D__R__V__H__P__V__H__A__G__P__I__A__P__G__Q__M__R__E__P__R__G__S__D__I__A__G__T__T__S__T__L__Q__E__Q__I__G__W__M__T__N___N__P__P
A1.AU.03.PS1044_Day0  ----------C----------A----------------C--GT-A--------A----------C----------A-----------------------G----------C--A-----T--CA---------GGC....C-----  764
A1.CH.03.HIV_CH_BID_V3538  --------G--T-----T----------------------GACA--------A----------CAG--------C----------G----T-------G----A----------A-----G--G----------GC....C-----  764
A1.CY.08.CY236     ----------C----------A----------G-----C--GT--------CCA----------CC---------A----------------------------C--A-----T--CA---------GGC....C-----  779
A1.ES.05.X1608_8   ----------T--T----------------------C--GC-A--------A----------C----------------------G----------Y-C--A-----T--CA---------GC....C-----  992
A1.KE.11.DEMA11KE001  ----------T----------G----------C--TA-A--A--A----------C----------------------A----------C--A----------------------GC....C-----  898
A1.RU.11.11RU6950  -----G-----T----------T----------GC-A--------CA-G--------T--C------G--A----------------------A----------------------GC....C-----  1083
A1.RW.11.DEMA111RW002  -------G--T----------------C--GT-A--------A-C----------G--C----------A--A-----C----------G--------C---TC---A--G-----A---------GC....C-----  910
A1.SN.01.DDI579    ----------T----------A----------C--GA---------A----------C----------------------G----------A----------------------GC..........  760
A1.UG.11.DEMA110UG001  ----------T-----T----------------C--GT-A--------CA--G----------C----------A-----C----------A---A--------ACC--------GGC....C-----  914
A1.ZA.04.04ZASK162B1  ----------T----------ATG--------C--G--A--C----A----------CC---------T----------A---A----A----------G..........  1000
A2.CM.01.01CM_1445MV  --------G--T----------------C--G--A--------A----------C----------------------C-GC....C-----  752
A2.CY.94.94CY017_41  ----------T----------------C--G--A--------A----------C----------T-------C-GC..G------  910
B.BR.10.10BR_MG029  ----------------------------------C--C-------C--A--G--------G--------T----------T----------G---C----------G- 1015
B.CA.07.502_1191_03  ----------A----------------------T-------C----G--------G----------C----------G----------TC---------- 980
B.CH.08.M2_0803101_NFLG8  ----------------------A----------T-A--------A----------C----------C---------- 1013
B.CN.10.DEMB10CN002  ----------A----------T----------T----C--CAA--G----------C---------- 925
B.ES.10.DEMB10ES002  ----------------------------------A---C--G----------C---------- 918
B.FR.11.DEMB11FR001  ----------------------------------T---C----------C----------A----------G--C----------G-------G---------- 802
B.GB.05.MM45d213_GN1  ----------------------------------T----------G----------G----------C---------- 954
B.HK.06.HK003      ----------T----------C---T---------A--C--------CG----------A-------G----------C----------C----------C---------- 911
B.HT.05.05HT_129389  ----------------------------------A------A----------G----------C----------C----------G---------- 758
B.JP.12.DEMB12JP001  ----------A--T----------C---T----C----------ACC----------C----------C---------- 931
B.KR.07.HP_18_07JHS10_3909  ----------------------------------T----------C---------- 1217
B.PE.07.502_0525_wg5  ----------------------------------T---------- 1003
B.RU.11.11RU21n    ----------A----T----------T----------G----------T---------- 1112
B.TH.08.MERLBDTRC10  ----------T--A----A----------A----------G----------C----C----------C---------- 765
B.US.11.ES38       ----------A----------T--------CA--G----------C----------C---------- 1476
C.AR.01.ARG4006    ----------T----------G-------G----C---T-A-------R----------G-C----------A----------C---A----------C----------C---------- 752
C.BR.07.DEMC07BR003  ----------T----------A-G----------GT-A--------A----------G-C----------AG----------C----------GG----------C----- 900
C.BW.00.00BW5031_1  ----------T--T-------A--G--C----------C--GT-A--------A-C--------G-G----T----A----------C----------G---------- 921
C.CN.10.YNFL19     --A--------TGT--------A--G----------T-A--------A----------T--A----------C----------GGG----------C----- 912
C.CY.09.CY260      ----------T----------G----------GT-A--------A----------G---C----------C---------- 767
C.ES.08.X2363_2    ----------C----------G----------GT-A--C----C----------T--A----------T----------C-G----------C----------G----------C----- 981
C.ET.02.02ET_288   ----------T----------A-G----------GA-A--------A----------CA--------A----------C----------G---------- 752
C.IN.09.T125_2139  ---C------T----------A--G----------T-A--------A----------A-GT----G----------C----------GGG----------C----- 1544
C.KE.00.KER2010    ----------T----------G----------T-A----------G----A-C--------A----------C----------C---------- 749
C.MW.09.703010256_CH256.w96  ----------T----------G------G----------T-A--------A--G----------G----------A-A--G----------C----------G----------C----- 1545
C.TZ.08.707010457_CH457.w8  ----------C----------G----------A----------A-C--------G-C----------AT----------C----------G---C----------GCC----C---G- 1542
C.YE.02.02YE511    ---C------T----------G----------T-A--------A--C--------G------T----A--------G----------G----------C----------GG----------C----- 749
C.ZA.08.705010534_CH534.w12  ----------T----------T-G----------C-A--------C--------A--G-C----------C----------T----------C----------C---------- 1537
C.ZA.10.DEMC10ZA001  ----------T----------G----------A--------A----------G----------T-A----G--G--T-A----------T----------AC----------G----------C----- 911
C.ZM.11.DEMC11ZM006  ----------T----------G----------T-A--------A--G----------G----G-------AG----------C----------G----------C----- 902
D.CM.10.DEMD10CM009  ----------CT----------------C--GT-A----------G----------T--C------C----------G----------C--------T-G---------- 916
D.CY.06.CY163      ----------T----------C--GC-A----------T--C----------AC----------T----------AC----------G---------- 764
D.KE.11.DEMD11KE003  ----------T----------GC-A----------G----CACC----------A----------C----------C--T-G----------C---------- 903
D.KR.04.04KBH8     ----------T----------GC----------C----------T--AC----------T----------C----------G----------C----- 1512
D.SN.90.SE365      ----------A----------T-A----------T----------T---------- 1552
D.TZ.01.A280       ----------GC-A----------A----------T----------C---------- 748
D.UG.10.DEMD10UG004  ----------GT-A----------A--G----------A----------T----------G----------C---------- 891
D.UG.11.DEMD11UG003  ----------C-----T-------G----------GC-A--------A--G----------G----------A----------T----------G----------GC---------- 905
D.YE.02.02YE516    ----------G----------C--GC-A------CA--A----------A----------T--T-T----C----------G----------A----------GC---------- 758
D.ZA.90.R1         ---C----------GC-A----------T----------T---------- 899
F1.AO.06.AO_06_ANG32  ----------C----------G----C-----A----CC---G----------C----------A--G----------T----------G---CA----------GC....C----- 755
F1.AR.02.ARE933    ----------C----------C---T-A------C----------CC---C----------G----T----------T----T----------T-C----------CA----------GC....C----- 850
F1.BR.10.10BR_RJ015  ----------C----------C--GA-A----CCA--G----A----C------T----A--G----G----------T----G----------CA----------GC....C----- 1020
F1.CY.08.CY222     ----------C----------G----T---C---T-A--------A----------C----------G----T----------T----------G--CA----------GC....C----- 752
F1.ES.02.ES_X845_4  ----------C----------A----------C---T-A----CA--G----------T-CC----------C-A--G----T----------T----------G--CA----------GC....C----- 1002
F1.RO.96.BCI_R07   ----------C----------C---T-A----------C----------C----------G----T----------T----------CA----------GC...CC----- 1580
F1.RU.08.D88_845   ----------T----------G----------C-----A--------A----------C----------T----------T----------CA----------GC....C----- 1015
F2.CM.02.02CM_0016BBY  ----------T----------C--GT-A----------G----A----CC-----T----A----------T----------T----------C----------GC....C----- 746
F2.CM.10.DEMF210CM001  ----------C-----T----------C----------A--------A----CC-----T----------G----T----------T----------C----------CG----------GC....C----- 796
F2.CM.10.DEMF210CM007  ---------T----------C--G--A----------G----------CC-----T----A--G----C----G----T----------A----G----CC----------GC....C----- 906
G.BE.96.DRCBL      -------G----T-------T----------C--GC-A----CA--G----------A--G--C----C----------T----------G----------G---A----------C-GC....C----- 1517
G.CM.10.DEMG10CM008  ----------T--T----------G----C--GACA-----CAA--G----------A----------C----T----------C-AT--G--A-----G---C----------TGGC....C----- 937
G.CN.08.GX_2084_08  ---C---G--T--T-------A---------G-----GC-A----CCA--G--------T--C----------A-----C----------C-A--G----------C----------C-G-....C----- 781
G.CU.99.Cu74       ---C---G--T--T----------G----C--GC-A--C--CC---G----------CC----------A----------T----------A----------T--C---------- 1166
G.ES.09.X2634_2    ---C---G--TG-T----------G----G------GACA-----CCA--G----------CC----------A----------T----------G----------C-GC....C----- 1027
G.GH.03.03GH175G   ------G--T--T----------G----C--GT-A----CA---G----------A----T---CA----T----------A-G----G----A-A----C-G-....C----- 1593
G.KE.09.DEMG09KE001  ---C------T--T----------G----C--GA-----CAA--G----------A----------T----------G----------CA----C-GC....C----- 937
G.NG.09.09NG_SC62  ---C------T--T-----T----------G----C--GC-A----CAA--G----------A----------T----------C----------AC----T--C---C----- 745
G.ZA.01.TV546      -------G--T--T--Y----------G-------GA-A----CCA--G----------C----------A----------T----------Y----------G----------A----------Y-GC....C----- 758
H.BE.93.VI991      ----------T--A----------------GC-A--------A----------C----------T----------G---C----------GGC..........C-- 945
H.CF.90.056        ----------T--A----------C--G--A----------C----------A----------C--T----------G----------C----------GGC..........G- 900
H.GB.00.00GBAC4001  --R--------T--A----------------GACA----------C----------C--T----------Y-A----------C----------RGY..........C-- 1074
J.CM.04.04CMU11421  ----------T--T----------C----------GK-A--------A----------CCR----------GC---------G----------T----------C----------C----------GGC....C----- 1071
J.SE.93.SE9280_7887  ----------T--T----------C--G--A--------A----------G----------T----------C----------GGC..........  876
J.SE.94.SE9173_7022  ----------T--T----------C--G--A--------A----------G----------T----------A--C----------GGC....C-----  877
K.CD.97.97ZR_EQTB11  ----------T--T----------C--GA----C--------A----------CC---------A--A--------T----G----T----------C--T----------AC----------GC....C-----  752
K.CM.96.96CM_MP535  ----------T--T-------T----------C--GT-A--C----------C----------A----------G---G----------C--C----------C----------GC....C-----  752
U.CA.01.TV749      C-A--------T----------C--GT-A--------A----------C----------C-------T----------G---C----------CGG..........  1561
U.CA.99.TV721      --A--------C----------C--GA----------G---G----------C----------G----------C----------C-GC..........  1563
U.CD.83.83CD003_Z3  ----------T--T--A----------T----C--GT-A----------C----------G----------T----------C----------GC..........  1051
U.CD.90.90CD121E12  ----------T--T--A----------T----C--GT-A----------C----------T----------C----------GC..........  755
U.CY.05.CY090      --------T----T----------C--G--A----------C----------T----G----T----------C--T----------G--G---C----------C----C-----  752
U.CY.08.CY223      -------G----T-----T----------C--GT-A----A----------G----------C----------G----------C----------GC..........  767
U.ES.10.DEURF10DZ001  ----------T----------C--G--A--C----A--A----------C------A----A--G----G--A----------C----------A--G--G---AA----------C..........  835
U.GR.99.99GR303    ----------T----------A----------GT-A----ACA--A----------T--C----------C----------A----------C----------GC..........  842
U.NL.95.U_NL_95_H10986_D1  ----------T--T----------C--G--A--C----A----------CC---------T----G----T----------C----------C----------GC....C-----  1441
01_AE.AF.07.569M   ----------C-------C----------GACA--------A----------C----------A----G----------A----------A----------C...-C---T--  765
01_AE.CF.90.90CF11697  ----------T----------GC-A--C----A----------C----------S----------C-A----A----------GC....C-----  1508
01_AE.CN.10.YNFL03  ----------A----------GT-A--C----A----------C----------G----------A----------GC..........  940
01_AE.HK.04.HK001  ----------A----------GT-A--C----A----------C----------A----G----------C----------A----------GC..........Y  930
01_AE.IR.10.10IR.THR48F  ----------A----------G--A--C----A--G----------C----------G-A--G----------T----------A----------C----------C..........  770
01_AE.JP.x.JRC77AE  ----------A--T-------T-------G--C-------GC-A--C----CA--A--------T--C----------G----------A----------GC..........  1571
01_AE.TH.09.AA090a_WG11  ----------A----------GT-A--C----A----------C----------A----G----------A-----A----------C----------GGC..........A-  982
01_AE.TH.90.CM240  ----------A----------C----------G--A--C----A----------C----------G----------C----------C..........  1118
01_AE.US.05.306163_FL  ----------A----------GA-A--C----------CA----------C-A--G----------TA----A----------C..........  753
01_AE.VN.98.98VNND15  ----------A----------G--A--C----A----------C----------G----------A-----A----------C..........  817
```

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

B.FR.83.HXB2 TATC.....CCAGTAGGAGAAATTTATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAATGTATAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAACTCTAAGAGCCGAGCAAG 1714
Gag __I_._._._P_V_G_E_I_Y_K_R_W_I_I_L_G_L_N_K_I_V_R_M_Y_S_P_T_S_I_L_D_I_R_Q_G_P_K_E_P_F_R_D_Y_V_D_R_F_Y_K_T_L_R_A_E_Q__
A1.AU.03.PS1044_Day0 925
A1.CH.03.HIV_CH_BID_V3538 925
A1.CY.08.CY236 940
A1.ES.05.X1608_8 1153
A1.KE.11.DEMA11KE001 1059
A1.RU.11.11RU6950 1244
A1.RW.11.DEMA111RW002 1071
A1.SN.01.DDI579 921
A1.UG.11.DEMA110UG001 1075
A1.ZA.04.04ZASK162B1 1161
A2.CM.01.01CM_1445MV 913
A2.CY.94.94CY017_41 1071
B.BR.10.10BR_MG029 1176
B.CA.07.502_1191_03 1141
B.CH.08.M2_0803101_NFLG8 1174
B.CN.10.DEMB10CN002 1086
B.ES.10.DEMB10ES002 1079
B.FR.11.DEMB11FR001 963
B.GB.05.MM45d213_GN1 1115
B.HK.06.HK003 1072
B.HT.05.05HT_129389 919
B.JP.12.DEMB12JP001 1092
B.KR.07.HP_18_07JHS10_3909 1378
B.PE.07.502_0525_wg5 1164
B.RU.11.11RU21n 1273
B.TH.08.MERLBDTRC10 926
B.US.11.ES38 1637
C.AR.01.ARG4006 913
C.BR.07.DEMC07BR003 1061
C.BW.00.00BW5031_1 1082
C.CN.10.YNFL19 1073
C.CY.09.CY260 928
C.ES.08.X2363_2 1142
C.ET.02.02ET_288 913
C.IN.09.T125_2139 1705
C.KE.00.KER2010 910
C.MW.09.703010256_CH256.w96 1706
C.TZ.08.707010457_CH457.w8 1703
C.YE.02.02YE511 910
C.ZA.08.705010534_CH534.w12 1698
C.ZA.10.DEMC10ZA001 1072
C.ZM.11.DEMC11ZM006 1063
D.CM.10.DEMD10CM009 1077
D.CY.06.CY163 925
D.KE.11.DEMD11KE003 1064
D.KR.04.04KBH8 1673
D.SN.90.SE365 1713
D.TZ.01.A280 909
D.UG.10.DEMD10UG004 1052
D.UG.11.DEMD11UG003 1066
D.YE.02.02YE516 919
D.ZA.90.R1 1060
F1.AO.06.AO_06_ANG32 916
F1.AR.02.ARE933 1011
F1.BR.10.10BR_RJ015 1181
F1.CY.08.CY222 913
F1.ES.02.ES_X845_4 1163
F1.RO.96.BCI_R07 1741
F1.RU.08.D88_845 1176
F2.CM.02.02CM_0016BBY 907
F2.CM.10.DEMF210CM001 957
F2.CM.10.DEMF210CM007 1067
G.BE.96.DRCBL 1678
G.CM.10.DEMG10CM008 1098
G.CN.08.GX_2084_08 942
G.CU.99.Cu74 1327
G.ES.09.X2634_2 1188
G.GH.03.03GH175G 1754
G.KE.09.DEMG09KE001 1098
G.NG.09.09NG_SC62 906
G.ZA.01.TV546 919
H.BE.93.VI991 1106
H.CF.90.056 1061
H.GB.00.00GBAC4001 1235
J.CM.04.04CMU11421 1232
J.SE.93.SE9280_7887 1037
J.SE.94.SE9173_7022 1038
K.CD.97.97ZR_EQTB11 913
K.CM.96.96CM_MP535 913
U.CA.01.TV749 1722
U.CA.99.TV721 1724
U.CD.83.83CD003_Z3 1212
U.CD.90.90CD121E12 916
U.CY.05.CY090 913
U.CY.08.CY223 928
U.ES.10.DEURF10DZ001 996
U.GR.99.99GR303 997
U.NL.95.U_NL_95_H10986_D1 1602
01_AE.AF.07.569M 926
01_AE.CF.90.90CF11697 1669
01_AE.CN.10.YNFL03 1101
01_AE.HK.04.HK001 1091
01_AE.IR.10.10IR.THR48F 931
01_AE.JP.x.JRC77AE 1732
01_AE.TH.09.AA090a_WG11 1143
01_AE.TH.90.CM240 1279
01_AE.US.05.306163_FL 914
01_AE.VN.98.98VNND15 978

HIV-1 Genomes

```
                                                                                                                                                        Gag p24 Capsid end  Gag p2 start
B.FR.83.HXB2                CTTCACAGGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCGAACCCAGATTGTAAGACTATTTTAAAAGCATTGGGACCAGCGGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGGCAAGAGTTTTGGCTGAA 1884
Gag                         A__S__Q__E__V__K__N__W__M__T__E__T__L__L__V__Q__N__A__N__P__D__C__K__T__I__L__K__A__L__G__P__A__A__T__L__E__E__M__M__T__A__C__Q__G__V__G__G__P__G__H__K__A__R__V__L__A__E_
A1.AU.03.PS1044_Day0        --A---------------G------------G--A--AC-------------A--T------------T-C---C----------A---T---GA------T----------------------C--------G---------A-------------G------------ 1095
A1.CH.03.HIV_CH_BID_V3538   --A----A--T---------------------------C----T-----------T-----------AT---------G------A---G---G-------T-------------------------------G-----------------A-----G-----------G 1095
A1.CY.08.CY236              --A----------------------------T-----AC-------------A--T--------------C--------------A---T---GA---T--T----G-----------------C--------G----------C------A-----G-----------G 1110
A1.ES.05.X1608_8            ----------T-----G-C---------------A--AC--A-T--------A--T------------T-C-----G-G------A---G---G-------T----------------------C--------G---------A-------A-----G--C--A-----G 1323
A1.KE.11.DEMA11KE001        --A-C------------GG---------------A--AC--A----------A--T------------T-C--C----G---CC-A---A---G-------T----------------------C--------G-----------------A-----------------G 1229
A1.RU.11.11RU6950           --A-G-----T-----G--C-----------C---C--C-A-----------A--T-----------A--C-------G------A---T---G-------T----------------------C--------G-----------------A-----G--------C--G 1414
A1.RW.11.DEMA111RW002       -------A---------GGC-----------C------C--A----------A--T-----C------T-C-------G------A-------G----T--T----------------------C--------G---------A-------A-----G-----A-----G 1241
A1.SN.01.DDI579             --A-------T-----G--C------------------C----------------T-----C------T-C-------G------A---A---G-------T----------------------C--------G-----G--------C--A-----G-----------G 1091
A1.UG.11.DEMA110UG001       -CA--------------GG---------------TC-AC--A----------A--T-----------AT-C-------G------A---G---G-------T----------------------C--A-----G---------A-------A-----G-----A-----G 1245
A1.ZA.04.04ZASK162B1        --A----A--A-----G--C----------G---T--AC----------------T-----C------CAG--------------A---A---G----T--T--------G-------------C--------G---------A-T-----A-----G---C-------G 1331
A2.CM.01.01CM_1445MV        --A-------T--------------------C------C----------------------------A-----CC-G-G---G--A---------------T-------------------------------G-----------------A-----G--C--------- 1083
A2.CY.94.94CY017_41         -CA----------------C--------G--C------C---------------------------GAT-C--C--G-G------A-------G---CT--T-------------------------------G---------A-------A-----G------------ 1241

B.BR.10.10BR_MG029          -------A--T---------------------------------------T----------------A-------------------------GA--------------------------------------G-----------T-----A--------------G--- 1346
B.CA.07.502_1191_03         ----------A--------------------------------------------------------A--C----------------------GA--------------------------------------G--------T--T-----A------------------ 1311
B.CH.08.M2_0803101_NFLG8    ----------T--------------------------------------------------------A-----------------A-------GA------------------------------------------------A----C--A-----G------------ 1344
B.CN.10.DEMB10CN002         --A-------A-----------------------------------------A--------------------------------A--------A---------------------T-------------------G--T---A-------A-----------------G 1256
B.ES.10.DEMB10ES002         ----------T-----------------------------------------A--------------A------C-------------------A--------------------------------------G--G--------------A------------------ 1249
B.FR.11.DEMB11FR001         ----------------------------------------------------A--T--------------------------------------A--------------------------------------G-----------T-----A-----G--------G--- 1133
B.GB.05.MM45d213_GN1        --A----------------------------------A------------------------------------C----------A---A---GA--------------G-----------------------G--------------C--A--------------G--- 1285
B.HK.06.HK003               -------A-----------------------------------T-----------------------A--------------------------A--------------------------------------G----------C------A------------------ 1242
B.HT.05.05HT_129389         -----------------GG---------------------------------A--------------A------C----------A--G----GA-----------------------------------------G-----T--------A--------C--------- 1089
B.JP.12.DEMB12JP001         -------------G--------------------T-----------------------------------------------------------A-----------------------------------------G--------------A------------------ 1262
B.KR.07.HP_18_07JHS10_3909  -G-----A--------------------------------------------A-----------------------------T----------GA--------------------------------------G--------------C--A------------------ 1548
B.PE.07.502_0525_wg5        ----------T--------------------------------------------------------------------------A--------A-----C-----G--G--------------C--------G--G-----T--------A------------------ 1334
B.RU.11.11RU21n             ----------A--------------------G--------------------A-----------------------------------------A------T----------------------------------G--------------A-----G--------G--- 1443
B.TH.08.MERLBDTRC10         ----------T---------------------------------------T----------------A--------------------------A--------------------------------A-----G--G------A-------A--------C--------- 1096
B.US.11.ES38                ----------T-----------------------------A--T--G-------------------GA---------------A---------GA--C---T-------------------------------G-----T-----------A--------------A--- 1807

C.AR.01.ARG4006             --A-C--A-----------------------G--------------------------------------C--C----G------A--G----G----T-GT----------------------------------------T-----C--A--------G--------G 1083
C.BR.07.DEMC07BR003         --A-C--A-----------------------T--------------------A-----------C----GC-------G------A-------G-------T----G--------------------------G--------T-----C--A--------G--------G 1231
C.BW.00.00BW5031_1          --A----A--T--------------------C-------------------G-----------------AC-------G---C--A-------G-------T-------------------------------G--------T-----C--A-----G-----------G 1252
C.CN.10.YNFL19              --A----A--T--------------------C--------A-----------A--T--------------C-------G------A-------G----T--A-------G-----------------------G--------T-----C--A--------G--------G 1243
C.CY.09.CY260               --A-C--A--------------------------------------------------------------C-------G------A-------G-------T-------------------------------G-----G--TA----C--A--------A--A------ 1098
C.ES.08.X2363_2             --A----A--------GGG------------G--------------------------------------C-------G------A-------G-------T----------------------------------------TA----C--A-----G--G--------G 1312
C.ET.02.02ET_288            --A----A--T-----------------------------------------T-----------------C--C----G------A--G----G----T--T-------------------------------G--------T-----C--A--------G--------G 1083
C.IN.09.T125_2139           --A----A--T--------------------C--------------------------------------C-------G------A-------G-------T-------------------------------G--------TA----C--A--------A--------G 1875
C.KE.00.KER2010             --A----A--A--------------------C-----------T--------------------------C-------G------A-------G-------T-----------------T-------------G--------T--------A--------GC-------G 1080
C.MW.09.703010256_CH256.w96 --A----A--T--------------------C---C----------------------------C-----C-------G----C-A-------G-------T----G--------------------A-----G--------TA----C--A--------G--A--C--G 1876
C.TZ.08.707010457_CH457.w8  --A----A-----------------------C---------A-------------T--------------C--------------A--G----G-------T-------------------------------G--------TA----C--A--------G--------G 1873
C.YE.02.02YE511             --A----A--T--------------------C-----A--------G-----------------------C-------G------A-------G----T--T-------------------------------G--------T-----C--A-----G-----------G 1080
C.ZA.08.705010534_CH534.w12 -CA----A-----------------------G-----------------------------------AGGC-------G------A------CAT------A-------------------------------G--------TA----C--A-----------------G 1868
C.ZA.10.DEMC10ZA001         --A----A--A--------------------------------------------------C--------C-------G--G-C-A-------G-------T-------------------------------G--------T-----------T-----A--------G 1242
C.ZM.11.DEMC11ZM006         G-A----A--T--------------------C--------------------------C-----------C--------------A-------G----T--T----G--------------------------G--------TA----C--------G--G--------G 1233

D.CM.10.DEMD10CM009         -------A--T--------------------------------T--------A--------------A-----C-----G-------------G-------T-------------------------------G-----G---A-------A--------A--------G 1247
D.CY.06.CY163               -------A--T--------------------------------T--------A--------------A-----C-----G-------------G-------T-------------------------------G-----C-----------A--------A--------G 1095
D.KE.11.DEMD11KE003         ----T-----T-------GG--------T-----------A--------------------C-----A-A---C---------A---------G-------T-------------------------------G--G------A-T-----A------------------ 1234
D.KR.04.04KBH8              ----------------------------C--------------------------------------------C---------------G---GA---T--T-------------------------------G--G------A-T-----A------------------ 1843
D.SN.90.SE365               ----------------------------T--C-----------------------------------A-A---C-----------A-------G-------T-------------------C--------G--G--G------A-T-----A--------C--------G 1883
D.TZ.01.A280                -------A--T-----------------T--------------------------------------A-----C---------------------------T-------------------------------G--G------A-T---------------C-A-----G 1079
D.UG.10.DEMD10UG004         ----------T-----------------T--------------------------T-----------A-----C--------------G--C---------T-G--------------------------------G------A-T-----A-----G--C--------G 1222
D.UG.11.DEMD11UG003         GC--T-----T-----------------T-----------------------------------C--A--C--C---------A-------G...------T-------------------------------G--G------A-T-----A-----------------G 1233
D.YE.02.02YE516             -------A--A-----------------------------------------A--------------A-----C-----G-------------G----T--T----------------------C--------G-----G---A-------A-----------------G 1089
D.ZA.90.R1                  ----------T--------C--------------T-----------------A--------------A--C--C--G--------A--G---CA---------------------------------------G-----G--T--------A-----------------G 1230

F1.AO.06.AO_06_ANG32        --A-------A-----GGG------------C--------------------------------------C--------G--------GT---G--G-T-GA----------------------C--------G--------T--------------------------G 1086
F1.AR.02.ARE933             --A-------T--------------------C--------------------------------------C----------G-------T---G---------------------------------------G--------TA-------------------------G 1181
F1.BR.10.10BR_RJ015         --A-------A------GG------------C------C-------------------------------C---------------------------T--Y-------------------------A--R--G--------T--------------------------G 1351
F1.CY.08.CY222              --A-------A-----GGG------------C-----------T--------------------------C--C-------------------G---------------------------------------G--------T---------------A----------G 1083
F1.ES.02.ES_X845_4          --A-------A-----------------G--C--------------------------------------C--AC----G-------------G---------------------------T-----------G--------T--T--C--------------------G 1333
F1.RO.96.BCI_R07            --A-------A-----GGGA-----------C---------------C----------------------C--C--------T----------G---------------------------------------G--------TA--------------A----------G 1911
F1.RU.08.D88_845            GCA-------A-----GGGC-----------C--------------------------------Y-G---C-----------------T----G---------------------------------------G-----G--TA-------A------A----------G 1346
F2.CM.02.02CM_0016BBY       --A-------A------GGC-----------C--------------------------------------C------------C-A-------G---------------------------------------G--------T--------A------A----------G 1077
F2.CM.10.DEMF210CM001       --A----A--A--------------------C------C----------------T--------------C-----------------------------C--------------------T-----A-----G-----G--TA-------A------A----------G 1127
F2.CM.10.DEMF210CM007       --A-------A-----GGGC-----------C--------------------------------------C--------------A-------G---------------------------------------G--------T--------A-----------------G 1237

G.BE.96.DRCBL               --A-------A-------GC-----------C---------A----------A----------------TC--C-------G---A----A--GA--------------------------------------G---------A-------A-----------A-----G 1848
G.CM.10.DEMG10CM008         --A-------T-----G--C-----------C--------------------------------------C--C----G------A-------G------------G--------------------------G--------T--------A--------G--A-----G 1268
G.CN.08.GX_2084_08          -CA-------T--------C--------------------------------A-----------------C--C----G--G-C-A---A---GA--------------------------------A-----G---------A----C--A-----------A-----G 1112
G.CU.99.Cu74                --A-------A------GGC-----------C------------------------------------G-C--C----G------A-------GA--------------------------------------G---------A-------A-----------A------ 1497
G.ES.09.X2634_2             --A-------A--------C-----------C--------------------------------------C--C----G------A-------GA------T-----------------T-------------G--------T--T--C--A-----------A-----G 1358
G.GH.03.03GH175G            --A-------A------C--------------------C--A----------------------------C--C------T----A---A---GA--------------------------------------G--------T--------A--------A--A-----G 1924
G.KE.09.DEMG09KE001         ----------T-------GC-----------C---C-------------------T--T--------------C-----------A----A--GA-----G-----------------------C--------G---------A-------A-----------A-----G 1268
G.NG.09.09NG_SC62           -CA-------A------GGC-----------C---C----------------------------------C--CC---G------A-------GA--------------------------------A-----G--------TA-------A-----------A-----G 1076
G.ZA.01.TV546               --A-------A--------------------C---C---------------------------A------C--C-----------A----A--GA------------------------------------------------A-------A-----------A-----G 1089

H.BE.93.VI991               --A-------T--------C-----------C-----------------------T--------C-G------------------A----G--G----T--A-------------------------------G--------AA-------A-----------------G 1276
H.CF.90.056                 -CA-------T--G--G-----------------------------------A--T--------C--------A----G------A----A--G----T--A-------------------------------G--------TA-T-----A-----------------G 1231
H.GB.00.00GBAC4001          -CA-------C--G--G--------------C--------------------A--T--------C--------------------A----A--G----T--A-------------------------------G--------T--------A-----------------G 1405

J.CM.04.04CMU11421          --A-------A--------------------T---------A----------A--T--------C-GA--C--Y-----------A-------GA-----------R-----------------Y--------G--------T--T-----A--------------W--- 1402
J.SE.93.SE9280_7887         --A-------T--------------------T---C----------------A--T--------C--------C-----------A---T---GA--------------------------------------G--------T--T--------G--------------- 1207
J.SE.94.SE9173_7022         --A-------T--------------------T--------------------A--T--------C--------C-----------A---T---GA--------------------------------------G--------T--T-----A------------------ 1208

K.CD.97.97ZR_EQTB11         -CA-------A-----------------------C--------------C--A-------------G---C-------G---G------T---G-------T-------------------------------------G--T--------A-----G-----------G 1083
K.CM.96.96CM_MP535          -CA-------A-----G--------------C--C--------------C--A-----------------C-----------G----------G----T--T-------G-----------------------G-----G--TA-------A------A----------G 1083

U.CA.01.TV749               ----------T-----G--C-----------C------C----------------T------------T-C--CC-G--------A---A---G------CT----------------------C--------G---------A-------A------A----------- 1892
U.CA.99.TV721               ----------T-----G--C-----------C------C----------------T-----C----G-T-C--CC-G--------A---G---G------CT----------------------C--------G---------A-------A------A----------- 1894
U.CD.83.83CD003_Z3          ----------A--------------------C---------A-T--------A-----------------C--------------A---A---G------------G--------------------------G---------A-------A--------G--------G 1382
U.CD.90.90CD121E12          ----------A--------------------C--------AACT--------A-----------------C--------------A---A---G-------T----G--------------------------G--------T--------A-----------------G 1086
U.CY.05.CY090               --A-------A--C-----------------C--------A---------------------------CA-----------------------G---------------------------------------G--------TAC-------------A----------G 1083
U.CY.08.CY223               --A----------------------------C--------------G-----T-----------------C--C----G--------------G---------G---------------------------------------A-------A-----------------G 1098
U.ES.10.DEURF10DZ001        -AA-------A--------C------------------C----------------T-----C--------C-------GG-------------G-------T----------------------------------------TA-------A-----G-----------G 1166
U.GR.99.99GR303             --A-----------------------------------C----A--G--------T-----C--------C--------------A---T---GA------T----------------------C--------G--------T--------A-----G------------ 1167
U.NL.95.U_NL_95_H10986_D1   -------A--T--------------------T-----------------C--A-----------------C-----------G-----------A---T----------------------------------G--------T--------------GA----------G 1772

01_AE.AF.07.569M            --A----------------C--------------------A--------------T-----C----G-T-C--------------A--GT---GA------T----------------------C--A-----G--------TA-T-----A---------C-------G 1096
01_AE.CF.90.90CF11697       --A----------------------------------AC-A--------------T-----C------T-C--C-----------A---A---G-------T----------------------C--------G--------T--------A-----------------G 1839
01_AE.CN.10.YNFL03          --A----------------C------------------C-A--------------T-----C------T-C--------------A---A---GA------T--------------GT---------------G--------A--------A-----G--------A--G 1271
01_AE.HK.04.HK001           --A-------A---------------------------C-A--------------T-----C------T-C--------------A---G---GA-----TT----------------------C--------G--------TA-------A-----G-----------G 1261
01_AE.IR.10.10IR.THR48F     --A----A------------------------------C-A-----------A--T-----C-----AT-C--------------A---T---GA--C---T----------------------C--------G--------TA-------A-----G-----------G 1101
01_AE.JP.x.JRC77AE          ----------A--------C------------------C-A--------------T-----C------T-C--------------A---A---GA------T----------------------C--------G--------TA-------A-----G-----------G 1902
01_AE.TH.09.AA090a_WG11     --A-------------G--C--------------T---C-A--------------T------------T-C--C-----------A---A---GA------T----------------------C--------G--------TA-------------G-----------G 1313
01_AE.TH.90.CM240           --A----------------C------------------C-A--------C-----T-----C------T-C--------------A---A---GA------T----------------------C--------G--------TA-------A-----G------------ 1449
01_AE.US.05.306163_FL       --A----------------C------------------C-A--------------T-----C------T-C-------G--G---A---T---GA------T----------------------C--------G--------TA-------------G-----------G 1084
01_AE.VN.98.98VNND15        --A----------------C------------------C-A--------------T-----C------T-C--------------A---A---GA------T----------------------C--------G--------TA-------A-----G-----------G 1148
```

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

Gag p24 Capsid end Gag p2 start

```
B.FR.83.HXB2                CTTCACAGGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCGAACCCAGATTGTAAGACTATTTTAAAAGCATTGGGACCAGCGGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGGCAAGAGTTTTGGCTGAA 1884
Gag                         A__S__Q__E__V__K__N__W__M__T__E__T__L__L__V__Q__N__A__N__P__D__C__K__T__I__L__K__A__L__G__P__A__A__T__L__E__E__M__M__T__A__C__Q__G__V__G__G__P__G__H__K__A__R__V__L__A__E_
02_AG.CM.10.DE00210CM013    --A---------------C-----------------C---------------T-----C------T-C-------G------A--G----G--------------------------------------------C--A----G--------T--------A----G---C-------G 1257
02_AG.CY.09.CY256           --A-G-------------C-----------------C---------------T-----C--------C-------G------A-------G-------T--------------------------------------------G--------TA-------A----G-----------G 1086
02_AG.ES.06.P1423           --A-------A-------C-----------------C-------------A--T-----C------T-C---------------A---A---GR--CW--T------------------------------C--------G--------T--------A----G---C-------G 1328
02_AG.GW.05.CC_0048         ------------------C-----------------A-------------A--------C-----A--C-------G--G---A-------G-------T----G---------------------------------A----G--------TA-------A----G--C--A-----G 1246
02_AG.KR.12.12MHR9          --A---------------C---------------AC-TC----------T---T-----C------G-C-------G------A-------G-------T-------G--------------------------------G--------TA-------A----G-----------G 1537
02_AG.LR.x.POC44951         --A-G-------------C-----------------C---------------T-----C------T-C---C----------A---T---G-------T--------------------------------------------G--------T---------------G-----------G 1885
02_AG.NG.09.09NG_SC61       --A-G-------------C-------------C--C----------------T-----C-------AC-------G-----------A--G-------T-----------------------------------------------------T--------A---A-----------G 1082
02_AG.NG.x.IBNG             --A-G-------------C-----------------C---------------T-----C------T-C-------G------A---A---G-------T--------------------------------------------G--------T--------A----G-----------G 1415
02_AG.SE.94.SE7812          --A-------T-------C-----------------C---------------T-----C------T-C-------G------A-------G-------T--------------------------------------------G--------TA-------A----G-----------G 1262
02_AG.SN.98.98SE_MP1211     --A-G-------------C-----------------C-----------T---T-----C---G---C---C-----------A-------G-------T--------------T-----------------------------------------T--------A----G-----------G 1087
03_AB.RU.97.KAL153_2        --A-------T-----G--C---------------C--C-------------A--T----------A--C-------G------A---T---G-------T-----------------T---C---------G---------------------A----G--------C--G 1116
04_cpx.CY.94.94CY032_3      --A-C--------G--------------------C--C----------------T-----C------T-C--C-----------A---A---G-------T--------------------------------------------G---------A-------A----------A----G 1250
05_DF.BE.x.VI1310           --A-------A-------C-----------------------------------A-------------A-----C-----------A----A-CA-----------------------------------------------G--G--G---A-------A---------------G 1263
06_cpx.AU.96.BFP90          --A-------A-----------------C-------------------------A-------------------C-------G------A-------GA---------------------------------------------G--------T-------------------A----G 1909
07_BC.CN.98.98CN009         --A----A--T-----------------C-----------------------------------------C-------G------A-------G----T-CA-------------------------------------------G--------TA-------A--------G--------G 1222
08_BC.CN.97.97CNGX_6F       --A----A--T-----------------C--------A---------------------------------C-------G------A---T---G----T--T-------------------------------------------------TA----C--A--------G--------G 1082
09_cpx.GH.96.96GH2911       --A-------T------------A--------------C---------------T-----C--------C--C----G------A-------G-------T-------G---------------------------------------G--------T--------A--------------Y--- 1095
10_CD.TZ.96.96TZ_BF061      ----G-----T-------------------C----------------------------------A-----C------------------------------T----------------------------------------G--G------A-T----A-----------A-----G 1263
11_cpx.CM.95.95CM_1816      --A---------------C-A-------------G----C-------G--------T-----C--C--AT----------G------A---A---GA-------T--------------------------------------------------------A-----G-----------G 1260
12_BF.AR.99.ARMA159         --A-------A-----G-----------------C-----------------------------------A--C-------------------T---G-----------------------------------------------G--------T--------------------------- 1892
13_cpx.CM.96.96CM_1849      --A-------T-------C-----------------C--------------T---------------------C--C----G------A-------G-----------------------------------------------G-----------------A-----------A------ 1262
14_BG.ES.05.X1870           --A-------A-----GGC-----------------C-------------------------------------C--C----G------A-------GA----------------------------------------------G-------------T--C--A-----------A----G 1357
15_01B.TH.99.99TH_MU2079    --A-----------------------------------A---C-A-C-----------T-----C------T-C-------------A---T---GA---T--T---------------------------------------------G--------TA-------A---------------G 1292
16_A2D.KR.97.97KR004        --A-------A----------------------------C-------------------------------AT-C--C--G-G---C--A-------G----T--T-------------------------------------------G---------A-------A----G----------- 1245
17_BF.AR.99.ARMA038         ----------A-----GGG-----------------C-------------------T--------A--C-----------------------G---ACA---G-------------------------------------------G--------T-----------------------G 1101
18_cpx.CU.99.CU76           --A----------------------------AC---------------T-----------AC------G------A-------G------------G---------------------------------------------G---------A-------A--T--G------------ 1201
19_cpx.CU.99.CU7            ----------T----------------T--------------------------A------------A----CC-----------------G-------T--------------------------------------------G--------TA-T-----A---------------G 1110
20_BG.CU.99.Cu103           --A--------------GGC-----------------C-------------------------------C-------C--C----G------A-------GA-------------------------------------C--------G--------------------A--------A------ 1359
21_A2D.KE.99.KER2003        ----------A-----------------C-------C------------------------------AT-C--C--G-------A-------G-------T----------------------------------C--------G---------A-T----A--------GC-A-----G 1089
22_01A1.CM.01.01CM_0001BBY  --A-------T--G---------------------C-A-------------------C-----AT-C-------G------A--GA---G-------T-G-------------------------------C--------G--------G-------A-----G-----A-----G 1089
23_BG.CU.03.CB118           --A-------A-----GGC-----------------C-------------------------------C--C----G------A-------GA---GT-----------------------------------------------------T--------A-----------A------ 1333
24_BG.ES.08.X2456_2         --A-------A-----GGC-----------------C-------------------C-------------C--C----G------A-------GA---T-----------------------------------------G--------T--------A-----------A------ 1344
25_cpx.CM.02.1918LE         --A-------A-----GGC-----------------T--T--------------A--T-------------C--C----G---G--A-------GA---T-------------------------------------C--------G-------------------A-------C-A--A--G 1092
26_AU.CD.02.02CD_MBTB047    --A-------A-----GCC-----------------C--TC--------------T-------------C------G--G---A----A--G----T--T-G----------------------------------------G----------------------G------------ 1895
27_cpx.FR.04.04CD_FR_KZS    --A----A---------------------AGC------------------A--T-------------C--C----G------A-------GA----------------------------------------------G--G----TA-T-----A--------G--A-----G 1878
28_BF.BR.99.BREPM12609      G-A-------A-----GG-----------------C-------------------------------C--------------------G-----------------------------------------------G--------TA------------------------G 1253
29_BF.BR.01.BREPM16704      G-A-------A-----GGG----------------C------------------------C--C----C--------------------G-----------------------------------------------G--------TA-------------G-----------G 1288
31_BC.BR.04.04BR142         --A-C--A--------G----------------T-------A------------A------------T-C-------G------A-------G-------T--------------------------------------G--------T-----C--A--------G--------G 1338
32_06A1.EE.01.EE0369        --A----C--T-----G--C---------------C--C-------------A--T----------A--C-------G------A-------G-------T----------------------------------C--------G------------------A----G--------C--G 1525
33_01B.ID.07.JKT189_C       --A----------------------------------AC-A-------------T-----C------T-C---C-----------A---T---GA------T-G-----------------------------C--------G--------TA-------A----G-----------G 1194
34_01B.TH.99.OUR2478P       --A-------A-------C-----------------C-A-------------T-----C------T-C----------------A---A---GA------T----------------------------------C-----------------TA-------A----G-----------G 1089
35_AD.AF.07.169H            --A--------------GG-------------A--A--------------A--T-----------T-C--CC--------------A---A---G-------T----------------------------------C--------G--------TA-------A-----------A-----G 1093
36_cpx.CM.00.00CMNYU830     --A-------T-------------------------C-A------------------C------T-C-----------------A---A---G-------T----------------------------------C--------G--------T-----------G-----A-----G 1092
37_cpx.CM.00.00CMNYU926     --A---------------C---------------C--C-A-------------T-----C------T-C-----------------A---A---G-------T----------------------------------C-----------------T--------A----G-----A-----G 1083
38_BF1.UY.03.UY03_3389      --A-------A-----------------C-----AC-------------------------T-C------------------T----G-----------G-----------T---------------------------G--------T--------------G----------CG 1298
39_BF.BR.04.04BRRJ179       ----C-----------GCC-----------------T------C-------------A-------------C-----------------------GA----------------------------------------------G--G------A-------A------------------ 1358
40_BF.BR.05.05BRRJ055       --A-------A-----------------C-------------------------------C----------------------G-----C---------------------------------------------G--------TA---------------A----------G 1370
42_BF.LU.06.luBF_18_06      --A-------A-----GGG----------------C------------------------C---------------------------G----T--------------------------------------------G----G--T-----------------------A--G 1429
43_02G.SA.03.J11223         -CA-------T-----------------C------------------------------AC--CC---G------A---T---G------------------------------------------A-------G---------A---C---------G-----A-----G 1409
44_BF.CL.00.CH80            --A-------A-----G------------------C-----------------------A--C---C--G--------------------G-----------------------------------------------G--------T--------------G--G--------G 1329
45_cpx.FR.04.04FR_AUK       --A-------------------------------------C-------G----A--T-----C--------C--C----G------A-------G-------T-------------------------------------T--------TA-------A---AGA----------- 1876
46_BF.BR.07.07BR_FPS625     --A-------A-----GGG----------------C------------------------C---------------------------G-----------------------------------------------G--------TA-------------A----------G 1325
47_BF.ES.08.P1942           --A-------T------------------------------------T--------------C-------------------------CA-----------------------------------C--A----G----G-----------A---------------A--- 1326
48_01B.MY.07.07MYKT021      --A--------------GC-----------------AC-A-------------T-----C------T-C-----------------A---A---GA--A---T-----------------------------C-----------G--------TA-------A----G------------ 1095
49_cpx.GM.03.N26677         --A---------------------------------C---------------T-----C--------C-------G--G---A-------G----T--T-----------------------------------C--------G--------T--------A----G-----A------ 1319
50_A1D.GB.10.12792          ----------T-------------C-----------------------------A---------------------------------T---------------------------------------C--------G-------------------A----G--------C--G 1298
51_01B.SG.11.11SG_HM021     ----------T-------------------------A-----G----T-----------------C----------C---G----A----C------------------------------------A-------G--------A-T------A--------C----------- 1091
52_01B.MY.03.03MYKL018_1    --A--------------Y------------------C-A-------------R--T-----C------T-Y--------------A---T---GT------T---------------------------------C--------G--------TA-------A----G-----------G 1201
53_01B.MY.11.11FIR164       --A---------------C---------------A--AC-A-------------T-----C------T-C--C-----------A---A---GA------T-G--------------------------------C--------G--------TA----------------G---C-------G 1230
54_01B.MY.09.09MYSB023      -------A--T---------------------------------G--------T---------------------------------GA-----------------------------------------------G--G------A-------A------A------------ 1283
55_01B.CN.10.HNCS102056     --A------------G--C---------------C--C-A-------------T-----C------T-----C-----------A---A---GA------T----------------------------------C--------G--------TA-------A----G------------ 1200
56_cpx.FR.10.URF5_patient_A -CA-C-------------C-----------------C---------------T-----C--------C---C--G--G---A-------G-------T--------------------------------------------G--------T--------A----G-----------G 1235
57_BC.CN.09.09YNLX19sg      -------A--T--------------------------------------------------------------------------A-----------------------------------------------------G--------TA---C--A--------G--------G 1086
58_01B.MY.09.09MYPR37       --A---------------C-----------------C-A-------------T-----C------T-C-----------------A---G---GA------T-G-----G--------------------------C--------G--------TA-------A----G-----------G 1235
59_01B.CN.09.09LNA423       --A---------------C-----------------C-A-------------T-----C------T-C-----------------A---A---GA------T---------------------------------C--------G--------TA-T-----A----G--C-------G 1083
60_BC.IT.11.BAV499          --A-C--A-----------------G--GT------------------T--------------------C-------G------A--G----G--S----T--------------------------------------------G--------T-----C--A--------G--------- 1859
61_BC.CN.10.JL100010        --A----A--T-----------------C-----------------------------------------C-------G------A-------G-------T----------------------------A------G--------TA---C--A--------G--------G 1323
62_BC.CN.10.YNFL13          --A----A--T-----------------T-----------------------------------------C-------G------A-------G----T--T--------------------------------------------G--------TA---C--A--------G--------- 1231
63_02A1.RU.10.10RU6637      --A-G-------------C-----------------C---------------T-----C------G-C-------G------A-------G-------T--------------------------------------------G--------T--------A----G--G--------G 1431
64_BC.CN.09.YNFL31          --A----A--T-----------------C-------A---------------------------------C-------G------A-------G----T--A---------------------------------------------G--------TA-------A--------G--------G 1244
65_cpx.CN.10.YNFL01         --A----A--T-----------------C-------A-------------A-------------------C-------G------A-------G----T--A------G--------------------------------G--------TA---C--A---C----G--------G 1273
67_01B.CN.11.ANHUI_HF115    --A----------G-----C-----------------C-A-------------T-----C------T-C-----------------A---A---GA-----TT---------------------------------C--------R--G-----TA-------R-----G-----------G 1239
68_01B.CN.11.ANHUI_WH73     --A----------G-----C-----------------A--------------T-----C------T-C-----------------A---A---GA-----TT---------------------------------C--A----G--G-----TA-------A-----G-----------G 1230
70_BF1.BR.10.10BR_PE004     --A-------A-----GG---------G--C--------------------------------------C-------G------A------A------A----A------------------------------A--------------T-------------------G--------G 1340
71_BF1.BR.10.10BR_PE008     --A----A--A-------------G--C-----------------------------C--G-C-------G--------------G---------------------------------------------------G--T-----------------------A--------G 1554
72_BF1.BR.10.10BR_MG002     --A----A--A-----GGG----------------C-------------------------------C-------G------A-------G-------------------------------------------G--G--G--TA-------A-----------------G 1392
O.BE.87.ANT70               --A-T--A--A---------------------TC--C-T--T--------C--T--------C--ACAG-----G---T----A--G----GA-----CT-------G------GT---C---------------------AACT--------C-----GC-A--A--- 1932
O.CM.98.98CMA104            -A--T--A--A---------------------TC--C-T--T--------C--------------ACAG-----G------C-A--G----GA-----CT----------------GT---C---------------------AACA--------C-----G--A--A--- 1364
O.CM.98.98CMABB141          -A--T--A--A-----------------------C--C-T--T--------C--------C--ACAG---------T---A--G----GA-----CT-------G------GT---C---------------------AACT-----A--C-A-A-A--A--A--- 1363
O.CM.98.98CMABB212          ----T-----A-------G-------------TC--C-T--T--G---T-C--------------ACAA-----R--------A------GA-----TT---------------GT---C-----R-----R--R--G--AACT--C-------A-A-AC-A--A--G 1368
O.CM.98.98CMU5337           -AA-T-----A-----------------------C--C-T--T--------C--------C--ACAG---------T---A------GA-----CT----------------GT---C---------------------AACT--------C-----GC-A--A--- 1365
O.CM.99.99CMU4122           -A--T--A--A-------------------------C-T--T--------C--------C--ACAG-----G---T----A--G----GA-----CT----G--G-----GT---C---------------------AACT--------C--G--GC-A--A--- 1363
O.FR.92.VAU                 -AA-T--A--A-----------------------C--C-T--T--------C--------C--ACAA-----G--------A--G----GA----GC--------------GT---T-----------G-------AACT--------T-----CC-A--A--- 1441
O.GA.11.11Gab6352           -AA-T-----A-------------------------C-T--T--------C--------C--ACAA-----G--------A--G----GA--CT-C-----G---------GT---C-----------G-------AACT--------C-----GC-A--A--- 1089
O.SN.99.99SE_MP1299         -AA-T--A--A-----------------------C--C-T--T--------C--------C--ACAG-----G---T----A--G----GA-----CT----------------T---T---------------------AACT--------C--G--A--A--A--- 1934
O.US.10.LTNP                -A--T--A--A-----------------------C--C-T--T--------C--------C--ACAA-----G---T----A--G----GA-----CT-------G------GT---C---------------------AACT--------C---A-AC-A--A--- 1853
N.CM.02.DJO0131             -A-------T-----G----------------AC-C--A--A--------A--T------------ACAG--AC-----------A--G-----A-----CT-------------------G--C-----------G--G-----A-CG--------------GC-A--A--G 1369
N.CM.02.SJGddd              -AA-------A-----G----------------AC-C--A--A--------A--------------ACAG--CC-----T---A--G-----A-----CT-------G---------G--C--C--------G--G-----A--G--------------GC-A--A--G 1362
N.CM.04.04CM_1015_04        -AA-------T-----G----------------AC-C--A--A--------A--T-----------RCAG--CC-----------A------T-A----GCT-------G---------G--C--C--A-----G--G-----A--G--------------GC-A--A--G 1362
N.CM.06.U14296              -A-------W-----G-------------T--AC-C--A--A--------A--------C------ACAG--CC-----------A--G---GA-----CT-------G---------G--C--C--A-----G--G-----A-CG--------------AC-A--A--- 1360
N.CM.06.U14842              -AW----A--A--------------------C--AC-C--A--A--------A----------------CAGC-CC-----------A-------GA-----CT-------G---------G--C--C--A-----G--G--W--A-CA--------------GC-A--A--- 1361
N.FR.11.N1_FR_2011          -AA-------A-----G--------------C--AC-C--A--A--------A--------------ACAGC-CC-----------A--G---GA-----TT-------------------G--C--C--------G--G-----A-CG--------------GC----A--G 1258
P.CM.06.U14788              -CAGT-----A-------G------------C-----AC-A--A--------T--T--------Y---CAG-----G-----TC---------GT--C--CT-----------------AT--C-----A--------G-----AACA--------C--G--C-----A--- 1349
P.FR.09.RBF168              --AGT-----A-----G-G------------C--A--A--AA-A--------T--T-------------CAG-----G-----TC---------G---C--CT-----------------AT--C-----A--------G-----AACA--------C--G--C--A--A--- 1916
CPZ.CD.06.BF1167            ----C---CCA--------------------T--AC-----A--------T------------------C--AC-G------------------G---A--CT-G---------C-----T--C-------------G-----A-CA-------------G--CC-A--A--- 1965
CPZ.CM.05.SIVcpzMT145       -A-------T--T-----------------C---C-----A-A--G-----T--------C-----A-AC--C----G------A-------GA--C--T----------------------T-----A--G--------------TCA--------T-----GC----A--- 1423
CPZ.GA.88.GAB1              ----------A--G-------------T--C--TC-C--A--G--G-----A---------------CAA---C----------------T-G---C---C-----G--------------T-----A-------------------ATCT-----------G---C-A------ 1945
CPZ.TZ.06.SIVcpzTAN13       -C--T---CCA--G----CA-------------------C--A-A--------A--T-------C---CA----C-------C--A----AG-GA--A---T-G--------C-A-----C-------G-------C--TAC------------A-A-CC--------- 1525
CPZ.US.85.US_Marilyn        ----T-CA--A-------CC------------G---C-T-----A-----C--T--T------------A----------G---C--------T-GA--C--C---------------G---------------A-----G-----------CT-----------G--C--A------ 1932
GOR.CM.07.SIVgor2139_287    -AAGT--A--T---------------------T-----AA-T--G-----T--T--T---------CAA---C-------CC-A--C----G---A--C-----------------AT--T------------G--G-----TACT--C--A--------C-----A--G 1421
GOR.CM.07.SIVgorCP2135con   -AAGT--A--T--G-------------------T---C-AA-T--G-----T--T--T---------CAA---C-------CC-A--G--G-G---A--C-----------------AT--T------------G--------TACT--C--A--------C--R--A--G 1420
```

Gag-Pol fusion TF protein start
Gag p2 end / Gag p7 nucleocapsid start
B.FR.83.HXB2
Gag

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

Gag-Pol fusion TF protein start
Gag p2 end Gag p7 nucleocapsid start

```
B.FR.83.HXB2               GCAATGAGCCAAGTAACA..................AAT......TCAGCTACC...ATA...ATGATG..................CAG......AGAGGCAATTTTAGG...AACCAAAGAAAG......ATTGTTAAGTGTTTCAATTGTGGCAAAGAAGGGCACAC 1991
Gag                        _A__M__S__Q__V__T__.__.__.__.__.__.__N__.__.__S__A__T__.__I__.__M__M__.__.__.__.__.__.__Q__.__.__R__G__N__F__R__.__N__Q__R__K__.__.__I__V__K__C__F__N__C__G__K__E__G__H__T
02_AG.CM.10.DE00210CM013   --------T------CA-..................C-G.........A-A-AT...---...------..................---......---A-----------...GG--GG----CA.........A-A----------C-----T--------A---CT 1358
02_AG.CY.09.CY256          --------T------CA-..................C-G.........C-C-A-...---...------..................---......---------------...GG-----A--CA.........A-A----------C--------------A---CT 1187
02_AG.ES.06.P1423          --------T------CA-..................C-G..........--AGGA...G--...------..................---......--------------A...GG---G----CA.........A-A----------C--------------A---CT 1429
02_AG.GW.05.CC_0048        --------T----C-CAG..................C-G.........T-C-A-...---...------..................---......---------------...GG---G----CA.........A-A----------C--------------A---CT 1347
02_AG.KR.12.12MHR9         --------T------CA-..................C-G.........C-C-A-...---...------..................---......---------------...GG---G----CA.........A-A--A-------C--------------A---CT 1638
02_AG.LR.x.POC44951        --------T------CA-..................C-G..........--C-A-...---...------..................--A......---------------...GG--GG----CA.........A-A----C-----C--------------A---CT 1986
02_AG.NG.09.09NG_SC61      --------T------CA-..................C-G.........A-C-A-...---...------..................---......---------------...GG---G----CA.........A-A----------C--------------A---CT 1183
02_AG.NG.x.IBNG            --------T------CA-..................C-G..........--C-A-...G--...------..................---......---------------...GG---G----CA.........A-A----------C--------------A---CT 1516
02_AG.SE.94.SE7812         --------T------CA-..................C-G.........C-C-AT...---...------..................---......---------------...GG---G---CCA.........A-A----------C--------------A---CT 1363
02_AG.SN.98.98SE_MP1211    --------T----C-CA-..................C-G.........T-C-A-...---...------..................---......---------------...GG---G----CA.........A-A----------C--------------A---CT 1188
03_AB.RU.97.KAL153_2       --------T------CA-..................---..........--A-A-...---...------..................---......-A-A-T----------...GG--C--A--GA.........A-------------C--------------A---CT 1217
04_cpx.CY.94.94CY032_3     -----------G-C-T--..................---...GCAG----AG--...---...------..................---......-A-A---A----A-...GG--------GA.......-C-A------------C--------G-----A--TCT 1360
05_DF.BE.x.VI1310          -------------C----..................---...TCAG-T-----A...GC-...------..................---......--------------A-...GG--C-----A.........A-------------C-----T------------T 1373
06_cpx.AU.96.BFP90         -----------G-C-T--..................GT-.........-GAG--...---...------..................---......-A-A---------A-...GG--CG-A--GA.......-G-A------C----C--------G-----A--TCT 2013
07_BC.CN.98.98CN009        ------------AC--AC..................-G-...............---...C-----..................---......---A--------AA...GG-TCT-A--GA.......---------A-------C--------G---------T 1323
08_BC.CN.97.97CNGX_6F      ------------AC--AC..................---...............---...C-----..................---......---A--------AA...GG-TCT-A--GA.......---------A-------C--------G---------T 1183
09_cpx.GH.96.96GH2911      --------T------CA-..................---.........A-A-A-...---...C-A---..................---......--------------A-...GGT--G-A--GA.........A-------------C--------------A--MYT 1196
10_CD.TZ.96.96TZ_BF061     -----------G-C----..................GG-......GG-AA---T...---...------..................---......--------------A-...GG--C--AG--A.......-GCA------------C------------------ 1370
11_cpx.CM.95.95CM_1816     ---------------CA-..................C-G.........A-A-AT...---...------..................--A......--------------A-...GG---G-A--GA.........A-------------C--------------A---CT 1361
12_BF.AR.99.ARMA159        -------------C----..................-G-.........A-AG-T...G--...C-----..................---......-A-A-T-GC----AT...GG-------GA.......---------A-----------------AC---A----T 1996
13_cpx.CM.96.96CM_1849     -----------G-CCT--..................-G-GCAGCAG----AG--...---...------..................---......-A-A---------A-...GG--C--A--GA.......---A-------------C-------GG-----A--TCT 1375
14_BG.ES.05.X1870          ---------------T--..................GGG..........--AT-A...G--...------..................--A......-A------------A-...GGT-------GA.......-ACA-------------C--------G-----A---CT 1461
15_01B.TH.99.99TH_MU2079   -----------T---CA-..................C-G.........A-A-AT...G--...------..................---......--------------A-...GG---G-A--GA.........A-------C-----C--------------A---CT 1393
16_A2D.KR.97.97KR004       ---------------CA-..................---...ACAAATT-A-A-ATA--G...------..................---......-A--------------...GGG---A---A..........--------------C--------G-----A---CT 1354
17_BF.AR.99.ARMA038        ------------------..................---.........C-A--T...G--...C-----..................--A......-A-A-T--C---A-...GG--C-----A.......-A-----A-----------------------A----T 1205
18_cpx.CU.99.CU76          --------T--T---CA-..................---.........A-C-A-...---...------..................---......-----A----C-A-...GG---G-A--GA.........A------------------------------A----T 1302
19_cpx.CU.99.CU7           -------------C----..................---...TCACAT------...---...------..................---......-A---T------A-...GG--T------A......G----------------C------------------T 1220
20_BG.CU.99.Cu103          -----------G---T--..................GG-...GCAGG---AT--...---...C-----..................--A......-A-A---------A-...GG--C--A--GA.......-A-A----------T--C--------------A---CT 1469
21_A2D.KE.99.KER2003       -------------C----..................--C...TCAG-T-T---A...G--...------..................---......--------------A-...GG--C--A--GA.......---A----------T--C------------------T 1199
22_01A1.CM.01.01CM_0001BBY --------T------CA-..................C--.........A-AGGT...---...------..................---......--------------A-...GG---G-A--GA.........A-------------C--------------A---CT 1190
23_BG.CU.03.CB118          -----------GA--C--..................GG-...GCAGG---AT--...---...------..................--A......---A----C---A-...GG--C--A--GA.......-A-A----------T--C-----------T--A---CT 1443
24_BG.ES.08.X2456_2        --------T--GA--T--..................GG-...GCAGGG--AT--...---...------..................--A......-A-A--------------...GG--C-----GA.......-A-A----------T--C--------G-----A---CT 1454
25_cpx.CM.02.1918LE        -----------G-C-T--..................GG-...ACAG---AG--...---...------..................---......--------------A-...GG--C-----GA.......-A-A-C--A---------C--------G-----A--TCT 1202
26_AU.CD.02.02CD_MBTB047   --------------GCA-..................---..........--A-A-......T--G--..................---......G-----------AA...GG---G---GGA...AAA---A----A---------C--------------A---CT 2002
27_cpx.FR.04.04CD_FR_KZS   -----------G-C-T--..................GG-...GCAGT-A-AG--...---...------..................---......-A--------C-A-...GG---G-A--GA.......---A-------------C--------G-----A--TCT 1988
28_BF.BR.99.BREPM12609     -------------C----..................---.........A-A--T...G--...------..................---......-----T--C---A-...GG---------A.......---------A-----T------------------A----T 1357
29_BF.BR.01.BREPM16704     -------------C----..................---.........A-A--T...G--...------..................---......-A-A-T--C---A-...GG--------GA.......---------A-----T------------------A----T 1392
31_BC.BR.04.04BR142        ----------------AC..................---.........A-A-A-...---...------..................---......---A--------AA...GG--CT-A--GA.......-C-A---A--C-----C--------G----------TT 1442
32_06A1.EE.01.EE0369       --------T----C-CA-..................---..........--A-A-...---...------..................---......-A-A-T------------...GGA-C--A--GA.........A-------------C--------------A---CT 1626
33_01B.ID.07.JKT189_C      ---------------CA-..................C-A.........A-AGGT...---...------..................---......------------C-A-...GG---G-A--GA.........A-------C-----C------------..---CT 1292
34_01B.TH.99.OUR2478P      ------------AC-CA-..................C-G..........--A-ATATG--G...------..................---......--------------A-...GG---G-A--GA.........A-------C-----C--------------A---TT 1193
35_AD.AF.07.169H           --------T----C-CA-..................C--.........A-A-A-....--C...------..................---......--------------A-...GG---G-A--GA.........A-------------C--------------A---CT 1194
36_cpx.CM.00.00CMNYU830    ---------------CA-..................C--..........--AGGT...---...------A.................---......---A--------------...GG--------CA.........A-A--A---------C--------G-----A---CT 1193
37_cpx.CM.00.00CMNYU926    -------CT----C-CA-..................C-G.........A-A-G-...---...------..................--A......---A---------A-...GGT-C--A--GA.......-A-A---A---------C--------G-----A--TCT 1187
38_BF1.UY.03.UY03_3389     ------------------..................---.........T-AG-T...G--...------..................--A......-A---T--C---A-...GG--C-----GA.......-C------A-----------------------A----T 1402
39_BF.BR.04.04BRRJ179      -------------C--AT..................C-A.........A---A-....---....---G--................---......--------------A-.......-C-----G-.......-C-------------------------------T 1459
40_BF.BR.05.05BRRJ055      -------AT---------..................---.........A-AG-T...G--...------..................---......-A-A-T--C---A-...GG--------GA.......---------A-----T------------------A----T 1474
42_BF.LU.06.luBF_18_06     -------------C----..................GG-.........A-AG-T...---...C-----..................---......---A-T--C---A-...GG--C------A.......---------A-----T--C--------G-----A----T 1533
43_02G.SA.03.J11223        -----------G-C-T--..................GG-...GCAG----AG--...---...------..................---......-A-----------A-...GG--C-----GA.......-A-A----------T--C--------G-----A--TCT 1519
44_BF.CL.00.CH80           --------T---------..................---.........A-AG-T...G--...------..................---......-A-A-T--C---A-...GG--------GA.......---------A-----------------------A----T 1433
45_cpx.FR.04.04FR_AUK      --------T----C-CA-..................-G-...ACAAACCTA-TG...--G...------..................---......--------------A-...GG---G-A----.........A-------------C--------------A---CT 1983
46_BF.BR.07.07BR_FPS625    ------------------..................C--.........A-A-AT...G--...C-----..................---......-A-A-T--C---A-...GG--------GA.......---------A-----T------------------A----T 1429
47_BF.ES.08.P1942          -------------C----..................GC-.........--AA-GT-...---...------..................---......-A--------A----...GG---G------.......-----C------------------G------A----T 1433
48_01B.MY.07.07MYKT021     -----------G-C-CA-..................C-A..........-GAGGT...---...------..................---......------------C-A-...GG---G-A--GA.........A-------C-----C--------------A---CT 1196
49_cpx.GM.03.N26677        --------T------CA-..................---.........T-C-A-...---...------..................---......-A-----------A-...GG---G-----A.........A-------------C--------------A---CT 1420
50_A1D.GB.10.12792         --------T--G--GCA-..................TC-.........T-A-A-...---...------..................---......--------------A-...GG---G-A--G-.........A-------------C--------------A---CT 1399
51_01B.SG.11.11SG_HM021    -------------C---T..................-G-......G--A---G-...---...------..................---......------------------...----------GA.......-C-----------T------------------T 1198
52_01B.MY.03.03MYKL018_1   -------------C-CA-..................C-A......CATR-A-AT...---...------..................---......--------------A-...GG--C--A--GA.........A-------C-----C--------------A---CT 1305
53_01B.MY.11.11FIR164      ---------A-T-C-CA-..................C--..........--A--T...---...------..................---......--------------A-...GG---G----GA.........A-------C-----C--------------A---CT 1331
54_01B.MY.09.09MYSB023     ------------------..................--C......C--C------...---...-----A..................---......------------------...------G---GA.......-C---C-------------------------TGT 1390
55_01B.CN.10.HNCS102056    -----------CA-GCA-..................TC-.........A-A-AT...---...------..................---......--------------A-...GG---G-----A.........A-------C-----C-------G-----A---CT 1301
56_cpx.FR.10.URF5_patient_A --------T----C-CA-.................C-G..........--C-A-...---...-----A..................---......---------------A...GG---G---TCA.........A-A--A-----T--C--------------A--TCT 1336
57_BC.CN.09.09YNLX19sg     ------------AC--AC..................--G...............---...------..................---......---A--------AA...GG--CT-A--GA.......---------A-------C--------G---------T 1187
58_01B.MY.09.09MYPR37      -----------C-C-CA-..................C-A.........A-AGGT...---...------..................--R......--------------A-...GG-A-C-R--RA.........A-------C-----C-------------------TT 1336
59_01B.CN.09.09LNA423      -------------C-CA-..................C-G.........A-A--T...---...------..................---......--------------A-...GG---G-A--GA.........A-------C-----C-------G-----A---CT 1184
60_BC.IT.11.BAV499         ------------C---AC..................---..........--A-A-...---...------..................---......--GA--------AA...GG--CT-A--GA.......-C-A---A--C-----C--------G---------CT 1963
61_BC.CN.10.JL100010       ------------AC--AT..................-G-...........G--...---...C-----..................---......---A--------AA...GG--CT-A--GA.......---------A-------C--------G--G-------T 1424
62_BC.CN.10.YNFL13         ------------AC--AC..................-G-...........G--...---...------..................---......---A--------AA...GG--CT-A--GA.......---------A-------C--------G---------T 1332
63_02A1.RU.10.10RU6637     --------T------CA-..................C-G.........T-C-A-...---...------..................---......------------------...GG---G-A--CA.........A-A--A---------C------------------CT 1532
64_BC.CN.09.YNFL31         ------G------C--AC..................---...........GT-...---...C-----..................---......--------------A-.....--T------------.......-----C-----------------------------T 1345
65_cpx.CN.10.YNFL01        -------------C--AC..................---...............---...------..................---......--------------AA...GG--CT-A--GA.......---------A-------C--------G---------T 1374
67_01B.CN.11.ANHUI_HF115   ---------ARC---CA-..................C--.........A-A-AT...---...------..................---......-A-----------A-...GG--CG-A--GA.........A-------C-----C--------------A---CT 1340
68_01B.CN.11.ANHUI_WH73    ---------A-K---CA-..................C-G..........--A-GT...G--...------..................---......-A-----------A-...GG--CG-A--GA.........A-------C-----C--------------A---CT 1331
70_BF1.BR.10.10BR_PE004    -------------C--AT........................--A---...---...------..................---......-A-A-T--C---AA...GGT---G---GA.......---------A-----T-----------C--A-A--T-T 1441
71_BF1.BR.10.10BR_PE008    ----------------AT..................................A-A--T...---...------..................---......-A-A-T--C---A-...GG---------A.........--AA---A-----------------G-----A--TCT 1655
72_BF1.BR.10.10BR_MG002    -------------C----..................--A.........A-AG-T...G--...------..................--A......-A-A-T-------A-...GG---------A......G-------A-----T--C-----------------A----T 1496
O.BE.87.ANT70              ------GCTAC--CCCAG.........CAAGATTTG--AGGAGGA-ACA-AG-A...G--...T-C---..................--A......-----GC-AAA-CCA...-TTAGG-A-GGA.......-C-A-A--A--------C-----A-----G--A--T-T 2054
O.CM.98.98CMA104           ------GC-TC--CCCAG.........CAAGATTTA--AGGAGGA-ACA-AG-A...G--...T-C---..................--A......-----GC-AAA-CCA...-TTAGG---GGA......CC---A-----------C-----G-----G--A--T-T 1486
O.CM.98.98CMABB141         ------GCTTC--CCCAG.........CAAGACCTG--AGGAGGA-ACA-AG-A...G--...T-C---..................--A......-----GC-AAACCCA...-TTAGG-AGGG-......CC-A-A--A--------C-----A-----G--A--TCT 1485
O.CM.98.98CMABB212         ------GTTTCT-CCCAG.........CARGATTTA--GGGAGGA-ACA-AT-A...---...T-C---..................--A......-----GC-GAG-CCA...--TA------G-.......-C-A-A--A--------C-----A-----G--A--TGT 1490
O.CM.98.98CMU5337          ------GCTAC--CCCAG.........CAAGATCTG--GGGAGGA-ATA-AG-A...G--...T-C---..................--A......-----GC-AAACCCA...-TTAG--A-GGA......CC-A-A--A---------------G-----G--A--TCT 1487
O.CM.99.99CMU4122          ------GCTAC--CCCAG.........CAAGATCTG--AGGAGGA-ACA-AG-A...G--...T-T---..................--A......-----GC-AAA-CCA...GGAAGG-A-GGA......CC-A-A--A---------------A-----G--A--TCT 1485
O.FR.92.VAU                ------GCTTC--CCCAG.........CAAGATCTA--AGGAGGA-ACA-AT-A...G--...T-C---..................--A......-----GC-AAAACCA...-TTAG--A-GGA......CC-A-A--A---------------A-----G--A--TGT 1563
O.GA.11.11Gab6352          ------GCTAC--CCCAG.........CAAGATCTG--AGGAGGA-ATA-AT-A...G--...T-C---..................--A......-----GC-AAACCCA...GTTA-----GG-......CC---A--A--C--T--C-----A-----G--A----T 1211
O.SN.99.99SE_MP1299        ------GCTGC--CCCA-............GATCTG--GGGAGGA-ACA-AG-A...G--...T-T---..................--A......-----GC-AAACCCA...-GTAGG-A-GG-......CC-A-A--A---------------A-----G--A--TCT 2053
O.US.10.LTNP               ------GCTACT-CGCAG.........CAAGATTTG-GAGGGGGA-ACA-AG-T...G-C...T-T---..................--A......-----GC-GAACCCA...--TAGG---GGA......GC-A-A--A--------C-----A-----G--A--T-T 1975

N.CM.02.DJO0131            --T---TCA--G---CAG..................-CA......G--A---GT...G-C...T-TGCA..................--A......G-GA-A--C----AA...GG-AT---G-GA.......-CCA---GA-------------------G--C--TTT 1476
N.CM.02.SJGddd             --T---TCA------CAG..................GCA......A--A---GT...G-C...T-TG--..................--A......--G--A--C----AA...GG-AT------A.......-CCA---A---------------------G--C--TTT 1469
N.CM.04.04CM_1015_04       --T---GCA--G-C-CAG..................GCA......G--A---GT...G-C...T-TG-A..................--A......-----A--C----AA...GG-AC---G-GA.......-CCA---A--------C------------G--C--TTT 1469
N.CM.06.U14296             --T---TCA--G---CAG..................GCA......A--A---GT...G-C...T-TGCA..................--A......--G--A--C----AA...GG-AT---G-GA......CCCA---A---------------------G--C--TTT 1467
N.CM.06.U14842             --T---TCA--G---CAG..................GGA......G--A---GT...G-C...T-TG-A..................--A......--G--A--C----AA...GG-ATG-----A.......-CCA---A---------------------G--C--TGT 1468
N.FR.11.N1_FR_2011         --T---TCA--G---CAG..................CCA......A--A---AT...--C...T-TGC-..................--A......--G--A--C----AA...GG-AT---G--A......CCCA------------------------G--C--TCT 1365

P.CM.06.U14788             --T---GCAGC--CT-ATCAAGCTAGCCAAGAATTA--AGGAGGG-ATA-A--A...G-T...T-T---..................---AGT.....................GGA--G----------......CCA--------C--T--C-----A----T---A----T 1468
P.FR.09.RBF168             --T---GCAGC--CT-ATCAAGCTAACCAAGATTTA--AGGAGGG-ATA-A--T...G-T...T-C---..................---AGT.....................GGA--G---W--......CCA--------C--T--C-----A---AT---A----T 2035

CPZ.CD.06.BF1167           --C---CAG---AC-CAG..................C-G......AAT-T--AT..........---G-A..................---GGG...-----CAGCGCC-CAA...GGAA-----GG-GCAGGA-ACA-A--A-----T--------A--GAT------T-- 2078
CPZ.CM.05.SIVcpzMT145      --C--------G----AC..................--A.........A-AGAA...--C...T-T---..................--AAGA...G-G--AC-AAA---A...CCT-CT-AT--A.......-AGA----A--C--T---------A--------A---CT 1530
CPZ.GA.88.GAB1             --C---TCAATG--TCAG..................---CAAGGGAG---AGAT...G-T...T-CT-C..................---......-AG--AC-AGG-GC-...GG--C--A--GA.......-AAA-A--A--C--T---------G--------T---CT 2058
CPZ.TZ.06.SIVcpzTAN13      ------GCAAC--CT--T..................---ACAGTAGG--TA-AT..........---C--..................---......G----T--GAAACCA...TTAAG--AGGG-.......CAGC--C-A-----T---------A---TT------T-- 1635
CPZ.US.85.US_Marilyn       --C---T-----A-G-A-..................--C......C-CT-A-G-...G--...T-TC-A..................---......-A---A--CGC-G--......--G-C-G-G-GA.......-AAA-C--A-----T----------------G--A--TCT 2039

GOR.CM.07.SIVgor2139_287   ------GCAGC--CCCAC.........TCAGATTTA--AGGAGGA-ATA-AG-A...G-C...T-T---..................--AAAGGGGG--AA----CAAG--....-TTA-G-AGGGA.......CCA--A-----C-----------G----T---A--T-T 1549
GOR.CM.07.SIVgorCP2135con  ------GCAGC--CCCAC.........TCAGATYTA--AGGGGGA-ATA-AG-A...G-C...T-T---..................--AAAGGGGG--AR----CAAG--....-TTA-G-AGGGA.......CCA--A--A--C-----------G----T---A--T-T 1548
```

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

Gag-Pol -1 ribosomal slip site
Pol start
Gag p7 nucleocapsid end
Gag p1 start
B.FR.83.HXB2 AGCCAGAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGATTGTACT GAG ... AGACAGGCTAATTTTTTAGGGAAGATCTGGCCTTC ... CTACAA GGGAAGGCCAGGGAATT 2131
Gag A R N C R A P R K K G C W K C G K E G H Q M K D C T E R Q A N F L G K I W P S Y K G R P G N
Pol F F R E D L A F L Q K G K A R E F
02_AG.CM.10.DE00210CM013 1498
02_AG.CY.09.CY256 1327
02_AG.ES.06.P1423 1569
02_AG.GW.05.CC_0048 1487
02_AG.KR.12.12MHR9 1778
02_AG.LR.x.POC44951 2126
02_AG.NG.09.09NG_SC61 1323
02_AG.NG.x.IBNG 1656
02_AG.SE.94.SE7812 1503
02_AG.SN.98.98SE_MP1211 1328
03_AB.RU.97.KAL153_2 1357
04_cpx.CY.94.94CY032_3 1500
05_DF.BE.x.VI1310 1513
06_cpx.AU.96.BFP90 2153
07_BC.CN.98.98CN009 1463
08_BC.CN.97.97CNGX_6F 1323
09_cpx.GH.96.96GH2911 1336
10_CD.TZ.96.96TZ_BF061 1513
11_cpx.CM.95.95CM_1816 1501
12_BF.AR.99.ARMA159 2136
13_cpx.CM.96.96CM_1849 1515
14_BG.ES.05.X1870 1601
15_01B.TH.99.99TH_MU2079 1530
16_A2D.KR.97.97KR004 1494
17_BF.AR.99.ARMA038 1345
18_cpx.CU.99.CU76 1442
19_cpx.CU.99.CU7 1360
20_BG.CU.99.Cu103 1609
21_A2D.KE.99.KER2003 1339
22_01A1.CM.01.01CM_0001BBY 1339
23_BG.CU.03.CB118 1586
24_BG.ES.08.X2456_2 1594
25_cpx.CM.02.1918LE 1342
26_AU.CD.02.02CD_MBTB047 2142
27_cpx.FR.04.04CD_FR_KZS 2128
28_BF.BR.99.BREPM12609 1497
29_BF.BR.01.BREPM16704 1532
31_BC.BR.04.04BR142 1585
32_06A1.EE.01.EE0369 1766
33_01B.ID.07.JKT189_C 1432
34_01B.TH.99.OUR2478P 1333
35_AD.AF.07.169H 1334
36_cpx.CM.00.00CMNYU830 1336
37_cpx.CM.00.00CMNYU926 1327
38_BF1.UY.03.UY03_3389 1542
39_BF.BR.04.04BRRJ179 1599
40_BF.BR.05.05BRRJ055 1614
42_BF.LU.06.luBF_18_06 1673
43_02G.SA.03.J11223 1659
44_BF.CL.00.CH80 1573
45_cpx.FR.04.04FR_AUK 2123
46_BF.BR.07.07BR_FPS625 1569
47_BF.ES.08.P1942 1573
48_01B.MY.07.07MYKT021 1336
49_cpx.GM.03.N26677 1560
50_A1D.GB.10.12792 1542
51_01B.SG.11.11SG_HM021 1338
52_01B.MY.03.03MYKL018_1 1445
53_01B.MY.11.11FIR164 1471
54_01B.MY.09.09MYSB023 1533
55_01B.CN.10.HNCS102056 1441
56_cpx.FR.10.URF5_patient_A 1476
57_BC.CN.09.09YNLX19sg 1327
58_01B.MY.09.09MYPR37 1476
59_01B.CN.09.09LNA423 1321
60_BC.IT.11.BAV499 2103
61_BC.CN.10.JL100010 1564
62_BC.CN.10.YNFL13 1472
63_02A1.RU.10.10RU6637 1672
64_BC.CN.09.YNFL31 1485
65_cpx.CN.10.YNFL01 1514
67_01B.CN.11.ANHUI_HF115 1480
68_01B.CN.11.ANHUI_WH73 1471
70_BF1.BR.10.10BR_PE004 1581
71_BF1.BR.10.10BR_PE008 1795
72_BF1.BR.10.10BR_MG002 1639
O.BE.87.ANT70 2197
O.CM.98.98CMA104 1629
O.CM.98.98CMABB141 1628
O.CM.98.98CMABB212 1633
O.CM.98.98CMU5337 1630
O.CM.99.99CMU4122 1628
O.FR.92.VAU 1706
O.GA.11.11Gab6352 1354
O.SN.99.99SE_MP1299 2196
O.US.10.LTNP 2118
N.CM.02.DJO0131 1622
N.CM.02.SJGddd 1615
N.CM.04.04CM_1015_04 1615
N.CM.06.U14296 1613
N.CM.06.U14842 1614
N.FR.11.N1_FR_2011 1511
P.CM.06.U14788 1611
P.FR.09.RBF168 2178
CPZ.CD.06.BF1167 2224
CPZ.CM.05.SIVcpzMT145 1676
CPZ.GA.88.GAB1 2198
CPZ.TZ.06.SIVcpzTAN13 1787
CPZ.US.85.US_Marilyn 2197
GOR.CM.07.SIVgor2139_287 1692
GOR.CM.07.SIVgorCP2135con 1691

```
                               Gag p1 end_Gag p6 start
B.FR.83.HXB2                  TTCTTCAGAGCAGACC..................................AGAGCCAAC............................AGCCCCACCAGAAGAGAGCTTCAGGTC..............TGGGGTAGAGACAACAAC......TCCCCC....... 2207
Gag                           F  L  Q  S  R  P__.__.__.__.__.__.__.__.__.__.__.__. E  P  T__.__.__.__.__.__.__.__.__.__. A  P  P  E  E  S  F  R  S__.__.__.__.__.__. G  V  E  T  T  T__.__.__P  P__.__._
Pol                           __S__S__E__Q__T__.__.__.__.__.__.__.__.__.__.__.__._R__A__N__.__.__.__.__.__.__.__.__.__.__S__P__T__R__R__E__L__Q__V__.__.__.__.__.__.__W__G__R__D__N__N__.__.__S__P__.__.__
A1.AU.03.PS1044_Day0          -C-C------------..................................---------............................-----------CG---CT---TG--AT..............G---A----T..------......CT----....... 1412
A1.CH.03.HIV_CH_BID_V3538     ---C-----A-----T..................................G--------............................-----------C---------TG--GT..............GA---A---A-..-----......CT--T-....... 1409
A1.CY.08.CY236                ---CC-----------..................................-------T-............................-----------C-----T---TG--AT..............G---A----T..------......CT----....... 1424
A1.ES.05.X1608_8              -C-CA-----------..................................-------T-............................-----------C-----K---TG--AT..............G---A----T..G--...CT-A--....... 1640
A1.KE.11.DEMA11KE001          ---CA----A-----T..................................G--------............................-----------C-----T--GTG--AT..............G---AC-G--T..-G-...CTT---....... 1543
A1.RU.11.11RU6950             ---C------------..................................-------T-............................-----------CG--A-A---TG--AT..............G---A----T..------......C---T-....... 1728
A1.RW.11.DEMA111RW002         ---CC-----------..................................-------T-............................-----------TG--ACT---GG--AT..............G---A----T..-G-......CT----....... 1555
A1.SN.01.DDI579               ---C------------..................................---------............................-----------C---TCT-C-GG--AT..............G--A-A----T..-----......CT--T-....... 1405
A1.UG.11.DEMA110UG001         ---C---A--------..................................---------............................---------G-C----CT---TG--AT..............G---A----T..------......CT----....... 1559
A1.ZA.04.04ZASK162B1          ---C-----A-----T..................................G--------............................-----------C--G----ATG--AT..............G---A----T..------......CT----....... 1648
A2.CM.01.01CM_1445MV          ---CC---------A-..................................---------............................------G----C-----A-G-TG-AAT..............G---A----T..-G-......-T--T-....... 1403
A2.CY.94.94CY017_41           ---C----------A-..................................---------............................-----------C-----A---G--AAT..............G---A----T..------......CT--T-....... 1561
B.BR.10.10BR_MG029            ----C-----------..................................---------............................--------------------------T..............-----AG-----------......------....... 1669
B.CA.07.502_1191_03           ---------A------..................................-------T-............................-----------------------A--T..............-----AG-----------......------....... 1634
B.CH.08.M2_0803101_NFLG8      --T-C------C----..................................---------............................------------G-------------T..............-----A--A---------......C---AG......T 1665
B.CN.10.DEMB10CN002           ---------A---GA-..................................-------T-............................--------------------------T..............-----AG-------G--......---AG-....... 1579
B.ES.10.DEMB10ES002           ----------------..................................---------............................--------------------------T..............-----AG---A----G-......---A--....... 1569
B.FR.11.DEMB11FR001           -C--------------..................................---------............................--------------------------T..............-----AG-----------......----T-....... 1456
B.GB.05.MM45d213_GN1          ----------------..................................---A-----............................------------G-------------T..............-----A------------......----T-....... 1608
B.HK.06.HK003                 ----------------..................................---------............................------------G-------------A..............-----A------------......------....... 1565
B.HT.05.05HT_129389           -C-------A---GA-..................................-------T-............................------------G-------------T..............-----A---..G------......------....... 1408
B.JP.12.DEMB12JP001           -C-----------G--..................................---------AGCCCCACC..................--------------------------T..............-----A-A----GAG--......----T-....... 1594
B.KR.07.HP_18_07JHS10_3909    ----------------..................................-------T-............................------T-----G-------------T..............-----AG-----------......----T-....... 1871
B.PE.07.502_0525_wg5          -------------GA-TCAGAGCAGGCC......................---------............................------------G-----------A-T..............--T--AG-----------......C---T-....... 1672
B.RU.11.11RU21n               -C--------------..................................---------............................------------G-----T-------T..............-----AG-----------......----T-....... 1766
B.TH.08.MERLBDTRC10           ---------A------..................................---------............................------T-------------------T..............-----A------------......------....... 1440
B.US.11.ES38                  -C--------------AGAGAGCAGACC......................-------T-............................------------G-------------T..............-----A--G----G----......----T-....... 2142
C.AR.01.ARG4006               -Y--C-----------..................................---------............................-----------C--------------T..............C-A--A..----------......C---T-....... 1397
C.BR.07.DEMC07BR003           -C-CC----------T..................................---------............................-----------C--------------T..............C-A--A..--T-------......C---T-....... 1545
C.BW.00.00BW5031_1            -C--C-----------..................................---A-----AGCCCCACC..........AATGCCAAC-----------C--------------T..............C----A..----------......C---G-....... 1584
C.CN.10.YNFL19                ---------A----A-..................................---------............................---T-------C--------------T..............-----A------------......---AT-....... 1560
C.CY.09.CY260                 -C-------A---G--..................................---------AGCCCCACCGACAGT.....GGTGCCAAC-----------C-----------A--T..............C-A--AGAG---------......C---G-....... 1439
C.ES.08.X2363_2               -C--------------..................................---------AGCCCCACCAGC........AGAACCAAC-----------C--------------T..............C-A--G.......-G-......----G-....... 1644
C.ET.02.02ET_288              -C-------A------..................................--C------AGCCCCACT..........AGAGCCAACG----------C--------------T..............C-A--A...-G-------......C--T--....... 1415
C.IN.09.T125_2139             -C--C-----------..................................---------............................-----------C--------------T..............C-A--A..-------CCC..C--AG-....... 2189
C.KE.00.KER2010               -C-------A---G--..................................---------............................-----------C--------------T..............C-A--A..-----..-G-......C-TTG-....... 1394
C.MW.09.703010256_CH256.w96   -C-------A---G--..................................---------AGCCCCACCACT........AGAACCAAC-----------C--------C--C..........................................C---G-....... 2193
C.TZ.08.707010457_CH457.w8    -C-------A----A-..................................---------............................-----------C----------GA--T..............C-A--A..----------......C-A---....... 2187
C.YE.02.02YE511               -C--------------..................................---------............................----------TT--------------T..............C-A--A..A---------......C---G-....... 1394
C.ZA.08.705010534_CH534.w12   -C-----------G--..................................---A-----AGCCCCACCAGC........AGAGCCAAC-----------C------T-------T..............C-A--G..----------......C--TG-....... 2203
C.ZA.10.DEMC10ZA001           -C--------------..................................---------............................-----------C--------------T..............C-A--A..----------......C--TG-....... 1553
C.ZM.11.DEMC11ZM006           -C--C-----------..................................---------AGCTCCCCT..........TCAGCCAAC-----------C--------------T..............C-A--A..----------......C---G-....... 1562
D.CM.10.DEMD10CM009           ----------------AGAGCCAACAGCCCCACT................---------............................-----------C----------G---T..............-----A----T..-----......C-T-T-....... 1582
D.CY.06.CY163                 ---------A------..................................---------............................-----------C----------G---T..............-----AG---T..-----......C---T-....... 1415
D.KE.11.DEMD11KE003           -C-----G----G--...................................---A--C--............................-----------C-------C-TG---T..............-----AG---T..-----......A---T-....... 1557
D.KR.04.04KBH8                -C-C----------A-..................................---------............................C----------C----------G---T..............-----AG---T..-----......C---T-....... 2166
D.SN.90.SE365                 ---C------------..................................---------............................----------AC----------G---T..............-----AG---T..G--......C---T-TCAGAAA 2213
D.TZ.01.A280                  -C--------------..................................---------............................-----------C-----T----G---T..............-----AG---T..--A......A---T-....... 1402
D.UG.10.DEMD10UG004           -C--------------..................................---------............................-----------C----T-----G---T..............-----AG-------A......C---T-....... 1545
D.UG.11.DEMD11UG003           -C--------A--G--..................................---------............................-----------C----------G---T..............-----A----T..-----......A-----....... 1556
D.YE.02.02YE516               ----------------..................................---------............................T----------C----------G---T..............-----A----T..-----......C---T-....... 1412
D.ZA.90.R1                    ----------------..................................---------............................-----------C----------G---T..............-----AG---T..-----......C---T-....... 1553
F1.AO.06.AO_06_ANG32          ----------------..................................---------............................-----------C----------G---T..............CA---AG---T..GC-..................... 1403
F1.AR.02.ARE933               -C--------------..................................---------............................------G----C----------G---T..............CA-A-AG---T..-----......C---T-....... 1498
F1.BR.10.10BR_RJ015           -CA-C----A---G--..................................---A---T-............................------G----C----------G---T..............CA---AG---T..CC-......AT-T--....... 1668
F1.CY.08.CY222                ----------------..................................-------T-............................------G----C----------G---T..............CAAA-AG---T..-G-......C---T-....... 1397
F1.ES.02.ES_X845_4            ----------------..................................---------............................-----------C----------G-A-T..............CA-A-AG--A--..G--......C---T-....... 1650
F1.RO.96.BCI_R07              ---------A------..................................---------............................---T--G--C-C---------TG---T..............CA---AG--A--..-----......----T-....... 2228
F1.RU.08.D88_845              ---C---R-------T..................................-------T-............................------G--G-CG---------G---T..............C----AG---T..-----......C---T-....... 1663
F2.CM.02.02CM_0016BBY         -CA-------------..................................---------............................G-----G----C----G-----G---T..............C--A-A----T..-----......----T-....... 1394
F2.CM.10.DEMF210CM001         ----------------..................................-----AG-GCAGACC..........AGAGCCTTC------G----C---------TG---T..............C----AG---T..-----......C---T-....... 1468
F2.CM.10.DEMF210CM007         ---------A------..................................---------............................------G----C----------G---T..............---A-AG---T..-----......----T-....... 1551
G.BE.96.DRCBL                 -C-------A---G--..................................---------............................---------C-C-----A----G---T..............C----AG---T..-G-......C---T-....... 2171
G.CM.10.DEMG10CM008           ---------A------..................................---------............................---------C-C---------TG---T..............C----AG---T..-G-......C---AA....... 1588
G.CN.08.GX_2084_08            -C-------A---G-T..................................---------............................---------C-C----------G---T..............C--A-AG-G--T..-----......C---T-....... 1432
G.CU.99.Cu74                  -------A-A---G--..................................-------T-............................---------T-C----------G---T..............C--ACAG---T..-G-......C---T-....... 1820
G.ES.09.X2634_2               ----C----A---G--..................................---------............................-----------C----------G---T..............CA-A-AG---T..-G-......C---T-....... 1675
G.GH.03.03GH175G              ---------A---G--..................................-------T-............................---T-------C----------G---T..............C---AGG------..-G-......C---T-....... 2244
G.KE.09.DEMG09KE001           ----C----A---G--..................................---------............................---------T-C----------G--CT..............C--A-AG---T..-G-......C---T-....... 1585
G.NG.09.09NG_SC62             ---C-----AG--G-T..................................---------............................-----------C----------G---T..............C----AG---T..-G-......C---T-....... 1399
G.ZA.01.TV546                 ------------G--...................................---------............................-----------C----------G---T..............C----AG---T..-G-......C---T-....... 1409
H.BE.93.VI991                 -C-CC----AG--G-T..................................---------............................-----------C----------G---T..............C--A-AG---T..C--......C---T-....... 1599
H.CF.90.056                   ----C--------G--..................................---A-----............................-----------C----------G---T..............C--A-AG---T..G--......C---T-....... 1554
H.GB.00.00GBAC4001            ----C--------G--..................................---A-----............................-----------C---AT-----G---T..............C----AG---T..G--......A-----....... 1728
J.CM.04.04CMU11421            ----Y--------R--..................................---------............................-----------C----------G---T..............C----A------------......CY-YY-....... 1719
J.SE.93.SE9280_7887           ----C-----------..................................---------............................-----------C-------C--G--CT..............C--A-AG---T.............C---T-....... 1521
J.SE.94.SE9173_7022           ----C-----------..................................---------............................-----------C-------C--G---T..............C--A-AG---T.............C---T-....... 1522
K.CD.97.97ZR_EQTB11           ---------A------..................................---------............................-----------C---------TG---T..............C----AGA--T..-----......C---T-....... 1400
K.CM.96.96CM_MP535            ----------------..................................---------............................-----------C----------G---T..............C----AG---T..-----......C---T-....... 1400
U.CA.01.TV749                 -C-CA-----------AGAGCCAACAGCCCCACCATTCCCTCAGAGCAGCCT---------............................-----------C-------C-GG--AA..............G-C--A--G--T..-----......CT--T-....... 2242
U.CA.99.TV721                 ---C------------..................................--C------............................-----------C-------C-TCA-AG..............CA-ACC----C-..-----......AG----....... 2208
U.CD.83.83CD003_Z3            ---------A------..................................---------............................-----------C----------G---T..............C--A-AG------------......C---T-....... 1696
U.CD.90.90CD121E12            ---------A----A-..................................---------............................---T-------C----------G---T..............C--A-AG--AGT.....---CAC..C---T-....... 1406
U.CY.05.CY090                 -C--C-----------..................................-------T-............................-----------C----------G-A-T..............C--A-AG---T..-----......CT---T....... 1403
U.CY.08.CY223                 ---C------------..................................---A-----............................-----------C------TC-GG---T..............C--A-AG---T..-G-CGC..C---T-....... 1421
U.ES.10.DEURF10DZ001          ---C------------..................................---A-----............................----------ATG--------GG--AT..............G---A-------------......CT--T-....... 1480
U.GR.99.99GR303               ---C------------..................................------G--............................-----------C----------G---T..............C--A-AG---T..-GT......C---T-....... 1481
U.NL.95.U_NL_95_H10986_D1     -C--------------..................................---------............................-----------CG---------G---T..............C--A-AG---T..-----......C---T-....... 2089
01_AE.AF.07.569M              ---C------------..................................---------............................------G----C---....--GGG--AT..............G---A----T..-----......CT-ATT....... 1410
01_AE.CF.90.90CF11697         ---C-----------T..................................---------............................-----------C---------GG--AT..............G---A----T..-----......CT--TT....... 2153
01_AE.CN.10.YNFL03            ---C------------..................................---------............................-----------C---A-A--GGG--AT..............G---A----T..-C-......AT--TT....... 1582
01_AE.HK.04.HK001             ---C------------..................................-----TT--............................-----------CG--A-A--GG---GA......................--T..-----......CT--TT....... 1566
01_AE.IR.10.10IR.THR48F       ---C------------..................................G------T-............................------G----C---A-A--GGG--AT..............G---A----T..-G-......CT--T-....... 1415
01_AE.JP.x.JRC77AE            ---C------------..................................---A---T-............................-----------C---A-A--GGG-AAT..............G---A----T.......................... 2207
01_AE.TH.09.AA090a_WG11       ---C------------..................................---------............................-----------C-----A--GGG--AT..............G---A---AGT..-----......CT----....... 1627
01_AE.TH.90.CM240             ---C------------..................................---------............................-----------C---A-A--GGG--ATGGGGGAAGAGATAACG---A----T..-----......CT--TT....... 1778
01_AE.US.05.306163_FL         ---C---------G--..................................---------............................-----------C-----A--GGG--AT..............G---A-------------......CT--TT....... 1398
01_AE.VN.98.98VNND15          ---C----------A-..................................---------............................-----------C---A-A--GGG--AT..............G---A----T..-----......CT--TT....... 1459
```

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

Gag end
Gag p6 end
Gag-Pol TF end
Pol protease start

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

B.FR.83.HXB2 GGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTA...GAAGAAATGAGTTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGTGGACATAAAGC 2464
Pol G G Q L K E A L L D T G A D D T V L E E E M S L P G R W K P K M I G G I G G F I K V R Q Y D Q I L I E I C G H K A
A1.AU.03.PS1044_Day0 1672
A1.CH.03.HIV_CH_BID_V3538 1669
A1.CY.08.CY236 1675
A1.ES.05.X1608_8 1912
A1.KE.11.DEMA11KE001 1803
A1.RU.11.11RU6950 1988
A1.RW.11.DEMA111RW002 1815
A1.SN.01.DDI579 1662
A1.UG.11.DEMA110UG001 1819
A1.ZA.04.04ZASK162B1 1905
A2.CM.01.01CM_1445MV 1663
A2.CY.94.94CY017_41 1821
B.BR.10.10BR_MG029 1926
B.CA.07.502_1191_03 1891
B.CH.08.M2_0803101_NFLG8 1933
B.CN.10.DEMB10CN002 1836
B.ES.10.DEMB10ES002 1826
B.FR.11.DEMB11FR001 1710
B.GB.05.MM45d213_GN1 1865
B.HK.06.HK003 1834
B.HT.05.05HT_129389 1665
B.JP.12.DEMB12JP001 1851
B.KR.07.HP_18_07JHS10_3909 2128
B.PE.07.502_0525_wg5 1929
B.RU.11.11RU21n 2035
B.TH.08.MERLBDTRC10 1697
B.US.11.ES38 2399
C.AR.01.ARG4006 1648
C.BR.07.DEMC07BR003 1796
C.BW.00.00BW5031_1 1835
C.CN.10.YNFL19 1817
C.CY.09.CY260 1687
C.ES.08.X2363_2 1895
C.ET.02.02ET_288 1666
C.IN.09.T125_2139 2440
C.KE.00.KER2010 1645
C.MW.09.703010256_CH256.w96 2444
C.TZ.08.707010457_CH457.w8 2438
C.YE.02.02YE511 1639
C.ZA.08.705010534_CH534.w12 2454
C.ZA.10.DEMC10ZA001 1804
C.ZM.11.DEMC11ZM006 1813
D.CM.10.DEMD10CM009 1839
D.CY.06.CY163 1672
D.KE.11.DEMD11KE003 1814
D.KR.04.04KBH8 2423
D.SN.90.SE365 2481
D.TZ.01.A280 1665
D.UG.10.DEMD10UG004 1802
D.UG.11.DEMD11UG003 1819
D.YE.02.02YE516 1675
D.ZA.90.R1 1810
F1.AO.06.AO_06_ANG32 1675
F1.AR.02.ARE933 1758
F1.BR.10.10BR_RJ015 1937
F1.CY.08.CY222 1657
F1.ES.02.ES_X845_4 1916
F1.RO.96.BCI_R07 2488
F1.RU.08.D88_845 1923
F2.CM.02.02CM_0016BBY 1654
F2.CM.10.DEMF210CM001 1728
F2.CM.10.DEMF210CM007 1811
G.BE.96.DRCBL 2428
G.CM.10.DEMG10CM008 1845
G.CN.08.GX_2084_08 1704
G.CU.99.Cu74 2068
G.ES.09.X2634_2 1947
G.GH.03.03GH175G 2504
G.KE.09.DEMG09KE001 1839
G.NG.09.09NG_SC62 1656
G.ZA.01.TV546 1666
H.BE.93.VI991 1853
H.CF.90.056 1811
H.GB.00.00GBAC4001 1982
J.CM.04.04CMU11421 1970
J.SE.93.SE9280_7887 1778
J.SE.94.SE9173_7022 1779
K.CD.97.97ZR_EQTB11 1660
K.CM.96.96CM_MP535 1660
U.CA.01.TV749 2481
U.CA.99.TV721 2480
U.CD.83.83CD003_Z3 1959
U.CD.90.90CD121E12 1663
U.CY.05.CY090 1663
U.CY.08.CY223 1681
U.ES.10.DEURF10DZ001 1740
U.GR.99.99GR303 1738
U.NL.95.U_NL_95_H10986_D1 2349
01_AE.AF.07.569M 1670
01_AE.CF.90.90CF11697 2413
01_AE.CN.10.YNFL03 1842
01_AE.HK.04.HK001 1826
01_AE.IR.10.10IR.THR48F 1654
01_AE.JP.x.JRC77AE 2461
01_AE.TH.09.AA090a_WG11 1887
01_AE.TH.90.CM240 2038
01_AE.US.05.306163_FL 1655
01_AE.VN.98.98VNND15 1719

```
B.FR.83.HXB2                  GGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTA...GAAGAAATGAGTTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGTGGACATAAAGC 2464
Pol                           _G__G__Q__L__K__E__A__L__L__D__T__G__A__D__D__T__V__L_..._E__E__M__S__L__P__G__R__W__K__P__K__M__I__G__G__I__G__G__F__I__K__V__R__Q__Y__D__Q__I__L__I__E__I__C__G__H__K__A
02_AG.CM.10.DE00210CM013      -----A--G----TA-------------C--------------------G---...--------A-A---A-------A---------------------------------------------------A-------------T--G-----T-----GA-A--G-- 1831
02_AG.CY.09.CY256             -----A--G-C--TA-------------C-------------------------...-----C--AGA---A-------A--------------------------------------------------------------YCT-------T-----GA-A----- 1660
02_AG.ES.06.P1423             -----A--G----RA----C--------C-----G-------------------...--------AGA---A-------------------R---------------------------------------A-------------T-------T------A-A--G-- 1902
02_AG.GW.05.CC_0048           -A---A--GT---TA----C--T-----C-------------------------...--------A-A-C-A-----------------------------------------------------------------------T-------T------A-A--G-- 1814
02_AG.KR.12.12MHR9            -----A--GT---TA----C--------C-------------------------...--------A-A---A-------A-----G-------------------------------------------------------T-------T-----GA-A----- 2123
02_AG.LR.x.POC44951           -----A--G----TA----C--------C-------------------------...-----C--AGAG--A---------------------------------------------------------A-------------CT-------T------A-A--G-- 2459
02_AG.NG.09.09NG_SC61         -----A--G----TA-------------C-------------------------...--------A-----A--------------------------------------G-------------------------------AGT-------T------A-A--G-- 1656
02_AG.NG.x.IBNG               -A---A--G----TA----C--------C-------------------------...-----C--A-A---A-------A---------------------------------------------------A--------A----T-------T------A-A--G-- 1989
02_AG.SE.94.SE7812            -----A--G----TA----C--------C-------------------------...--------A-A---A-------A-----------------------------------------------------------------T-------T------A-A--G-- 1836
02_AG.SN.98.98SE_MP1211       -A---A--G----TAA------------C---------A---------------...--A-----A-A---A-------A------------------------------------T--------------------------T-------T------A-A--G-- 1661
03_AB.RU.97.KAL153_2          --A--A--G-----A---------------------------------------...-----C--A-A-----------A--------------------G--------------G---------------------------T-------T------A-A--G-- 1690
04_cpx.CY.94.94CY032_3        -----A--GA---G---G-----T------------------------------...--------A-A-----------A------G-----------------C-------------------------A-------------CT-------T------A-A--G-- 1830
05_DF.BE.x.VI1310             --------G-------------------------G-------------------...-----C--A-A-----------A---------------------------------------------------------------TG----------------G----- 1846
06_cpx.AU.96.BFP90            -----A--G----TA----C--------C-------------------------...-----C--A-A---A-------A-------------------------------------------A--------A----T-----G-T------A-A--G-- 2492
07_BC.CN.98.98CN009           ---------T--------------------------------------------...-----C----A--------G-A-----------------------------------------------------A------C------------C-----C----- 1796
08_BC.CN.97.97CNGX_6F         -----C--GA---A--G----C-----C--C-----------------------...------G-A-A-----------A---------------------A---------------------------A-----G--A---CT-------T------A-A--G-- 1644
09_cpx.GH.96.96GH2911         --A--A---T-----------G-----C--------------------G-----...--------A-A-----------A-------------------------------------------------------A-----T-----G--T--------A--G-- 1666
10_CD.TZ.96.96TZ_BF061        -----A--G----------T----------------------------------...----------A-----------A-----------------G--A----------------------------------A-----T-------A------T------- 1843
11_cpx.CM.95.95CM_1816        -CA-----G--G-GA---------------------------------------...--------A-A---A-------A--------------------------------------------------GG-A-------------TGAA---A-A--G-- 1828
12_BF.AR.99.ARMA159           -----A--G---------------------------------------------...-----C--A-A-----------A-------------------------------------------A---------A-CG-----------T--------C--G-- 2469
13_cpx.CM.96.96CM_1849        ----AA--G----TA----C-----------G----------------------...A-------A-A---------------------------------------------------------------A-C--GAG---------TGAG---A-A--G-- 1869
14_BG.ES.05.X1870             -----A--G----TA----C--G-------------------------G-----...-----T--A-A---A-------A----------------------A-------T--A----------------A--------A----T-------T------A-A--G-- 1931
15_01B.TH.99.99TH_MU2079      --A--A--G--G--A--------------C-------------C----------...-----T--A-A---C-----G-A-----C------------------------------------G-----G-A-------------CT-------T------A-A--G-- 1863
16_A2D.KR.97.97KR004          -AA--A--G----GA---------------------------------G-----...-----T--A-A-C-A-------A------------------------------------------------G-A-------------ACT-------T----A-A-A-G--- 1827
17_BF.AR.99.ARMA038           -----A--G---------------------------------------------...-----C--A-A---A-------A----GG-----------------------------------------A---------G-T-------------T--------C--G-- 1678
18_cpx.CU.99.CU76             -----A--G----TA----C--------C-------------------------...-----C--A-A---A-------A--------------------------------------------------A--------A----T-------T------A-A-GG-- 1769
19_cpx.CU.99.CU7              -----A--G---------------------------------------------...-----C--A-A---A-------A-----G----------------------------------------------------A----C--------T--------G----- 1690
20_BG.CU.99.Cu103             -----A--G----TA----C--------C-------------------------...-----T--A-A---A-------A--------------------------------------------------------A----T-------T------A-A--G-- 1930
21_A2D.KE.99.KER2003          -----A--G--------------------------G------------------...-----------A-----------A--------------------------G-------T--------------------A----C-C------------------------ 1669
22_01A1.CM.01.01CM_0001BBY    --A--A--G-----A--------C------------------------------...-----C--AGA-----------A------------------------------G--------A--------A--------A----T-------T------A-A--G-- 1672
23_BG.CU.03.CB118             -----A--G----TA----C--------C-------------------------...-----T--A-A---A-------A--------------------------------------------------------A----CT-------T------A-A--G-- 1907
24_BG.ES.08.X2456_2           -----A--G----TA--G-C--------C-------------------------...-----T--A-A---A-------A--------------------------------------------------------A----T--G-----T------A-A--G-- 1915
25_cpx.CM.02.1918LE           --A--A---A---TA----C--------C-------------------------...-----T--A-A---A-----G-A-----G-------------------------------------------------G-----G--T-------TA-----A-A--G-- 1672
26_AU.CD.02.02CD_MBTB047      -AA--A-------GA---------------------------------------...-----T--A-A---T-------A------------------------------G--------A--------AG--AC------T---------GA-A--G-- 2469
27_cpx.FR.04.04CD_FR_KZS      -A---A--G----G----------------------------------------...--------A-A-----A-----A-----G-------------------------------------------G-A----GG-AG--ACTG----C--T------A-A--G-- 2455
28_BF.BR.99.BREPM12609        -----A--G----G----------------------------------------...-----C--A-A-----------A------------------------------------------A---------G-C-----------C--T--------C--G-- 1830
29_BF.BR.01.BREPM16704        -----A--G---------------------------------------------...-----C--A-A-----------A-----------------------------G------------A---------G-C-----------T--------C--G-- 1865
31_BC.BR.04.04BR142           --------G-------G----C-----C--------------------------...--------A-AA----------AT--------------------A-----G---------------------------A----T-------T------A-A--G-- 1936
32_06A1.EE.01.EE0369          ----AA--G----TA----C----------------------------------...-----T--A-A---A-------A-----------G-----------------------------------------A---AT-----G--T------A-A--G-- 2099
33_01B.ID.07.JKT189_C         ---------T---G----------------------------------------...-----G--A-A-----------A-------------------------------------G--A----------------T-----G--T------A-A--G-- 1765
34_01B.TH.99.OUR2478P         --A--A--G--G--A---------------------------------------...-----T--A-A-----------A----------------------------------------------------------CT------------C-----C----- 1666
35_AD.AF.07.169H              -----A--G-A--G---------C------------------------------...-----C--A-A-----------A--------------------------C-------------A--------A--GCT-------T-----GA-A--G-- 1664
36_cpx.CM.00.00CMNYU830       --------GT---TA----C--------C-------------------------...-----C--A-A---A-------A--------------------------------------------A-------------CT-------T-----GA-A--G-- 1669
37_cpx.CM.00.00CMNYU926       -----A--G----TA----CT-------C-------------------------...-----C--A-A-----------A---------C--------------------A-----------A--------A---CT-------T------A-A--G-- 1654
38_BF1.UY.03.UY03_3389        -----A--G----------T----------------------------------...-----C--A-A---A-------A------------------------------------------A---------A-C---TG--------T----AC--C--G-T 1887
39_BF.BR.04.04BRRJ179         -----A--G----G----------------------------------------...-----T--A-A-----------A----C---------------------A--G-------------------GA-C----C--------T--------C--G-- 1941
40_BF.BR.05.05BRRJ055         -----A--G-----A--------------------------A------------...-----C--A-A---A-----G-A---------------------------------G---G-A---------GG-C----C--------T------------------ 1947
42_BF.LU.06.luBF_18_06        ----AA--GA-----------G-----C--------------------------...-----------A-----------------------------------------T-------A---------T-C---AG-----T--T--------C------ 2018
43_02G.SA.03.J11223           ----AA--G----TA-------------C-------------------------...---C---AGA---A-----------------------------------------A-----------A-------------CT--------TGAA---A-A--G-- 1989
44_BF.CL.00.CH80              -----A--GA---A----------------------------------------...-----C--A-A-C---------------C-----------------------------------A------A--G-C---AC---------T--------C--G-- 1906
45_cpx.FR.04.04FR_AUK         --A--A--G-----A-------------C-------------------------...--------A-A-----------A-----------------------------A-------------------------G-T--------T------A-A--G-- 2450
46_BF.BR.07.07BR_FPS625       --------G----G----------------------------------------...-----C--A-A-----------A------------------------------------------A---------G-C-------------T--------C--G-- 1902
47_BF.ES.08.P1942             -----A--GA-G------------------------------------------...-----C--A-A-----------A-------------------------------------------------------A--A---------T------A-G----- 1903
48_01B.MY.07.07MYKT021        ---------T-G-------C----------------------------------...-----------A-----------------G--------------------------------------------------C-----------------C----- 1669
49_cpx.GM.03.N26677           --------G--G------------------------------------------...--------AGA-----------A-----G----------------------------------------------A--G-T-------------TGAG---A-A--G-- 1887
50_A1D.GB.10.12792            -----A--G--G---------C--G-----------------------------...-----------A-----A-----A------------------------------C--------------------------C---------------------------- 1872
51_01B.SG.11.11SG_HM021       --------G------------C--------------------------------...--------AGA---A----G------------------------------------------------------------G--TCAG----T------------------ 1671
52_01B.MY.03.03MYKL018_1      --A--A--G--G-GA---------------------------------------...-----T--AGA-----------A------------------------------------------G--A--------R-----T-------T------A-A--G-- 1778
53_01B.MY.11.11FIR164         -C-------T--------------------------------------------...-----G-----A--------------------------------------------------G--A----G--------TC-------T------A-A--G-- 1804
54_01B.MY.09.09MYSB023        ---------GA-------------------------------------GC----...-----C--A-A------------------------------------------------------------------C--------T-----G--C----- 1866
55_01B.CN.10.HNCS102056       -AA--A--G-----A---------------------------------------...-----T--A-A-----------A------G-------------------------------G--C--G--A-------------T-------T------A-A-GG-- 1774
56_cpx.FR.10.URF5_patient_A   -----A-------TA----C----------------------------------...-----C--A-A---A-------AG-----------------------------------C---------------TCTG-------T------A-A-GG-- 1815
57_BC.CN.09.09YNLX19sg        -----C--GA---A--------C-----C-------------------------...--------A-A-----------A-----------------------A-----------------G--A--------A---CT-------T------A-A-GG-- 1648
58_01B.MY.09.09MYPR37         ---------T--------------------------------------------...-----T--AGA-----------A----------------------------------------------------------R---C------------C-----C----- 1809
59_01B.CN.09.09LNA423         --A--A--G-----A---------------------------------------...-----T--A-A-----------A---------G-------------------C-----------G--A--------------TGT-------T-----GA-A--G-- 1648
60_BC.IT.11.BAV499            -----A--G-------G----CC-----C-------------------------...-------CA-AA----------AT------------------A-----G---C-------------------------A---GCT-------T------A-A--G-- 2421
61_BC.CN.10.JL100010          -----C--GA---A--G----C-----C--------------------------...--------A-A---A-------A----------------------A------------------------A--------A---CT-------T------A-A--G-- 1885
62_BC.CN.10.YNFL13            -----C--GA-G--A--G----C---------------------------G---...--------A-A-----------A----------------------A--------------T---------A--------A---CT--T----T------A-A-GG-- 1793
63_02A1.RU.10.10RU6637        -A---A--G----TA----C--------C-------------------------...-----C--AGA---A-------A---------------------------C-------------------------------T--G-------A----A-A--G-- 2005
64_BC.CN.09.YNFL31            -----C--GA---A--G----C-----C--------------------------...-----------A-----------A---------------------------------------------A-------------CT-------T--------C--G-- 1818
65_cpx.CN.10.YNFL01           -----C--GA---A--G----C-----C--------------------------...--------A-A---A-------AG---------------------A-----------------------------A--------A---GCT-------T------A-A--G-- 1835
67_01B.CN.11.ANHUI_HF115      --A--A--G-----A---------------------------------------...-----T--A-A-----------A------------------------------------G-----G--A---------------T-------T------A-A--G-- 1807
68_01B.CN.11.ANHUI_WH73       --A--A--G-----A---------------------------------------...-----T--A-A-----------A------------------------------------G-----G--A---------------T-------T------A-A----- 1798
70_BF1.BR.10.10BR_PE004       -----A--GA---R-----T----------------------------------...R----C--A-A---A--RRG-A--R------------R---R-R-----R--R-----------------A-------A-T--------------T---R----C--G-- 1920
71_BF1.BR.10.10BR_PE008       -----A--G-------------------------G-------------------...-----C--AGA-----------A------------------------------G--------------------A---------GA-C--G-----------T--------C--G-- 2128
72_BF1.BR.10.10BR_MG002       -----A--G---------------------------------------------...-----C--A-A-----------A-----------------------------------------------A------------A---------------------A----- 1972
O.BE.87.ANT70                 -----C--C---TGT----T-T-GC-G---------------C------C--...A-CA-C--ACAA---GA-----A-----------------------T--A-----------A-------A-G-A------A-TG-GACAG-------AGAA---AGGG-G-T 2518
O.CM.98.98CMA104              -----C--T---TGT----TCT-GC-G-------G-------------------...A-T--C--ACAA---GA-----A--------------------A----A--G----G---A-GG----A-G-A-T-A--A-TG--ACAG-------AGAA---AGGG---T 1950
O.CM.98.98CMABB141            -AA--C--T---TGT----T-T--C-G-------G--------------C--...A-CA-C--ACAA---GA-----A----C-----------------T--A----------A-------A-G-A---AC-A-TG-GACA------G-AGAG--AGA----T 1949
O.CM.98.98CMABB212            -----C--T---TGT------T-GC---------G-------------------...A-TA-C--ACAA--AGA----------C-----------------T--A-----C----A--G----A-G-A---A-CA-TG-GACAG-----G-ACAA---A-GG---T 1954
O.CM.98.98CMU5337             -----C--T---TGT----T-T-GC-G-------G--------------GC--...A-CA-C--ACAA---GA-----A----C-----------------T--A-----C--C-A----G-A-G-A---A--A--G-GGCAG-------AGAG--AGGG-R-T 1951
O.CM.99.99CMU4122             -----C--T---TGT----T-T-GC-G--C----G---------------C--...A-CA-T--ACAA---GA-----A----C-----------------T--A---------------A-------A-G-A---T--R-TG-GAAAG----G-AGAG--AGAG---T 1949
O.FR.92.VAU                   -----C--G---TGT--G-T-T--C-G-----G--G--------------C--...A-CA-C--ACAA---GA-----A---GC-----------------T--A---------------A--G----A-G-A---T--A-TG-GGCTG-----T-AGAG--AGAG---T 2030
O.GA.11.11Gab6352             -----C--T---TGT----T-T-GC-G--C----G-------C-------C--...A-CA-T--ACAA---GAG--GGA----C-----------------T--A-----A----A-----A-G-A--CA--A-TG-GACAG-------AGAG--AGG----T 1675
O.SN.99.99SE_MP1299           -----C--T---TGT----T-T--C-G-------G---------------C--...ACCA-C--ACAA---GA-----A----C-----------------T--A---------------A--G----A-G-A---A---AG-G-CAG-------AGAG--AGGG---T 2517
O.US.10.LTNP                  -----C--T---TGT----T-T--C-G-------G---------------C--...A-CA-C--ACAA---GA----------C-----------------T--A----A-----A----A-------A-G-A-----CA-TG-GGCAG-------AGAA---AGGG---T 2439
N.CM.02.DJO0131               --AAAAG--G---GA--------TC-----------T-----------GA--...------G--ACAA--AGA-----A-----G-------------A-----------A-------------------A------A-TG--ACA-----C--ACAA--AGG----- 1970
N.CM.02.SJGddd                ----AAG--G---GA--------T------------T------------A--...----------ACAA--AGAG--------------------------A-----------A-------------------A------A-TG--ACAG-------ACAG--AGA----- 1954
N.CM.04.04CM_1015_04          --AAATG--G---GA--------T--R---------T------------A--...----------ACAA--AGA---------------------------A-----------A-------------------A------A-T---ACA-----C--ACAA--AGG----- 1969
N.CM.06.U14296                --AAAAG--G---GA--------T------------T------------A--...-------T-ACAA--AGA----A-----------------------------------A-------------------A------A-TG--ACAG----C--ACAA--AGG----- 1967
N.CM.06.U14842                --AAAAG--A---GA--------T------------T------------A--...-------T-ACAA--AGA----A----------------------A-----------A-------------------A------A-TR--ACA-----C--ACAA--AGG----- 1971
N.FR.11.N1_FR_2011            --A-A-G--GA--TA--------T------------T------------A--...------CT-ACAA--AGA----------G----------------------------A-------------------A------A-T---TCAG-G--C--ACAG--AGA----- 1865
P.CM.06.U14788                --------TA-TTGT--G-----------------G--T-----C------G--...--TA-CT-ACC---AGA-------------------------------A--R-------A------A-G-A-T---AG-AG-TAAGR-------AGAA---AGAC---T 1935
P.FR.09.RBF168                --------TA-TTGT--G---T--------------T-----C------G--...--TA--T-ACC---AGA---------------T-------A----A--G-------A-C-----A-S-A-T---AGWAG-TAAGG-------AGAA---AGAG---T 2499
CPZ.CD.06.BF1167              CA---A--GG-CTGCAC---C--T-----C-----G--T------------T...-C---CC-T-AG--AAA----TC---C--C--T-GG-CA--T--A-----A--G--A---G--CCT---CA-G-A---T-CA-TG---CA--TACC--AGCA--GA-A-C--T 2566
CPZ.CM.05.SIVcpzMT145         -AA--A---T---TT-----CT-GC----C-----G--T-----C-----A-T...-----CT-A-A----------A-----GT----------------------------------T------A---A-T---AAGAG--AA---------AGAA---A-A---GAT 2006
CPZ.GA.88.GAB1                -AA---------TGT------T-GC-----------T------------A--...--GAG---ACAA--A-A----CTT---------------------A---------------------C-A---A-T----A-TG---A------G--AGAA--GAGA----T 2525
CPZ.TZ.06.SIVcpzTAN13         CAA--A--T--CTGTC-------G-----C-----T--T-----C-GT--T--C...TGCA-TT-ACAGC-AAA----CAG------------CC-------T--A--G--G-----ACC------TG----CT--A-T--T-CGG-GC----TG-CAATA-AG---T 2105
CPZ.US.85.US_Marilyn          -CA-----GA--GTA--------T-----C--------T------------C--...--CA-C--ACAAA-TGA---G-C-----G-----------G-------------------A-------A---A--------CG-CAAT--------AGAG---AGA----- 2524
GOR.CM.07.SIVgor2139_287      -----A--T---TGT--G---T-G-----C-----G--T------------G--...---C-CT-ATC---AGA---T--G---------G--------A--C--A----A----T-------A-G-A-----AG-TG--GAA--------AGAACATAGG----T 2001
GOR.CM.07.SIVgorCP2135con     -----A--T---TGT--G---T-G-----C-----G--T------------G--...---C-CT-ATC---AGA---T--G---------G--------A--C--A----A--C--T-------A-G-A-----AG-TG--GAA--------AGAACATAGA----T 2000
```

Pol Protease end Pol p66 and p51 RT start

```
B.FR.83.HXB2                  TATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCCATTAGCCCTATTGAGACTGTACCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAG 2634
Pol                           __I__G__T__V__L__V__G__P__T__P__V__N__I__I__G__R__N__L__L__T__Q__I__G__C__T__L__N__F__P__I__S__P__I__E__T__V__P__V__K__L__K__P__G__M__D__G__P__K__V__K__Q__W__P__L__T__E__
A1.AU.03.PS1044_Day0          ---------------------------------------------------CA--W----Y--------C--T-----------Y--A-----T--------------------------------------------G-----G---------------C------C-- 1842
A1.CH.03.HIV_CH_BID_V3538     ---------------------------------T------------------A----------------C--T--------------A-----T--C-----A-------------C---G-----------------------G------------------------- 1839
A1.CY.08.CY236                ----------------------------------------------------A-------C-----------T--C--------C--A-----T--------------------------------------------T----G----------------C--------- 1845
A1.ES.05.X1608_8              ----------------------------------------------------A-------C-----------T-----------C--A-----T--------------------------------------------------G------------------------- 2082
A1.KE.11.DEMA11KE001          ------------------------C--------------------------CA-------C-----------T--C--------C--A-----T--------------------------------------------------G-----G------------------- 1973
A1.RU.11.11RU6950             ---------G-----------------C------------------------A-----------C----------------------A--A--T--------A-------------CG-----------------------C--G-----------------A------- 2158
A1.RW.11.DEMA111RW002         ------------------------------------------------G---A-------------------T--------------A--C--T-----------------------------------------------------A---------------------- 1985
A1.SN.01.DDI579               ---------------G--T-----------C--------------------CA-------------------T--------------A-----T------A-A-------------C---------------------A------------------------------- 1832
A1.UG.11.DEMA110UG001         ------------G--G------------------------------------A-------C-----------T--------------A-----T--------C-------------C------A--------------------G------------------------- 1989
A1.ZA.04.04ZASK162B1          A--------------G------------------------------------A-------------------T--C-----------A-----T------A-A-------------------G---------------------G------------------------- 2075
A2.CM.01.01CM_1445MV          C---------------------------------------------------A-A---GT----C-------T--------------A-----T--------A-----------------------------------------G------------------------- 1833
A2.CY.94.94CY017_41           C--------------------------C--------------C---------A-----GT----C-------T--------------A-----T--------A-----------------------------------T-----G------------------------- 1991
B.BR.10.10BR_MG029            ------------------------------------------------------------------------T--C--------C-----C--T--------A------------------------------------------------------------------- 2096
B.CA.07.502_1191_03           ---------------------------------------------------CT-----------C----A--T--C-----------------T--------A------------------------------------------------------------------- 2061
B.CH.08.M2_0803101_NFLG8      -------------C--A-------C--------------------------CG--A----C---C----------------------------T--------A--------------------A---------------------------------------------- 2103
B.CN.10.DEMB10CN002           ------------G---A-------------------T------------------A--------C----C-----------------------T--------A------------------------------------------------------------------- 2006
B.ES.10.DEMB10ES002           ----------------A----------------------------------C--AA-------------A--T--------------------T--------A-----------------------------------------------G------------------- 1996
B.FR.11.DEMB11FR001           ------------------------------A--------------------------------------------------------------T--------A-------------G-C--------------------------------------------------- 1880
B.GB.05.MM45d213_GN1          ----------------------------------------------------T--------G-------G--------G--------------T-----------------------------------------------------------------C---------- 2035
B.HK.06.HK003                 C---------------A----------------------------------------------------------------------T-----T--------A------------------------------------------------------------------- 2004
B.HT.05.05HT_129389           ---------------------------------------------------------------------------------------------T--------A--------------------------R-----------------W--------------R------- 1835
B.JP.12.DEMB12JP001           ---------------------------------T-----------------------A------C----C--------------C--------T--------A--------------------------------------------G---------------------- 2021
B.KR.07.HP_18_07JHS10_3909    --C------------------------------------------------------------------C-----------------------T--------A-----------------------------------------G------------------------- 2298
B.PE.07.502_0525_wg5          --CG--------G------------------A--------------------T-----------C----------------------------T--------A--------C-------------------------------G-------------------------- 2099
B.RU.11.11RU21n               ------------------G------------------------------------------------------------------------------------T--------A----------------------------------------G----------------C--------- 2205
B.TH.08.MERLBDTRC10           ----------------A----------------A------------------T-----------C-----------C----------------T--------A------------------------------------------------------------------- 1867
B.US.11.ES38                  ----------------A-------------A--------------------------------------------C-----------------T--------A---A--------------------------------------------------------------- 2569
C.AR.01.ARG4006               ---------------------G-----------------------------CA-----------C----A-----AC-C--------A-----T--C-----A-----------------------------------------G--C---------------------- 1818
C.BR.07.DEMC07BR003           ------------GC----G--------------------------------CA-----------C----A-----AC----------A-----T--C-----A-----------------------------------------G--C------------C--------- 1966
C.BW.00.00BW5031_1            ----------------------------------------------------A----------AC----A--T--AC----------A-----T--------A-----------------------------------------G--------G--------A------- 2005
C.CN.10.YNFL19                ------------------------------------------------------A---------C----------------------------T--------A-----------------------------------------G------------------------- 1987
C.CY.09.CY260                 -------T-R-----G--------------------------------G---A-------G---C----M-----AC----------A-----TT-C-----A--------------------------G--------------G------------------------- 1857
C.ES.08.X2363_2               -------T--------------------------------------------A-----------C-A--A-----AC----------A--------C-----A-----------------------------------------G------------------------- 2065
C.ET.02.02ET_288              -------------C--------------------------------------------------C----A-----AC----------A-----T--C-----A-----------C--------------------------------C--------------A------- 1836
C.IN.09.T125_2139             ------------------------C---------------------------A-----------C----A-----AC----------A--C--T--------A---A----------------------G--------------G--C---------------------- 2610
C.KE.00.KER2010               ---------------------------------T------------------A-A---------C----A-----AC----------A-----T--C-----A-----------------G-----------------------G-----------------------G- 1815
C.MW.09.703010256_CH256.w96   ---------------G-----------------------------------CA-----------C----A-----CC----------A-----T--C-----A-----------------------------------------G------------------------- 2614
C.TZ.08.707010457_CH457.w8    -------T----------G-----A-----G---------------------A-----------C----A-----AC----------A-----T--C-----------------G--------------------------------------------------T--G- 2608
C.YE.02.02YE511               ------------------------A---------------------------A--C--------C----A--T--AC----------A--------C-----A-----------------G--A-------------------GG------------------------- 1809
C.ZA.08.705010534_CH534.w12   -------G----------G------------A----------A---------A-----------C----A-----AC----------A-----T--C-----A------------------------------------------------------------------- 2624
C.ZA.10.DEMC10ZA001           ---------------------G-----------------------G-----CA-----------C----A--T--AC----------A-----T--------A------------C----G-----------------------G------------------------- 1974
C.ZM.11.DEMC11ZM006           ----------------------------------------------------------------C----A--T--AC----------A-----T--C-----A-----------------------------------------G--C---------------------- 1983
D.CM.10.DEMD10CM009           ---------------------------------------------------CT-------C--------C------C----------A-----T--------A---A--------------------------------------------------------------- 2009
D.CY.06.CY163                 ------------------------G--------------------------CA-------C--------C-----CC----------A--------------A--------------------A---------------------------------------------- 1842
D.KE.11.DEMD11KE003           ----------------------------------------------------T----------A-----------A-----------A-----T--------A--------------------------G--------------------G------------------- 1984
D.KR.04.04KBH8                ------------------------------------------------G--------------------C-----------------A-----T-----------------------G----------TG--------G----G-------------------------- 2593
D.SN.90.SE365                 ---------------G--------------C--------------------CA-------------------T--------------A-----T------------A----------------------G-----C-----------------------G--A------- 2651
D.TZ.01.A280                  ------------G--------------------------------------CT----------------C--T--------C-----------T-------CC-----------------G--------G-----------------------------G---------- 1835
D.UG.10.DEMD10UG004           ------------G--------------------------------------CT----------------------A-----------A-----T--------A--------------------------G--------T--------------G---------------- 1972
D.UG.11.DEMD11UG003           ----------------A-------------------T-----------G---T----------------------------------A-----T--------A--------------------------G---------------------------------------- 1989
D.YE.02.02YE516               ---------------------------------------------------------------------A--T---C-------C--A-----T--------A------------------------------------------------------------------- 1845
D.ZA.90.R1                    ----------------A----------------------------------------------AC----C-----------------A-----T--------A--------------------------G---------------------------------------- 1980
F1.AO.06.AO_06_ANG32          --C---------G---A-----------------------------------A----------------------------------A--------------A---A-------------G--A--------------T-----G------------------------- 1845
F1.AR.02.ARE933               --C-T-------G--G------------------------T------T----A-------------------T------C-------A-----T------A-A--------C--------G-----------------------G------------------------- 1928
F1.BR.10.10BR_RJ015           --C---------G--------------G------------------------A-------------------T--------------A-----T--------A-----------------G-----------------------G------------------------- 2107
F1.CY.08.CY222                --C---------G--------C--C--G------------------------A-------------------T--------------A-----T--------A-----------G-----G-----------------------G------------------------- 1827
F1.ES.02.ES_X845_4            ------------G-----------------A---------------------A----------------------------------A-----T--------A--------------------A--------------------G------------------------- 2086
F1.RO.96.BCI_R07              --C---------G-----G--------G------------------------A----------------------------------A-----T--------A---------------C-------------------------G------------------------- 2658
F1.RU.08.D88_845              ---------------G------------------------------------A----------------------------------A--A--T--------A-----------------G----------------------GG------------------------- 2093
F2.CM.02.02CM_0016BBY         ---------------------G-----G------------------------A----------------------------------A-----T--------A-----------------------------------------G------------------------- 1824
F2.CM.10.DEMF210CM001         ---------------G-----------G--------------------G---A----------------------------------A-----T--------A---------------C-------------------------G---------------C--------- 1898
F2.CM.10.DEMF210CM007         ---------------------G-----G--------------------G--CA-----------------------C----------A-----T--------A-----------------------------------G----GG------------------------- 1981
G.BE.96.DRCBL                 ------G------------------A----T--------G------A----------------------------------A-----T--------A-----G-----------------------------G---CGG------------------------- 2598
G.CM.10.DEMG10CM008           ------G--------------------------------------G------A-------------------T--------------------T--------A-----C-----------------------------A-----G-----------------------G- 2015
G.CN.08.GX_2084_08            ------G------C-----------------A--------------------A-------------------T--------------A-----T--------------------------------------------------G------------------------- 1874
G.CU.99.Cu74                  ------A-----T------------------A-------------G------A-------------C-----T--------------A-----T--------A-----------------------------------------GA----------------A------- 2238
G.ES.09.X2634_2               C-----G------------------------A--------------------A----------A--------T---C----------A-----T--------A-----C-----------------------------T-----G-----------------A------- 2117
G.GH.03.03GH175G              ------G------------------------A-------------G------A----------------C--T--------------A--A--T--------A-----C----------------------------------GGA----G------------------- 2674
G.KE.09.DEMG09KE001           ------G------------------------A-------------G------A----------------C--T--------------A-----T--------A-----------------------------------------G------------------------- 2009
G.NG.09.09NG_SC62             ------A--T---C-----A-----------A-------------G------A-------C--------C--T--------------A-----T--------------------------------------------------G---------------------A--- 1826
G.ZA.01.TV546                 ------G------------------------A----------C--G--G---A-A-----A-----------T---C----------A--C--T--------A-----------------------------------------G-----------------A------- 1836
H.BE.93.VI991                 ------A-----------------------------T-----------G---A-A--------A--G--------C--------G--A-----T--------A-------------C---------------------------G---------------C--------- 2023
H.CF.90.056                   ------------------------------------T-----------G---A-A--------A-----------C-----------A-----T--------A-----------------------------------------G------------------------- 1981
H.GB.00.00GBAC4001            ------------------------------------T-----------G---A-A--------A-----------C-----------A-----T--------A-------------C---------------------------G--------------------G---- 2152
J.CM.04.04CMU11421            Y--R--------------------------------------------G--CA-----------C----C--T--------------A--Y--T--------A--------------------A--------------S-----------------------------R- 2140
J.SE.93.SE9280_7887           ------------G---A-------------------------------G--CA-----------C-------T--------------A-----T--------A------------------------------------------A------------------------ 1948
J.SE.94.SE9173_7022           ----------------A----------------------------------CA-----------C-------T--------------A-----T--------A------------------------------------------A------------------------ 1949
K.CD.97.97ZR_EQTB11           ------------G--------------------------------------CA----------------G--T--------------A-----T--------A-----------------------------------------G------------------------- 1830
K.CM.96.96CM_MP535            ----------------------------------------------C----C--------------------T--------------A-----T--------A-----------------------------------------G-----------------A------- 1830
U.CA.01.TV749                 ------------------------------------T-----------G--CA-------C-----------T-----------C--A-----T--------A----------------------------------------GG------------------------- 2651
U.CA.99.TV721                 ---------------------------------T--T-----------G--CA-------A-----------T-----------C--A-----T-A------A----------------------------------------GG------------------------- 2650
U.CD.83.83CD003_Z3            ----------------------------------------------------A----------A-----------------------A-----T--------A-----------------------------------------G--G---------------------- 2129
U.CD.90.90CD121E12            -------------C-------G------------------------------A----A-----------------------------A--C--T--------A-----------------------------------------G--A-------------A------- 1833
U.CY.05.CY090                 ---------G--G------A----C---------------------------A-------C---------------C----------A-----T--------A--C--------------------------------T-----GA----------------A------- 1833
U.CY.08.CY223                 ------------------------A--------------------G--G---A-------------------T---C----------A-----T--------A-----------------------------------G-----G-----------------A------- 1851
U.ES.10.DEURF10DZ001          ------------G---------------------------------C----CA-------A-----------T--------------A-----T--------A-----------------------------------------GA------------------------ 1910
U.GR.99.99GR303               ----------------------T-----------------------C-----G-------C-----------T-----------C--A-----T--------A------------GC------------G--------T-----GA----------------R-Y----- 1908
U.NL.95.U_NL_95_H10986_D1     ------------------------------------------------G--CA-------------------T--------------A-----T--------A-----------------------------------------------G-----------A------- 2519
01_AE.AF.07.569M              ----------------------------------------------C-----A-------------------T-----------C--A-----T------A-C-------------C---------------------G--C--G------------------------- 1840
01_AE.CF.90.90CF11697         ----------------------------------------------C-----A-------------------T-----------C--A--A--T--------C-------------C---------------------A-----G----------------A-------- 2583
01_AE.CN.10.YNFL03            ---------------------------T-----------------GC-G---A-----------C-------T--------------A-----T--------C-----------------------------------A-----G------------------------- 2012
01_AE.HK.04.HK001             ---------------------------------------------GC-----A----------------C--T-----------C--A--C--T--------C-----------------------------------A-----G--------G---------------- 1996
01_AE.IR.10.10IR.THR48F       -G-------------G------------------------------C-----A-------------------T---C-------C--A-----T--------C-------------C------------T--------A-----G--C-----G---------------- 1824
01_AE.JP.x.JRC77AE            ---------------------------------------------GC-----A-------------------T-----------C--A-----T--------C-------------C------A-----G--------A-----G--------G---------------- 2631
01_AE.TH.09.AA090a_WG11       ------------G---------------------------------C-----A-------------------T-----------C--A-----T--------C---A---------C---------------------A-----G--------G---------------- 2057
01_AE.TH.90.CM240             ----------------------------------------------C-----A-------------------T-----------C--A-----T--------C-------------C---------------------A-----G--------G---------------- 2208
01_AE.US.05.306163_FL         ------------G---------------------------------C-----A-------------------T-----------C--A-----T--------C-------------C------------G--------A-----G--------G---------------- 1825
01_AE.VN.98.98VNND15          ------------G---------------------------------C-----A-------------------T-----------C--A-----T--------C-------------C---------------------A-----G--------G---------------- 1889
```

```
                                                                                   Pol Protease end  Pol p66 and p51 RT start
B.FR.83.HXB2                 TATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCCATTAGCCCTATTGAGACTGTACCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAG 2634
Pol                          __I__G__T__V__L__V__G__P__T__P__V__N__I__I__G__R__N__L__L__T__Q__I__G__C__T__L__N__F__P__I__S__P__I__E__T__V__P__V__K__L__K__P__G__M__D__G__P__K__V__K__Q__W__P__L__T__E__
02_AG.CM.10.DE00210CM013     ------------G---------------------------------C-----A-------------------T--------------A-----T--------A-----------------------------------T-----G------------------------- 2001
02_AG.CY.09.CY256            C---------------------------------------------C-----A-R--------A--------T--------------A-----T--------A--C--G-----------G----------------------GG------------------------- 1830
02_AG.ES.06.P1423            ------------G--------------W------------R-----C-----A-------------------T-----------C--A-----T--------AR--------------C---M---------------------G-------G------R---------- 2072
02_AG.GW.05.CC_0048          C-----------------------C--------------------GC-----A----------A--------T--------------A-----TT----C--A------------------------------------------------------------------- 1984
02_AG.KR.12.12MHR9           C---------------------------------------------C-G--CA----------------G--T--------------A--------------A-----G-----------------------------T-----G------------------------- 2293
02_AG.LR.x.POC44951          C-----------------G---------------------------C-----A-----------------------------C----A-----T------A-A-----G--T--------------------------------------G------------------- 2629
02_AG.NG.09.09NG_SC61        C---------------------------------------------C-----A-------------------T--------------A-----T--------A-----G-------G---------------------------G--C---------------------- 1826
02_AG.NG.x.IBNG              C---------------------------------------------C-----A-------------------T--------------A-----T--------A-----------------------------------------G------------------------- 2159
02_AG.SE.94.SE7812           C---------------------------------------------C-G---A-------------------T--------------A-----T--------A--C-----------------------C--------------G------------------------- 2006
02_AG.SN.98.98SE_MP1211      C--G---T---------------------------------------A----A-T--------A--------T--C-----------A-----T--------A-----G--T--------------------------------G------------------------- 1831
03_AB.RU.97.KAL153_2         ---------G-----------------C------------------------A-----------C-------T--------------A--A--T--------A-------------C---------------------------G-----------------A------- 1860
04_cpx.CY.94.94CY032_3       C-----C-----G---------------------------------C----CA-----------C-------T--------------A-----T--------A-----------------------------------------G------------------------- 2000
05_DF.BE.x.VI1310            ----------------------------------------------------T-------C--------C-----------------A--------------A-----------------------------------------------------------A------- 2016
06_cpx.AU.96.BFP90           -------------------------------A--------------C----CA-------C-----------T--------------A-----T--------A---------------C----------G--------------G------------------------- 2662
07_BC.CN.98.98CN009          ----------------------------------------------------------------C----------------------A--C-----C-----A-----------------------------------------G------------------------- 1966
08_BC.CN.97.97CNGX_6F        ------------------G-----C-----------------A---------A-------C---C----A-----AC----------A--C--T--C-----A-----------------------------------------G------------------------- 1814
09_cpx.GH.96.96GH2911        G---------------------------------------------C-G--CA----------A--------T--------------A-----T--------A------------------------------------------A-------------C---------- 1836
10_CD.TZ.96.96TZ_BF061       ---------------------------------------------G------T----------------------------------A-----T--------A--G-----------------------------------------------------G---------- 2013
11_cpx.CM.95.95CM_1816       ------------G-----------------C-----------------G--CA----------------A--T--------------A-----T--G-----C-----------------------------------------G------------------------- 1998
12_BF.AR.99.ARMA159          ------------G---A-------------G---------------------------------C----------------------------T--------A------------------------------------------------------------------- 2639
13_cpx.CM.96.96CM_1849       ----------------A-----------------------------------A-------------------T--------------A-----T--------A------------------------------------------------------------------- 2039
14_BG.ES.05.X1870            ------G------------------------A-------------G------A-------------------T--------------A-----T--------A-----------------------------------T-----GA------------------------ 2101
15_01B.TH.99.99TH_MU2079     ------------G---------------------------------C-----A----------AC-------T-----------C--A-----T--------A-----------------------------------A-----G--------G---------------- 2033
16_A2D.KR.97.97KR004         C---------------------------------------------------A-----GT---AC-G-----T--------------A--A--T--------A--------------------------------------------C---------------------- 1997
17_BF.AR.99.ARMA038          ---------------------------------T-----------------------------AC----------C-----------------T--------A-----------------------------------------------G------------------- 1848
18_cpx.CU.99.CU76            ----------------------------------------------------A-------------------T--------------A-----T--------A-----------------------------------------G------------------------- 1939
19_cpx.CU.99.CU7             ---------------------G------------------------------T----------------C-----------------A--------------A--------------------------G-----------------G---------------------- 1860
20_BG.CU.99.Cu103            ------G------------------------A-------------G------A-------------------T--------------A-----T-----C--T-----G--------------------------------------G---------------------- 2100
21_A2D.KE.99.KER2003         ------C-----G---------------------------------------T-------C--------C-----------------A--A-----------A-----------C--------------G-----------------------G---------------- 1839
22_01A1.CM.01.01CM_0001BBY   ------------G--------------------T------------C-----A----------------C--T--C--------C--A--A--T--------A-----------GC----------------------A-----GA------------------------ 1842
23_BG.CU.03.CB118            ------G------------------------AC------------G------A-------------------T--C-----------A-----T-----C--A--------------------------------------------G---------------------- 2077
24_BG.ES.08.X2456_2          ------G------------------------A-------------G------A-------------------T--C-----------A-----T--------A-----G--------------------------------------A--------------------G- 2085
25_cpx.CM.02.1918LE          ------G------C-----------------A--------------------A-A--------------C--T--A-----------A-----T--------A-----------------------------------------G------------------------- 1842
26_AU.CD.02.02CD_MBTB047     ------------------------------------------------G---A----------------C--T--------------A-----T--G-----A---------------C----------G--------------G---------------C--------- 2639
27_cpx.FR.04.04CD_FR_KZS     C-----------G-----------------------T-----------G---T-A--------------------------------A-----T--------A-----------------------------------------G-----G------------------- 2625
28_BF.BR.99.BREPM12609       --C---------G--------------------------------------CA-------------------T--------------------T--------A------------------------------------------------------------------- 2000
29_BF.BR.01.BREPM16704       ------------G--G------------------------------------A-------------------T--------------------T--------A--A-----------------------------------------------G---------------- 2035
31_BC.BR.04.04BR142          ------------------G--------------------------------CA-----------C----A-----AC----------A-----T--C-----A-----------------------------------------G--C---------------------- 2106
32_06A1.EE.01.EE0369         ----------------------------------------------C----CA-------C-----------T-----------G--A--A--T--------A----------------------------------------GG------------------------- 2269
33_01B.ID.07.JKT189_C        ----------------------------------------------C-----A-C-----C-----------T--------------------T--------A-----------------------------------T--------------G---------------- 1935
34_01B.TH.99.OUR2478P        ----------A----------------------------------------------A------C-------T--------------------T--------A--------------------------------------------------------G---------- 1836
35_AD.AF.07.169H             ---------------G------------------------------------A-------C-----------------------C--A-----T--------T-------------GG--------------------------G------------------------- 1834
36_cpx.CM.00.00CMNYU830      ------------G------------------A--------G----GC-----A-------------------T--------------A-----T--------A-----------GC----------------------A-----G------------------------- 1839
37_cpx.CM.00.00CMNYU926      ---------------G--------C---------------------C-----A-------------------T--------------A--C--T------A-C-----------TC-------------G--------------G------------------T------ 1824
38_BF1.UY.03.UY03_3389       --C-------------A----------------A------------------------------C----------------------------T--------A--------------------------------------------------------------C---- 2057
39_BF.BR.04.04BRRJ179        --------------------------------C---T-----------------C--------------C-----------------------T--------A----------------------------------------G-------------------------- 2111
40_BF.BR.05.05BRRJ055        ------------------------------G--T--T------------G----C--------------C-----C-----------------T--------A------------------------------------------------------------------- 2117
42_BF.LU.06.luBF_18_06       --C-------------A-------------------T--------------C---C--------C----C--------------C--------T--------A-----------------------------------------G------------------------- 2188
43_02G.SA.03.J11223          ------G------C-----------------A-------------C------A-------------------T--------------A-----T--------------------------------------------------G----------------------C-- 2159
44_BF.CL.00.CH80             ------------------------------------------------------------------------------------------C--T--------A----------------------------------------------------------------G-- 2076
45_cpx.FR.04.04FR_AUK        -------G----------G-----C--------T------------C-----A----A--------------T--------------A-----T--------A-----------------------------------------G------------------------- 2620
46_BF.BR.07.07BR_FPS625      ---------------G-----------G------------------------A-------------------T--------------A-----T--------C-----------------G----------------------GG------------------------A 2072
47_BF.ES.08.P1942            ---------------------------------------------------------------A-----------------------------T--------A-----------------G-----------------------G--C--------------A--G---- 2073
48_01B.MY.07.07MYKT021       ---------------------------------------------------------------------------------------------T--------A------------------------------------------------------------------- 1839
49_cpx.GM.03.N26677          -G----------G--------------------------------------CA-----------C-------T--C-----------A-----T--------A--------------------A--------------T------------------------------- 2057
50_A1D.GB.10.12792           ---C------------A----------------------------------C--------C-----------T-----------C--A-----T--------------------------------------------------G------------------------C 2042
51_01B.SG.11.11SG_HM021      ------------G-----------------------------------G-----------C-----------T--C-----------------T--------A--------------------A---------------------------------------------- 1841
52_01B.MY.03.03MYKL018_1     ------------G---------------------------------C-----A-------C-----------T-----------C--A-----T--------C-----------------------------------A-----G--------G---------------- 1948
53_01B.MY.11.11FIR164        ---------------G------------------------------C-----A-----------C-------T-----------C--A--Y--T------A-T-----------------------------------A-----G------------------------- 1974
54_01B.MY.09.09MYSB023       ------------T--------------------T--T---------------------------C----------------------------T--------A------------------------AAC-----------G-G---------------T---------- 2036
55_01B.CN.10.HNCS102056      ------------G---------------------------------C-----A-------------------T-----------C--A-----T--------C-----------------------------------G-----G------------------------- 1944
56_cpx.FR.10.URF5_patient_A  C---------------------------------------------C-G---A-------------------T--C-----------A-----T-A------A-------------GC--------------------A-----G------------------------- 1985
57_BC.CN.09.09YNLX19sg       ------------------------C---------------------------------------C----G-----------------------T--------A--------------G--G-----------------------G------------------------- 1818
58_01B.MY.09.09MYPR37        ---------------G-----G------------------------------------------M----------------------------Y--------A------------------------------------------------------------------- 1979
59_01B.CN.09.09LNA423        ----------------------------------------------C-----A-------------------T--------C-----------T--------A--------------------------G---------------------------------------- 1818
60_BC.IT.11.BAV499           ------------G--------------------------------G-----CA-A---------C----A-----AC----------A-----T--------A-----------------------------------------G--C---------------------- 2591
61_BC.CN.10.JL100010         ------------T--------------------------------------C------------C----------------------T-----T--------A-----------G-----------------------------G------------------------- 2055
62_BC.CN.10.YNFL13           ---------------------T--C--------------------G------A-------C---C----G-----AC----------A--C--T--C-----A--------------G--------------------------G--------------C---------- 1963
63_02A1.RU.10.10RU6637       C-----------------------------------T--------GC-----A----A--C-----------T--------------A--A--T--------A-----------------------------------------G------------------------- 2175
64_BC.CN.09.YNFL31           ----------------A-----------------------------------------------C----G-----------------------T--------A-----------------------------------------G------------------------- 1988
65_cpx.CN.10.YNFL01          ------------------------C---------------------------A-A---------C----A-----AC----------A--C--T--C-----T-----------------------------------------G------------------------- 2005
67_01B.CN.11.ANHUI_HF115     ----------------------------------------------C-----A-------------------T------TG---C--------T--------C-----------------------------------A-----G--------G---------------- 1977
68_01B.CN.11.ANHUI_WH73      ----------------------------------------------C-----A-------------------T--------------------T--------C-----------------------------------T-----G--------G---------------- 1968
70_BF1.BR.10.10BR_PE004      --C---------G--GA-----------------------------------A-------C--------C--T--A-----------A-----TT-------------------------G-----------------------G------------------------- 2090
71_BF1.BR.10.10BR_PE008      ------------G--G--------C--G--------------------G---A----------------A--T--------------A-----T--------A-----------------G-----------------------------------------A------- 2298
72_BF1.BR.10.10BR_MG002      ------------G--------------------------------------------------------------------------------T--------A-----------------------------------------G------------------------- 2142
O.BE.87.ANT70                ACAG--A--------G--G--------T-----T--T--T------------A-A-----AGGAT-A-----T--AC----C--C--T--A-----C--A-CCC-A--G---------C----A--------------A--------A-----------CC-AT-TA--- 2688
O.CM.98.98CMA104             ACAG--A--------G-----------T-----T-----T-----------CA-A-----AGGAT-A--------AC-------C--T--A---T-C--A-CCC-A----------G-C----A--------------A--------A-----------CC-AT-TA--- 2120
O.CM.98.98CMABB141           ACAG--A--------G-----------T-----T-----C------------A-A-----AGGAT-A-----T--AC-------C--T--A----AC--A-CCC-A--G---------C----A--------------A------A-A-----------CC-AT-TA--- 2119
O.CM.98.98CMABB212           ACAG--A--------G--G--------T-----T--T--TC----G------A-A-----AGGAT-A--A--T--AC-------C--T--A--T--C--A-CCC-A-----------GC----A--------------A--------A-----------CC-AT-TAG-- 2124
O.CM.98.98CMU5337            AC----A--------G--G--------C-----T--T--T-----------CA-A-----AGGAT-A-----T--AC-------C--T--A-----C--A-CCC-A--G---------C----A-----M--------A--------A-----------CC-AT-TA--- 2121
O.CM.99.99CMU4122            ACAG--A--------G--G--------T-----T--T--T------------A-A-----AGGAT-A--C--T--AC-------C--T--A-----C--A-CCC-A--G--------RC-G--A--------------A------A-A-----------CC-AT-TA--- 2119
O.FR.92.VAU                  A-GG--G--------G--G--------T-----T-----TC-A---------A-A-----AGGAT-A-----T--AC-------C--T--A-----C--AACCC-A------------C----A--------------A--------A-----------CC-AT-CAG-- 2200
O.GA.11.11Gab6352            GCAG--AGT------G--G--------T-----T--T--T------------A-A-----AGGAT-A-----T--A-----C-----T--A-----C--A--AC-A--G---------C----A--------------A--T-----A-----------CC-AT-TA--- 1845
O.SN.99.99SE_MP1299          AC-G--A--------G--G--------T-----T--T--T-----------CA-A-----AGGAT-G--------AC-------C--T--A--T--C--A-CCC-A------------C----A--------------A------A-A-----------CC-AT-TA--- 2687
O.US.10.LTNP                 ACAG--A--------G--G--------T-----T--T--C------------A-A-----AGGAT-G-----T--AC-------C--T--A-----C--A-CCC-A--G---------C----A--------------A------A-A-----------CC-AT-TAG-- 2609
N.CM.02.DJ00131              AG-T--------G-----------A--------T--T--T--A--------CT-T--A--A--------C--T--------------A--A--T--------A--------------------A-----G--------------G--A-----------T-------C-- 2140
N.CM.02.SJGddd               AG-T--------------------A--------T--T--T--A-----------T--A--A--------C--T--------------A--A--T--------A--------------------A--------------------G--A-----------T-------C-- 2124
N.CM.04.04CM_1015_04         AG-T--------G--G--------A--------T--T--T--A--------CT-T--A--A--------C--T--CC----------A--A--T------A-A--------------------A-----------------------A-----------T-------C-- 2139
N.CM.06.U14296               AG-T--------G-----------A--------T--T--T--A--------C--T--A--A--------C--T--------------A--A--T--------A--------------------A-------------------GG--A-----------T-------C-- 2137
N.CM.06.U14842               AG-T--------G-----------A--------T--T--T--A--------CT-T--A--A--------C--T--------------A--A--T--------A--------------------------G--------------G--A-----------TC------C-- 2141
N.FR.11.N1_FR_2011           AG-T--------------------A--------T--T--T--A-----------T--A--A--------C--T--------------A--A--T--------A--------------------A-------------------GG--A-----------T-------C-- 2035
P.CM.06.U14788               -TAT--R--T--------G--Y--A--C--A--A--T-----A---------A-T-----ATTA-----A--T-----G-G---C-----C-----C--A--A-TA-----------------AG-------------G--T-GG--G--G--------C---T--A-G- 2105
P.FR.09.RBF168               -TAC--A--T--------R--C--A--C--A--A--Y-----A---------A-T-----ATTA-----A--T---C---G---C-----C---T-C--A--A-TA--------R-W-------G------W------G--T--G--A--G-M------C--RT--A-R- 2669
CPZ.CD.06.BF1167             G-A-ACA-A----C--A-------A--C--A--A--T---------------G-T--A-GA--AC-------T---C-------C--T-----TA-AG-A--A-TA---AA---G-GC--C---GA------------G--T--G--A------------C-AT-CAG-- 2736
CPZ.CM.05.SIVcpzMT145        A-CCA----T--------------A--C--------T-----A---------G-CC-A--AA-AT-A--G--T---C-------C--A--A--T--A--A--C--A---AA---GTTTC----------G-----C--G--C--G--A-----G-----TC--T--A--- 2176
CPZ.GA.88.GAB1               AG----G-----------------A--------A--T---------------A-T-----A--AT-G-----T---C--GTG--C--A-----TT-A-----A-----C-----C-----G--A--------------T-----G--A--G--------C--AT---C-- 2695
CPZ.TZ.06.SIVcpzTAN13        A-G--CC--T--C-----G--GGA------CC-A--T-----A--------CA-T--ACAA--AT-G--C--T-----------C--------TAAAG----AGTA---AA----G--C----A-A---C-----C--G--------A-----------TC--T-TA--- 2275
CPZ.US.85.US_Marilyn         ACAG---T----T-----G-----A-----A--A--T--T-----C--G---A-T--A--A--A-----A--T--------------T--C---T-----A-A-------------G------A-----------------T--G--A--G----------------C-- 2694
GOR.CM.07.SIVgor2139_287     AG----G-----TC----T-----A--C-----A--T--T--A-----G---A-T--A--AA-A-----A--T--CC-------------A---T-C--A--TGTA-----C---TCT-----------G--------A--C-----C--G--------T---T--A--- 2171
GOR.CM.07.SIVgorCP2135con    AG----G-----T-----T-----A--C-----A--T--C--A-----G---A-T--A--AA-A-----A--T--C--------------A---T-C--A--TGTA-----C---TCT-----A-----G--------A--C-----C--G--------T---T--A--- 2170
```

```
B.FR.83.HXB2          AAAAAATAAAAGCATTAGTAGAAATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGAAAATTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAA 2804
Pol                   E  K  I  K  A  L  V  E  I  C  T  E  M  E  K  E  G  K  I  S  K  I  G  P  E  N  P  Y  N  T  P  V  F  A  I  K  K  K  D  S  T  K  W  R  K  L  V  D  F  R  E  L  N  K  R  T  Q
A1.AU.03.PS1044_Day0  ----------------AC----------C----------A-----------------------------------A-----T----------------C----------G----------G--C-----A---------- 2012
A1.CH.03.HIV_CH_BID_V3538 ----------------AC----------AT---------A-----------------------------------T----------T-C----------------C----------T-------C-----A---------G 2009
A1.CY.08.CY236        ----------------AC----------A----------G--A-----------A-----C----------A-----T----------G-----C----------G----------G--C-----A---------- 2015
A1.ES.05.X1608_8      ----------------AC----------T----------A-----------------------------------A-----T----------G-----C----------G----------G--C-----A---------- 2252
A1.KE.11.DEMA11KE001  ----------------AC----------T----------A-----------------------------------A-----A----------T-C----------G----------G--C--C--A----A---- 2143
A1.RU.11.11RU6950     -G--------------AC----------AG---------A----A-----------C----------------------T----------G-----G----------G--C-----A----------G 2328
A1.RW.11.DEMA111RW002 ----------------AC--C-----G---A----------A-----------------------------------T----------C--C----------G----------G--C-----A---------G 2155
A1.SN.01.DDI579       ----------------GAC--C-------AG--------G------A----------C----------------------T----------G-----G----------G--C-----A-----A--G 2002
A1.UG.11.DEMA110UG001 ----------------AC----------T----------C---GG----------G----------CA-----T----------C----------G----------G--C-----A---------G 2159
A1.ZA.04.04ZASK162B1  ----------------AC----------A----G------A-----------------------------T----------G----C--C------G----------G--C-----A---------- 2245
A2.CM.01.01CM_1445MV  ----------------AC----------A--C----------A-----------------------G----T----------C----------------------------C---------------- 2003
A2.CY.94.94CY017_41   ----------------AC------C---A--A----------A-----------C----------G----T----------C----------T----------C---------------- 2161

B.BR.10.10BR_MG029    ----------G-----------------A----------A-----------------------------T----------C----------------------------C---------------- 2266
B.CA.07.502_1191_03   ----------------------------A----------A-----------------------------------------G----------G----------G---------------- 2231
B.CH.08.M2_0803101_NFLG8 ----------------------C----AC---------A-----------------------------A----------TG---------A--A-G---------------------- 2273
B.CN.10.DEMB10CN002   ----------------------------A----------A-----------------------------------------A----------G---------------------------- 2176
B.ES.10.DEMB10ES002   ----------------------------A----------A-----------------------------T----------G----------G----------G---------- 2166
B.FR.11.DEMB11FR001   ----------A-----------------A----------A-----------------------------------------G----------G---------------------------- 2050
B.GB.05.MM45d213_GN1  ----------------AC----------A----------A-----------------------------A----------C----------C----------G---------------- 2205
B.HK.06.HK003         ----------A-----------------A----------A-----------------------------A----------G----------T----------A---------- 2174
B.HT.05.05HT_129389   ----------R----------Y----A------R----A-----------------------------------------------------M------Y---------- 2005
B.JP.12.DEMB12JP001   ----------A----------C----------A-----C-----------------------------------G----------G----------A---------- 2191
B.KR.07.HP_18_07JHS10_3909 -------G----------A----------A----------A-----------------------T----------G--A----A--G----------C---------------- 2468
B.PE.07.502_0525_wg5  ----G----------A----------A----------A---C-------------------------------------------C----------T---------------- 2269
B.RU.11.11RU21n       ----------A----------A----------A----------A-----------------------T----------T----------T---------------- 2375
B.TH.08.MERLBDTRC10   ----------A----------A----------A--G---------A----------A----------C---------------- 2037
B.US.11.ES38          -G---------A----------A----------T----------G----C----------C---------------- 2739

C.AR.01.ARG4006       ----------AC---------GA--A----------A-----A----------T--T--C----------T----A--G--G----------G----C-----A--------G 1988
C.BR.07.DEMC07BR003   ----------AC--C------GAT--A----G-G----A----------T--C----------A--G----------G----C-----A---------- 2136
C.BW.00.00BW5031_1    ----------AC--C------GA---A----------A-----A----------T--C---------C----A--G--G----------G----C-----A---------- 2175
C.CN.10.YNFL19        -G---------AC----------AT--A----------A-----A----------T--C----A-----T----A--G--G----------G----C-----A---------- 2157
C.CY.09.CY260         -G----------C--AY--C----GA---A----------A--R--W----------T--C----------T----A--G--G----------G----C-----A---------- 2027
C.ES.08.X2363_2       -G----------AC--CC----GA---A----G------A----------C----------T--C----------T----A--G--G----------G----------T--G----A-----A---------- 2235
C.ET.02.02ET_288      ----------AC---------GAG--A----G------A----------C----T--C----------A--G--G----------G----C-----A---------- 2006
C.IN.09.T125_2139     -G----------C--AC--C------GAG--A----------A-----A----------T--C----A----------A--G----------G----C-----A---------- 2780
C.KE.00.KER2010       ----------AC---------GA---A----G------A-----GA---------C----T--C----------A--G----------G----C-----A---------- 1985
C.MW.09.703010256_CH256.w96 ----------AC--C------GA---A----------A-----A----------C----T--C----------T--A--G--G-------A--G----------G----------A--G----- 2784
C.TZ.08.707010457_CH457.w8 ----------AC--C------GAG--T----------A-----A----------C--T--C----------C----A--G----T----------G----------T----C-----A---------- 2778
C.YE.02.02YE511       ----------AC---------GA---C----G------A--G--A----------G----T--C----------A--G----------G----C-----A---------- 1979
C.ZA.08.705010534_CH534.w12 -G----------AC--C--A---GA---A----G------A----A----------T--C----------A--G--G----------G----C-----A--G----- 2794
C.ZA.10.DEMC10ZA001   -G----------AC--C------GA---A----G------A----A----------A-----C----T--C----A----T----A--G----------G----C--C--A---------- 2144
C.ZM.11.DEMC11ZM006   -G----------AC--C--A---GAC--A----------A-----A----------T--C----A---C------A--G----------G----C-----A---------- 2153

D.CM.10.DEMD10CM009   ----------AC----------A----------A---------GG----------T----------A-----T----------G----------C---------------- 2179
D.CY.06.CY163         ----------G---AC----------T----------A-----G----------T----------T----------G--G----G----------C---------------- 2012
D.KE.11.DEMD11KE003   ----------AC-C----------A------TA---A-----G----------C----------A----------G----------C---------------- 2154
D.KR.04.04KBH8        ----------AC----------A----------A-----G----------A----------G----------G----------C---------A---------- 2763
D.SN.90.SE365         ---------G-----AC---G----------A----------A----------TA-----T----------G----------C----A---------G 2821
D.TZ.01.A280          ----------C--AC----------A----------A-----G----------A-----T----------G----------C---------------- 2005
D.UG.10.DEMD10UG004   -----G-------CC--AC----------G---A-----G------A-----G------T----------A----------G------GC---------T----------G---- 2142
D.UG.11.DEMD11UG003   ----------C--AC----------T------A---A-----G----------A----------G----------G----------C---------------- 2159
D.YE.02.02YE516       ----------A---------C---A----------A-----G----------T----C--A-----T----------G----------T---------------- 2015
D.ZA.90.R1            ----------G---AC----------A----------A----------A----------G----------G----------G---------G 2150

F1.AO.06.AO_06_ANG32  ----------C--AC---G--A---GA---A-------A-G----------G----------C--T----------A---------- 2015
F1.AR.02.ARE933       -----C---G-----TC------A---TT--A-------A----A----------N----------T----------G----------C----------A---------- 2098
F1.BR.10.10BR_RJ015   ----------AC-----A---TT---A-------A----A----------G----------G----------A---------- 2277
F1.CY.08.CY222        ----------C--AC-----A------A------A---A----------T----------C----------T----------A---------- 1997
F1.ES.02.ES_X845_4    ----------C--AC-----A------A----------A----------C----------T----------A---------- 2256
F1.RO.96.BCI_R07      -G----------AC-----A--G---C------A---A----------C----------G----------G----------T----------A---------G 2828
F1.RU.08.D88_845      -------TG--C--AA-----A------A------A---A----------T--T----------C--A---------G 2263
F2.CM.02.02CM_0016BBY ----------C--AC-----C----------A---A----------A----------G----------C----------G----------A---------- 1994
F2.CM.10.DEMF210CM001 ----------C--AC-----C----------A----------A----------C----G----------C----------A---------- 2068
F2.CM.10.DEMF210CM007 ----------AC----------A----------A---------G--------A----------G----------C----------G----------A---------- 2151

G.BE.96.DRCBL         -G----------AC----------AT----------A---------G----------C----T--C----A----------G----------G--C-----A--G----- 2768
G.CM.10.DEMG10CM008   ----------AC----------A----------A---------G----------T--C----A-----T----------C----------G--C-----A---------- 2185
G.CN.08.GX_2084_08    ----------AC----------A----------A----------C----------G----------C-----A---------- 2044
G.CU.99.Cu74          -G----------AC---C------G---T----------A--G----------C-----A----------A--G----------G--C-----A---------- 2408
G.ES.09.X2634_2       -G----------AC----------A----------A----------C----A----------G----------G--C----------G--C-----A---------G 2287
G.GH.03.03GH175G      -G----------AC----------A---C----------A----------C-----C----------C----------G--C-----A---------- 2844
G.KE.09.DEMG09KE001   -G--G-----G-----AC----------A----------A----G----------A----------G----------G--C----------G--C----------C--- 2179
G.NG.09.09NG_SC62     G-----G-------AC---------TG--A----------A----------C--G---A----------G----------G--G----------G--C-----A---------- 1996
G.ZA.01.TV546         -G----------AC---------C----A----G------A----------G----------C-----A----------G----------G----G------T---G--C-----A---------- 2006

H.BE.93.VI991         ----------AC---------TT---A----------T----------A----C-G---------C----A----------A--G--GA--------G----------G----------A---------- 2193
H.CF.90.056           ----------ACG----------A----A---A----C---G---A----G---------GC-----A----------A--G--G--T----------G----------C-----A---------- 2151
H.GB.00.00GBAC4001    ----------AC----------A----------A---------G---A----------G---------C-----A----A----T----A--G--G----------G----------G--C-----A---------- 2322

J.CM.04.04CMU11421    ----------AC----------TT---A----------A----------R----------T--C----------G----------T-------C-----R---------- 2310
J.SE.93.SE9280_7887   ----------AC-C--------G----A-----G------A---------G-G----------T--C----------T-------C-----A---------- 2118
J.SE.94.SE9173_7022   ----------AC-C--------G----AC----G---G--A---------G----------T--C-C----------T-------C-----A---------- 2119

K.CD.97.97ZR_EQTB11   ----------------------G---------A------A----A----------C--C-----G----------T----------T-------C-----A-------C- 2000
K.CM.96.96CM_MP535    ----------AC----------A----------A----A----------G----T--C----------G----------T-------C-----A---------- 2000

U.CA.01.TV749         -G-------G------C----------A------A----A---------G----------T-----C-----G----------C----------G-----C-----A---------- 2821
U.CA.99.TV721         -G-------G------A----------C------A----A---------G----------C-----G----------C----------C-----A---------- 2820
U.CD.83.83CD003_Z3    ----------AC----------CT---A----------C----------C----A--G--G----------G----------C-----A--G----- 2299
U.CD.90.90CD121E12    ----G----------AC----------A------A----A----------C----------T----A--G--G----------C-----A---------- 2003
U.CY.05.CY090         -G-------G------AC-----A-------A---------A--A-A----------T----------A-----A----A--G--GA-T--C----------A----------G--------C-A---------G 2003
U.CY.08.CY223         ----------C--AC---------C-AT--A----G-A----A-----C----------T----------T----------A----------G-----C-----A---------G 2021
U.ES.10.DEURF10DZ001  -R-----C-G----C--AC----------G----C----------A----------A-----C----------A-----T----------T----------A---------- 2080
U.GR.99.99GR303       -R-----C-G----C--AC----------A----------A----------T----------G----C--G----G----------G----C-----A---------G 2078
U.NL.95.U_NL_95_H10986_D1 --------G------AC---C----------A----------A----------C----------G----T----------C--T----------A---------- 2689

01_AE.AF.07.569M      ----------AC----------A--AC---GGC-----A----C----------A-----T----------G----C--C----------G----------A---------- 2010
01_AE.CF.90.90CF11697 -------T--------AC----------A---------G------A----A----------G----------T--T--G---------C----------G----------G--C-----A---------G 2753
01_AE.CN.10.YNFL03    -G----------AC----------A---------G------A----A----------T----------G----C----------G--C-----A---------G 2182
01_AE.HK.04.HK001     ----------AC----------A---------G------A--G--C----------T----------G----C--C----------G----------G--C-----A-----C--G 2166
01_AE.IR.10.10IR.THR48F ----------AC----------A----------A-----C----------T----------T----------G----C--C----------G----------T----G--C-----A---------G 1994
01_AE.JP.x.JRC77AE    ----------AC----------A---------G------A-----C----------T----------T----------G----C--C----------G----------C-------G--C-----A---------G 2801
01_AE.TH.09.AA090a_WG11 ----------AC----------A---------G------A----------A--A--A----------T----------G----C--C----------G----------G--C-----A---------- 2227
01_AE.TH.90.CM240     ----------AC----------A---------G------A-----C----------T----------T----------G----C--C----------G----------G--C-----A---------G 2378
01_AE.US.05.306163_FL ----------C--AC----------AG--------G------A-----C----------C----------T----------G----C--C----------G----G----------G--C-----A---------G 1995
01_AE.VN.98.98VNND15  ----------C---AC----------A----------A-----C----------T----------G---G-C--C----------G-------G--C-------G--C-----A---------G 2059
```

```
B.FR.83.HXB2                AAAAAATAAAAGCATTAGTAGAAATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGAAAATTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAA 2804
Pol                         E__K__I__K__A__L__V__E__I__C__T__E__M__E__K__E__G__K__I__S__K__I__G__P__E__N__P__Y__N__T__P__V__F__A__I__K__K__K__D__S__T__K__W__R__K__L__V__D__F__R__E__L__N__K__R__T__Q_
02_AG.CM.10.DE00210CM013    ----------G-----------G--------------------A-----A-----------------------------------C-----------------------G-----T------------------------------------T-A-----A--------- 2171
02_AG.CY.09.CY256           -------------C---AC---C------YT------------------A--------------------------------------------G-----Y--------------T--------------------------------------C-----R--G------ 2000
02_AG.ES.06.P1423           --------R-G------AC---T--------M--C--------------A----------G-------S----------C--------------GK-------------------T--------G--------------C--------------C-----A--------- 2242
02_AG.GW.05.CC_0048         -----------------AC---C-------A---A--------------A--------------A--------G--------------------------------------G--T--------------------------------------C--------------- 2154
02_AG.KR.12.12MHR9          -----------------AC---C-----C--------------------------------------C-----------C--------------------------------G--T--------------------------------------C--------------- 2463
02_AG.LR.x.POC44951         -----------------AC---C------G-------------------A--------------------------C--------------------------------------T-----C--------------------------------C--------------- 2799
02_AG.NG.09.09NG_SC61       -----------------AC---C--------------------------A-----------------------------------------------------------------T--------------------------------------C--------------- 1996
02_AG.NG.x.IBNG             -----------------AC---C--------------------------A-----------------A--------------------------------------------G--T--------------------------------------C--------------- 2329
02_AG.SE.94.SE7812          -----------------AC---C-----------A--------------A----------G-------------------------------A----------------G---A-T-------G------------------------------C--------------- 2176
02_AG.SN.98.98SE_MP1211     -----T----GCAT---AC---C------G-------------------A--------------------------------------------------T-------G------T-----A--------------------------------C--------------- 2001
03_AB.RU.97.KAL153_2        -G---------------AC---C-------AG-----------------A-------------------------------------------------------------------------------------------G---------------------------- 2030
04_cpx.CY.94.94CY032_3      -------------C---AC---G--A--------C--------------C--G---------------------------------------A-------T-----------------C-----------------------------------C-----A--------G 2170
05_DF.BE.x.VI1310           -----------------AC---------------T--------------A-----C----G---------------C---------------A-------------------G----------GG--------------------------------------------- 2186
06_cpx.AU.96.BFP90          -G--G------------AC-----------A---A--------A-----A--------------------------------------------------T--------------T--------------------------------------------A--G------ 2832
07_BC.CN.98.98CN009         -G---------------AC--C-------GAT--A-----G--------A------A-------------------------T--C------A-------------A--G--G-----------G-----------------------G-----C-----A--------- 2136
08_BC.CN.97.97CNGX_6F       -----------------AC--C-------GAT--A-----G--------A------A----------------C--------T--C------A------------GA--G--G-------G---G-----------------------G-----C-----A--------- 1984
09_cpx.GH.96.96GH2911       -----------------AC----------G-------------------A--------------------------------T-----------------T-----------G-----C-----G-----------------------------------A--------G 2006
10_CD.TZ.96.96TZ_BF061      -G------------C--AC-----------A---A--------------A--------G-GG--------------------T---------A-------------------------------G--------------------------------------------- 2183
11_cpx.CM.95.95CM_1816      -----------------AC------C--------A--------------A--------------------------C-----T--C--------G-----T-----------------C--------------------T-----T--------C-----A--------- 2168
12_BF.AR.99.ARMA159         -----------------AC------A--------A--------A-----A--------------------------------------------------------------------------------G-----------------------------A--------- 2809
13_cpx.CM.96.96CM_1849      -----------------AC------C----A---A--------------A-----------------A--------------T--C--------G-----T---------------------------------C----T-----T--G-----C-----A--------- 2209
14_BG.ES.05.X1870           -----------------A----------------A-------G------A-----------------A-----------------C--C---A----------------------------------------------------------G--C-----A--------- 2271
15_01B.TH.99.99TH_MU2079    -----------------AC-----------A----------G-------A-----C------------------------------------A-------T-----------G-----C--C--------G--------------------G--C-----A--G-----G 2203
16_A2D.KR.97.97KR004        -----------------AC-----------A---A-----------G--A--G--------------------------------C--------G-----T-----------------C---------------C-------------------C--------------- 2167
17_BF.AR.99.ARMA038         -------------------------A--------A--------A-----A------------------------------------------------------------------------------------C-------------------------A--------- 2018
18_cpx.CU.99.CU76           ----G------------AC------A-----------------------A-----------G-----------------------C--C---A-------------------------------------------------------------------A--------- 2109
19_cpx.CU.99.CU7            -----------------AC---------------T--------------A--------C--------------G--------T---------A-------------------------------G-----------G-----------------------------C--- 2030
20_BG.CU.99.Cu103           ----------------------------------A--------------A-----------------------------------C-----------------------G-----------------------------------------G--C-----A--------- 2270
21_A2D.KE.99.KER2003        -----------------AC-----------TG--T--------------A----------G-------------------------------A-------T-----------G--------C--G---------C-------C--------------------------- 2009
22_01A1.CM.01.01CM_0001BBY  -G-----------T---AC-----------A------------------A-----------G-----------G------------------A-------T-----------------C-----------------------------------C-----A--------G 2012
23_BG.CU.03.CB118           ----------------------------------AT-------------A-----------------------------------C--------------------------GA----------G-----------------------G--------------------- 2247
24_BG.ES.08.X2456_2         ----------------------------------A--------------A--G--C-----------------------------C--------------------------------------G-----------------------G-----A--------------- 2255
25_cpx.CM.02.1918LE         -G---------------AC-----------AT--T--------------A----------G---A-----------------------C---A----------------G-----T-----------------------------------G--A-----A--------- 2012
26_AU.CD.02.02CD_MBTB047    --------------A--AC------A----AT--A--------------A------------G-------A--------------C------A-------T-----------------C--------------------------T-----G--C-----A--------- 2809
27_cpx.FR.04.04CD_FR_KZS    ----------------------------------A--------------A----------G---------------------T--C--A-----G----------G---------------------------------T-----T--G-----C-----A--------- 2795
28_BF.BR.99.BREPM12609      -------------------------C--------A-----------------------------------------------T--------------------------------------------------------------------G------------------ 2170
29_BF.BR.01.BREPM16704      -----------------A-G-----C--------A-----------------------------------------------T-------------------------------------------------------------------------------A------- 2205
31_BC.BR.04.04BR142         -G---------------AC-----------AT--A-----G-G------A--------------------------C-----T--C--------------------A--G--G-----------G-----------------------G-----C-----A--------- 2276
32_06A1.EE.01.EE0369        -G-------G-------AC---------------A--------A-----A------------------------------------------A-------T-----------G--T--------------------------------------------A--G------ 2439
33_01B.ID.07.JKT189_C       -----------------A-------C--------A-----------------------------------------------------------------------------G-----C--------------G-----G-----T--------------A--------- 2105
34_01B.TH.99.OUR2478P       ------------------C----------G----A--------------A------------------------------------------A-------T-----------G-----C--A--------G--------------------G--C-----A--------G 2006
35_AD.AF.07.169H            -----------------AC---------------A--------------A--------------------------C---------------A-------T-----------G-----C--------------------G-----T--------C-----A--------- 2004
36_cpx.CM.00.00CMNYU830     --------------C--AC----------G-------------A-----A--------------------------------------C----------------G---------T--------------------------------------C-----A--------- 2009
37_cpx.CM.00.00CMNYU926     -G---------------AC---------C-AT--A--------------A----------G-G-------------------T--C------A----------------------------------------------------------G--C-----A--------- 1994
38_BF1.UY.03.UY03_3389      -----------------A-G-----A---G----T--------A-----A----------------------------------------------------------------------------------------------------------------A------- 2227
39_BF.BR.04.04BRRJ179       -----------------C-----------G-----T-A---G---CT--A---------------------------------------------------------------------------------------A-G---C-------------------------- 2281
40_BF.BR.05.05BRRJ055       -----------------AC------A--------A--------A-----A--------------------------------------------G-----T-----------GA---------G---------------------------------------------- 2287
42_BF.LU.06.luBF_18_06      -----------------AC------A----TG--A-----G--A-----A--------------------------------------------G-----T-----------------------------G--------------T------------------------ 2358
43_02G.SA.03.J11223         -G---------------AC----------TT------------------A-----------------------------------C------A----------------------------------------G--G-----------R--G--C-----A--------- 2329
44_BF.CL.00.CH80            -----------------A-------------------------------------------------------------------------------------------------------C--------G-----------------------------A--------- 2246
45_cpx.FR.04.04FR_AUK       -----------------AC------------------------------A-----------G------------------------------A-------------------GA-T--------------------------------------------A--------- 2790
46_BF.BR.07.07BR_FPS625     -----------------AC------A---G----A--------A-----A--G-----------------------------------------------------------------------------------------------------------A--------- 2242
47_BF.ES.08.P1942           -----------------A-------C--------A--------------------C-----------------------------------------------------------------A-----------------------T--------------A--------- 2243
48_01B.MY.07.07MYKT021      -----------------A-G--------------A--------------A--------------C-----------------T-----------------------------------C--------------------------T--------------A---G----- 2009
49_cpx.GM.03.N26677         -----------------AC---------------A--------------A-----------------------------------------------------------------T--------------------------------------------A--------- 2227
50_A1D.GB.10.12792          -G------------C--AC----------G----T------G-------A------A-----------------------------------A-------A-----------G--T--C----------------------------.---G--C-----A-----A--- 2211
51_01B.SG.11.11SG_HM021     -G--------------------------------A--------------A------------------------------------------A-------T-----------G-----------------------------------------G--------------G 2011
52_01B.MY.03.03MYKL018_1    -----------------A----------------A--------------A--R-----------------------------------------------T-----------G-----C--C--------G--------------------G--C-----A--------C 2118
53_01B.MY.11.11FIR164       -----------------AC-----------A----------G-------A-----C--C---------------------------------A-------T-----------G--TG----C--------G--------------------G--C-----A--------G 2144
54_01B.MY.09.09MYSB023      -----------------A-------------C--AC----------------------------------------------------------------------------GA---------G---------------------T--------------A--------- 2206
55_01B.CN.10.HNCS102056     -----------------AC------C----A---A------G-------A-----C--------------------C-----------------------T-----------------C--C--------G-----------C--------G--C-----A--------G 2114
56_cpx.FR.10.URF5_patient_A -G---------------AC-----------A---TG------------------------G-------------------------------------------------------------------------C----------------------------------- 2155
57_BC.CN.09.09YNLX19sg      -G---------------AC--C-------GAT--A-----G--------A------A-------------------------T---------A-------------A--G--G-----------G-----------------------G-----C-----A--G------ 1988
58_01B.MY.09.09MYPR37       ----------G------A----------------A--------------A-----C--------------------------------------------T-----------G--------C--------G--------------------G--C-----A--------G 2149
59_01B.CN.09.09LNA423       ----------------------------------A--------------A-----C--------------------------------------------T-----------G-----C--C--------------------------------C-----A-----G--- 1988
60_BC.IT.11.BAV499          -G--------G------AC--C-------GAT--A-----G--------A------A-------------------------T--C-----------------------------------------------------------T------------------------ 2761
61_BC.CN.10.JL100010        -G-----------G---AC--C-------GAT--A--------------A--G---A-------------------------T--C------A-------T-----A--G--G-----------G--------------------T--G-----C-----A--------- 2225
62_BC.CN.10.YNFL13          ----G--------C---A----------------A--------------A-----------------------------T--------------------------------G-----------------------------------G-----------A--------- 2133
63_02A1.RU.10.10RU6637      -----------------AC--CC-----------A--------------A-----------------A--------------T---------A----------------------T--------------------------------------C--------------G 2345
64_BC.CN.09.YNFL31          -G---------------AC--C-------GAT--A-----G--------A------A-------------------------T--C------A-------------A--G--G-----------G-----------------------G-----C-----A--------- 2158
65_cpx.CN.10.YNFL01         -G---------------ACG-C-------GAT--A--------------A------A-------------------------T--C------A-------------A--G--G-----------G-----------------------------C-----A--------- 2175
67_01B.CN.11.ANHUI_HF115    -----------------AC---------------A--------------A--------------------------------------------------T-----------G-----C--C--------G--------------------G--C-----A--------G 2147
68_01B.CN.11.ANHUI_WH73     -----------------AC------------------------------A-----C-----G--------------------T-----------------T-----------G-----C--C--------G--------------T--------C--------------- 2138
70_BF1.BR.10.10BR_PE004     -----------------AC------A---G----A--------------A-----C--------------------------------------------------------------------------------------------------------A--G-----G 2260
71_BF1.BR.10.10BR_PE008     -----------------AC------A--------C--------A-----A--------------------------------------------------T-----------G--T--------------G-----------------------------A--------- 2468
72_BF1.BR.10.10BR_MG002     -----------------AC------A----TG--A-----------------------------------------------------------------T--------------------------------------------------G--------A--------- 2312
O.BE.87.ANT70               --------G----C--GAC--C---A---CAG--A------C-A-----A----------G---A--A-----------T--T-----A--TA-C-----T-----A--G-----TG-------------------G--------T--G---T-A-----------A--- 2858
O.CM.98.98CMA104            --------G----C--GAC--C---A---CAG---------C-A-----A----------G-G-A--A-----------T--T-----A--TA-T-----T-----A--G-----T--------G--------------------T------T-A--C----A---A--- 2290
O.CM.98.98CMABB141          --------G----CC-GAC--C---A---CAG---------C-------A----------G---A--A-----------T--T-----A--CA-C-----T-----A--G-----T--------G--------G--G-----C--T--G---T-A--C--------A--- 2289
O.CM.98.98CMABB212          -G------G-----C--AC--C---A--CCA---A------C-------A--G--C----G---A--A-----------T--T-----A--TA-T-----T-----A--G-----T--C--C--G-----------------C-----------A-----A-----A--- 2294
O.CM.98.98CMU5337           --------G-------GAC--C---A---CAG--A------C-A-----A--G-------G---A--A-----------T--T-----A--T--C-----T-----A--G-----T--------G--------G--G--------T--G---T-A--C--------A--- 2291
O.CM.99.99CMU4122           --------G----C--GAC--C---A---CAG--A------C-A-----A----------G-G-A--A-----G-----T--T-----A--TA-C-----T-----A--G-----T----G---G--------GC-G-----C--T--G---T-A--C--------A--- 2289
O.FR.92.VAU                 -G------G----CC-GAC------A----A---AT-----C----G--A----------G---A--A-----------T--T-----A--TA-C-----T-----A--G-----TG-A-----G-----------G-----C---------T-A--C--A-----A--- 2370
O.GA.11.11Gab6352           --------G----C---AC--C---A---CAG--AT-----C-A-----A----------G---A--A-----------T--T-----C--TA-C-----T-----A--G-----T--------G--------------------T--G---T-A--C-----G--A--- 2015
O.SN.99.99SE_MP1299         --------G----C--GAC--C---A---CAG--A------C-A-----A----------G---A--A-----------T--T-----A--TA-C-----T-----A--G-----T--C-----G--------GC-G-----C--T--G---T-A--C--------A--- 2857
O.US.10.LTNP                -G------G----C--GAC--C---A--CCGG--A------C-A-----A----------G---A--A-----------T--T-----C--TA-------T--------G-----T--------G-----------------C--T--G---T-A--C--------A--- 2779
N.CM.02.DJO0131             --------G-G------AG-C--------G----A------C-------A--------T-G---A--------------------C------A-C-----T-----A--G-----T--C-----------------------------G---T-A-----A--G--C--- 2310
N.CM.02.SJGddd              --------G-G------AG---------------A--------------A--------T-G---A--------G--------T--C------A-T-----T----GA--G-----T--C-----------------------------G---T-A-----A--G--C--- 2294
N.CM.04.04CM_1015_04        --------G-G------AG---------------A--------------A--------T-G---A-----------------T--C------A-T-----T----GA--G-----T--C--C------------------------------T-A-----A--G--C--- 2309
N.CM.06.U14296              --------G-G------AG---------------A--------------A--------T-G---A--------G--------T--C------A-T-----T----GA--G-----T--C-----G-----------------------G---T-A-----A--G--C--- 2307
N.CM.06.U14842              ----G---G-G------AG---------------A--------------A--------T-G---A--------G--------T--C------A-T-----T-----A--G-----T--C-----------------------------G---T-A--C--A--G--C--- 2311
N.FR.11.N1_FR_2011          --------G-G------AG---------------A-----------------------T-G---A-----------------T--C------A-W-----T----GA--G--R--T--C-----------------G---------------T-A-----A--G--C--- 2205
P.CM.06.U14788              --------G----T---AG--CC--C---CA----------C-------A-----AA-------------------------T--C--C--CA-T-----A-----A--G-----TG-C-GC--G-----G-----------C-----------A-----A-----A--- 2275
P.FR.09.RBF168              --------G----T---AG--ST------CA----------C-------A-----AA-M-----------------------T--C--C--CA-C-----A-----A--G-----T--C-GC--R-----G---------K-C-----------A-----------A--- 2839
CPZ.CD.06.BF1167            -G-----TG----CC--AC----------GA-A-TT-AAG-GCA-----------CATTG--G-A--A-----C-----T-----------CA-C-----T--T--A--G-----------A-----------------------T-----G--C-----A--G--A--- 2906
CPZ.CM.05.SIVcpzMT145       --------G-G------ACT-----A----A---A--------------A--------------A-----A--------T--------A--TA-T-----A--T--------G-----C--C------------C-T--T-----------------------G------ 2346
CPZ.GA.88.GAB1              -------T-----C---AC----------CA---A--------------------A--G-----A--A-----------------------TA-T-----A--C--A--G--------------------------------C-----------A-----------A--- 2865
CPZ.TZ.06.SIVcpzTAN13       -G-----TG----C---AC------A----GTACTT-A-----------A-----C--TGCC--A--A--A--------T--------A--CA-------A--C--------G--TC-T-----------G--G-----G--C---------T-A-----A-----A--G 2445
CPZ.US.85.US_Marilyn        -------T--G------AC---------CCA---A--------A-----A--G---A---G-G-A-----A--------T--------A---A-T-----T--C--A--G-----T--C-----------------------C--T------T-A-----A--G--A--- 2864
GOR.CM.07.SIVgor2139_287    --------G----C---AC--CT------CAG--A------C-------C--G--A--------------------------T-----A--TA-T-----------A--G--------------------------G--------T------T-A-----AC----A--G 2341
GOR.CM.07.SIVgorCP2135con   --------G----C---AC--CT------CAG--A------C-------C--G--A--------------------------T-----A--TA-T--C--------A--G-----------K--------------G--------T------T-A-----A--G--A--- 2340
```

```
B.FR.83.HXB2               GACTTCTGGGAAGTTCAATTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAAACAATGAGACACCAGGGAT 2974
Pol                        _D__F__W__E__V__Q__L__G__I__P__H__P__A__G__L__K__K__K__K__S__V__T__V__L__D__V__G__D__A__Y__F__S__V__P__L__D__E__D__F__R__K__Y__T__A__F__T__I__P__S__I__N__N__E__T__P__G__I
A1.AU.03.PS1044_Day0       -----T--------G--------G-----G-----A--G--CC----------------------------A--------G--C--------------G--T---------A----T--A-----------G--C-------------C------------------A-- 2182
A1.CH.03.HIV_CH_BID_V3538  -----------------G-----G-----G-----A--G--T-----------------------------A--------G--------------------T---C----GAG------A-----------G--C--------------------------------A-- 2179
A1.CY.08.CY236             -----T--------------G--------G-----A--G--CC----------------------------A--------G--C-----------------T--GC----------T--A-----------G--C-------------C------------------A-- 2185
A1.ES.05.X1608_8           -----T-----G-----G-----------G-----A--G--CC-------------------------G--A--------G--C-----------------T---------AG------------------G--C-------------C------------------A-- 2422
A1.KE.11.DEMA11KE001       --T--T-----------------------------A-----T-----------------------------A--------G--------------------T---------AG---T--A--A-----------C-------------C--------C---------A-- 2313
A1.RU.11.11RU6950          -----T------------C----------C-----A--G--T-----------------------------A--------G--------------------T---------AG------A--------------C--T-----A----A--------------------- 2498
A1.RW.11.DEMA111RW002      -----------------G--------C--------A-----T-----------------------------A--------G--C-----------------T----------GT--T--A-----------G--C--------------------------------AG- 2325
A1.SN.01.DDI579            -----------------------------G-----A--G--T--G-----------------------------------G--C-----------------T---------AG---T--A--------------C--------C----C------------------AG- 2172
A1.UG.11.DEMA110UG001      -----T-----------------------G-----A-----T-----G-----G-----------------A--------A--------------------T---------AG---T--A--------C--G--C-------------C------------------A-- 2329
A1.ZA.04.04ZASK162B1       -----------------------------------A-----T-----------------------------A--------G--C-----------------T--------------T--A--------------C-------------C------------------AG- 2415
A2.CM.01.01CM_1445MV       -----------------G-----------------A-----A-----------------------------A--------G--C-----------C---------T-------------A--A--------------A-----------------------------AG- 2173
A2.CY.94.94CY017_41        -----------------G-----------------A-----A---------------G-------------T--------G--C-----------C---------C-------------A--A-----------C-------------CC-----------------AG- 2331

B.BR.10.10BR_MG029         -----------G-----------G--------C-----------------------------------------------------------C-------------------------------------------------------A---T----------------- 2436
B.CA.07.502_1191_03        -----------------------------------T------------C----------------------------------------C--------------GT-----------------------------------------GC---T----------------- 2401
B.CH.08.M2_0803101_NFLG8   -------------------------------------G----C----------C-----------------------A-----------------------------------------A----------------------------C--------------------- 2443
B.CN.10.DEMB10CN002        -----T-----------------------------T--------------------------G--------A------------------------------------A------------------------------------------------------------- 2346
B.ES.10.DEMB10ES002        -----------------------------------T--------------------------------------------------------C------------C----------------------------C-------------C--------------------- 2336
B.FR.11.DEMB11FR001        -----------------------------------T--G-----------------------G---------------------------------------------A-------------A----------------------------------------------- 2220
B.GB.05.MM45d213_GN1       ------------------------------------------------------------------------------------------------------------A----A-------------------------------------------------------- 2375
B.HK.06.HK003              -----------------------G--------------------------------------------G--------------------------------------------------A--A--------------------Y----C---------A----------- 2344
B.HT.05.05HT_129389        -----------------------------------TK-------------------------------------------------------Y--------T------A----------------Y--------Y--------------------------------A-- 2175
B.JP.12.DEMB12JP001        -----------------------------------T--------------------------------------------A---------------------------A----T--T--A--A--------------------C---T---------------------- 2361
B.KR.07.HP_18_07JHS10_3909 -----------------------------------A-------------G-------------------------------------------------------A--A--A-------------------------------C-------------------------- 2638
B.PE.07.502_0525_wg5       ----------------R-------------------------------C---G-------------------------------------------------------A----A-----A-------------------------------------------------- 2439
B.RU.11.11RU21n            --------------------------------C--A-----------------T--------G--------A------------------------------------CC---T-----A----------------------------C--------------------- 2545
B.TH.08.MERLBDTRC10        -----------------------------------T------------C----------------------------------------C--C--------T--G---A----A---------------------------------G--------------------G- 2207
B.US.11.ES38               --------------------------------C--T-----------G--------------------------C------------------------------TG----AG---------------------C----------------------------------- 2909

C.AR.01.ARG4006            -----T--------C--------G--------C--A--------------------------G-----------C-----G-----------------A--T--------------T-----A-----------C-----------------------A----------- 2158
C.BR.07.DEMC07BR003        -----T-----------------G--------C--A--------------------------G-----------------G-----------------A--T---C-----AG---T--A--A-----------C-----------------------A----------- 2306
C.BW.00.00BW5031_1         -----T-----G--------------------C--A-T---------G--------------G--------A--------G------------------------------A----------------------C------------G----------A----------- 2345
C.CN.10.YNFL19             --T-----------------------------C--A---------C----------------G-----------------G--------------------T---C-------G--------A-----------C-------------G---------A--T-------- 2327
C.CY.09.CY260              -----T--------------------------C--G--------G-----------------G-----G--A--------G--------------------T----------R---------A-----------C----------------------CA----------- 2197
C.ES.08.X2363_2            -----T-----------G-----G--------C--A--G-----G-----------------G--------A--------G-----------------C--T---------A-Y--------A-----------C--T-------------------------------- 2405
C.ET.02.02ET_288           -----T-----------------G--------C--A--------------------------G------T-A-----A--G-----------C--------T---------AG---------A--C--------C-------------C---------A----------- 2176
C.IN.09.T125_2139          --T--T--------------------------C--A--------------------------G-----------------G--------------------T---T----------------A--------C--C-------------C---------A----------- 2950
C.KE.00.KER2010            -----T-----G--------------------C--A--------G-----------------G-----------------G--------------------T------A--A----------A-----------C-------------G---------A----------- 2155
C.MW.09.703010256_CH256.w96 -----T--------------------------C--A--------------------------G-----------------G--------------------T------A----T--------A-----------C-----------------------A--------A-- 2954
C.TZ.08.707010457_CH457.w8 -----T--------------------------C--A--------G--G--------------G--------A--------G--------C-----------TC----------T--------------------C-----------------------A----------- 2948
C.YE.02.02YE511            -----T--------------------------C--A--------------------------G-----------------G--------------------A------A----A--------A-----------C-----------------------A----------- 2149
C.ZA.08.705010534_CH534.w12 -----T--------------------------C--A--------------------------G-----G-----------G--------------------TC--------A----------A-----------C--------C--C-----T-----A----------- 2964
C.ZA.10.DEMC10ZA001        -----T-----G--------------------C--A--------------------------G-----G-----------G-----------------C--T----------G---------A-----------C-----------C-----------A----------- 2314
C.ZM.11.DEMC11ZM006        -----T--------------------------C--A-----------------------------------A-----A--G--------------------T---------AGT--------A-----------C-----------------------A-----G----- 2323

D.CM.10.DEMD10CM009        --------------C--------------------T------C-G-----------------------G------A-A-----------------------A------A----A--T--A-----C--A-----C-------------C--------------------- 2349
D.CY.06.CY163              -----------------------------------G------C-G--------------------------------A-----------------------------------T--T--------C--------C-------------C--------------------- 2182
D.KE.11.DEMD11KE003        --------------G---C----------------T-----A--------------------------------------------------C------------T----------T--A--A--C--------C--------------------------------A-- 2324
D.KR.04.04KBH8             -----------------G-----------G--C--T------C-------------------------------------------------C-----------------------T--A--------C-----C--------------C-------------------- 2933
D.SN.90.SE365              -----------------------------G-----A--G--T--G-----------G--------------A---------------------------------C-------T--T--A-----------------T----------C--------------------- 2991
D.TZ.01.A280               --T---------------C----------G-----T------C--------------------------T-------------------C------------------A----T--T--A--------------C-----------C-C--------------------- 2175
D.UG.10.DEMD10UG004        -----T------------C----------------T------C-------------------------G------------------------------------T------G---T--A--A------------------------------------G----G--A-- 2312
D.UG.11.DEMD11UG003        ------------------C-------------C--T------C-------------G------------------------------------------------------A----T--A--A-----------C--T-----------------------------A-- 2329
D.YE.02.02YE516            -----------------------------------T------C-G-----------------G-----------------------C--------------T---------AGT--T--A-----C--------C-------------C--------------------- 2185
D.ZA.90.R1                 --T--T--------A--G-----------G--C--T------C-G------------------------------------------------------------TG---------T--------------------------------------------G-------- 2320

F1.AO.06.AO_06_ANG32       --T--T-----G-----------------------T--------------------G-----------------------G-----------C---------------A-G--G-----------C--------C------------G-T-----------------A-- 2185
F1.AR.02.ARE933            --T--T-----G-----------------------T-----A--------------G-----------------------G-----------------------------G--T-----------C--------C------------G-C-----------------A-- 2268
F1.BR.10.10BR_RJ015        --T--T-----G-----------------------T--------------------G-----------------------G-----C-----------C---------A-G--G--T--------C--------C-------------CT-----------------A-- 2447
F1.CY.08.CY222             --T--T-----------------------------T--------------------G------------T----------G--------------------T------A-G--G--------A--C-----------T---------C-C-----------T-----A-- 2167
F1.ES.02.ES_X845_4         --T--T-----G-----------------------T-----------------G--G-----------------------G--------------------T------A-G--G--------------------C--A-----------C----------A------A-- 2426
F1.RO.96.BCI_R07           --T--T-----------------------------T--------------------G------------T----------G---------------------------A-G--G-----A--------------C------------C-C-----------------A-- 2998
F1.RU.08.D88_845           --T--T--T--G-----------------------T--------------------G-----------------------G--------------------T------A-G--G-----------C-----------------------C-------------------- 2433
F2.CM.02.02CM_0016BBY      --T--T-----G-----G--------------C--T-----T---------------------------T----------G---------------------------A-G--A-----------C--------C--------------T-----------------A-- 2164
F2.CM.10.DEMF210CM001      --T--T--------C-----------------C--T--------------------------------------------G-----------------G-----G---A-G--G-----------C--------C--------------C-----------------A-- 2238
F2.CM.10.DEMF210CM007      --T--T-----G------C-------------C--G--------------------------------------------G--------------------T------A-G--G--T--A-----C--------C--------------C-----------------A-- 2321

G.BE.96.DRCBL              -----------G--C-C------------T-------GG---------C------G---------------A--------G--------------------T---------A----T--A--A-----------C--T----------C---T--------------A-- 2938
G.CM.10.DEMG10CM008        -----------G--C--------------T-----T--G---------C----------------------A--------A-----------------C--T---------AGT--T--A--------------C--T---------G----T-----A----------- 2355
G.CN.08.GX_2084_08         -----------G--C--------T-----T--------G----------------G---------------A--------A------------------------------A-T-----A--------A-----C--T--------------T--------------R-- 2214
G.CU.99.Cu74               -----T-----G--C--------------T--C-----G--------------------------------A--------G---------------------------A-------T--A--------------C--T--------C-C---T----------------- 2578
G.ES.09.X2634_2            -----------G--C--------------T-----------T-----------G-----------------A--------G--------C--------G------------AG---T--A--A-----------C--T--------------T----C------------ 2457
G.GH.03.03GH175G           --------------C--------------T--------G-----G----------G---------------A--------A--------C---------------------AG---T--A--------------C--T--------------T----------------- 3014
G.KE.09.DEMG09KE001        -----------G--C--------------T--------G--T-----------------------T-----A--------G--------C---------------------AG---T--A--------A-----C--T---------G---------------------A 2349
G.NG.09.09NG_SC62          -----------G--C--------C-----T--------G---------------------A----------A--------A------------------------------A----T--A--------A-----C--T--------------T----------------- 2166
G.ZA.01.TV546              -----------G--C--------------T-----T--G--------------------------------A--------G-----------------C---------A-------T--A--------------C--T--------------T----------------- 2176

H.BE.93.VI991              --------------A--G-----------------A--------------------G------T----------------G-G---------------C--T---C----------------------------C-------------C--------------------- 2363
H.CF.90.056                -----------------G--------------C--A---------------------------T----------------G-----------------C--T------A----A-----A--------------C----------------------------------- 2321
H.GB.00.00GBAC4001         -----------------GC-------------C--A-----------------C--------------------------G-----------------C--T---T-------------A--------------C-------------C---------A----------- 2492

J.CM.04.04CMU11421         -----Y-----R--G--G-----------------A-----------G--------R-----C-----------------G--Y--C-----------C--T---------R-------------C--------C--T-----------------------------RR- 2480
J.SE.93.SE9280_7887        -----------------G-----------------A--------------------------C-----------------G-----------------C--T---T-------T--------A-----------C--T-------------------------------- 2288
J.SE.94.SE9173_7022        -----------------G-----------------A--------------------------C-----------------G-----------------C--T---T-------T--------------------C--T-------------------------------- 2289

K.CD.97.97ZR_EQTB11        -----------------------G-----------A-----------------------------------------A--G-----------------C--T------A----T--------------------C--T-----------T-----------------A-- 2170
K.CM.96.96CM_MP535         -----------------------------------A-----A--------------------------------------G-----------------C--T------A-------------------------C--T-----------T--------A---------G- 2170

U.CA.01.TV749              -----------------G-----G-----C--C--A-----T-----------------------------A--------G--C--------------C------------A-T--T--A--A-----------C--T-------------------------------- 2991
U.CA.99.TV721              -----T-----------GC----------C-----A-----T-----------G-----------------A--------G--C-----------------------------T--T--A--A-----------C--T-------------------------------- 2990
U.CD.83.83CD003_Z3         --T-----------------------------C--A--------------------------------------------A-----------------C--G---------A-T--------A--------C--C-----------------T----------------- 2469
U.CD.90.90CD121E12         --T--------------G--------------C--A--------------------------------------------A-----------------C--G---------A-T--------A--------C--C-----------------T----------------- 2173
U.CY.05.CY090              A-T--T-----------------------------T------C-------------------------------------GA--------------------------A-G--G--------A--C---A-------------C----CC-------C---------A-- 2173
U.CY.08.CY223              -----------------------------------A-----T-----------G-----------------A--------G--------------------T--GC----------T--A--A--------------T----------C---T-----------G----- 2191
U.ES.10.DEURF10DZ001       -----T-----------------------------T--T---C-G--G--------------G------T-A--------A-----------------C--------------A--T--A-----C--A-----C--T---------G----T----------------- 2250
U.GR.99.99GR303            -----------------------------G-----A-----T-----------------------------A--------G--C--------------C--T------A----T--T--A--A--C--------C--T--------C-C---------A----------- 2248
U.NL.95.U_NL_95_H10986_D1  --T--T--------G-----G--------------A--------------------------------G----G------G--------------------T------CC--GG--------A--C--------C--T-----------C-----------------A-- 2859

01_AE.AF.07.569M           -----T-----------------------G-----A-----A-----------------------------A--------A--------------------T-----C---AG---T--A--------------C-------------C------------------A-- 2180
01_AE.CF.90.90CF11697      -----T------------------G----G-----A-----T-----------------------------A--------A--------------------T---T------G---T--A--------------C-----------------------A--------A-- 2923
01_AE.CN.10.YNFL03         -----T-----------------------G-----A-----T-----G-----------------------A--------A--------C--------C--T---------AG---T--A--------------C--------------------------------A-- 2352
01_AE.HK.04.HK001          -----T------A-------------------C--A-----T-----------------------------A--------A--------------------A------A--Y-T--T--A--------A-----C--------------------------------A-- 2336
01_AE.IR.10.10IR.THR48F    -----T-----------------------G-----A-----T---------------------------T-A--------A--------------------T---------AG---T--A--------------C-------------C-----------A------A-- 2164
01_AE.JP.x.JRC77AE         --T--T-----------------------G-----A-----T-----------------------------A--------A--------------------T---------AG---T--A--------------C--------------------------------A-- 2971
01_AE.TH.09.AA090a_WG11    -----T-----------G-----------------A-----T-----------------------------A--------A--------------------T---------AG---T--A-----------------------------------------------A-- 2397
01_AE.TH.90.CM240          -----T-----------------------G-----A-----T-----------------------------A--------A--------------------T---------AG---T--A--------------C--------------------------------A-- 2548
01_AE.US.05.306163_FL      -----T-----------------------G-----A-----T-----------------------------A---A-A--A-----------------C--T---C-----AGT--T--A--------------------------------T--------------A-- 2165
01_AE.VN.98.98VNND15       -----T-----------------------G-----A-----T-----------------------------A--------A--------------------T---------AG---T--A--------------C--------------------------------A-- 2229
```

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

HIV-1 Genomes

```
B.FR.83.HXB2               TAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATA 3144
Pol                        __R__Y__Q__Y__N__V__L__P__Q__G__W__K__G__S__P__A__I__F__Q__S__S__M__T__K__I__L__E__P__F__R__K__Q__N__P__D__I__V__I__Y__Q__Y__M__D__D__L__Y__V__G__S__D__L__E__I__G__Q__H__
A1.AU.03.PS1044_Day0       C--G-----------------------A-----------------------------G------------------C-------C-A----GC-A--------------A----------------------C-----------------T--------------A---- 2352
A1.CH.03.HIV_CH_BID_V3538  C--G-----------------------A--------------------------T--GGC------------------------C-------C---G--------A--------------------------C-----------G-----T-GG-------A-------- 2349
A1.CY.08.CY236             C--G-----------------------A-----------------------------GT----T--------G-------C---C------TC-A-------------------------G-----------------------------T--------------A---- 2355
A1.ES.05.X1608_8           C-----------------------------------------------------T--G--------------------------C------TC-A-G--------A---A----------------------C-----------------T-----------------C- 2592
A1.KE.11.DEMA11KE001       C-----------------------------------G--------------------G--------------------------C------TC-A----------A---A-C--------------------C-----------------T--------------A---- 2483
A1.RU.11.11RU6950          C-----------------A--------------------------------------G--------------------G-----C------TT-A----------T--------------------------C-----------C-----T--------------A---- 2668
A1.RW.11.DEMA111RW002      ---G--------------------------------------------------T--G-----------------A--------C------TT------------A---A----A-----------------C-----------------T------------------- 2495
A1.SN.01.DDI579            C--G-----------------A-----------------------------------GTA------------------------C------TC------------A---A----------------------C---------------A-T------------------- 2342
A1.UG.11.DEMA110UG001      C--G-----------------------------------------------------G--------------------------C-----GTC-A-G--------G--GA----------G-----------------------------T------------------- 2499
A1.ZA.04.04ZASK162B1       C--G-----A-----------------------------------------------G--------------------------C------TC------------------C--------------------C--A--------------T--------------A---- 2585
A2.CM.01.01CM_1445MV       ---G--------------A--------------------------------------G----------T---------------C-------C-A----------G---A-C-----C--------------C-----------------T-----------T------- 2343
A2.CY.94.94CY017_41        ---G--------------A--------------------------------------G--------------G-----------C------TC-A-G---A----AT--A-C-----C--------------C-----------------T---------A-C------- 2501
B.BR.10.10BR_MG029         A-----------------------------------------------------T------------------------------------------C-----------A-C-----------------------A-------------------------AA------- 2606
B.CA.07.502_1191_03        A--------A--------------------------G---------------------TA---------------------A-----------------------------------------------------A---------------------------G------ 2571
B.CH.08.M2_0803101_NFLG8   C-----------T---------------------------------------------GC---------------------------------------------A-----------------TG--------------------------------------------- 2613
B.CN.10.DEMB10CN002        C--------------------------A-------------------------------------------------------------A---------------------------------T------------CTA-----------------------AG------ 2516
B.ES.10.DEMB10ES002        ---------------------A--------------------------------T---T----------------------------------------------A-----------C---------------C-----------------C----------A------- 2506
B.FR.11.DEMB11FR001        ---------------------C-----------------------------T--------C--------------T--------------------G-----T-----G--C-----------------------A-----------G---------------------- 2390
B.GB.05.MM45d213_GN1       ------------------------------G-----------------------------------------G--------A--------------------------------------------------C--------------------------------A---- 2545
B.HK.06.HK003              ---G---------------------------------------------------------TC------------------------C----G----C------------------------------------------------------------------------ 2514
B.HT.05.05HT_129389        ---------------------K--W------------------------------------------------------------------R-------------------A--------------------------------------------R-----R------- 2345
B.JP.12.DEMB12JP001        ------------------------------------------------------T---T-------------G------------------------T-------------C--------------------C--A-------------------------A-------- 2531
B.KR.07.HP_18_07JHS10_3909 ---------A--------------------G-----------------G--------------------------------------------------------------------------------------A--------------------------A------- 2808
B.PE.07.502_0525_wg5       ---------------------------------------------------------------------------T-----------------------------A---------------------------------------------C--------AAAG-A---- 2609
B.RU.11.11RU21n            ---------------------------------------------------------------------------------T-----------------------A--------------------------------------------------G------------- 2715
B.TH.08.MERLBDTRC10        ------------------------------------------------------T----------------G---T---------------------------------------------------------C----------G----------------A-------- 2377
B.US.11.ES38               -----------------------------------------------------------------------G---T-----------------------------A--------------G-----------------------------T------------------- 3079
C.AR.01.ARG4006            ---------A--T---------------------------------T-G--------G-----Y-----------------A--C-----GGC------------A-----------------T-----------------------------------------A---- 2328
C.BR.07.DEMC07BR003        ---------A--T--------G--------------------------G--------G--C--T--------------------C-----GGC------------A-----------------T--------C--------------------------------A---- 2476
C.BW.00.00BW5031_1         ---G-----A--------A-----------------------------------T--G--------------------------C------GC-A----C-----A-----------------T--------C-----------------T-----------T--A---- 2515
C.CN.10.YNFL19             ---G-----A--T--C--------------T-----------------------T--G----------T---------------C-----GGCG-----------A-----C-----------T--------C-----------------------------T--A---- 2497
C.CY.09.CY260              ------Y--A--T--------------------------------------------GT-------------------------C------GC------------------C-----------T--------------------------T--------------A---- 2367
C.ES.08.X2363_2            ---------A--T-----------------------G--G------T----------GGC---T-----------T-----T--C------GC-A----------A-----C-----------T--------C------------------C------------------ 2575
C.ET.02.02ET_288           ---------A--T-----------------------------------------T--G--------------------------C------GC-A----C-----A-----C--------------------C-----------------------G--------A---- 2346
C.IN.09.T125_2139          ---G-----A--T--------------------------------------------G-------------G------------C-----GGC--G---------A-----C-----------T-----------------------------------------A---- 3120
C.KE.00.KER2010            ---------A--T-----------------T--------------------------G------------------C-------C------GC-A-G--------A-----C-----------T--------C--------------------------------A---- 2325
C.MW.09.703010256_CH256.w96 ---------A--T--------------A-----------------------------G--------------------------C------GC-A----C-----A-----C-----------T--------C--A--------G--------------------A---- 3124
C.TZ.08.707010457_CH457.w8 ------C--A--T-----------------------------G--------------G--------------------------C--C--GGT-AG---C-----------------------T--------C-----------------T--------------A---- 3118
C.YE.02.02YE511            ---------A--T--------------------------------G-----------G--------------------------C------GC------C-----A-----------------T--------C--A--------------T----------AA--A---- 2319
C.ZA.08.705010534_CH534.w12 ------C--A--T--------------A--------------------------T--G-------------G----C-------C-----G-TGA----------A--G--C--T--------T--------C--------------------------------A---- 3134
C.ZA.10.DEMC10ZA001        ---------A--T--------------------------------------------G--------------------------C------GC------------A-----C-----------T--------C--A-----------------------------A---- 2484
C.ZM.11.DEMC11ZM006        ---------A--T--------------A--------------------------T--G--------------------------C------GC------------A-----C-----------T--------C-----------G--------------------A--C- 2493
D.CM.10.DEMD10CM009        ------------------------------------------------G---------T-C--T-------------------------AG--G-----------AC-------------------------------------------T-----G------------- 2519
D.CY.06.CY163              ---G--------------------------------------------G---------T----T-------------------------A---G-----------AC-------------G-----------------------------T------------------- 2352
D.KE.11.DEMD11KE003        ---------------------------------------G--------------T--------------------------A-----------------C-----A--G-----------------------C------------------------------------- 2494
D.KR.04.04KBH8             ------------------------------------------------------------------------------------C--------------------A--------A------------------C------------------------------------ 3103
D.SN.90.SE365              ---------------------C----------------------AG--------------------------------------C--------TA----------A-----------C--------------------------------------G-----A------- 3161
D.TZ.01.A280               ------------------------------------------------------T--------------------------A--C--------------------A--------------------------------------------------G-----A------- 2345
D.UG.10.DEMD10UG004        ------------------A--------------------------G--------------------------G--------A-----------G-----------A--G--------------------------A--------------T------------------- 2482
D.UG.11.DEMD11UG003        ---------------C-----------A-----------------------------------------------------A-----------------------A--G--------------------------A---------------------------------- 2499
D.YE.02.02YE516            ---------------------------A--G-----G---------------------T-C--T--------------------C--------------------AC-------------G-----------------------------T------------------- 2355
D.ZA.90.R1                 ---------------------------------------------G------------T-------------------------C--------------------T-----------------------------------------------------------A---- 2490
F1.AO.06.AO_06_ANG32       C--------------------G------------------------------------T----------------A--------C-------C-AG---------A-----------C---------------C----------G-----------------A------- 2355
F1.AR.02.ARE933            ------C--------------------A--G-----------------------T---T-------------------------C-------C-A-------------G--------C--------------------------G------------------------- 2438
F1.BR.10.10BR_RJ015        C--G--C--A-----------------A------------------------------T----------------------T--C-------C------------------------C--------------------------G------------------------- 2617
F1.CY.08.CY222             C-----C--------------------A--------------------------T--------T--------------------C------GT-A----------------------C--------------------------G-----------------A------- 2337
F1.ES.02.ES_X845_4         C-----C---------------------------------------------------T----------------A----------------T-A----------------------C---------------C-------------A--T-----------A------- 2596
F1.RO.96.BCI_R07           C-----C--------------------A------------------------------TAC-----------------------C-------C-AG------------CA----T--C--------------------------G------------C------------ 3168
F1.RU.08.D88_845           ------Y-----------A--C-----A-----------G------------------TC-------------------A----C-------C-A----------------------C--------------------------G---------A-------A------- 2603
F2.CM.02.02CM_0016BBY      ---------------------------------------G--------------------------------------------C------GC-AG---------G-----G-----C--------------------------G------------------------- 2334
F2.CM.10.DEMF210CM001      ---------------------C-----------------G-------------------------------G------------C-------C-A-----T----A-----A-----C--------------C------------------------C-------A---- 2408
F2.CM.10.DEMF210CM007      ----------------------------------------------------------------------C-T-----------C------GC-A----------A-----------C--------------------------G------------------------- 2491
G.BE.96.DRCBL              --------------T.......--------------------------------T--G--------------------------C-------C------------A-----G-----C-----------------A-----------A---------------------- 3101
G.CM.10.DEMG10CM008        ------------------------------------------------------T--GT------------G------------C--------GA----------A--G--G-----C-----------------A--------G--------------------A---- 2525
G.CN.08.GX_2084_08         ---------------------------A--R--------------------------G-------------R---------A----------T-A----------A--G--G-----C--G--------------A--------------------G-CC---------- 2384
G.CU.99.Cu74               ------------T-----------------------------------------T--G--C--T--------------------C--------------------A----AA-----C-----T-----------A--------------------------A------- 2748
G.ES.09.X2634_2            ------------------A--C--------------------------------T---T--------------------A----C-------T-A---------------AG-----C-----T-----------A------------------A--------------- 2627
G.GH.03.03GH175G           ---------------------C--------------------------------T--G----------T------------A--C-------C-A----------------G-----C-----------------A--------G-----------------A------- 3184
G.KE.09.DEMG09KE001        G--G--------T--------G-----------------------------C--T--G--------------------------C-------T-A-------------G--G-----C-----T-----------A--------------------G------------- 2519
G.NG.09.09NG_SC62          C-----------------------------------------------------T--G-----------------------A--C-------C-A----C-----AT-G--G-----C--G--------------A---------------------------------- 2336
G.ZA.01.TV546              ------------------A-----C-----------------------------T--G-----T--------------G-----C-------C-A----------A--G--G-----C-----T--------C--A---------------C---------A-------- 2346
H.BE.93.VI991              ------------T-----------------------------C--------------G--------------G-----------C-----------------T--AG-GA----T--C--------------------------------------------A--A---- 2533
H.CF.90.056                ------------T--------------------------------------------G----------------------C---C------G----------T--A--G-----T--C-----------------------------------------------A---- 2491
H.GB.00.00GBAC4001         ------------T--------------------------------------------G-A------------------------C-----G-----G-----T--A--G-----T--C-----------------------------------------------A---- 2662
J.CM.04.04CMU11421         C--------------Y-----A-----------------------------------GT----------------A-----A-----C----C-A----C-----MC----C--------G-----------C--------G--------T----------AA-----C- 2650
J.SE.93.SE9280_7887        C--------------C-----A--------------------C-----------T--GT--------------------A-A---------G--AG---C-----A-----------C--G-----------C--------G-----------G-------AA------- 2458
J.SE.94.SE9173_7022        C--------------C-----A--------------------------------T--GT--------------------A-A---------G--AG---C-----A-----------C--G-----------C--------G-----------G-------AA------- 2459
K.CD.97.97ZR_EQTB11        ------------------A--A--------------------------------T---T-------------------------C-------GGA----------T--G---T-A--C--------------------C-----G--------------------A---- 2340
K.CM.96.96CM_MP535         ---------------------A-----A------------------------------CA------------------------C-------T-A----------A--G-----A--C-----------------A--------G--------------------A-C-- 2340
U.CA.01.TV749              ---------------C---T-A--------------------------------T--GT----------------T-----A--C-------G------C-----AC----------C-----------------------G-----------------------A---- 3161
U.CA.99.TV721              ---------------C-----A--------------------------------T--GT----------------T-----A--C--------------C-----AC----------C--------------C--------G--------------G--------A---- 3160
U.CD.83.83CD003_Z3         ---------------------------------------------------------G--------------------------C---------G----------G-----------C--------------------------C------------------------- 2639
U.CD.90.90CD121E12         ----------------------------------------------T----------G--------------------------C-----G-C-A----------A----C------C--G-----------------------C------------------------- 2343
U.CY.05.CY090              ------------T--------C-----------------C------------------T-------------------------C---C---C-A----------A-----------C---------A----------------G---A-----A-------A--A---- 2343
U.CY.08.CY223              ------------------------------------------C-----------T--G-----------------T-----T--C--C---GC-A----------A--G--G--A--------------------A--------------T------------------- 2361
U.ES.10.DEURF10DZ001       ---------------------------------------------------C--T--GGCA-------------------------------C-A----------GG----G-----------------------A--------------T-----G--------A---- 2420
U.GR.99.99GR303            ---------------------------------------------------------G-----T-----------------A--C-------T-A----------A---A-C--A--C--------------------------G-----T--------------A---- 2418
U.NL.95.U_NL_95_H10986_D1  ---------A--------------------------------------------T---T----------------------A--C-A----GC-A----------AG-GA----------------------C--A-----G-----A-------------AA------- 3029
01_AE.AF.07.569M           C-----------------A--G-----------------------T-----------G--------------------------C----A--T-A----------G-----C--------------------C--------R--------T-----------------C- 2350
01_AE.CF.90.90CF11697      C--------------------G-----------------------------------G--------------------------A-A----GC-A----------A--------------------------C-----------------T------------------- 3093
01_AE.CN.10.YNFL03         C--------------------G-----------------------G-----------G----------T---------------C-------T-A----------A--G-----------------------C-----------------T-----------------C- 2522
01_AE.HK.04.HK001          C--------------------G-----------------------G-----------G-----------------------A---------GC-A-C--------A---A----------------------C-----------------------------A------- 2506
01_AE.IR.10.10IR.THR48F    C--------------------A-----------------------------------G--------------------------C----AG-T-A----------A--GA----------C-----------C-----------------T-----------------C- 2334
01_AE.JP.x.JRC77AE         C-----C--------------G-----------------------G-----------G-----------------T--------C-----G-T-A----------AG-G--C--------------------C-----------------T-----------------C- 3141
01_AE.TH.09.AA090a_WG11    ---------------------G-----------------------T-----------G------------------C-------C------CT-A-----T-------------------------------C-----------------------------------C- 2567
01_AE.TH.90.CM240          C--------------------G-----------------------G-----------G--------------------------C-------T-A----------A--G----------------A------C-----------------T-----------------C- 2718
01_AE.US.05.306163_FL      C--------------------G-----------------------------------G--------------------------C-------C-A----------A---A----------------------C-----------------T-----------------C- 2335
01_AE.VN.98.98VNND15       C--------------------G-----------------------G-----------G--------------------------C-------T-A----------A--------------------------C--A--------------T-----------------C- 2399
```


Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

HIV-1 Genomes

B.FR.83.HXB2 GAACAAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAAT 3314
Pol R T K I E E L R Q H L L R W G L T T P D K K H Q K E P P F L W M G Y E L H P D K W T V Q P I V L P E K D S W T V N
A1.AU.03.PS1044_Day0 2522
A1.CH.03.HIV_CH_BID_V3538 2519
A1.CY.08.CY236 2525
A1.ES.05.X1608_8 2762
A1.KE.11.DEMA11KE001 2653
A1.RU.11.11RU6950 2838
A1.RW.11.DEMA111RW002 2665
A1.SN.01.DDI579 2512
A1.UG.11.DEMA110UG001 2669
A1.ZA.04.04ZASK162B1 2755
A2.CM.01.01CM_1445MV 2513
A2.CY.94.94CY017_41 2671
B.BR.10.10BR_MG029 2776
B.CA.07.502_1191_03 2741
B.CH.08.M2_0803101_NFLG8 2783
B.CN.10.DEMB10CN002 2686
B.ES.10.DEMB10ES002 2676
B.FR.11.DEMB11FR001 2560
B.GB.05.MM45d213_GN1 2715
B.HK.06.HK003 2684
B.HT.05.05HT_129389 2515
B.JP.12.DEMB12JP001 2701
B.KR.07.HP_18_07JHS10_3909 2978
B.PE.07.502_0525_wg5 2779
B.RU.11.11RU2ln 2885
B.TH.08.MERLBDTRC10 2547
B.US.11.ES38 3249
C.AR.01.ARG4006 2498
C.BR.07.DEMC07BR003 2646
C.BW.00.00BW5031_1 2685
C.CN.10.YNFL19 2667
C.CY.09.CY260 2537
C.ES.08.X2363_2 2745
C.ET.02.02ET_288 2516
C.IN.09.T125_2139 3290
C.KE.00.KER2010 2495
C.MW.09.703010256_CH256.w96 3294
C.TZ.08.707010457_CH457.w8 3288
C.YE.02.02YE511 2489
C.ZA.08.705010534_CH534.w12 3304
C.ZA.10.DEMC10ZA001 2654
C.ZM.11.DEMC11ZM006 2663
D.CM.10.DEMD10CM009 2689
D.CY.06.CY163 2522
D.KE.11.DEMD11KE003 2664
D.KR.04.04KBH8 3273
D.SN.90.SE365 3331
D.TZ.01.A280 2515
D.UG.10.DEMD10UG004 2652
D.UG.11.DEMD11UG003 2669
D.YE.02.02YE516 2525
D.ZA.90.R1 2660
F1.AO.06.AO_06_ANG32 2525
F1.AR.02.ARE933 2608
F1.BR.10.10BR_RJ015 2787
F1.CY.08.CY222 2507
F1.ES.02.ES_X845_4 2766
F1.RO.96.BCI_R07 3338
F1.RU.08.D88_845 2773
F2.CM.02.02CM_0016BBY 2504
F2.CM.10.DEMF210CM001 2578
F2.CM.10.DEMF210CM007 2661
G.BE.96.DRCBL 3271
G.CM.10.DEMG10CM008 2695
G.CN.08.GX_2084_08 2554
G.CU.99.Cu74 2918
G.ES.09.X2634_2 2797
G.GH.03.03GH175G 3354
G.KE.09.DEMG09KE001 2689
G.NG.09.09NG_SC62 2506
G.ZA.01.TV546 2516
H.BE.93.VI991 2703
H.CF.90.056 2661
H.GB.00.00GBAC4001 2832
J.CM.04.04CMU11421 2820
J.SE.93.SE9280_7887 2628
J.SE.94.SE9173_7022 2629
K.CD.97.97ZR_EQTB11 2510
K.CM.96.96CM_MP535 2510
U.CA.01.TV749 3331
U.CA.99.TV721 3330
U.CD.83.83CD003_Z3 2809
U.CD.90.90CD121E12 2513
U.CY.05.CY090 2513
U.CY.08.CY223 2531
U.ES.10.DEURF10DZ001 2590
U.GR.99.99GR303 2588
U.NL.95.U_NL_95_H10986_D1 3199
01_AE.AF.07.569M 2520
01_AE.CF.90.90CF11697 3263
01_AE.CN.10.YNFL03 2692
01_AE.HK.04.HK001 2676
01_AE.IR.10.10IR.THR48F 2504
01_AE.JP.x.JRC77AE 3311
01_AE.TH.09.AA090a_WG11 2737
01_AE.TH.90.CM240 2888
01_AE.US.05.306163_FL 2505
01_AE.VN.98.98VNND15 2569

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546


```
B.FR.83.HXB2                  GAACAAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAAT 3314
Pol                           R__T__K__I__E__E__L__R__Q__H__L__L__R__W__G__L__T__T__P__D__K__K__H__Q__K__E__P__P__F__L__W__M__G__Y__E__L__H__P__D__K__W__T__V__Q__P__I__V__L__P__E__K__D__S__W__T__V__N_
02_AG.CM.10.DE00210CM013      --G--------------T-----G-------AC---AA------T-------------T--------------------------T-----------A-----------------C-----------C----A----CA---------------T--------------- 2681
02_AG.CY.09.CY256             --G--------R-----T-----GRW-----AC-R-R-------T----T-----R-----G--------------------R--T-----------A-----------------C--G--------C----A----CA------------------------------- 2510
02_AG.ES.06.P1423             -GG--------------T-A---GG------A----AA------T----------------------------G-----------T-----------A-----------------C-----------C-----C---CAA--------C--------------------- 2752
02_AG.GW.05.CC_0048           --G--------------T-----G-------A----AA-----------------------G--------A--G-----------T-----------A-----------------C-----------T---------CA---A-----------------------T--- 2664
02_AG.KR.12.12MHR9            --G----------TA--T-----G-----T-AC----A------T----------------------------------------T-----------A-----------------C-----------C---------CAA------------------------------ 2973
02_AG.LR.x.POC44951           --G-----------C--T-----G-C-----AC---AA------T----------------------------G-----C-----T-----------A-----------------C-----------C---------CA------------G------------------ 3309
02_AG.NG.09.09NG_SC61         --G--------------T-A---G-------AC----A-----------T-----------------------------------T-----------A-----------------C-----------T-----C---CA---------C--G--A--------------- 2506
02_AG.NG.x.IBNG               --G--------------T-----GG------AC---AA------T-------------------G--------------------T-----------------------------------------C------G---AA--A--------------------------- 2839
02_AG.SE.94.SE7812            --G--------------T-----G-------AC----A------T-------------TG----G--------------------T-----------A-----------------C-----------C------------------------------------------ 2686
02_AG.SN.98.98SE_MP1211       --G--------------T-----G-------AC---AA------T----------------------------G-----------T-----------A-----------------C-----------C---------CA------------------------------- 2511
03_AB.RU.97.KAL153_2          -----G----------A------G--------C-----------T-------------------------------------------------------------------------------T---------------T----------------------------- 2540
04_cpx.CY.94.94CY032_3        --G----------A-----A---G-G-----A-----A------T-C--------------------------------C-----T-----------G-----------------C-----------G---------CAA-C-G-------G--T--------------C 2680
05_DF.BE.x.VI1310             -GG--------A-----T-A---G-------AC-GCA------GT----T-----------------------------------------------G-----------------------------G---------CAAT-------C--G------------------ 2696
06_cpx.AU.96.BFP90            --G--------------T-A---G-G-----A----AA--------------------T--------------------C-----T-----------G-----------------C-----------G---------CAA--C-----C--G--G--------------- 3342
07_BC.CN.98.98CN009           ----------------A------------T--------------T----------------G-----------------------T-----------G-----------------C---------------------CA---------------T--------------- 2646
08_BC.CN.97.97CNGX_6F         ----------------A------G----------A-A-------T----------------G-----------------------T-----------G-----------------C---------------------CA------------G--T--------------- 2494
09_cpx.GH.96.96GH2911         --G--------A----A--A---G-------A----AA------T----T--------------G--------------C-----T-----------G-----------------C-----------G---------CAA--------C--G--A--------------- 2516
10_CD.TZ.96.96TZ_BF061        C--T------------A--A--GGG---C--A--A-A-------T-----G-------------G-----------G--------T--G--------G---------------------------T-----------CA-------A----GA-T--T------------ 2693
11_cpx.CM.95.95CM_1816        --GA---GG----A--AT-A--GA-G-----A------------T-C-----------T--------C--------------------C--------A-----G-----------C--------T-----A-------A---------C--G--AT-T------------ 2678
12_BF.AR.99.ARMA159           ----G-----------A--A--------------A---------T-------------------------------------------------------------------------------G--------------------------------------------C 3319
13_cpx.CM.96.96CM_1849        --GA---GG-------AT-A---AG-------------------T-------------T--G-----------------------------------A-----G-----------C--------------AG-----CA-T-------C--G------------------ 2719
14_BG.ES.05.X1870             --G--------------T-A---A----------A-AC-----GT-------------T--------------------------------------A-----G-----------C--------------A------CA---------C-----A--------------- 2781
15_01B.TH.99.99TH_MU2079      ----------------A--A---GCT-----A-----C-----GT----T--G--------------------G-----------------------A-----------------C-G---------C----------AA------------------------------ 2713
16_A2D.KR.97.97KR004          ----------------AT-A---A-T-----AC---AA------T--TA---------T--------------------------T----------AA-----------------C-----------C---------AA------------G------------------ 2677
17_BF.AR.99.ARMA038           ----------------A--A--------------A-A-------T-A--------------------------------------------------------------------------------G---------A-A-----------G------------------ 2528
18_cpx.CU.99.CU76             ----------------A--A-----------A----AA------T----------------------------------------T-----------A-----------------C-----------C-----G---CA------------------------------- 2619
19_cpx.CU.99.CU7              -----------------T-A--G-------------A-------T-------------------------------------G--------------G--------------------G------------------TCA--------------A--------------- 2540
20_BG.CU.99.Cu103             ----------------A----------------G--A-------T--TA---------T-----------------G--C-----------------A-----G-----C-----C--------------A------CA---------C--------------------- 2780
21_A2D.KE.99.KER2003          ----------------AT-A---G-------A------------T-G-----------T--------------------A-----T---------------------------------------------------AAA-----------G--G--------------- 2519
22_01A1.CM.01.01CM_0001BBY    ---T---------C---T-----GCT-----A-----C-----GT----T--------------G-----A--------------T--C--------A--C--------------C-----------C--A------CA-------A-C--------------------- 2522
23_BG.CU.03.CB118             --G-------------A-----------------A-A-------T----------------------------------------T--------------------A--------------------------------------------------------------- 2757
24_BG.ES.08.X2456_2           --G-------------A-----------------A-A-------T-C----------------------------------------AG-----------------------------------------------------------G-----A--------------- 2765
25_cpx.CM.02.1918LE           ----------------AT-A---G-----T-A----AA------T----T--------------G--------G-----------T-----------A-----G-----------C-----------C--A------AAA--------G--------------------- 2522
26_AU.CD.02.02CD_MBTB047      -GG--------------T-----GCT-----A----AA------T----A--------T-----G--------G-----------T-----------A-----------------C-----------C---------AAA--------------A-A------------- 3319
27_cpx.FR.04.04CD_FR_KZS      --G--------------T-A---GCC---T------A-------T----T--G--------------------------C-----T-----------A-----G-----------C--------------A------CA------------G--A--------------- 3305
28_BF.BR.99.BREPM12609        --------G-------A-----------------AGA------GT-------------------------------------------------------------------------------------G----------------G----TT------------ 2680
29_BF.BR.01.BREPM16704        ----------------A---------------------------T-------------------G--------------------------------------------------------------------------------------------------------- 2715
31_BC.BR.04.04BR142           ----------------A-----------C-------A-------T-CGAT-----------G-----------------C-----T-----------G-----------------C----------------------A---------T--G--G--------------- 2786
32_06A1.EE.01.EE0369          -GT--------------T-A--GG-----T-A----A-------T-C-----------T--------------------C-----T-----------G-----------------C-----------G----------AA--------C--G--A--------------- 2949
33_01B.ID.07.JKT189_C         ----G--G---------------GCT-----A----AC------T----T--------------G--------G-----C-----------------A-----------------C-----------C----------AA------------------------------ 2615
34_01B.TH.99.OUR2478P         -------------------A---G-T-----A------------T----T--------------G--------G-----------T-----------A-----------------C-----------C----------AA------------------------------ 2516
35_AD.AF.07.169H              ---T---------A---T-A---GCT-----A----A-------T----------------G--G--------------------T---------------------------------------------------AC---------------A--------------- 2514
36_cpx.CM.00.00CMNYU830       --G--------------T-----A-------AC---AT------T----------------------------------------T-----------A-----------------C-----------C---------CAA----------------A--------G---- 2519
37_cpx.CM.00.00CMNYU926       --G--------------------G-C--C--AC---AA------T-------------T-----------------------------C--------A-----------------------------C--A------CA-T-------T-------T------------- 2504
38_BF1.UY.03.UY03_3389        ----------------A---------------------------T-------C-----------------------------------------------------------------------------------------------G--G------------------ 2737
39_BF.BR.04.04BRRJ179         --G-------------A----------------G----------T--TA--------------------------------------A-------------------------------------------------A-------------G------------------ 2791
40_BF.BR.05.05BRRJ055         -------------A---T-A---G-----T-AC-A-AA-----GT--TTT--G-----GC----------A--------C-----------------G---------------------------------------A----------C--G------------------ 2797
42_BF.LU.06.luBF_18_06        ---T------------A-----G--G------C----------C--C--------------------------------------------------A-----------------------------G---------CAA------A-C--G------------------ 2868
43_02G.SA.03.J11223           --G--------------T-A---G----C--A----RA------T-------------T--------------------------T-----------A-----G-----------C--------------A------CA---------C--G--A--------------- 2839
44_BF.CL.00.CH80              ----------------A-------------------A-------T-A--------------------------------------------------------------------------------------------A-----------G------------------ 2756
45_cpx.FR.04.04FR_AUK         --G-------------A--A---G-----T-A----AA------T--------------C----G--------------C-----T-----------G-----------------C---------------------CAA--------C--G--A--------------- 3300
46_BF.BR.07.07BR_FPS625       ----------C------T-----A-C-----AC---AA------T----T--------------------A--G-----C-----T-----------G---------------------------------------CAAT-----A-C--G--G--------------- 2752
47_BF.ES.08.P1942             ----------------A--------------AC---AA-----G-----T--------------------------------C-----------------------A-------------------------------------A------G----A------------- 2753
48_01B.MY.07.07MYKT021        -------------------A---GCT--------A-AC------T----T-----------------------G-----------------------A-----G-----------C-----------C---------CAA------------------------------ 2519
49_cpx.GM.03.N26677           ----------------A--A---GC------A----AA------T-------------------G-----------G--C-----T-----------G-----------------C-----------G---------CAA--------C--G--AGC------------- 2736
50_A1D.GB.10.12792            --G----------A---T----GGCT-----A-----C------T----T--C-----------G-----------------G--------------A-----G--A--------C--G--------C--------------------------A--------------- 2721
51_01B.SG.11.11SG_HM021       -------GG-------A---------------------------T-C-----------T-----------------------------------------------------C-----------------A-------A-T----------G------------------ 2521
52_01B.MY.03.03MYKL018_1      ----------------A---------------------------T----------------------------------------------------------------------C-G---------C----------AA------------------------------ 2628
53_01B.MY.11.11FIR164         -------------------A--GGCT-----A-----C------T----T--------T-----G-----A--G--------C--------------A-----------------C-GG--------C----------AA----------G------------------- 2654
54_01B.MY.09.09MYSB023        ----------------A--------G-----------T------T-CTT----------C----------A--------C---C-------------A-----------------C-G---------C----------AA------------------------------ 2716
55_01B.CN.10.HNCS102056       -------------------------G---T------A------------------------G-----------------C----------------------------------------------------------C------------G------------------ 2624
56_cpx.FR.10.URF5_patient_A   --------G-------A--A---------T------AA--------A--------------------------------------T-----------A--------T--------C-----------C---------CA---A-----C--------------------- 2665
57_BC.CN.09.09YNLX19sg        --G--------------T-A---G----------A-A-------T-C-----G--------------------------C----------------------------------------------------------A-------------G------------------ 2498
58_01B.MY.09.09MYPR37         --------G----------A---GCT-----A-----C------T----T--G-----T-------------------------------------------------------------------------------A------------------A------------- 2659
59_01B.CN.09.09LNA423         ---T----G----A-----A---GCT-----A-----C------T----T--------------G--------G-----------------------A-----------------C-----------C------G---AA------------------------------ 2498
60_BC.IT.11.BAV499            --G--------------T-A---A-------A---CA---------CG-T-----------------------------C-----T-------C---G-----------------C---------------------CA---------C--G--T--------------- 3271
61_BC.CN.10.JL100010          --G-------------A--A---G----------A-A-------T----------------G-----------------------T-----------G-----------------C---------------------CA------------G--T--------------- 2735
62_BC.CN.10.YNFL13            --G-------------A-------------------A-------T-----------------------------C----------T-----------G-----------------C---------------------CA------------G--T--------------- 2643
63_02A1.RU.10.10RU6637        --G--------------T-----GC------A--A-AT------T----T--------------G--------------------T--C--------A-----------------C-----------C---------CA---------C--------------------- 2855
64_BC.CN.09.YNFL31            --G--------------T-A---G----------A-A-------T-C--------------G-----------------------T-----------G-----------------C---------------------CA------------G--T---------A----- 2668
65_cpx.CN.10.YNFL01           --G--------------T-A---G----------A-C------GT----T-----------------------------C-----T-----------G---------------A-C---------------------CA------------G--T--------------- 2685
67_01B.CN.11.ANHUI_HF115      -G-T------------A-----G-------------A-------T-A--------------------------G-----------T-----------A--------------A--C-G---------C----------A------------------------------- 2657
68_01B.CN.11.ANHUI_WH73       -G-T------------A-----G---------C---A-------T-A--------------------------G-----------T-----------A--------------A--C-G---------C----------A------------------------------- 2648
70_BF1.BR.10.10BR_PE004       --G--------------T-A---G-------AC---AA------T----T-----------------------G-----C-----------------G-----------------------------G---------CAAT-----A-C--------------------- 2770
71_BF1.BR.10.10BR_PE008       --G--------------T-A---G-------AC-A-AA-----------T--------------G-----A--------C-----------------G-----------------C-----------G---------CA-T----------G------------------ 2978
72_BF1.BR.10.10BR_MG002       -----------------T-----G-------A----AA------T----T--------------------A--G-----C-----------------G--------------A--C-----------G---------CAAT-A--------G-----------A------ 2822
O.BE.87.ANT70                 ---A--GG--T--ATT---T---G----C--A-ATCA-------T-C--T--C--T--------G-----A--G--------C--T--A--------G-----G--------A--C---------------T-C--CCAAT----TA-C--G--TGTG-----A--A--- 3368
O.CM.98.98CMA104              -G-A--GGG-T--ATT---T--GG-----T-A-AT-A-------T-C--T--C--T--T-----G--------G--------C--T--G--------G-----G--------A--C--G-----------A--C--CCA-T----T--T--G--AGTA-----A--T--- 2800
O.CM.98.98CMABB141            ---A--GGG-T--ATT---T---G----CT-A-ATCA-------T-C--T--C--T--T-----G-----A--G--G-----C--TT-G--------G--------------A--C-----------------C--CCAAT----TAGC--G---GTG-----A--A--- 2799
O.CM.98.98CMABB212            ---AG-GGG-T--ATT---T--GG-----T-A-ATCA-------T-C--T--C--T--T-----G-----A--G--------C--TT-A--------A-----G--------A--C--G--------G-----C--CCAAT----TA-C-----AGTG-----A--A--- 2804
O.CM.98.98CMU5337             -G-AC-GGG----AA----T---G-G--CT-A-ATCA-------T-C--T--C--T--T--G--G-----A--G-----------T--G--------G-----G--------A--C-----------------C--CCAAT-A--TA-C--G--GGTG-----A--A--- 2801
O.CM.99.99CMU4122             ---A--GGG----ATT-------G----CT-A-ATCA-------T-C--T--C--T--T-----G-----A--G--------C--T--G--------G-----G-----C--A--C-----------------C--CCAAT----TA-C--G--G-TA-----A--A--- 2799
O.FR.92.VAU                   -G-A--GGG-T--ACTA--C---A-----T-A-AC-A-------T--TAT--C--T--T-----------A--G-----A--C--T--AG-------G-----G-----C--A--C--G--------------C--CCAAT----T--C--G--AGTG-----A--A--- 2880
O.GA.11.11Gab6352             ---A--GGG-T--ATT---C---G----CT-A-AT-A------GT--TAT--T--T--T-----G-----A--G--------C--T--C--------G-----G--------A--C--------------------CCAAT----T--C--G--GGTA-----A--A--- 2525
O.SN.99.99SE_MP1299           ---AG-G-G----ATT---T---G----CT-A-ATCA-------T-C--T--C--T--T-----G-----A--G--------C--TT-G--------G-----G--------A--C-----------G-----C--CCAAT----TA-C--G--GGAA-----A--A--- 3367
O.US.10.LTNP                  ---A--GGG-T--ATCA--T---G-----T-A-ATCA------GT-C--T--C--T--T-----------A--G--------C--T--G--------G-----G--T-----A--C-----------C--A--C--CCAAT-A--TA-C--G--TGTG-----A--A--- 3289
N.CM.02.DJO0131               --GAG-C-G-G--A--A--T---AGT-----T----A------CT----G--C--T--------------A--G--G--A--G-----C--------A-----G--------A-----------G--C---A-A---AA-T-A------C-G--TGTA------------ 2820
N.CM.02.SJGddd                --GAG-C-G----A--A--T---AGG-----T---GA------CT----G--C--T-----G--------A--G--G--A--G--T--C--------A-----G--------A-----------G--C-----A---AA-T----------G--TGTA------------ 2804
N.CM.04.04CM_1015_04          --GAG-C-G----A--A--T---GGT-----T----A------CT----A--C--T--------------A--G--G--A--G-----C--------A-----G--------A--------------C-----A---AA-T-A--------G--T-CA------------ 2819
N.CM.06.U14296                --GAG-C-G----AA-A--T---G-G-----T---CA------CT-C--G--C--T--------------A--G--G--A--------C--------A-----G--------A-----------G--C-----A---AA-T-A-----G--G--TGTA------------ 2817
N.CM.06.U14842                --GAGGC-G--C-A--A--T---G-T-----T----A------CT----G--C--T--------------A--G-----A--------C--------A-----G--------A--------------C-----AG--AA-T-A-----G--G--AGTA------------ 2821
N.FR.11.N1_FR_2011            --GAGGC-G----A--A--T--GGG------TC---A------CT----G--C--T--------------A--G-----A--------C--------A-----G-----C--A--------------C-----A---AA-T-A--------G--TGTA------------ 2715
P.CM.06.U14788                --CAGCGGG-----A-A--Y--GG-R--C--C-ATGTA-----GT-C--A--T--T--------R-----A--G--G-----T-----C--------A-----G-----------C--G-----K--G--A--C--CAAAT-A--------G--A--T-----A--A--- 2785
P.FR.09.RBF168                --CAGCGGG-----A-A--T--GG----C--C-ATGTA-----GT----A--T--T--------------A--G--G-----T-----C--------A-----G-----------C--G-----G--G--A--C--C--AT-A--------G--AGCT-----A--A--- 3349
CPZ.CD.06.BF1167              A-GA-GTGG-TC-----T-A--A--C--CT--AAAGCA-----CT-AGAA--C--C--T------T----AG-GA----A-----T--C--------A------T-A--C--A--G--G--------G----AG---AAAT-A-----GCC---A-CA-----A--T--- 3416
CPZ.CM.05.SIVcpzMT145         ---AG-G-G----A-----C---GC---------A-AA------T----A-----------G--G--------------C-----T--C--------A-----GT-A-----A-----G--------C----AA---AAAT--------C-G--AGTG--------T--- 2856
CPZ.GA.88.GAB1                ---A---GG-G--A--A--A---------T--C-T-AA-----GT-C--A--C-----------------A--G--G--A--C--TT-A--------A-----G-----C--A--C-----------C--------TCAAT-A-----------GGTA------------ 3375
CPZ.TZ.06.SIVcpzTAN13         ---AGCT---GA-A-----A-----G--CT---AGCA-------T-AGAA--C------------T----AG--A-----------T-A--------C-----G--A-----A--------------G---AAC---CACC--------CC---ACAG-----A--G--- 2955
CPZ.US.85.US_Marilyn          -GGA---GG----ACT---C------T--T--C-T-CT------T-C--T--C--T--T--G--G------G-G-----A------T-A--------C--------------A-----G--------G-----C---CA-T-A---C-------A-TT-----A------ 3374
GOR.CM.07.SIVgor2139_287      -GCAG--GG----AA-A--C--GG-----T-A-ATGCA------T-C-----T--T--------G-----A--G--------C--TT-G--------G-----G--G--C--------G--------G---------AA-T-A--------G--A--T-----A------ 2851
GOR.CM.07.SIVgorCP2135con     -GCAG---G----AA-A--C--GG-----T-A-ATGCA------T-C-----C--T--------G-----A--G--------C--TT-G--------G-----G--G--C--------G--------G--A------AA-T-A--------G--A--T-----A------ 2850
```

B.FR.83.HXB2 GACATACAGAAGTTAGTGGGGAAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAAACTCCTTAGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGAGAGATTCTAAAAGA 3484
Pol D I Q K L V G K L N W A S Q I Y P G I K V R Q L C K L L R G T K A L T E V I P L T E E A E L E L A E N R E I L K E
A1.AU.03.PS1044_Day0 2692
A1.CH.03.HIV_CH_BID_V3538 2689
A1.CY.08.CY236 2695
A1.ES.05.X1608_8 2932
A1.KE.11.DEMA11KE001 2823
A1.RU.11.11RU6950 3008
A1.RW.11.DEMA111RW002 2835
A1.SN.01.DDI579 2682
A1.UG.11.DEMA110UG001 2839
A1.ZA.04.04ZASK162B1 2925
A2.CM.01.01CM_1445MV 2683
A2.CY.94.94CY017_41 2841
B.BR.10.10BR_MG029 2946
B.CA.07.502_1191_03 2911
B.CH.08.M2_0803101_NFLG8 2953
B.CN.10.DEMB10CN002 2856
B.ES.10.DEMB10ES002 2846
B.FR.11.DEMB11FR001 2730
B.GB.05.MM45d213_GN1 2885
B.HK.06.HK003 2854
B.HT.05.05HT_129389 2685
B.JP.12.DEMB12JP001 2871
B.KR.07.HP_18_07JHS10_3909 3148
B.PE.07.502_0525_wg5 2949
B.RU.11.11RU2ln 3055
B.TH.08.MERLBDTRC10 2717
B.US.11.ES38 3419
C.AR.01.ARG4006 2668
C.BR.07.DEMC07BR003 2816
C.BW.00.00BW5031_1 2855
C.CN.10.YNFL19 2837
C.CY.09.CY260 2707
C.ES.08.X2363_2 2915
C.ET.02.02ET_288 2686
C.IN.09.T125_2139 3460
C.KE.00.KER2010 2665
C.MW.09.703010256_CH256.w96 3464
C.TZ.08.707010457_CH457.w8 3458
C.YE.02.02YE511 2659
C.ZA.08.705010534_CH534.w12 3474
C.ZA.10.DEMC10ZA001 2824
C.ZM.11.DEMC11ZM006 2833
D.CM.10.DEMD10CM009 2859
D.CY.06.CY163 2692
D.KE.11.DEMD11KE003 2834
D.KR.04.04KBH8 3443
D.SN.90.SE365 3501
D.TZ.01.A280 2685
D.UG.10.DEMD10UG004 2822
D.UG.11.DEMD11UG003 2839
D.YE.02.02YE516 2695
D.ZA.90.R1 2830
F1.AO.06.AO_06_ANG32 2695
F1.AR.02.ARE933 2778
F1.BR.10.10BR_RJ015 2957
F1.CY.08.CY222 2677
F1.ES.02.ES_X845_4 2936
F1.RO.96.BCI_R07 3508
F1.RU.08.D88_845 2943
F2.CM.02.02CM_0016BBY 2674
F2.CM.10.DEMF210CM001 2748
F2.CM.10.DEMF210CM007 2831
G.BE.96.DRCBL 3441
G.CM.10.DEMG10CM008 2865
G.CN.08.GX_2084_08 2724
G.CU.99.Cu74 3088
G.ES.09.X2634_2 2967
G.GH.03.03GH175G 3524
G.KE.09.DEMG09KE001 2859
G.NG.09.09NG_SC62 2676
G.ZA.01.TV546 2686
H.BE.93.VI991 2873
H.CF.90.056 2831
H.GB.00.00GBAC4001 3002
J.CM.04.04CMU11421 2990
J.SE.93.SE9280_7887 2798
J.SE.94.SE9173_7022 2799
K.CD.97.97ZR_EQTB11 2680
K.CM.96.96CM_MP535 2680
U.CA.01.TV749 3501
U.CA.99.TV721 3500
U.CD.83.83CD003_Z3 2979
U.CD.90.90CD121E12 2683
U.CY.05.CY090 2683
U.CY.08.CY223 2701
U.ES.10.DEURF10DZ001 2760
U.GR.99.99GR303 2758
U.NL.95.U_NL_95_H10986_D1 3369
01_AE.AF.07.569M 2690
01_AE.CF.90.90CF11697 3433
01_AE.CN.10.YNFL03 2862
01_AE.HK.04.HK001 2846
01_AE.IR.10.10IR.THR48F 2674
01_AE.JP.x.JRC77AE 3481
01_AE.TH.09.AA090a_WG11 2907
01_AE.TH.90.CM240 3058
01_AE.US.05.306163_FL 2675
01_AE.VN.98.98VNND15 2739

```
B.FR.83.HXB2                 GACATACAGAAGTTAGTGGGGAAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAAACTCCTTAGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGAGAGATTCTAAAAGA 3484
Pol                          _D__I__Q__K__L__V__G__K__L__N__W__A__S__Q__I__Y__P__G__I__K__V__R__Q__L__C__K__L__L__R__G__T__K__A__L__T__E__V__I__P__L__T__E__E__A__E__L__E__L__A__E__N__R__E__I__L__K__E
02_AG.CM.10.DE00210CM013     --T--------A--------A---C-A--------------------TG----A----------A----C-G----G------C--G---G-A--------------TA--G-------G--T-----------AT-----T-------G-----G--A----------- 2851
02_AG.CY.09.CY256            --T--------AC----T--A---C-A--------------------TG----A--C-----G-A----C-G-----------C--G---G----G-----------TA--G-GA----GT-T--G--------AT-----T-------G-----G--A----------- 2680
02_AG.ES.06.P1423            --T--------A--------A---C-A--------------------TG----A----------AA--GC-G-----------C--G---G--------T-G-----T---G-----T-G--T--G--G-----AT-----T-------G-----G-------------- 2922
02_AG.GW.05.CC_0048          --T--------A--------A---C-A--------------------TG----A----------A----C------G------C--G---G----------------GA--G--A----G--T--G--------AT-----T---------------G--A--C-------- 2834
02_AG.KR.12.12MHR9           --T--------A-----A--A---C-A--------------------T-----A----------A----C-G-----------C--G---G----------------TA--G-------G--T--G--------AT-----T-------G-----G--A----------- 3143
02_AG.LR.x.POC44951          --T--------A--------A---C-A-----------------C--TG----A----G-----A------G-----------C--G---G----------------TA--G--A----G--T--G--------AT-----T-------G--T--G--A--------A- 3479
02_AG.NG.09.09NG_SC61        --T--------A--------A---C-A--------------------TG----A---------------C-G-----------C--G---G-------C--------TA--G-------G--T--G--------AT-----T---------------G--A----------- 2676
02_AG.NG.x.IBNG              --T--------A--------A---C-A--------------------TG----A------A---A----C-G----G------C--G---G----------------TA--G--G----G--T--G--------AT-----T-------G-----G--A----------- 3009
02_AG.SE.94.SE7812           --T--------A--------A---G-A------CA-GTCAG------TG----A----------A------G-----------C------G-----A----------TA--G--A----G--T--G--------AT-----T-------G-----G--A----------- 2856
02_AG.SN.98.98SE_MP1211      --T--------A--------A---C-A--------------------T-----A----------A----C-------------C--G------------T-------TA--G-------G--T--G--------AT-----T-------G-----G--A----------- 2681
03_AB.RU.97.KAL153_2         ------------C-------A--------------------------TG-------------------------------------G---G---------------------------------C------------------------------G-------------- 2710
04_cpx.CY.94.94CY032_3       --T--C--------------A---C-A--------------------T----------------A---------------T-----G---GT------C--------CA--G----------TAC---G------T-----T-A-----G-----G-------------- 2850
05_DF.BE.x.VI1310            --T--------------A--A---C-A--------------------T----------------A---------------------G-------------------G---------------------------AT-------------------G--A----------- 2866
06_cpx.AU.96.BFP90           --T--------A--G-----A---C-A--------------------T-----------------------------------C--G---G-A--------------CA--G----------T-C---------A------T-------G-----G--A--T-------- 3512
07_BC.CN.98.98CN009          --T-----A-----------A-----A--C-----------------T--T--A---------------C-T--------------G--GG----------------CA--G----------T-----------AT-----T-------------G--A----------- 2816
08_BC.CN.97.97CNGX_6F        --T------------------A-----A--C-----------------------A---------------C-T--------------G--GG----------------CA--G----------T-----------AT-----T-A-G---------G--A--T-------- 2664
09_cpx.GH.96.96GH2911        --T--------A-----A--A-----A--C-----------------TG---------------A----C-G-----------C------G-T-G-----G-TA--G--A----G--T--G--G-----AT-----T-------G-----G-------------- 2686
10_CD.TZ.96.96TZ_BF061       --T-----------------A--G--A--C--------------------------------G------C-G--------------G---G------T-T-------CA--G-------G--T-----G-----AT-----T-------------G--A----------- 2863
11_cpx.CM.95.95CM_1816       --T-----A--A--------A---C-A--------------------T------------A--------C----------------G--G---------T-------CA--G-------G--T-C----------T-G--GT-------------G--A----------- 2848
12_BF.AR.99.ARMA159          --------------------A-----A-----------------------------------G-A-----------G---------G--------G---------------------------A--------------------------------G--A----------- 3489
13_cpx.CM.96.96CM_1849       --T-----A-----------A---C-----------------A--C--T-----A----------A----C----------------G--GG----------------CA--G--T----G--TAC---------AT-----T-------G--------A--T-------- 2889
14_BG.ES.05.X1870            --T-----A-----------A---C-A-----------C-----A--T-----A----------AA---C----------------G---G----------------CA--G----------T-T---------AT-G---T-------G-----G--A-----G----- 2951
15_01B.TH.99.99TH_MU2079     --T--------A--------A---C-A---------------A-----TGG--------G----A----C-G-----------C--G---GT---------G-----CA--G-----T-G--T-----------AT-----T-------G-----G-------------C 2883
16_A2D.KR.97.97KR004         --T--------A--------A--G--A--------------------TG--A-----------A----C-G-----G-------G---GT---------G-----CA--G-------G--TA----------AT-----T-A-A---G-----G-------------- 2847
17_BF.AR.99.ARMA038          --------------------A-----A-----------------------------------G-A---------------------G--------G------------A--G-----------A-------------------------------G--A----------- 2698
18_cpx.CU.99.CU76            --T--------------A--A---C-----------------A-----T----------------A---GC-----------A----G--G----------------CA--G-------G--T-T---------AT-G---T-A-----------G--A----------- 2789
19_cpx.CU.99.CU7             --T--------------------GC-------------C--------T-----A-----------A--GC----------------G---G-------------------------------------------AT------------------G--A----------- 2710
20_BG.CU.99.Cu103            --------------------A--------------------------T---------------------C----------------G------------T-----------G-----------AC----------------T-------G-----G--A----------- 2950
21_A2D.KE.99.KER2003         --T-----------------A-----A--C--------C--------T-----A--------------G-----------------G-----------------------------------------------AT------------------G--A--------G-- 2689
22_01A1.CM.01.01CM_0001BBY   --T--------A--------A-----A--------------------TG-------------------GC-G--------A-----G---G--------------------G-----T-G--T-----------AT-----T-------G--T--G----------G--- 2692
23_BG.CU.03.CB118            --------------------A--------------------------T----------------A----C----------------G------------------------G-------G--T-C---------AT-G---T-------------G--A----------- 2927
24_BG.ES.08.X2456_2          --TC----A-----------A--GC-A--------------------T-----------------A---C-------G--------G---G--------------------G-----------AC----------------T-------------G--A--------G-- 2935
25_cpx.CM.02.1918LE          --T--------A--------A---C-A--------------------TG--------C-------A-----------------------GG-A--------------CA--G-------G--T-T---------AT-G---T-------G-----G--A---------A- 2692
26_AU.CD.02.02CD_MBTB047     --T--T-----A-----A--A---C-A--------------------TG-------A--G----A----C-G-----------C--G---G----------------CA--G-CA----G--TA--G-------AT-----T-------G-----------CT-G----- 3489
27_cpx.FR.04.04CD_FR_KZS     --T-----A--------A--A---C-A--------------------T--------C-------A----C------G---------G--GG----G----------CA--G--T--T-G--T-----------AT-G---T-------------G--A-------G--- 3475
28_BF.BR.99.BREPM12609       --T-----------------A-----A-------------------------------------A---------------------G--------G---------------G-------------------------------------------G--A----------- 2850
29_BF.BR.01.BREPM16704       --------------------A-----A---------------------G----A----------A---------------------G-----T--G---------------G-----------A----------------G--------------G--A--A-------- 2885
31_BC.BR.04.04BR142          --T-----------------A-----A--C-------------------------A-----G-A----C-G--------------G--------------------CA--G-------G--T-----------AT-----T-------G-----G--A----------- 2956
32_06A1.EE.01.EE0369         --T--------A--G--A--A---C-A---------------A-----T-------------------G--------------C--G---G-A-G-------------TA--G----G------T-C---------AT-----T-------G--------A--T-------- 3119
33_01B.ID.07.JKT189_C        --T--------AC-------A---C-A---------------A-----T-----------G----A----C-G-----------C--G---GT------T--------CA--G-------G--T-----------AT-----T-------G----AG--------G---AC 2785
34_01B.TH.99.OUR2478P        --T--------A--------A---C-A---------------A-----TG----------G----A----C-G-----------C--G---GT-------T-------CA--G-G-----G--T--------R--AT-----T-------G--------A---------AC 2686
35_AD.AF.07.169H             --T-----------------A-----A-----------C--------T-----A-----------A-----G--------------G--------------G-----TA--G--T----G---------------AT------------------G--AC---------- 2684
36_cpx.CM.00.00CMNYU830      --T--------A--G--A--A---C-A--------------------TG---------------A---GC-G--------A--C--G---G-A------------------G-----T-G--T--G--------AT-----T-------G--T--G-------------- 2689
37_cpx.CM.00.00CMNYU926      --T--------A--------A---C-A--------------------TG-------------------GC-G----G---T--C--G-----T-----GT-------T---G--A--AAG--TA-G--------AT-----T-------G----TTTT----G------- 2674
38_BF1.UY.03.UY03_3389       --------------------A-------------------------------------------A---------------T-----G------------T----------G--------------------------------------G-----G----------G--- 2907
39_BF.BR.04.04BRRJ179        --------------------A--------------------------T-----------G--G-A-----G-----G--------G------------------------G-----------C----------------T-------G--T----------------- 2961
40_BF.BR.05.05BRRJ055        --T---------C----A--A---C-A--------------------T--------------------G-----------------G---G----G-----------CA--G-G-----G--TAC----------T-----T-------G--T--G--A-----G--G-- 2967
42_BF.LU.06.luBF_18_06       --------------------A--------C-----------A-----TGG--A----------AA---------------------G------------------------G-------------------------------------------G-----T-------- 3038
43_02G.SA.03.J11223          --T-----A--A--------A---C-A--------------------T-----A----------A----C-G----G------C--G---G----------------T---G--A----G--T--G--------AT-----T-------G-----R--A----------- 3009
44_BF.CL.00.CH80             --------------------A-----------------------------------------G-A-----------G-T-A-----G---C-T--G---T-----------G-----------AG---------A-----------------T--G--A--T-------- 2926
45_cpx.FR.04.04FR_AUK        --T--------A--G-----A---C-A---------------------T----A-----G----A---------------------G---G-A--------------CA--G---T---G--T-C---------AT-----T-A-----G-----G--A-------G-AG 3470
46_BF.BR.07.07BR_FPS625      --T--------------A--A---C-A--------------------T----------G---G-A---------------------G---G----G-----------CA--G-GT----G--T-C----------T-----T-------G--T--G--A----------- 2922
47_BF.ES.08.P1942            --------------------A--------------------------------------------A--------C-G---------G--------G-----------CA--G-----------A----G--------------------------G--A--T-------- 2923
48_01B.MY.07.07MYKT021       --T--------A--------A---C-AG--------------A-----TG---------C-G----A------G-----------C--G-----T--------------CA--G-------G--T-----------AT----GT-------G-----G------------A- 2689
49_cpx.GM.03.N26677          --T--------A--G--A--A---C-A--------------------T----------------A----------------A----G---G-A--------------CA--G-----T-G--T-C---------AT-G-----------G-----G--A----------- 2906
50_A1D.GB.10.12792           --T--------A------------C-A---------------A-----TG---------------A----------------A-C--G---------------------A--G--A--T-G--T-CG--------AT----GT-------G-----G--A----------- 2891
51_01B.SG.11.11SG_HM021      --------------------A--------------------------TG---------------A---------C-----------G------------------------G-------G--------------------G--------------G-------------- 2691
52_01B.MY.03.03MYKL018_1     --T--------A--------A---C-AR--------------A-----TG----------RR---A----C-G--------------G---GT---------------C---G-------G--T-----------AT----GT-------G-----G-----------AC 2798
53_01B.MY.11.11FIR164        --T--------A-----A--A---C-A---------------A-----TG----------G----A----C-G-----------C------GT-------T-------CA--G---AGM-G--TR----------AT----GT-A-----G-----G-----------AC 2824
54_01B.MY.09.09MYSB023       --T--------A--------A--GC-A---------------A-----TG----------G----A----C-G-----------C--G---GTT------T-------CA--G----------T-C---------AT----GT-A-----G-----G-----------AC 2886
55_01B.CN.10.HNCS102056      --------A-----------A--G--A---------------A-----TG------C--G--G-A----C-G--C--G-----C--G---GT---G--G--------CA--G----------T-----------AT----GT-------G-----G-----------AC 2794
56_cpx.FR.10.URF5_patient_A  --T--------A--------A-----A-------------------------------------------------------------------------------------------T-------C------------------------------G-------------- 2835
57_BC.CN.09.09YNLX19sg       --------------------A--G--A--C--------------C--------A---------------C-C--------------G--------------------T---G----------T-----------AT-----T-------------G--A----------- 2668
58_01B.MY.09.09MYPR37        --------A-----------A-----A--------------------TG----------G-----G--------------------G---G--------------------G-----------AG---------RT----RY-------------G--A-----G----- 2829
59_01B.CN.09.09LNA423        --T-----A--A-----A--A---C-A---------------A-----TG---------------A---GC-G----G-G---C--G---GT---------------CA--G--A----G--T--G--------AT----GT-------------G------------A- 2668
60_BC.IT.11.BAV499           --T-----------------A-----A--C--------------------------A--------A---C-G--------------G--------G-----------CA--G-------G--T-----------AT-----T-------G-----G--A----------- 3441
61_BC.CN.10.JL100010         --T-----------------A--G--A--C-----------A-----------A---------------C-T--------------G--GG--------T-------CA--G-G---------T-----------AT-----T-------------G--A-------GG-- 2905
62_BC.CN.10.YNFL13           --T-----------------A-----A--C-----------------------A-----------A---C-T--------------G--GG----------------CA--G----------TAC---------AT-----T---AG--------G--A-------G--- 2813
63_02A1.RU.10.10RU6637       --T--------A--------A---C-A--------------------T--------------G--------G-----------C--G---G----------G-----TA--G-GA----G--T--G--------AT-----T-------G-------------------- 3025
64_BC.CN.09.YNFL31           --T---------C-------A-----A--C-----------------------A-----------A---C-T--------------G--GG-------T--------CA--G----------T-----------AT-----T----------T--G--A--T-------- 2838
65_cpx.CN.10.YNFL01          --T-----------------A---C-A--C-----------------------A-----G---------C-T---------A----G--GT----------------CA--G----------T-----G-----AT-----T-------------G--A----------- 2855
67_01B.CN.11.ANHUI_HF115     --T--------AC-------A---C-A--C------------A-----T----------------A---------------------G------------------------G-------------------------------------------G-----C-------- 2827
68_01B.CN.11.ANHUI_WH73      --T--------AC-------A---C-A---------------A-----T-------C--------A---------------------G------------------------G-------------------------------------------G-----C-------- 2818
70_BF1.BR.10.10BR_PE004      --T--------A-----A--A---C-A--------------------T-----------------A--------------T---------G----G-----------CA--G-GA----G--TAC----------T-----T-A-----G--T----------------- 2940
71_BF1.BR.10.10BR_PE008      --T--------------A--A-----A--------------------T----------------A---------------------G---G----G-----------CA--G-G-----G--T-C----------T-----T-A-----G--T--G-------------- 3148
72_BF1.BR.10.10BR_MG002      --T--------------A--A---C-A---------------A-----TGG--------------A-----------G---------G--------G-----------CA-------------T-C----------T-----T-------G--T--G-------------- 2992
O.BE.87.ANT70                --T-----A--AC--A-A--A--GC-A---------------A--C--T-A---A----G---G-AG-----G-----GT-AA-------C-----GT--T-----------G----TT---GTAG---G--------G---T-A-AG-----------AGGT------C- 3538
O.CM.98.98CMA104             --T-----A--A---A-A--A--G--A---------------A-----T-A---A----G-----AG------------GT-AA-------C-----GT--T-------G---G----TT---GTA----G-------------T-A-A---------G--AGG--------- 2970
O.CM.98.98CMABB141           --T-----A--AC--A-A--A--G--A---------------A--C--T-A---A----G-----AAG----G-----GT-AA-------C-----GT--T-G---------G----TT---GTA----G------T-G---T-A-A---G--------AG--G------- 2969
O.CM.98.98CMABB212           --T-----A--AC-------A-----A--C------------A-----T-A---A----G-----AAG----G--C--GT-AA-C-----------GT----------G---G-GG-CT---GTA----G-----A------T-A-AG-----------AGA--------- 2974
O.CM.98.98CMU5337            --T-----A---C----A--A-----A---------------A-----T-----A--C-G-----AG--------------AA-------C-----GT-----G-----R---G----TT---GTA----G-----A------T-A-AG--G---------GA--------- 2971
O.CM.99.99CMU4122            --T-----A--AC----A--A--G--A---------------A--C--T-----A----G-----AG-----G-----GT-AA-------C-----GT--T-G------A--G----TT---GTA-----------G---T-A-AG--G--------AGG--------- 2969
O.FR.92.VAU                  --T-----A--A--G--A--A---C-A--------C--------C--T-AT--A----G-----AAG----G-----GT-AA----------T--GT--T-G---------GG----T---GCA----G-----A------T-A-A---G-----G--AAG--G------ 3050
O.GA.11.11Gab6352            --T--------A---A-A--A-----A---------------C--T-A---A--C-G-----AAG-----------T-AA-------C-----GT--T-G---------G----T----GTA-----------T-----T-A-AG--G--------AGG--------- 2695
O.SN.99.99SE_MP1299          --------A--AC--A-A--A--G--A---------------A--C--T-A---A----G-----AAG-G--G-----GT-AA-------CG----GT--T-G------A--G----TT---GTA-----------------T-A-AG--G--T-----AAGT-------- 3537
O.US.10.LTNP                 -----------AC-CA-A--A--G--A-----------------C--T-A---A----G-----AAG--C-T-----GT-AA-------C------T--T--------A--G----CT---GT-----G-------G---T-A-AG------------AAT-------- 3459
N.CM.02.DJO0131              --T-----A--A-----A--A--G--A-----------------C--T-----A--C-------AA--GC-T----G-T-AA-C--------------TT-G---------G-CAACT-T---A----------AT-----T-A-----------G-----AT---G--- 2990
N.CM.02.SJGddd               --T-----A--A-----A--A-----A-----------------C--T-----A--C-----T-AA--GC-G------T-AA-C------G-T-----TT-G---------G-CAACT-T---A----------AT-------A----------AG-----AT------- 2974
N.CM.04.04CM_1015_04         --T-----A--A-----A--A--G--A-----------------C--T-----A--C--------A--GC-T------T-AA-C-----------G--TT-G---------G-TA-CT-T---------------AT-------A--------T--G-----AT------- 2989
N.CM.06.U14296               --T-----R--A-----A--A--G--A-----------------A--T-----A--C-------AA--GC-T----G-T-AA-C--------------TT-G---------G-CA-CT-T---A-----------T-------A--------T--G-----AT------- 2987
N.CM.06.U14842               --T-----A--A--G--A--A--G--A-----------------C--T-----A----------AA--GC-T----G-T-AA-C--------------TT-G---------G-CAACT-T---C----------AT-------A-----------G-----AT------- 2991
N.FR.11.N1_FR_2011           --T-----A--A-----A--A--G--A-----------------C--T-----A--C-------AA--TC-T----G-T-AA-C-----G--------TT-G---------G-TAACT-T---A----------AT-------A-----------G-----AT------- 2885
P.CM.06.U14788               -----T------C----A--A--G--A--------T-----A-----T-----A----GG----AAG-G-----Y-----AA----------T--RT-CT-----------G-TG-YT-T--TAG----------AT-------A-AG-----T-A---A--T-------- 2955
P.FR.09.RBF168               -----C------C----A--A--G--A--------------------T-----A--A-GG----AAG----R--C-----AA----------T---T-CT-----------G-TG-T--T--TAG----------AT-------A-AG-----T-A---A--T-------- 3519
CPZ.CD.06.BF1167             --T--T-----A--------T-----A---------------A-----T--------C-G-AC---AG--C-C--C---T-AA----G---GT-----G---------T---G---AGT-G--T-C---------AT-----T-A-AG-----T-A-C-A--T-------- 3586
CPZ.CM.05.SIVcpzMT145        --T--T-----------A--C-----A---------------A--A--T-----A------A-T-A---GC-T------T-AA----G---G-G--G-GT--G---------G--G-T--C--CCC------T--A-----------------TCA------A-------- 3026
CPZ.GA.88.GAB1               --T-----A--AC-GA-A--A--G--A--------------------T-----A--A---A---A------------G--GA-A--------A--GAA---G-----T---G-T--T--C---CC---------AT-----T-A---------T--G-----AG---GCAC 3545
CPZ.TZ.06.SIVcpzTAN13        C----C--A---C-G--A--C--G--A---------------A--C--TG----A--A--GAC--AAG-------C---T-AA-A------GTA----G------T--CCCTG-GGA-------AG---------AT---GT-A-A------T-A---A----------- 3125
CPZ.US.85.US_Marilyn         --T--T-----A-----A------C----C--------------A--T-----A--A---A---A---G-----C---T-AA-A-A----G-T-----T-----T-----TG--AATT-T---C-T--G-----AA-G--GT-A-A-------------A--------G-- 3544
GOR.CM.07.SIVgor2139_287     --T--C----GA--G--A--T---C-A--------T-----A--A--T--------A-G-----AAG--C-C--C---T-AA----G--T--T---T----T---------G-GA-C--T-GTA----------AT-------A-A---G-----G--A--AT------- 3021
GOR.CM.07.SIVgorCP2135con    --T--C-----AC-G--A--------A--------C-----A--A--T-----A--A-G-----AAG--C-C--C---T-AA----G--T--T---T----T---------G-GA-C--T-GTA----G-----AT-------A-A---G-----G--A--AT------- 3020
```

```
B.FR.83.HXB2                  ACCAGTACATGGAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCATTT.AAAAATCTGAAAACAGGAAAATATGCAAGAATGAGGGGTGCCCACACTAATGATGTAAAACAATTA 3653
Pol                           __P__V__H__G__V__Y__Y__D__P__S__K__D__L__I__A__E__I__Q__K__Q__G__Q__G__Q__W__T__Y__Q__I__Y__Q__E__P__F_#K__N__L__K__T__G__K__Y__A__R__M__R__G__A__H__T__N__D__V__K__Q__L_
A1.AU.03.PS1044_Day0          ------G--------A--C-----------------A--------------------A--A-------A------------C---------------------.--------A----------------------A----TC---TY---------------G------- 2861
A1.CH.03.HIV_CH_BID_V3538     C--T--G-----G--A--------T-----------A---G----------------A--A----C--AT---------------------------------.------T-A-------------------A--A-G--TC---T----------------GG------ 2858
A1.CY.08.CY236                G-----G--------A--C-----------------A---G-------G--------A--A-------AA---------------------------------.--------A--G-------------------AA---TC---T------------------------ 2864
A1.ES.05.X1608_8              C--T--G--------C--C--------------------------------------A--A-------A------------C-------T---G---------.--------A-------------------A--A-G--TC---T----------------G------- 3101
A1.KE.11.DEMA1IKE001          C--T--G------ACA--------------------A--------------------A--A-------A------------C---G-----------------.-----------------G-------------AA-A-TC---T------------------------ 2992
A1.RU.11.11RU6950             ---T--G-----T--A--------------------T---G----------------A--A--A----A------------------------G---------.--G-----A-------------------A--A-G--TC---T-----C------------------ 3177
A1.RW.11.DEMA111RW002         G--T--G-----G--A---------------------C--G-G-----------A--A--A-------A----------------------------------.-----------------G-------------A----TC---T------------------------ 3004
A1.SN.01.DDI579               C--T--G-----G-CA-----------------------------------------A--A------AA------------------C---------------.--G----------------------------A----TC-T-T----------------G------- 2851
A1.UG.11.DEMA110UG001         C--T--G-----G--A-----------------------------------------A--A-------AT---------------------------------.-------------------------------AA--ATC---T----------------G------- 3008
A1.ZA.04.04ZASK162B1          C--T--G-----G--A------------G--------C-------------------A--A-------A----------------------------------.--------A-------------------A--A----TC---T--------------------G--- 3094
A2.CM.01.01CM_1445MV          C--T--------G--A--C----------------------C------------------------G-A---------------------------A------.--------A--------G----------A--A----ACCA-T------------------------ 2852
A2.CY.94.94CY017_41           C--T--------G--A--C--------------------------------------A--A-------A------------------------G--A--C---.--G--------------G----------A--G----TCCA--------------A-------G--- 3010

B.BR.10.10BR_MG029            ---------------A--C--------------------------------------------------T---------------------------------.-------------------------------C------G--------A------------------ 3115
B.CA.07.502_1191_03           ----------------------------A--------------------T-------A---------------------------------------------.--------------------G-----------------------------------------G--- 3080
B.CH.08.M2_0803101_NFLG8      -----------------T-------------------C-------------------------------T---------------------------------.--------------------G-----------------CA----T--------C--------G--- 3122
B.CN.10.DEMB10CN002           ----------------------------------------------------------------G-------------------------------A----A-.--------------------GC-------------------T------------------------ 3025
B.ES.10.DEMB10ES002           --------------------------------------------------------------------------------------------------A----.--------------------G-------A---------------T-------------G------- 3015
B.FR.11.DEMB11FR001           G--------------------------C----------------------------------A---T------------------------------------.-----------------------------G------------------------------------ 2899
B.GB.05.MM45d213_GN1          ------------------------------------------------------------A-A-TGT----------------------------------A-.------T-----------------------------A----------------------------- 3054
B.HK.06.HK003                 ------G-----------C------------------------------T----------A-----G-----G------------------------------.------------------------------------------------------A----G------ 3023
B.HT.05.05HT_129389           ------------G-----C--------------------------G--------------------T------------------G-----------------.--------A-----------------------------------------------G--------- 2854
B.JP.12.DEMB12JP001           ----------------------------------------G----------------A-------T---------------------------G---------.--------------------G-------------A-------------------A-------G--- 3040
B.KR.07.HP_18_07JHS10_3909    ---G-----------A---------------------------------------------------------------------------------------.----------------------------A------------------------------G------ 3317
B.PE.07.502_0525_wg5          ------------------------T--C-------------------------------------T----------T----------A--------A------.--------------------G-------A---------------------------G-----G--- 3118
B.RU.11.11RU21n               ------------------------T---------------G--------C-------------A---------------------------------------.--------------------------------------------------------------G--- 3224
B.TH.08.MERLBDTRC10           -------------------------------------------------------------------------------------------------------.-----C--------------G--C--------------CA----T--------------------- 2886
B.US.11.ES38                  ------------------------------------------------------------------------G------------------------------.----------------------------A-------A--------------------------C-- 3588

C.AR.01.ARG4006               ---------------A--------T-----G--------------T-----------A----AA--G-----------------------------------C.----------------------------A-------AC--------------------G---G--- 2837
C.BR.07.DEMC07BR003           ---------------A--------------G-----T--------T-----------A------G-G--T--G-----------------------A-----C.--G-------------------------A--C----AC--------------------G---G--- 2985
C.BW.00.00BW5031_1            ---------------A--------T-----------T--------------------A------G-T-A---------------------------A-----C.--------------------G-------A---A---AC--------------------G---G--- 3024
C.CN.10.YNFL19                -A------------CA-----------------------G-----T--------A--A--A----TG-A---------------------C-----A-----C.-----------------G--G-------A-------AC------T-----------------G--- 3006
C.CY.09.CY260                 G--------------A--------------------------------G-----A--A--A---G-G-A---G-----------G-----C------------.--------A-----------G-------A-G-----AC------------------------G--- 2876
C.ES.08.X2363_2               ---------------A-----------G-----------G-----T-----------A------G-G-----------------------C-----A------.--------------------G-------A--G--A--C------------------------G--- 3084
C.ET.02.02ET_288              G--------------A--------T--------------------T-----------A------G-T-A---------------------C-----A------.-----------------G--G-------A--G----AC-----------------------TG--- 2855
C.IN.09.T125_2139             --------------CA-----------------------G-----T-----------A--------G-----------------------C-----A-----C.-----------G-----G--G-------A-------AC---------C--------------G--- 3629
C.KE.00.KER2010               ---------------A-----------------------G-----T-----------A--------T-A---------------------C-----A-----C.--------A-----------G-------A-------AC------------------------G--- 2834
C.MW.09.703010256_CH256.w96   ---------------A-----------------------------------------A-----AG-T-A---G-----------------C-----A-----C.--------A-------------------A-CA---AAC---------A------------------ 3633
C.TZ.08.707010457_CH457.w8    --------------CA-----------------------------T-----------A--------T-AT--------------------C-----A--T--C.-----------G----------------A-------TCC-----T--------------G--G--- 3627
C.YE.02.02YE511               ---------------A-----------------------------------------A------T-T-A---------------------C-----A-----C.------T----G----------------A---A---AC------------------------G--- 2828
C.ZA.08.705010534_CH534.w12   ------------G--A--C--------T-----------------T-----------A-------TT-A---------------------C-----A------.-----------------G--G-------A------AAC---------------------G--G--- 3643
C.ZA.10.DEMC10ZA001           ---------------A-----------T-----------G-----T-----------A-----AA-T-A---------------G-----C-----A-----C.--G-----A--------G--G-------A--G----AC------------------------G--- 2993
C.ZM.11.DEMC11ZM006           ---G-----------A--------------------------------------------------T-AT--------------------C-----A-----C.-----------------G----------A-------AC--------------------G---G--- 3002

D.CM.10.DEMD10CM009           ---------------A-----------------------------------------A--A------A----------C-GG-----C--------A------.--------A--G--------G-------A------A----------------------G------- 3028
D.CY.06.CY163                 ------T--------A------------------------------A----------A--A------A----------T-GG--------------A------.--------------------G-------A---------------T--------------------- 2861
D.KE.11.DEMD11KE003           --------------------------A------------------------------A--A-------A---------------G-----------A-A--A-.-----------G----------------A----------A-------------------------- 3003
D.KR.04.04KBH8                ------------------------T------------------------T-------A--A--------------------------A---------------.------T----G--------G-------A----------A-------------------------- 3612
D.SN.90.SE365                 ---------------------------------------------------------A--A-------A------------------------G---------.--------------------G-------A--G----A----T------------------------ 3670
D.TZ.01.A280                  -A------------CA-----------------------------------------A--A-A---------------------------------A------.--------------------G-------A----------A-------------------------- 2854
D.UG.10.DEMD10UG004           ------------------------T---G-------------------------A--A----------A------------------C--------A-A--A-.--------A-----------G----------A-------A----------C--------------- 2991
D.UG.11.DEMD11UG003           ---------------A---------------------------------G-G-----A--A----C--AT--------C-----------------A-AG-A-.--------A-----------G-------A-T--------A----T--------------G------ 3008
D.YE.02.02YE516               -------------------------------------------------G-------A--A-----------------G-GG-----------G--A------.--------------------G--------------------T------------------------ 2864
D.ZA.90.R1                    ---------------------------------------------------------A--A---G-T--------------C--------------A------.--------------------G----------CA--------------------------------- 2999

F1.AO.06.AO_06_ANG32          ------------G--A-----------------------------------------A--A-----C-AG-----------------C---------------.--------------------G-------A--G----TC---------------------G------ 2864
F1.AR.02.ARE933               ------------G--A-----------------------------------------A--A--------G------------------------------C--.--------A-----------G---------------TC---------------------------- 2947
F1.BR.10.10BR_RJ015           ------G-----G-CA-----------G---------------------T-------A--A-----G--A---------------------------------.--------------------G-------A-------TC---------------------G---C-- 3126
F1.CY.08.CY222                ------------G--A------------------------G-------------A--A--A--------G--------G-----------C----------G-.--------A-----------G-------A------ATC------------------G--------- 2846
F1.ES.02.ES_X845_4            ------------G-CA----------------------------T------------A--A--------G---------------------------------.--------------------G-------A-----A-TC---------------------------- 3105
F1.RO.96.BCI_R07              ------------G--A--------------------T--------------------A--A--------G--------------G------------A-----.--------A-----------G-------A-------TC---------------------------- 3677
F1.RU.08.D88_845              ------------G--A-----------------------------------------A--A--------G--G-----------------------A----A-.--------------------R-------A-------TC--------------------G------- 3112
F2.CM.02.02CM_0016BBY         ------------G--A--------------------T--------------------A--A-----T-AT--------------------------A---CA-.-G------------------G-------A--G----TC---------------------------- 2843
F2.CM.10.DEMF210CM001         ------------G--A--------------------A--------------------A--A-----G-A---------------------------AA--CA-.-----C--------------G-------AG-G---ATC---------------------------- 2917
F2.CM.10.DEMF210CM007         ------------G--A--------------------T-----------------------A-----T-AT------------------------------CA-.--------A-----------G-------A--GA---TC---------------------------- 3000

G.BE.96.DRCBL                 ---T-----------C---------------------C-----------G-------A--A----T-------------------G---------------A-.----------------------------A-GG-G--TC---------------------G------ 3610
G.CM.10.DEMG10CM008           C--T--G-----G--C--C-----------G-----------------GG-------A--A--AGT--AA--G----------------------------AC.----------------------------A--G-G--TC------------------------G--- 3034
G.CN.08.GX_2084_08            ---T-----------C--C--------------------R---------G-------A--A----CC-A--------------------------------AC.--------------------G-------A--G-G--TC---------------------------- 2893
G.CU.99.Cu74                  ---T-----------C--------------------AA-----------G-------A--A---A-C-A---------G----------------------AC.--G-------------------------A--G-G--TC---------------------------- 3257
G.ES.09.X2634_2               ---T-----------C--------------------A------------G-G-----A--A----T--A---G--------C--G--C-------------AC.----------------------------A--G-G--TC---------------------------- 3136
G.GH.03.03GH175G              G--T-----------C--------------------A------------G-G-----A--A----T--A--------------------------------AC.--G-----A------------------CA--G-G--TC-A----------------T--G--G--- 3693
G.KE.09.DEMG09KE001           G--T--------G--C---------------------------------G-------A--A--A-T--AT--------------------C----------AC.----------------------------A--G-G--TC---------------------------- 3028
G.NG.09.09NG_SC62             ---T-----------C--C------T----------A------------G-------A--A-A--CT--T--------------G----------------AC.--------C-------A---G-------A--G-G--TC---------------------GG-G--- 2845
G.ZA.01.TV546                 ---T-----------C--C-----------------A------------G-------A--A----T--AA-------------------------------AC.----------------------------A--A-G--TC---------------------------- 2855

H.BE.93.VI991                 --------------CA--------T-----------A-----------------------A-----C-A---G-----------------------------C.-----------G----------------A-------TC---------------------G------ 3042
H.CF.90.056                   ----A----------A--------T----------------------------G------A--------------------------------G---------.-----------G----------------A------AAC----------------A----------- 3000
H.GB.00.00GBAC4001            ------------G--A--------T-------------------------------------------AA--------------------------------C.--G--C--A--G----------------A------AAC----------------A----G------ 3171

J.CM.04.04CMU11421            --------------YR--C---------G-------A------------G-------A--A---G---A---------M------------------------.-----C-----R-----R----------A--G---------------------------------- 3159
J.SE.93.SE9280_7887           ------------G--A---------T--G-------A------------G-G-----A--A----TG-A------------------------G---------.-----C-----------G----------A--G----A----------------------------- 2967
J.SE.94.SE9173_7022           ------------G--A------------G---G---A------------G-G-----A--A--A-TG-A------------------------G---------.-----C-----------G----------A--G----A----------------------G------ 2968

K.CD.97.97ZR_EQTB11           ------G-----G--A--C-----------------------------------------A-----C--------------------A-------------A-.------T----------G--G-----------A---TC---------------------------- 2849
K.CM.96.96CM_MP535            G-----G-----G--A-----------------------------------------A--A---A-T-A------------------A------------CA-.-----------------G--G--------------ATC---------------------G------ 2849

U.CA.01.TV749                 ---T--------G--A--C--------------------------------------A--A---A-T-A---G------------------------------.----------------------------A--G--A-TC---------------------------- 3670
U.CA.99.TV721                 ---T--------G--A--C--------------------------------------A--A---A-T-A---G------------------------------.-----------------G----------A--GT-A-TC-A-------------------------- 3669
U.CD.83.83CD003_Z3            ---------------A------------------------G-------------------A-------A-------T-------------C------------.-----C----------------------A--G----TC-----------------------GC--- 3148
U.CD.90.90CD121E12            ---------------A--------------------T--------------------A--A-------A-------T-------------C------------.-----C-A--------------------A--G----TC-----------------------GC--- 2852
U.CY.05.CY090                 ---------------A--------------------------------------------A-----T-----G--------------------G---------.--------------------G-------A---A---TC-------C--------A----------- 2852
U.CY.08.CY223                 ---T--G-----G-CA--C---------G-----------G----------------A--A-------A-------------------------------CA-.--------------------G-------A--G---ATC----------------A---G---G--- 2870
U.ES.10.DEURF10DZ001          GG----G-----G-CA--------------------------------------A---------G----A--G--------------A-----G------CA-.-----C-----------G--------------A-A--C---------------------------- 2929
U.GR.99.99GR303               ------G-----G--A--------------------------------------------A-----C--------------------A-------------A-.-----------------G-----C------GA----TC----------------A----------- 2927
U.NL.95.U_NL_95_H10986_D1     ------------G--A--------------------T--------------------A--A-------A------------------A-------------A-.--------A--------G--G----------G--A-TC---------------------------- 3538

01_AE.AF.07.569M              C--T--G--------A------------------------G--------G-------A--A-----G-A---G------------------------------.--------A-----------------C----AA--ATC---T------------------------ 2859
01_AE.CF.90.90CF11697         C--T--G--------A--------------------A---G----------------A--A-------A----------------------------------.--------A----------------------AA---TC---T----------------G------- 3602
01_AE.CN.10.YNFL03            C--T--G-----G-CA---------------------------------G-------A--A-----G-A-----------------------------A----.------T-A-----------------T----AA---TC---T-----C----------G------- 3031
01_AE.HK.04.HK001             C--T--G--------A------------------------G--------G----------A--A----A----------------------------------.--------A----------------------AA---TC---T----------------G------- 3015
01_AE.IR.10.10IR.THR48F       C--T--G--------C------------------------G--------G-------A--A-----G-A----------------------------------.--------A--------G--------C----AA---TC---T----------------G------- 2843
01_AE.JP.x.JRC77AE            C--T--G--------A------------------------G--------G-------A--A-----G-A----------------------------------.--------A--------G--------C----AA---TC---T----------------G------- 3650
01_AE.TH.09.AA090a_WG11       C--T--G--------A--------------------T---T----------------A--A-----G-A----------------------------------.--------------------------C----AA---TC---T--T-------------G------- 3076
01_AE.TH.90.CM240             C--T--G--------A------------------------G--------G-------A--A-----G-A----------------------------------.--------A-----------------C----GAG--TC---T----------------G------- 3227
01_AE.US.05.306163_FL         C--T--G--------A--------T------------------------G-------A--A--C--G-A----------------------------------.--------A-----------------C----AA---TC---T------------------------ 2844
01_AE.VN.98.98VNND15          C--T--G--------A------------------------G--------G-------A--A-----G-A----------------------------------.--------A-----------------C----AA---TC---T----------------G------- 2908
```

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

```
B.FR.83.HXB2                ACAGAGGCAGTGCAAAAAATAACCACAGAAAGCATAGTAATATGGGGAAAGACTCCTAAATTTAAACTGCCCATACAAAAGGAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCTGAGTGGGAGTTTGTTAATACCCCTCCCTTAGTGAAATT 3823
Pol                         _T__E__A__V__Q__K__I__T__T__E__S__I__V__I__W__G__K__T__P__K__F__K__L__P__I__Q__K__E__T__W__E__T__W__W__T__E__Y__W__Q__A__T__W__I__P__E__W__E__F__V__N__T__P__P__L__V__K__L
A1.AU.03.PS1044_Day0        G----A-TG---------G-GGT--TG---------------------------------------T-A-----------A-----------C----------TG--C--------G--T--------C-----A-----------C-----------TC----A----- 3031
A1.CH.03.HIV_CH_BID_V3538   ----CA-TG---------G--G-------------------------------C--------C-G---A----------GA------------G---------TG-----------G--T--------------A-----------C-----------TC----A----- 3028
A1.CY.08.CY236              G----A-T---------G---G---TG-------------------------------------G---A-----------A------------CT--------TG--C--------G--T--------C-----A-----------C-----------TC----A----- 3034
A1.ES.05.X1608_8            G--A-T-TG--------GG--GT--TG-------------------------------------G---A----------------------------------TG--C-----------T--------------A-----------C-----------TC----A----- 3271
A1.KE.11.DEMA11KE001        -----A-TG--------GG-GGT--TG-------------------------T---------------A-----------A--G--------T----------TG--T--------G--T--------------A-----A-----C-----------TC----A----- 3162
A1.RU.11.11RU6950           ----CA-TG--------GG-GG------------------------------------------G---A-----------A------------G---------TG-----C-----G--T--------------A-----------C-----------TC----A---C- 3347
A1.RW.11.DEMA111RW002       -----A-TG--A------G------TG-----------------------------------------A-----------A-----------G----------TG--C--------G--T--------------A-----------C-----------TC----A---C- 3174
A1.SN.01.DDI579             GT---A-TG---------G-GGT------------------------G-----C----------G---A-----------A----------------------TG-----------G--T----AT--C-----A-----------C-----------TC----A----- 3021
A1.UG.11.DEMA110UG001       -----A-TG---------G-GGT---------------G---------------------------T-A-----------A-----------------------G--C--------G--T--------------------------C------------C----A----- 3178
A1.ZA.04.04ZASK162B1        -------T---------GG-GGTT-----------------------------C------------T-A-----------A----------------------TG----G------G--T--T-----------A-----------C-----------TC-G--A---C- 3264
A2.CM.01.01CM_1445MV        -----A-----A---------GT--TG--------------------------------G----G-T-A-----------A-----------G-----------------------G--T--------------------------C------------C----A---C- 3022
A2.CY.94.94CY017_41         -----A-----A-------------TG-----------G---------------------------T-A-----------------------G---------G-G-----------G--T--------------------------C-----------TC----A---C- 3180

B.BR.10.10BR_MG029          G--------------------GT-----------------------------------------G---A-----------A-----------------------------------------------------------------C-----------------A----- 3285
B.CA.07.502_1191_03         -----------------G---G----------------------------------------------A-----------------------------------G-----------------A-----------------------C----------------------- 3250
B.CH.08.M2_0803101_NFLG8    ---------------------G----------------------------------C-------G---A----------GA-----------C-----------G-----------------------------------A-----C------------C---------- 3292
B.CN.10.DEMB10CN002         ---------------------G-----C----------------------A-------------G-T-A------------------------G----------------------------------------------------C-----------------A----- 3195
B.ES.10.DEMB10ES002         ---------------------G---TG-----T-------------A---A-T---------C-G-T-A-----------A--G-------------------T---------------T--------------------------C----------------------- 3185
B.FR.11.DEMB11FR001         ---------------------G-TCA------T-----------------A-------------G---A------------------------G-G--------------------------------------------------C----------------------- 3069
B.GB.05.MM45d213_GN1        -------T-------------G-------G-----------------T--------------------A-----------A---------C--G-------------A--------------------------------------C-----------T----------- 3224
B.HK.06.HK003               --G------------------G------------------------------------------G---A-----------A----------R------------------------------------------------------C-----T-----------A--G-- 3193
B.HT.05.05HT_129389         ---------------------G--------G---------------------TA----------G---A-----------A-----------T-T---------------------G-----------------------------C-----------------A----- 3024
B.JP.12.DEMB12JP001         --------C-----G------G-------------------------G-GA-------------G---------------A--G---------G---------TG-----------------------------------------C-----------------A----- 3210
B.KR.07.HP_18_07JHS10_3909  ---------------------G---T------------------------------------------A-----------A------------G----------------------------T--A-----------------G--C----------------------- 3487
B.PE.07.502_0525_wg5        -----------A---------G----------------G---------------------------T-A--------G--A------------G----------G-----------------------------A-----------C---------------------C- 3288
B.RU.11.11RU21n             ---------------------G----T-----------G-----------------------------A--------G--A-----------------------------------------------------------A-----C-----------A-----A----- 3394
B.TH.08.MERLBDTRC10         --C--------A---------G-------------------------------C--C-----------A-----------A---G--------G-------------T--------------------C-----A-----------C-----------T----------- 3056
B.US.11.ES38                ----------------------------------------------------T-----------G---A--------G--A-----------------------------------------------------A-----------C--------------------G-- 3758

C.AR.01.ARG4006             ---------------------G--CAG---------A----------------C----------G-T-A-----C--G--A-----------GT-------------C---------------------------------------------------C----A--G-- 3007
C.BR.07.DEMC07BR003         ---------------------G--CAG--------------------G--A-------------G-T-A-----C--G--A------------T-C-----------C--------------------------------------------A-----------A--G-- 3155
C.BW.00.00BW5031_1          ---------------------G--CTG-------------------------T-----------G---A-----T-----------------G--------------C---------------------------------------------------C----A--GC- 3194
C.CN.10.YNFL19              G-------T-----G----------TG-------------------------------------G-T-A-----C-----------------G---------G----C--------------------------------A-----------------------A----- 3176
C.CY.09.CY260               ---------------------G---AG---T---------------------------------GG--A-----T--G--A-----------G--------------C--------------------------A------------------------C----A----- 3046
C.ES.08.X2363_2             ---------------------T---TG-----------G-------------------------G---A-----C-----------------C--------------C--------------------------------A------------------C----A----- 3254
C.ET.02.02ET_288            -----------------G---G--CT--------------------------T-----------G-T-A--T--C--G--A--------------------------C--------G-----------------A-----A------------------C----A----- 3025
C.IN.09.T125_2139           --------T-----G------G--CT----------------------------------------T-A-----C-----A-----------G--------------C--------------------------------A---------------------------C- 3799
C.KE.00.KER2010             ---------------------G---------------------------G--------------G-T----A--C-----A-----------G--------------C--------------------------------------------------------A---C- 3004
C.MW.09.703010256_CH256.w96 --T-----------G-----TG---TG--------------------G----T------G----G---A-----C-----------------G--------------C--C-----------------------------A---A--------------C----A----- 3803
C.TZ.08.707010457_CH457.w8  -----A--------G------G---TG--------------------G-----C----------G---A-----C-----A-----------G--------------C---------------------------------------------------C----A----- 3797
C.YE.02.02YE511             ---------------------G---TG-------------------------T-----------G---A-----C-----------------G--------------C--------------------C-----A-----------G------------C----A----- 2998
C.ZA.08.705010534_CH534.w12 G--------------------G-T-TG-------------------------------------G---A-----C----G--------------T------------C--------------------------------------------------TC----A----- 3813
C.ZA.10.DEMC10ZA001         ------------------G--G--CAG---T---------------------------------G---A-----T-----A-----------G--------------C---------------------------------------------------C----A----- 3163
C.ZM.11.DEMC11ZM006         -----------------G---G---TG---T-------G--------------C----------G-T-A-----C-----A-----------G--------------C---------------------------------------------------C----A----- 3172

D.CM.10.DEMD10CM009         ---------------------G---TG-----------C-------------------------G---A------------------------G---------GG--A--------------T-----------------------C-----------T-----A----- 3198
D.CY.06.CY163               ---------------------G---T------------G-------------------------G---A------------------------G---------TG-----------------T-----------------------C-----------T-----A----- 3031
D.KE.11.DEMD11KE003         -----------A-----------ACAC---T-T-----G-------------------------G---A--------G--------------C-----------------------------------------------------C-----------T-----T----- 3173
D.KR.04.04KBH8              ---------------------G--CA-----------------------G--------------G-T-A------------------------G-------------A--------G-----------------------------C--C--------T-----A----- 3782
D.SN.90.SE365               ---A----------G---------CA--------------------------------------G---A------------------------G---------T------------------------------------------C-----------T-----A----- 3840
D.TZ.01.A280                ---------------------G--CT----G-T-----G-------------T-----------G---------------------------------------------------------------------------------C-----------T-----T----- 3024
D.UG.10.DEMD10UG004         -----A---------------G--CA----T-T-----G-----------A-------------G---A--------------G-----------------------C--------G-----------------------------C-----------T-----T--G-- 3161
D.UG.11.DEMD11UG003         G---C---------G------T--CA----T-T-----G--------------------G----G---A-------------------------T------------T--------G-----------------------------C-----------T-----T----- 3178
D.YE.02.02YE516             -------T-------------G--CA------------G-------------------------G---A------------------------G---------TG-----------------T-----------------------C-----------T-----A----- 3034
D.ZA.90.R1                  G----------------TT--T----------------G-------------------------G---------------------------------------------------G-----------------------A-----C-----------T-----A----- 3169

F1.AO.06.AO_06_ANG32        -----A-----------GG--G-T-TG--------------------------------G------T-A-----------------------G-----------------------------------------A-----------C--------C--TC----A---C- 3034
F1.AR.02.ARE933             -----A-----------G---G-TCA---------------------------C-----G--------A-------T---A--G--G-----T----------TCAGTAT------------------------A-----------C--------C--TC----A---C- 3117
F1.BR.10.10BR_RJ015         -----A-----------GG--T-TCT---------------------C----T-----------G---A------TT---A--G--------T--------------A--C--------T--------------------------A--------C--TC----A---C- 3296
F1.CY.08.CY222              ---------------------GTTCAG--------------------G--A-T------G--------A------------A----------GG-------A-T---T--------------------------A-----------C--------C--TC----A---C- 3016
F1.ES.02.ES_X845_4          -----------------G---G-TCTG-----------------------A--------G--------A-----------------------GG----------------------------------------A-----------C--------C--TC----A---C- 3275
F1.RO.96.BCI_R07            -----------A-----G---G-T--G---T----------------------------G--------A-----------------------GG----------------------------T-----------A----------CC--C-----C--TC----A---C- 3847
F1.RU.08.D88_845            -----A-----------G---G-TGT---------------------------C-----G------T-A--------------G--------G---------G-------C-----------------------------------C--------C--TC----A---C- 3282
F2.CM.02.02CM_0016BBY       -----A-TG--A---------G-------G---------------------GTC----------G---A-----------A-----------------------------------------------------A-----------C-----------TC----A----- 3013
F2.CM.10.DEMF210CM001       -------T----------G--G------------------------------T-----------G---A---------------------------------G----A--------------------------A-----A-----C-----------TC----A----- 3087
F2.CM.10.DEMF210CM007       -----T-T-------------G-------G-------------------G-GT-----------G---A-----------------------------------------------G-----------C-----A-----------C-----------TC----A----- 3170

G.BE.96.DRCBL               -----A-T-------------G-------GG-----A---------------T-------------T-A--T---A----A------------GTG--------------------G-----T-----------------------C-----------T-----C----- 3780
G.CM.10.DEMG10CM008         -----A-T---A-----G---G-------G------A-------------A-T-----------G---A--T---AG---A-----------TG-------------C--------G--T--------------T-----------C-----T-----T---------C- 3204
G.CN.08.GX_2084_08          -----A---------------G-T-----G------A-------------AGT---------------A--T---AGG--A------------GT------------C--------G-----------------T-----A-----C-----------TC----A--G-- 3063
G.CU.99.Cu74                GT---A-T-------------T-------G--------------------A-----------------A--T----G---A------------GT------------C--------G-----------------T-----------C--------A--TC----A----- 3427
G.ES.09.X2634_2             -----A-T---------GG--T-------G------A-------------A-------------GG--A-------G--GA--------------------------A--------G-----------------C-----A-----C-----------TC----A----- 3306
G.GH.03.03GH175G            -----A-TG------------G-------GG-------------------AGT---------------A--T---AG---A------------GTG-----------------------T--------------------------C-----------TC----A----- 3863
G.KE.09.DEMG09KE001         -----A-T-------------G-------G-----------------------------G--------A-------G--GA------------GT------------A--C-----G-----------------T-----------C--C-----C--TC----A----- 3198
G.NG.09.09NG_SC62           -----A-T-------------T---T---G--------G-----------A-T-----------G---A--T---A-G--A-----------CGT------------A--C-----G-----------------T-----------C--------C--TC----A--G-- 3015
G.ZA.01.TV546               -------T-------------G-TGT--------------------------T---------------A--T---AG---A------------GT------------A--------G-----------------T-----------C-----------TC----A--G-- 3025

H.BE.93.VI991               -----A-T---A---------G-T--------------------------A-T-----------G-T-A--T--------A-----------G---------------C-------------A--------------------------C-------ATC---------- 3212
H.CF.90.056                 -----A-----------G---T-T--------------------------A-T-----------G---A--T--------A-----------G--C--------------------------A-----------A--------------C-------ATC----A----- 3170
H.GB.00.00GBAC4001          -----A---------------TGT--------------------------A-T-----------G---A--T--------A-----------------------------------G-----G--------------------------C-------ATC-G--A----- 3341

J.CM.04.04CMU11421          -----A---------------G--TTG-----------G-----------A-------------G---A-----------------------R--------------Y--------G-----------------------------C-----------TC----A----- 3329
J.SE.93.SE9280_7887         G----A-TG------------G--TTG---GC------------------A-------------G---A----------GA-----------G--------------C--C-----G-----------------------------C--------C--TC----A--G-- 3137
J.SE.94.SE9173_7022         T--C-A-TG------------G--TTG---GC------G-----------A-------------G---A-----------------------G--------------C--------G-----------------------------C--------C--TC----A----- 3138

K.CD.97.97ZR_EQTB11         -----A-T----------G--G---TG-----------------------A-------------G-T-A-----------A----------GG--------------A--------G-----------C-----A-----------C------------C----A---C- 3019
K.CM.96.96CM_MP535          -----A---------------G-------G--------------------A-------------G---A-----------A-----------G--------------A--------------------C-----A-----------C-----T------C----A---C- 3019

U.CA.01.TV749               ---A-T-----------G---G---TG---T-------G----------G------C-------G---C-----------A-----------TG---------TG--------------T--------C-----------------C-----------TC----A----- 3840
U.CA.99.TV721               ----CT-----------G---G---TG---T-------G-----------------C-------G---C-----------A-----------TG---------TG-----------G--G--------------------------C-----------TC----A----- 3839
U.CD.83.83CD003_Z3          -----A---------------T-T-TG-------------------------T---------C-----A-----------------------GT-------------T--------------------C-----------------C-----------TC----A--G-- 3318
U.CD.90.90CD121E12          -----A---------------T-T-TG--------------------------------------G--A-----------------------G---------G----C--------------------C-----------------C-----------TC----A--G-- 3022
U.CY.05.CY090               -----A-----A---------T-T-----------------------G--------------------A-----------------------C-T----G----G-----------G--T--------------A-----------C-----T--C--TC----A----- 3022
U.CY.08.CY223               GT---A---------------G--CA---G--------------------AGTC----------G---A------------------------G---------TG-----------G--T--------------A-----------C-----------TC----A----- 3040
U.ES.10.DEURF10DZ001        -----A-----------G---G--CT--------------------------------------GGT-A-----------A-----------G-T------------T--C-----------------C-----A-----A-----C------------C----A---C- 3099
U.GR.99.99GR303             -----A-T---------------------------------------------C----------G---A--A----------G---------G--------------C--------G---C-------------A---------...--------C--TC----A----- 3094
U.NL.95.U_NL_95_H10986_D1   -----A---------------G---TG-----------------------A--------G----G---A--T--------A-----------G--------------A--------------------C-----A-----------C------------C----A--GC- 3708

01_AE.AF.07.569M            -----A-T-------------G----------------G--------------C----------G---A-----------A----------------------TG-----------G--T-----------G--A--------------------G--TC-G--A----- 3029
01_AE.CF.90.90CF11697       G----A-TG--------GG-GG-------------------------------C----------G---A----------GA---------------------G-G-----------G--T--------------A-----------C--------C--TC----A----- 3772
01_AE.CN.10.YNFL03          G----A-TG------------G-T-----------------------------C----------G---A----------GA-----------------------------------G--T--------------A-----------------------------A----- 3201
01_AE.HK.04.HK001           G----A-TG------------GT----------------------------G-C--------------A----------GA----------------------TG-----------G--T--------------A-----------C-----------TC----A----- 3185
01_AE.IR.10.10IR.THR48F     -----A-T-------------G------------------------------------------G---A-----------A----------------------TG-----------G--T--------------A-----------------T--C--TC----A----- 3013
01_AE.JP.x.JRC77AE          -----A-TG--------G---G-------------------------------C----------G---A----------GA-----------------------------------G--T--------------A--------------------C--TC----A----- 3820
01_AE.TH.09.AA090a_WG11     -----A-TG---------G--G---TG--------------------------C----------G---A----------GA----------------------TG-----------G--T--------------A-----------C-----------TC----A----- 3246
01_AE.TH.90.CM240           -----A-TG---------G--G-------------------------------C----------G---A----------GA----------------------TG-----------G--T--------------A-----------------------TC----A----- 3397
01_AE.US.05.306163_FL       -----A-TG------------GT------------------------------C----------G-T-A----------GA--G-------------------TG-----------G--T--------------A-----A-----------------TC----A----- 3014
01_AE.VN.98.98VNNDI5        -----A-TG------------G-------------------------------C----------G---A-----------A----------------------TG-----------G--T--------------A-----------------------TC----A----- 3078
```

```
B.FR.83.HXB2                 ACAGAGGCAGTGCAAAAAATAACCACAGAAAGCATAGTAATATGGGGAAAGACTCCTAAATTTAAACTGCCCATACAAAAGGAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCTGAGTGGGAGTTTGTTAATACCCCTCCCTTAGTGAAATT 3823
Pol                          _T__E__A__V__Q__K__I__T__T__E__S__I__V__I__W__G__K__T__P__K__F__K__L__P__I__Q__K__E__T__W__E__T__W__W__T__E__Y__W__Q__A__T__W__I__P__E__W__E__F__V__N__T__P__P__L__V__K__L
02_AG.CM.10.DE00210CM013     -------T---------GG-GG-T--------------------------C---------G--C---------GA-----------G--------TG----------G--T-------------A----------C----------TC----A---C- 3190
02_AG.CY.09.CY256            -------T---------GG-GG-TTT------------------------C---------G--A---------G------------G--------TG----------G--T-------------A----------C----------TC----A----- 3019
02_AG.ES.06.P1423            GT-----T---R------G-GGT---------------------------C----G----G--A---------.A-----------G--------TG----------G--T-------------A----------C--------C--TC-G--A----- 3260
02_AG.GW.05.CC_0048          --G----T-------------GTT-TG--------------------A--C---------G-T-A--------------------TG--------TG----------G--T--T----------C-----A----C--------G--TC----A----- 3173
02_AG.KR.12.12MHR9           -------T-----------G-GG-T-------------------------C---------G--A---------------------G--------TG----------G--T-------------A----------C----------TC----A----- 3482
02_AG.LR.x.POC44951          -------T-----------G-GG-T-------------------------C---------G--A---------GA----------TG--------TG----------G--T-------------A----------C----------TC----A----- 3818
02_AG.NG.09.09NG_SC61        -------T---------GG-GT-T-TG-----------------------A---------G------------------------CG--------TG----------G--T-------------T----------C--------C--TC----A----- 3015
02_AG.NG.x.IBNG              -------T-----------G--G-T-TG----------------------C---------G--A---------GA--G---------------TG----------G--T-------------A----------C----------TC----A----- 3348
02_AG.SE.94.SE7812           -------T---------GG-GG-T--------------------------C---------G--A---------GA-----------G--------TG----------G--T-------------A----------C----------TC----A----- 3195
02_AG.SN.98.98SE_MP1211      -------T---------GG-GG-T--------------G-----------C--C------G--A---------GA-----------G--------TG----------G--T-------------A----------C------------C----A----- 3020
03_AB.RU.97.KAL153_2         ---------------------G----T---------------------------------A-----------A-----------------------------------------------------------A----C----------------A----- 3049
04_cpx.CY.94.94CY032_3       -----A----------G---G--TG---T-------------------------G----G-T-A--------------------C-------------A-------G----------C-----A----A----C----------TC----A----- 3189
05_DF.BE.x.VI1310            G--------------------G-A-------------G----------------------G--A------------------------T-------------------------------------------A---------------T-----A----- 3205
06_cpx.AU.96.BFP90           -----A---------------G--CT-----------------T----------------G--A---------A----------G--------------A---------------C----A----A----C-------------C--------C- 3851
07_BC.CN.98.98CN009          --------T-----G------G---TG-------------------------------G-T-A----C-----A----------G-T-----------C-------------------------A-------C---------------A----- 3155
08_BC.CN.97.97CNGX_6F        --------T-----G------G---TG---------------------T---------G-T-A--A-C-----A----------G--------------C-------------------------A---------------------A----- 3003
09_cpx.GH.96.96GH2911        -----A-TG--------GG-G-----------------------------C-----------T-A---------A-------------------G-------G--T-------------A----A----C----------TC----A----- 3025
10_CD.TZ.96.96TZ_BF061       --C------------------G--CA---T-T-----G-------------------G------A-------------------------T--------G--C-----------------T------------------------C----A--G-- 3202
11_cpx.CM.95.95CM_1816       -----A-TGA-A---------TG---G---------G-----------A-T---------G--A---------G----------G--------------C---------G----------C-----A----------------C--TC----A--G-- 3187
12_BF.AR.99.ARMA159          -----A---------------G--------G-------------------------G-------A------T--A--G-------T------------------------------------A----A----C--------C--TC----A---C- 3828
13_cpx.CM.96.96CM_1849       -----A-T---------GG-GG---------T-------------------CT---------------C------G--A----------GT-------------C---------G----------T----A----C----------TC----A---C- 3228
14_BG.ES.05.X1870            -----A-T-------------G-------G------R-------------A-------------G--A--T---G--A----------GT-------------A--------------------C----------C----------TC----A----- 3290
15_01B.TH.99.99TH_MU2079     -----A-TG------------G----------------------------C---------G--A---------A---------------------TG----------G--T-------------A------------------A--T-----A----- 3222
16_A2D.KR.97.97KR004         -----A-----A---------GT--TG---------------------A-T---------G-T-A---------A----------------------G----------G--T---------------------------C----------TC----A--GC- 3186
17_BF.AR.99.ARMA038          ---------------------G--------------------------------G----G--A-----TT-----G-------T--------------------------------------------------------C--TC----A---C- 3037
18_cpx.CU.99.CU76            GT---A-TG---------G-GG--------------------------------------T-A-------G--A----------G--------------A---------------------------------G------GTTTC--------C- 3128
19_cpx.CU.99.CU7             ---------------------T--TT------------G---------------T----------G--A---------------------G--------GG-----------------T----------------------C----------T-----A----- 3049
20_BG.CU.99.Cu103            G--------------------G--CA--------------------------------------T-A--T-----G--A---------------------TG--------------T-------------A----------C------------------------ 3289
21_A2D.KE.99.KER2003         -----A-----A---------G---TG---GC-----------------A--C---------G-T-A---------A----------G--------TG----------G--T-------------------------C--------C--TC----A---C- 3028
22_01A1.CM.01.01CM_0001BBY   -----A-TG---------G-GG-------G------------------------------G--A-------------------G--------TG----------G--T-------------A----------C----------TC----A--G-- 3031
23_BG.CU.03.CB118            -------T-------------G-ACA----------------------------------G-T-A---------A----------G--------TG-----------------------------A----------C--------------------A----- 3266
24_BG.ES.08.X2456_2          ---------------------G-ACA----------------------T---------G-T-A---------A----------G--------TG---C------G--T-------------A----------C--------------------A----- 3274
25_cpx.CM.02.1918LE          G---CA-----------G---G-T-----GG------------------------T---------G-T-------------A---------------------------C-------------A-----C----A----------C--C--------TC----A----- 3031
26_AU.CD.02.02CD_MBTB047     G----A-TG---------G-G----TG--G---------------------------G--A-----------------T--------G--------G--T-------------------------C------------T-----A---C- 3828
27_cpx.FR.04.04CD_FR_KZS     G----A-TGA---G------G---------------G---------------G----G--A------------------------------T------------G-------------------A----------C----------TC----A--G-- 3814
28_BF.BR.99.BREPM12609       -----------------G--G-------------------C---T----------C-T---------G--A---------------------T------------------------------------------C------------------------ 3189
29_BF.BR.01.BREPM16704       ---------------------------------G----------------G---------A----------A--G---------G---------------------------------A----A----C--------C--TC----A---C- 3224
31_BC.BR.04.04BR142          -----------------------CTG----------------------------------G-T-A----T--G-GA-------------------C-------------------------------T-----C----A----- 3295
32_06A1.EE.01.EE0369         -----A---------------G---TG------------G------------------G--A-------------G-----G------------A---------------C-----C----A----C-------------C----A---C- 3458
33_01B.ID.07.JKT189_C        -----A-TG------------G-----------T-----------------A-----------T-A--T-------A--G------------------TG--A-------G--T-------------A------------------G--TC----A----- 3124
34_01B.TH.99.OUR2478P        -----A-T-------------G----------------------------C---------G--A---------A---------------------TG----------G--T-------------A----------C------------T-----A----- 3025
35_AD.AF.07.169H             -------TG---------GG-GG--TG--G------------------A-----------G--A---------A-----------G--------TG--T-------G--T-------------A----A----C--------C--TC----A----- 3023
36_cpx.CM.00.00CMNYU830      -----A--G----------TG----------------------------T-------------A---------A-------T--------TG----------G-------------------A----------C----------TC----------- 3028
37_cpx.CM.00.00CMNYU926      G----A--------G---G--G----------T---A---------------C---------G--A---------GA-----------G--------GG--T---------G--T-------------A----------C----------TC----A----- 3013
38_BF1.UY.03.UY03_3389       -------T-------------T-AGT---G-----------------------------G----G--A----C----A-----------G--------TG----------G-------------------------C-----T-----A---------Y- 3246
39_BF.BR.04.04BRRJ179        GT---A-T---A-----GG--T-TTT------------------------------------T-A-----------A--G-------T-------------------------------------A----------C-----T--C--TC----A---C- 3300
40_BF.BR.05.05BRRJ055        ---------------------G------------------------------C--G----G---------T--GA--G-------T----------------------------------------A----------C--------C--TC----A---C- 3306
42_BF.LU.06.luBF_18_06       -----------------G--S--CT---GGC------C--------G-G--T-------------T-A-----------------G---------------T-------------------------------------C------------------------ 3377
43_02G.SA.03.J11223          ---C---TG--------GG-GG-T--G----------------------A--C---------G--A---------A------------------------------------G--T-------------A----------C----------TC----C-G--- 3348
44_BF.CL.00.CH80             -----------------------A--------------------------------G----G-T-A-------T--A--G-------T---------------------------------------A----------C--------C--TC----A---C- 3265
45_cpx.FR.04.04FR_AUK        -----A-----------G--G--TT-------------------------C---------G--A---------A----------G-------------A---------------C-----A----A-----------------C----A---C- 3809
46_BF.BR.07.07BR_FPS625      -----A-----------G---T-TCT-----------------G---T-----G-------A-------T--A--G-------T-------------------------------------A----------C--------C--TC----A---C- 3261
47_BF.ES.08.P1942            ------------------------------------------------G--A-------------G--A---------GA--G-------------------C---------------------------------C--------C--TC----A--G-- 3262
48_01B.MY.07.07MYKT021       -----A-TG---------G--G----------------------------C---------G--A--T-------GA----------------------G----------G--T---C----------A------------------G--TC----A----- 3028
49_cpx.GM.03.N26677          G----ACTG-------G-G--G---T----G-----------------------C--A-----C-G--A---------A----------------------C----------------------------A----------C----------TA-G--A----- 3245
50_A1D.GB.10.12792           GT---A-TG--------GG-G----GG-----T--------------------------------G---------A--G-------------------TG--C--C-----G--T-------------A----A----C----------TC----A----- 3230
51_01B.SG.11.11SG_HM021      ---------------------G----------------------------T-----G----G--A--T-------A-----------------------TG-----------------------------------C---------------------------- 3030
52_01B.MY.03.03MYKL018_1     -----A-TG-------G---G-----------------------------C---------G--A---------A----------T---------G----------G--T-------------A----------C----------TC----A----- 3137
53_01B.MY.11.11FIR164        -----A-TG--A---------G----------------------------C---------G--A---------A----------------------------------------------G--T-------------A----------------------TC----A----- 3163
54_01B.MY.09.09MYSB023       ---------------------G----------------------G--A-----------------------------G-------------------------------------------------------------------C---------------------------- 3225
55_01B.CN.10.HNCS102056      -----A-TG--A---------G----------------------------C---CG----G--A--T-------------------------------A---------------C----------------C----------TC----A---C- 3133
56_cpx.FR.10.URF5_patient_A  --------------------------------A------------------------------A---------A------------------------------------------------------------C------------C----A----- 3174
57_BC.CN.09.09YNLX19sg       --------T-----G------G---GG---T-----------------------------G-T-A----C-----A----------G--------------C--------------------------A---------------------A----- 3007
58_01B.MY.09.09MYPR37        G----A-TG------------G-T--------Y------------------C---------G--A---------GA----------------------TG----------G--G-------------A-----------------------TC----A----- 3168
59_01B.CN.09.09LNA423        -----A-TG------------G----------------------------C---------G--A---------A----------------------TG----------G--T-------------A-----------------------T-----A----- 3007
60_BC.IT.11.BAV499           ---------------------G--CAG-----T---------------------------GGT-A-----C--G-----------G-G--------------C-------------------------A-----------T-----C----A----- 3780
61_BC.CN.10.JL100010         --------T-----G------G-T-TG----------C----------------------G-T-A-----C-----A----------G--------------C--------------------------A---------------------A----- 3244
62_BC.CN.10.YNFL13           --------T-----G------G---TG-----T---------------------------G-T-A-----C-----A----------G--------------C--------------------------A-------------------------G-- 3152
63_02A1.RU.10.10RU6637       ----CA-TG----------G-GG-----------------------A-------------G-T-A---------GA----------CG---------G----------G--T-------------A----A----C-------------C----A---C- 3364
64_BC.CN.09.YNFL31           --------T-----G------G---TG-----T-------------T-------------G-T------C-----A----------G--------------C--------------------------A---------------------A----- 3177
65_cpx.CN.10.YNFL01          --------T------------T---TG-----------------------T-----G---G-T-A-----C-----A--------A-T--------------C-----------A-------------------C---------------------------- 3194
67_01B.CN.11.ANHUI_HF115     -----------A-----------------------G---------A-T--------------T-A---------A-------------------T----------------------------------------C----------------------G-- 3166
68_01B.CN.11.ANHUI_WH73      -----------A-----------YR-----------------------T--------------T-A---------A-------------------T----------------------------------------C---------------------------- 3157
70_BF1.BR.10.10BR_PE004      -----A--------G------G-T--------T---------------------G---------A------T--A--G-----------------------A--------------------A----------C--------C--TC----A---C- 3279
71_BF1.BR.10.10BR_PE008      G----A-----A--G--G---G-TCA-----------------------------------------A------T--A--G-------T-------------------------------------A----------C--------C--TC----A---C- 3487
72_BF1.BR.10.10BR_MG002      -----A-----------G---G-TCT------T-----------------------G----G------------G------------------------------------------------T----------C--------C--TC----A---C- 3331
O.BE.87.ANT70                G----A-T-A-C--G--GG-GT-TCA----TCT---A-T--C--------ATTG-----------G-----AG-CACT-GA-----T--------------G-G--C-------------------A--A-----A-----C-GC--A--C--A--GA-C----- 3877
O.CM.98.98CMA104             G----A-T-C-A--G---G-GT--CAG---TCT-----T--C--------ATTG----G-----G-----AG-CACT-G------T--------------G-G--C--------G--------------A--A-----R-----C-GC--A--C--A--GA-C----- 3309
O.CM.98.98CMABB141           G----A-T-C-C----GG-GT-TCA---GCT-----T--C--------ATTG-----------G-----AG-TACT-GA-----T-----G------G---C-------------------------A--A-----A-----C-G---A--C--A--GA-C----- 3308
O.CM.98.98CMABB212           G----A-T-C-C--G--GG-GGGTCA---GGCT-----T-------G--ATTA-------C--G-----AG-TACT-GA-----T-----G-T-----G---A---------------------------A-----A-----C-GC--A--C--A--GA-C----- 3313
O.CM.98.98CMU5337            -----A-T-A-C--G--GG-GT-TCA---GCT-----T--C-----G--ATTG-----------GT----AG-CACT-GA-----T--------------G-G--C-------------------A--A-----A-----C-GC--A--C--A----A-C----- 3310
O.CM.99.99CMU4122            G----A-T-A-C--G--GG-GT-TCA---GCT-----T--T-----G--ATTG-----------G----A--CACT-G---G--T--------------G-G--C-------------------A--A-----A-----C-GC--A--C--A--GA-C----- 3308
O.FR.92.VAU                  G--A-T-T-A-C--G--GG-GT-TCA---GCT---A-C-----------ATTG--C--------G-----AG-CACT-GA---GT---------------G---T----------------------------A-----A-----C-GC--A-----A--GA-C----- 3389
O.GA.11.11Gab6352            G----A-T-C-C--G---G-GT-TCA---GCT---A-C--C--------ATTG-----------G-----AG-CACT-GA-----T------T-------G-G--C--------G--------------A--A-----A-----C-GC--A--C--A--GA-C--G-- 3034
O.SN.99.99SE_MP1299          G----A-T-C-C--G-GGG-GT-TCA---GGCT---A-T--C--------ATTG-----------G-----A--CACT-GA-----T--------------G-G--C-------------------A--A-----A-----C-G---A--C--A--GA-C----- 3876
O.US.10.LTNP                 G------T-C-C--G--GG-GT-TCA----TC------C--C-----G--ATTG----------GG--A--A--CACT-GA-----T------T-------G---T-------------A-------A--A-----A-----C-GC--A--C--A--GA-C----- 3798

N.CM.02.DJO0131              GTT--A-TG--A------G--G-A------G-T--------T----------T-----G------T-A--AG-----------GTG-----GG----------T--CA-----------A----------------A--C--C--C--T------C-T--A----- 3329
N.CM.02.SJGddd               GTT--A-TG--A-G----G-GG-A--------T--------T-----------------G----G-T-A--AG-----------GTG-----GG----------T--TC-----------A--T---------------A--C--C-GC--------C-T--A----- 3313
N.CM.04.04CM_1015_04         --T--A-T---A-----GG--G-A--------T--------T-----------------G------T-A--MG-----------GTG-----GG----------T---A-----------A--T--------A------A--C-----C--T------C-T--A----- 3328
N.CM.06.U14296               G-T--A-TG--A------G--T-A--------T---R----T-----------------G------T-A--AG------------TG-----GG----------T---C-Y---------A--T-----------A----A--C--C--C--T------C-T--A----- 3326
N.CM.06.U14842               G-T--A-TG---------G-GG-A--------T--------T-----------------G------T-A--AG------A----TG-----G---------------AC-----------A-GT-----------T----A-----C--C--T--C--TC-T--A----- 3330
N.FR.11.N1_FR_2011           GTT----TG--A-G----G-GG-A-----------------T--------A-------------G-T-A--AG------A---GTG-----GG----------T---C-----------A--T-----------T----A-----C--C---------C-T--A---C- 3224

P.CM.06.U14788               G----R-T-A-A--G--GG--T-ACA-----------G----------ACTA--C--G----G---A--AG--A-T-GAA-TGTG-----G--T------T----C-----------------------------A----------A------GC--A--C--TC-TA-T--GC- 3294
P.FR.09.RBF168               G----A-T-A-A--G--G-C-T-ACA-----------------------ATTA--C--G----G---A--AG--A-T-GAAGT-TG-----G--T------T----C-----------------------------A----A-----A------G---A--C--TC-TA-T--GC- 3858

CPZ.CD.06.BF1167             G---G--T-----------TGGG-----------TA------------GAGTA---GT---CC----T--AG-TAC-------GTG---C--CA-------T-T--C---------G-T--G-----------A--A---------------C-----C--A---A-A--G-- 3925
CPZ.CM.05.SIVcpzMT145        G-T--AA----A---------G-T-----------T-------------A--A--A--------G---C--TG-T-----A----TG------G-C-----G----T--C-----G--T-------------------------A--CA-A--C-----A---C-----A--G-- 3365
CPZ.GA.88.GAB1               G-G--A-----A--------TG-T--T----------------T--------A--A---------GGT-A--AG----G--A----GT------G---------G----A--C----G--A--------C-----A-----A---A-C--------A--------C----- 3884
CPZ.TZ.06.SIVcpzTAN13        G---GTCTTA-A--G-------GGT-AT---------------T-----C-CTGTA--A--G---CTCT----AG-GACT-------------AGTCAG-----------T--C-------T--------T-------G-A-----A-----A--CA-----C-----C--A--GA-T-G-C- 3464
CPZ.US.85.US_Marilyn         GTT--A-TG------------G-T--T--G-----T--C--T---------GTA--A---------T---TAG-G--G-----GTC-----------------T----A---------G--------------A--T-----A-----C-------------C-C--A---C- 3883

GOR.CM.07.SIVgor2139_287     G----A-TGA-A-----GG-CT-ACA-----------T--C--------ACTG-----G--C-GG--A------TGTCG-AGT--C------CG-T-TT-----T--C---------G--TT-------C-----------A---A---G---A------C-CA-C--GC- 3360
GOR.CM.07.SIVgorCP2135con    G----A-T-A-A-----GG-CT-ACAG-----------T--C--------ATTG-----G--C-GG--A------TGTCGAAGT--C------CG-T-TT-----T--C---------G--TT-T-----------------A--------G---A--C----C-CA-T--GC- 3359
```

```
                               p51 end and p66 RT continue Pol p15 RNAse H start
B.FR.83.HXB2                   ATGGTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCTAACTGACACAACAAATCAGAAGACTGAGTTACAAGCAA 3993
Pol                            __W__Y__Q__L__E__K__E__P__I__V__G__A__E__T__F__Y__V__D__G__A__A__N__R__E__T__K__L__G__K__A__G__Y__V__T__N__R__G__R__Q__K__V__V__T__L__T__D__T__T__N__Q__K__T__E__L__Q__A__
A1.AU.03.PS1044_Day0           ---------------------C-----------------G--T--------------------C--T-----A-----GC----G--------G-----C---G-C--G-----------G---A-TT----------G-----------A--------AC-G--T---- 3201
A1.CH.03.HIV_CH_BID_V3538      ---------------------C------A----------G--T--------------------------A--------GC----G--------G---A-C---G-C-A------------G---A-TT----------G-----------A--------A-----T---- 3198
A1.CY.08.CY236                 ---------------A-----C-------C---------G--G--------------------C--T-----------GC----G--------G-----C---G-C--------------G-----TT----------G-----------A--------AC-G--T---- 3204
A1.ES.05.X1608_8               ---------A-----------C-------C---------G--T--------------------C--T------------C-------------G-----C--AG-C--------------G-----TT-----G----G-----------A--------A-----T---- 3441
A1.KE.11.DEMA11KE001           ------T--A-----A-----C-------A---------G--T--------------------C--T------------C----G--------G-----A---G-C--G-----------G-----TT----------G-----------A--------AC----T---- 3332
A1.RU.11.11RU6950              ------------G--------------------G-----G--T----------------------------------GGA-------------G---------G-C--------------G-----TC-T----A---A-----------A--------A--G--T---- 3517
A1.RW.11.DEMA111RW002          ---------A-----------C--T--------------G--T-----------------------T----GA--G--GC-------------G-----C---G-C-------A------G-----TT----------G-----------A--------A-----T--T- 3344
A1.SN.01.DDI579                ------------------G--C------A----------G--T-----------------------T-----------GC----G--------G-----C---G-C--------------G-----TT--T---TG--G-----------A--------A-----T---- 3191
A1.UG.11.DEMA110UG001          ---------------A-----C--T--------------G--T----------------------GT-----A-----GC-------------G-----C---G-C--------------G-----TT----------G-----------A------------T-T--T- 3348
A1.ZA.04.04ZASK162B1           ---------------A-----C------T----------G--T--T-----------------------A--------GC-------------G-----C---G----G-----------G-----TC----------G-----------A--------A--G--T---- 3434
A2.CM.01.01CM_1445MV           ---------------AC-------------------------T-----------------------T--A--------GC-------------G-----C---G-C---------------A-----T-----G-G--G-----------A-G------A-----T---- 3192
A2.CY.94.94CY017_41            G--------------A-------------C------------T-----------------------T--A---------C-------G-----G-----C---G-C---------------A-----T-----G-G--G-----------A--------A-----T---- 3350
B.BR.10.10BR_MG029             ---------------A-----------------G-----G--T--T--------------T-----------------G-C------------------C---G-C-----------------GA--CA------------------------A--------G------- 3455
B.CA.07.502_1191_03            ---------------A--G-----------------------T-----------------------T-----A-----------------G------------G-C-A------------G---A-T-A----------------------------------------- 3420
B.CH.08.M2_0803101_NFLG8       ------------------------------------------------------------------T-------G------------------G-----A---G-C----------------C-A--C-----------------------C----A------------- 3462
B.CN.10.DEMB10CN002            ------T-----------------------------------T-----------------------T-----C------------------------------G-C--------------G------T-----GAG---------------------------------- 3365
B.ES.10.DEMB10ES002            ---------------------C--------------------T-----------------------T-----C------------------------------G-C--------G------------C--A-G--A-----------------------A---------- 3355
B.FR.11.DEMB11FR001            -------------------------------------------------------------------------------------------------------G-C-A----------------A--T------------------------------------------ 3239
B.GB.05.MM45d213_GN1           ----------------------------A-------------T------------------------------------------------------------G-C-A-----------G------T--A-------------------A--------------G---- 3394
B.HK.06.HK003                  ---------------A--------T-----------------T--------------------C--------A------C-------------------------C--------------G------T-------------------------A-----A---------- 3363
B.HT.05.05HT_129389            ------T--------A--G-----------------------T-----------------------T--A---------------------------------G-----------------------T--T-------T------------------------------- 3194
B.JP.12.DEMB12JP001            ------------------------------------------T-----------------------T--------------------G---------------G-C-------A------G------T-------------------------------A---------- 3380
B.KR.07.HP_18_07JHS10_3909     ---------------A--------------------------T-----------------------T------------C-------------------------C--------G-----G------T-------G--T------------------------------- 3657
B.PE.07.502_0525_wg5           ------T--------A------------T-------------T--------------A--------T-----A------------------------------G-C-A-----A------G--G---T--A----------------------A--A-----------T- 3458
B.RU.11.11RU21n                ------------------------------------------T--------------A--------T------------------------------A-----G-C-A------------G------T-------------------------A-----A---------- 3564
B.TH.08.MERLBDTRC10            -----------------------------A------------T-----------------------T------------A-------------------------C-A-------------------------C------------------------------------ 3226
B.US.11.ES38                   ---------------A------------C-------------T------------------T-C--T--A--A------------------------------G-C-----G--------------T-AT---------------------------------------- 3928
C.AR.01.ARG4006                ----------C-G---------------AC------------T--------------A--------T-----A------A-G--G--------G---------G-C--------G--G---A----TG-T---------------------------------------- 3177
C.BR.07.DEMC07BR003            ----------C-G-----------------------------T-----------------------T--A--A------A-G-----------G-----C---G-C--------G--G---A-C--TC-T--------A------------------------------- 3325
C.BW.00.00BW5031_1             ----------C-------------G---C----T-T------T--------------A--------T-----A------A-------------G---------G-C--------G--G--GA----TT-T--------A--------------A-----A--------G- 3364
C.CN.10.YNFL19                 ----------C-G--------T--T----C---T-T------T--------------A--------TC----A------AA------------G---------G-----------A-G---A----TT-T--------A--------------------A--G------- 3346
C.CY.09.CY260                  ----------C-G----------------C------------T--T-----------A--------T-----C------A-------------G---A----AG-C--------G--G---A----TT-T----A---A--C---------------------------- 3216
C.ES.08.X2363_2                ----------C-G---------------A-------------T--------------A--------------A------C------G------G---------G-C--------G--G---A----T--T--------A--C-----------------A---------- 3424
C.ET.02.02ET_288               G---------C-------G----------C------------T--------------A--------T-----A------A-------------G---------G-C-----G-----G---A----TT-T--------A--------------------A---------- 3195
C.IN.09.T125_2139              ----------C-G--------T-------C---T-T------T--------------A--------T-----A------A-------------G---------G-C--------GA-G---A----TT-T--------A--------------------A--G------- 3969
C.KE.00.KER2010                ----------C-G--------T-------C------------T--------------A--------T-----A------G----G--------G---------G-C--------G--G--------TT-TA---G---A--------------------A--G------- 3174
C.MW.09.703010256_CH256.w96    ------T---C-G-----G----------C---G--------T--------------A--------T-----A------AAG-----------T---------G-C--------G--G--------TT-T-------GA-----------A------------------- 3973
C.TZ.08.707010457_CH457.w8     ----------C-G--------------------------G-----------------A--------T-----A------A-------------G---------G-C-----------G---A----TT-T-----A--A--------------------A---------- 3967
C.YE.02.02YE511                ---------AC-G--------------G-----T-T------T--------------A--------T-----A------A-G-----------------A---G-C--------G--G---A----TT-T----A---A---------------G----A-----G---- 3168
C.ZA.08.705010534_CH534.w12    ----------C------------------C------------T--------------A--------T-----A------A-------------G---------G-C-----------G---A----TC-------A--A--C------------G--------------- 3983
C.ZA.10.DEMC10ZA001            ------T-----G-----G---------A-------------T--------------A-------GT-----C--G---A-------------G---------G---A------G--G--------T--T--------A--------------------A---------- 3333
C.ZM.11.DEMC11ZM006            ---------AC-G--A--------------------------T--T-----------A--------T-----A------C-G-----G-----G---------G-------------G---A----TT-T---T----A--------------------A---------- 3342
D.CM.10.DEMD10CM009            ------T--------A--G--T-------C------------T-----------------------T--A--A-----------G------------------G-C-----------G----C----T-AT-T--------C-----------A---------------- 3368
D.CY.06.CY163                  ------T--------A--G---------A-------------T------------------T----T--TA--------------------------------G-C-----------G---------C-AT-T--------C--------------------G------- 3201
D.KE.11.DEMD11KE003            ------------------G-----------------------T-----------------------T--A--------------G------------------G-C-----------G--------AT----G-----------------------A--A--------C- 3343
D.KR.04.04KBH8                 -----------------GG---------A-------------T-----------------------T--A-------G--------G----------------G-C------------------A--C-TT--------------------------------------- 3952
D.SN.90.SE365                  ---------------A--G-----------------------T--------G-----A--------T--A--A-----------------------------AG-------G----------C-A--C-T-----------C---------------------------- 4010
D.TZ.01.A280                   ------------------G---------A-------------T-----------------G-----T--A-------G-------------------------G-C-A---------G---------T-T--------------------------------------C- 3194
D.UG.10.DEMD10UG004            ------------------G---------AC------------T-----------------------T--A---------------------------------G-------------G---------C-TT----------------------------A--------C- 3331
D.UG.11.DEMD11UG003            ---------------A--G---------A-------------T-----------------------T--A---------------------------------G-C--------------G------T-T---G-------------------------A--G-----C- 3348
D.YE.02.02YE516                ------T---------------------A-------------T-----------------------T--A---------------------------------G-C-----------G---------C-AT-T--------T--------------------G------- 3204
D.ZA.90.R1                     ---------A--------G-----------------------T-----------------------T--A---------G-----------------------G-C-----------G---------C-TT-T------------------------------------- 3339
F1.AO.06.AO_06_ANG32           ---------------A-C---G-------C------------------------------------T--AA-A--C---AA---G------------------G-C-------AGG------C----C----------A--C------------G--------------- 3204
F1.AR.02.ARE933                ------T--------A-CG--G-------C-------------------------------T----T--A--A--C---AA----------------------G-C--------------G-C----C----------G--C------------G--------A-----G 3287
F1.BR.10.10BR_RJ015            ------T--------A-C---G---------------------------------------T----T--A-G---C---GA------------------C---G-C--------------G-C----TTTT-------G--T------------G-C------G-G---- 3466
F1.CY.08.CY222                 C-----T--------A-C---G------AC---G---------------------------T----T--------C---A-----------------------G-C---------------------T----------G--T------------G--------------- 3186
F1.ES.02.ES_X845_4             ------T--------A-C---G-------C------------------------------------T--A--T--C---AA---G------------------G-C--------G-----G-C---AT----------G--C------------G-A------------- 3445
F1.RO.96.BCI_R07               ------T--------A-C---G-------C-------------------------------T----T--A-----C---CA----------------------G-C--------------G-C----T----------A--C------------G--------------- 4017
F1.RU.08.D88_845               ---------------A-C---G------CC--A---------T--T---------------T----T--A--A--C---AA----------------------G-C-A--------------C---------------G--------------AG--------------- 3452
F2.CM.02.02CM_0016BBY          ---------A-----A-C---G------A-------------T----------------------GT--A---------A-----------------------G-C--------------G------T----------G--------------------A---------- 3183
F2.CM.10.DEMF210CM001          ---------------A-CG--G--------------------T--------------A--------T--A--A------C-----------------------G-C-A---------G--G------T----------G--------------------A---------- 3257
F2.CM.10.DEMF210CM007          ---------------A-C---C-------------T------T-----------------------T-----A------C-----------------------G-----------------------T----------G-----------------A--A---------- 3340
G.BE.96.DRCBL                  ------T--------A-C----------CC-----T------T-A---------------------T--------A-----------G-----------G---G-C-A-----------TTA--A-T-----------A--------C--A---G----A-----T---- 3950
G.CM.10.DEMG10CM008            ------TG-------A-C----------CC------------T-A---------------------T-----A--A-----------G---------------G-C-------A------GA--A-TT----------A--------C--A--------A---------- 3374
G.CN.08.GX_2084_08             ---------A-----A-C----------CR---------G--T-A---------------------T-----A--A--G--------G---------A-----G-C-A-------------A--A-----A-CAG---A--------C--A--G-C---A-----T---- 3233
G.CU.99.Cu74                   ------TG---------C---G------CGG-----------T-AT--------------------------A--A---------------------------G-C-A-----A------G---A-T-----------T--------C--A-A------A---------- 3597
G.ES.09.X2634_2                ------TAG------A-C---G-------C------------T--T-T---------A-------GT-----A--A-----------G---------------G-C-------AG--G------A-T------T----A--------C-----AG----A---------- 3476
G.GH.03.03GH175G               G-----T--------A-C----------CC------------T-A---------------------T--------A-----------G-----G---------G-C-A-----A-------A--A-T---T-------A--T-----C--A---G----A-----G---- 4033
G.KE.09.DEMG09KE001            ------T--------AGC----------CC---------G--T-AT-----------A--------T-----A--A---G-------G---------A-----G-C-A-----A-------A--A-T-----------A--------C--A---T----AC----T---- 3368
G.NG.09.09NG_SC62              ------TG-------A-C-----------C------------T-A------------A--------T---A-A--A--------G--G-----------------C----A--A-------A--------A--CAA--A--------C-----A-----A-----T---- 3185
G.ZA.01.TV546                  ------TGA------A-C----------CC------------A-A---------------------T-----A--A---------------------------G-C-------A-------A----------------A-----------A--------A-----G---- 3195
H.BE.93.VI991                  ------T----------C---G-------A------------T-A---------------------T-----A------A-G--G--------------C---G-C-------A-------A----TT-------G--A--------------A-----A---------- 3382
H.CF.90.056                    ------T--------A-C---G-------C------------T-A----A----------------T-----A--------------------------C---G---------AG------------T-------G--A--------------------A---------- 3340
H.GB.00.00GBAC4001             ------T--------A-C------------------------T-A---------------------T-----A------A----G--G-----------C---G-C-----G-A-------A----TT-------G--A--------------------A---------- 3511
J.CM.04.04CMU11421             ---------A-----A--R---------AGG-----------T--Y-------------------------------R-AAR--R--R-----G---A-C---G-C-A---------G-----G--T--YT--------------------C-R-----RC-G-----C- 3499
J.SE.93.SE9280_7887            ---------A-----A--G---------AGG-----------T------------------T-----------------AC------G-----G---------G-C-A---------G-----A--T--------------------------------AC----C--C- 3307
J.SE.94.SE9173_7022            ---------A-----A--G---------AGG-----------T------------------T----------A------G-------G-----G---------G-C-A---------G-----AA-T--------------------------------AC-------C- 3308
K.CD.97.97ZR_EQTB11            ---------------A-C------------------------T--------------------C--T-----A------CAG---------------------G-C-A----------------A--T-AA-------A--------------A-----A---------- 3189
K.CM.96.96CM_MP535             ---------------A-C---------------G--------T-----C---------------C-T-----A------AAG----G----------------G-C--------------G------T--A-------A--------------AG----A---------- 3189
U.CA.01.TV749                  ------------G--A-----C------AGG-----------T--------G--------------T--A-----------------------------C---G-C--------------------TT--T-------A-----------A-----------G--T---- 4010
U.CA.99.TV721                  ------------G--A-----C--------------------T--------------A--------T--A---------C-------------------C---G-C---------------A----TT--T---A--GA-----------A-----------G--T---- 4009
U.CD.83.83CD003_Z3             ---------------A-----C------AG---------G--T--------------------------------A---AAG---------------------G-C------------------A-TT----------A--------------AG-A--AC--------C 3488
U.CD.90.90CD121E12             ---------------A-----G-------C------------T----------------------G------A--A---AAG---------------------G-CCA------------AC--A-TT----------A--------------AG-G--AC--------C 3192
U.CY.05.CY090                  ---------------A-C-----------C------------T-----------------------T--TA--------GA----------------------G----------G-----------TT-AT-------A----------CT--------A---------- 3192
U.CY.08.CY223                  ----------C------GG----------C---G--------T--------------A--------T--A--A------C-------------G-----C---G-C-A------------G-----TT----G-A---G-----------A--------A-----T---- 3210
U.ES.10.DEURF10DZ001           ----------------------------A-------------T-----------------------T-----A----G---------------G---------G-C--------------G-----TT--T----A--G-----------A--------A-----T---- 3269
U.GR.99.99GR303                ---------------T----G---------------------T--------------A--T-----T--A--A--A---AGG---------------------G-C--G--------W--G-----TG----------A--------------------A---------- 3264
U.NL.95.U_NL_95_H10986_D1      ------T-----G--A--G--C-------A------------T--------G--------------T-----A------A-G-----------G---------G-C--------------G------T----------A--------------------A---------- 3878
01_AE.AF.07.569M               ---------A-----A-----C-----------------R--T----------------------GT------------C-------------------C---G-Y--------------G--AA-T-----C-----G-----------A--------A-----T---- 3199
01_AE.CF.90.90CF11697          ---------A-----A-----C------C----------G--TC---------------------GT-------A---GC-------------------C---G-C--------------G--A--TT----------G-----G-----A--------A-----T--G- 3942
01_AE.CN.10.YNFL03             ---------A-----A-----C-----------------G--A----------------------GT-----A--C--GC-------------G--C--C--AG-C--------------G--A--TT-----G----G-----------A--------A-----T--G- 3371
01_AE.HK.04.HK001              ---------A-----A-----C--------G----------------------------------GT-----A-----GC-------------G-----C---G-C--------------G--A--TT----------A-----------A--------------T---- 3355
01_AE.IR.10.10IR.THR48F        ---------A-----A-----C-----------------G--T----------------------GT-----------GC-------------G-----C---G-C---------A---GG--A--TT----------G-----------A--A-----A-----T--G- 3183
01_AE.JP.x.JRC77AE             ---------A-----A-----C-----------------G--T----------------------GT-------A---GA-------------G-----C---G----------------G--A--TT----------G-----------A--------A-----T--G- 3990
01_AE.TH.09.AA090a_WG11        ---------A-----A-----C-----------------G--T----------------------GT-----------GC----------C--G-----C---G-C--------G--------A--TT----------GG----------A--------A-----T--G- 3416
01_AE.TH.90.CM240              ---------A-----A-----C-----------------G--T----------------------GT-----------GC-------------G-----C---G-C--------------G--A--TT----------G-----------A--------A-----T--G- 3567
01_AE.US.05.306163_FL          ---------A-----A-----C-----------------G-------------------------GT-----------GC-------------G-----C---G-C-----------------A--TT----------G-----------A--------A-----T--G- 3184
01_AE.VN.98.98VNNDI5           ---------A-----A-----C-----------------G--T----------------------GT-----------GC-------------G-----C---G-C--------------G--A--TT----------G-----------A--------A-----T--G- 3248
```

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

p51 end and p66 RT continue Pol p15 RNAse H start

```
B.FR.83.HXB2                 ATGGTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCTAACTGACACAACAAATCAGAAGACTGAGTTACAAGCAA 3993
Pol                          __W__Y__Q__L__E__K__E__P__I__V__G__A__E__T__F__Y__V__D__G__A__A__N__R__E__T__K__L__G__K__A__G__Y__V__T__N__R__G__R__Q__K__V__V__T__L__T__D__T__T__N__Q__K__T__E__L__Q__A__
02_AG.CM.10.DE00210CM013     ---------A----------C-------C------------T------------------C--T-----A-----C-----------G---------G-C-------------GGA--A-TT---------G----------A-------A-----T---- 3360
02_AG.CY.09.CY256            --------------------C------CC-----------T---------------------T----------GC-----------G-----C---G-C-A------G-----G----T----------G----------A-------A-----T---- 3189
02_AG.ES.06.P1423            --------------------T--A--------G-------T---------------------T--A----G---C-----------------C---G-C---------------------R-TT--T----C--G---------------------T---- 3430
02_AG.GW.05.CC_0048          ---------A----------C-------------------T----------C----------T-----A----GA-----------G-----C---G-C---------------G-----T----------G----------A-------A-----T---- 3343
02_AG.KR.12.12MHR9           ---------A----------C-------C-----------T---------------------T----------C------------G-----C---G-C---------------G-----TT---------G----------A-------A-----T---- 3652
02_AG.LR.x.POC44951          ---------------A----C-------------------T-------G-------------T----------GC-----------G-----C---G-C-A-------------G-----TT---------G----------A---------------T---- 3988
02_AG.NG.09.09NG_SC61        ---------A-----------------C---G--------T------------------C--T----------GC-----------G-----C---G----G------------G--G--TC---------G--C-------A-------A-----G---- 3185
02_AG.NG.x.IBNG              ---------------A----C-------------------T------------A--------T--A-------GA-----------G-----C---G-C---------------G-----TT---------G----------A-------A-----T---- 3518
02_AG.SE.94.SE7812           ----------C----A----C-------------------T---------------------T----------GA-----------G---------G-C---------------G-----TT-----G---G----------A-------A-----T---- 3365
02_AG.SN.98.98SE_MP1211      --------------------C------A---------G--T------------------C--T-----A----GC-----G-----G-----C---G-C---------------T-----TT---------G----------A-------A-----T---- 3190
03_AB.RU.97.KAL153_2         ----------------------------------------T------------A--------T-----------C--------------------G-C---------------G------T------------------------------------ 3219
04_cpx.CY.94.94CY032_3       ---------------A-C---C-------C----------T---------------------T--A--A--A---CAG--------------------G-------C------------------T-----T---A-----------------A---------- 3359
05_DF.BE.x.VI1310            ------------------G--------T------------T---------------------T--A--------------------------A----G--A-----------------C----T----------G--G-----------AG--------------- 3375
06_cpx.AU.96.BFP90           ---------------A-C---G------------------T----------------G-----T--A--A----GCA--------------------G-C---------------G-----T----------A----------------G--A---------- 4021
07_BC.CN.98.98CN009          ----------C-G--A-----T------------T-----T------------A--------T----------A------------G---------G-C--------GA-G--------TT-T--------A--------------------A--G--C---- 3325
08_BC.CN.97.97CNGX_6F        ---------AC-G--------T-------C-----T----T------------A--------T----------A---G--------G---------G-C--------GA-G---A----TT------------------------------A--G-------- 3173
09_cpx.GH.96.96GH2911        ---------------A-----C------A---------G--T---------------------T----------GC---G----G--G-----A----CC----------------GA----AT----------A--C-------A---------C-----T--C- 3195
10_CD.TZ.96.96TZ_BF061       ----------C-G--------------------------------------------------T--A----A----------------------------G-C-------------G-G--C-A--T-TA-------T---------------------A--------C- 3372
11_cpx.CM.95.95CM_1816       ---------------A--G--G------A-----------T---------------------T-----A-----G-----------G-----C---G-C------------G------G--T----------A----------A-------AC--G------ 3357
12_BF.AR.99.ARMA159          ------T--------A-C---G-------C--A---------------------T----T-AA----C---AA---------------------G-C-A-----------------------T---------A--C---------G---------T---- 3998
13_cpx.CM.96.96CM_1849       G-----T-G------A-C----------C-----------T-A--------------------T-----A--A----------G--------------G-C-------A-------A--A-T-----------A-------C--A-------A--G------- 3398
14_BG.ES.05.X1870            ------T--A-----A-C--------------------A-AT--------------------T-----A--A-----G---------------------C-A-----A-------A--A-T-----------A-------C--A--AG---A---------- 3460
15_01B.TH.99.99TH_MU2079     ---------A-----A-----C------------------G------------------------GT----------GC-----------G-----C---G-C-------------G--G--A--TT---------G----------A-------A---------- 3392
16_A2D.KR.97.97KR004         ------T--------A-C---G-----AC-----------T------------------C--T--A--------C------G----G-----C---G-C-A-------------------A----T----------T-----------------A-----T---- 3356
17_BF.AR.99.ARMA038          ------T--------A-C---G-------C--A------------------------CT----T--A----C--AA-----G--------------G------------------G-C----T----------G--C----------G---A-----T---- 3207
18_cpx.CU.99.CU76            ------T--------A-C---G-------C----------T---------------------T-----A-----A---G-------G-----C---G-C-A-------------G---A--T--A-------A-------------A----A---G------ 3298
19_cpx.CU.99.CU7             ---------------A--G---------A-G---------T---------------------T--A---------------------------------G-----------G----------T-T------------------------------------ 3219
20_BG.CU.99.Cu103            ------------------G---------------------T---------------------T--A---------------------------------G-C-------A-----G------T------------------------------------- 3459
21_A2D.KE.99.KER2003         ---------------A------------------------T--T------------------T--A-------C------G----G-----C--GG-C--------------A------T---G--G--G---------A-G------A-----T---- 3198
22_01A1.CM.01.01CM_0001BBY   ---------A-----A-----C-----A---------G--T----C---------------GT----------GC---G-------G-----C---G-C---------G------G--A---T----------A---T------A-------A-----T---- 3201
23_BG.CU.03.CB118            ----------------------------------------T---------------------T--A---------------------------------G-C-------A-----G------T------------------------------------- 3436
24_BG.ES.08.X2456_2          ------T-----------G---------------------T-------T-------------T--A--T----------------------------G-C-----G-------G------T------------------------------------- 3444
25_cpx.CM.02.1918LE          ---------------------------A---------G--T---------------------T-----A----GC-----G-----G-----C---G-C---------------G-----GT---------G----------A-------A-----T---- 3201
26_AU.CD.02.02CD_MBTB047     ---------------A-----C-----A-G----------T------------A--------T-------------------------G---------G-C-A-------------G--------T------G-----C----A-------A-----T---- 3998
27_cpx.FR.04.04CD_FR_KZS     ------T--------AC-G--G------------------------------C---------T-----A-----A-----------G-----C---G-C-A-----G--G--G-T---------A----------A-------A---------- 3984
28_BF.BR.99.BREPM12609       ---------------A------------------------T---------------------T-----A-----CA--------------------G-C---------------TC-TT----A---------------------------------------- 3359
29_BF.BR.01.BREPM16704       ------T----G--A-C----------A--------------------------T----T--A-----C---AA--------------------G-------G---------C----T----------G--C------------G------------- 3394
31_BC.BR.04.04BR142          ----------C-G-------------------G-------T------------A--------T-----A----A-G---------G---------G-C--------G--G---TG---TT-T--------A-------------------G-------- 3465
32_06A1.EE.01.EE0369         ---------------A-C---G-----------T------A----------------G-----T--A--A-----AA--------------------G-C--------G-----G-----------T---A-------------A--A---------- 3628
33_01B.ID.07.JKT189_C        ---------A-----A-----C-----A-G----------T----------------------GT-A--A----GC-----------G-----C---G-C---------------G--A--TT-T------G----------A-------A-----T--G- 3294
34_01B.TH.99.OUR2478P        ---------A-----A-----C---------------G--T----------------------GT----------GCG------------A-------C---------------G--A--GT---------G----------A-------A-----T--G- 3195
35_AD.AF.07.169H             --------------------C-----------G-----G--T------------------C--T----------GC-----------G-----C---G-C--G------------G-----TC----G---G----------A-------A-----T---- 3193
36_cpx.CM.00.00CMNYU830      ---------A-----A--G--C---------------G--T---------------------T----------GC-----------G-----C---G-------------------------TT---------G----------A-------A-----T---- 3198
37_cpx.CM.00.00CMNYU926      ---------------A--G---------------------T--T-----------------G---T-------C--GCA-----------G--C--C--GG---A-----------GA----TT--T-C----G----------A-------A-----T---- 3183
38_BF1.UY.03.UY03_3389       ------Y--------A-C---G-------C--A---------------------T----T--A-----C---AC---------------------G-C-----------G----C----T---------G--C-----------G---------T---- 3416
39_BF.BR.04.04BRRJ179        ------TG-------A-C---G------A--------------------A--------T-----A-----------------------------G-C-------G-----G-----T-----GAG------------------------------------ 3470
40_BF.BR.05.05BRRJ055        ------T--------A-C---G-------C----------------------------T--A--A-------------C--A-A---G-C---------------G------C-------G-T-------------A--------------- 3476
42_BF.LU.06.luBF_18_06       --------------------C-------------------T---------------------T-----A--S------------------------G-C-----------------C---T---------G--C-----------G--------------- 3547
43_02G.SA.03.J11223          ------T--------------C-------G----TG----T---------------------T-----AT---GCA------------------C--G-C---------------G-----TT---------G----------A-------A-----T---- 3518
44_BF.CL.00.CH80             ------T----G--A-C---G-------C----------------------------T----T--A---GC--AG-----------G-----C---G-C------------------C---TT---------G--C-----------G----------C---- 3435
45_cpx.FR.04.04FR_AUK        ----------C----A-C-----------C-----T----T---------------------T--A--A-----AAG--------------------G-C---------------G-----T----G--A--A-------------------A---------- 3979
46_BF.BR.07.07BR_FPS625      ------T--------A-C---G-------C------------------------T-------A---------AAG--------------------G-C------------------C----T---------G--C-----------T--------------- 3431
47_BF.ES.08.P1942            ----------------------------A-----------T-----------------------------------------------------------A---------G---------T-------------------C--------------------- 3432
48_01B.MY.07.07MYKT021       ---------A-----A-----C------A----------------------------------GT----------GCG-----------G-----C---G-C---------------G--A--TT---------G----------A-------A-----T--G- 3198
49_cpx.GM.03.N26677          ------------G--A-----C-------G--------G--A---------------------T----------C---------T--G-----C---G-C-A------------G--C--AT----G----G----------A---------------T---- 3415
50_A1D.GB.10.12792           G--------A-----------C-------A--------G--G------------------C--T----------C-----------G-----C---G-----G------------G-----TT-----A---G----------A-T------C-G--T---- 3400
51_01B.SG.11.11SG_HM021      ------T---------G----------G-----------T---------------------T--A-------------G--------------C---G-C---------------G------T---------------------A---------G---- 3200
52_01B.MY.03.03MYKL018_1     ---------A-----A-----C---------------G--T---------------------T----------GC-----------G-----C------C------------GA-A--TT------A---G----------AC------A-----T--G- 3307
53_01B.MY.11.11FIR164        ---------A-----A-----C---------------G--T---------------------T-----A---G----G-------G--A-C--G-C---------------G--A--TT---------G----------A-------A-----T---- 3333
54_01B.MY.09.09MYSB023       ----------------------------A--A--------T---------------------T--A--A--C---------------------C--C-A----------------------------C-A--------------AG------------ 3395
55_01B.CN.10.HNCS102056      ------T--------A------------------------T-------------------------------------------------C-A-------------G------T-------------------------------A---------- 3303
56_cpx.FR.10.URF5_patient_A  ------T--A-----------C-----T------------T---------------------T----------GA-----------T-----C---G-C---------------------TT---------A----------A-----A-------T---- 3344
57_BC.CN.09.09YNLX19sg       ----------C-G--------T-------C----------T------------A--------T-----A-----C-----------G---------G-C--------GAGG---A----TT-T-----A--A------------------A--G-------- 3177
58_01B.MY.09.09MYPR37        ---------A-----A-----C---------------G--T-------R-----R----Y-GT-A--------G-----------G-----C---G-C---------------G--A---T---------G----------A-------A-----T--G- 3338
59_01B.CN.09.09LNA423        ---------A-----A--G--C------------------T-------G-------------G--A--------GC-----------G-----C------C------------G-----A--TT---------G----------A-------A-----T--G- 3177
60_BC.IT.11.BAV499           ----------C-G---------------G-----T------------------A--------T-----A-----A-G---------G---------G-C--------G--G-----CA-TT-T-------A--M--------------------------- 3950
61_BC.CN.10.JL100010         ----------C-G--------T-------C----------T------------A--------T-----A--A--A-----------G---------G---------GA-G---A----TT-T--------A--------------------A--G-------- 3414
62_BC.CN.10.YNFL13           ----------C-G--------T-------C-----T----T------------A--------T-----A-----A-----------G---------G-C--------GA-T---A----TT-T----A--A-------------G-----A--G-------- 3322
63_02A1.RU.10.10RU6637       -----------G-------------------------G--T-------------------------A----GA-----------G-----C---G-C---------------G---A-TT---------G----------A--------C-----T---- 3534
64_BC.CN.09.YNFL31           ----------C-G--------T-------C-----T----T------------A--------T-----A-----A-----------G---------G---------GA-G---A----TT-T--------A--------------------A--G-------- 3347
65_cpx.CN.10.YNFL01          ------T---------------------AG----------T-A-------------------T-----A------------------------------C-A-----------G-------------------------------G-----C--------- 3364
67_01B.CN.11.ANHUI_HF115     ----------C----A------------------------------------C--T--A--C-----------------------------G-C---------------------------T------A------------------------------- 3336
68_01B.CN.11.ANHUI_WH73      ----------C----A------------------------------------C--T--A--C-----------------------------G-C-------------G--------G-----C------A------------------------------- 3327
70_BF1.BR.10.10BR_PE004      ------T-----------C--G-----AC----------------------------T-C--T--A----C--G---G-----------C---G--------------------G-----T--A------G--C-----------G--------------- 3449
71_BF1.BR.10.10BR_PE008      ------T--------A-C---G-------C----------------------------T---GT--A----C---AA------------------G-C--G---------------------T---------G--C-----------G--------------- 3657
72_BF1.BR.10.10BR_MG002      ------T--------A-C---G-------C---------------T-----------------T-------C---AA------------------G-C-A--------------G-CG---T-----TA--G--C-----------AG--------------- 3501
O.BE.87.ANT70                -------AG-C----A-GT-----T--TA-G--G----------AT----------A-------T--A----A---C-------G-------------AG-ACA---G-A---G--GA-AA-A-AAT--GA---G--C--C-----A--AG----A---ATG--G- 4047
O.CM.98.98CMA104             -------AG------A-GT-----T--TA-G--G-----C-----T----------A-------T--A--T--A--GC------G-------------AG-GCA---G-A---G---A-AA-A-AAT--GA---G--C--C-----A---G----A---ATG---G 3479
O.CM.98.98CMABB141           ---------------A-GT-----T--TA----G----------AT----------A-------T--A--T--A---C------G-------------AG-ACA---G-A---G---A-AA-A-AAT--GAG--G--C--C-----A---G----A---ATG---G 3478
O.CM.98.98CMABB212           ---------------A-C------T--T-----G---C------T----------A------GT----T--A---------G--G----------AG-ACA-----A---G---A-AA-AGAG----AA--G-----C-----A---G----A---ATG--TG 3483
O.CM.98.98CMU5337            ---------------A--T-----T--TA-G--G--------T-AT----------A---------AAA-T--A---C------G-------------AG-ACA---G-A---G---A-AA-A-AAT--GAG--G--C--C-----A---G----A---ATG---G 3480
O.CM.99.99CMU4122            ---------GA-----A-GT-----T--TA-G--G--------A-AT----------A--------T--AA-T--A---C------G-------------AG-ACA---G-A-----------GA-A-AAT--GAG--G--C--C-----A---G----A---ATG---- 3478
O.FR.92.VAU                  --------A-------A--T--G--T--CA----G-----------T----------A--------T--A--C--A---C------G-------A---AG-ACA---G-A---G---A-AA-A-AAT--GA---G--C--C-----A---G----A---ATG---G 3559
O.GA.11.11Gab6352            -------AGA-----A-GT-----T--TCCG--G----------AT----------A----------A--T--A---C----G--G-----G-------AG-ACA-----A---G---A-A--A-AAT-GGAG--G--C--C-----A---G----AC--ATG---- 3204
O.SN.99.99SE_MP1299          ---------------A-GT-----T--TA-G--G----------AT----------A----------A--T--A---C---------C----------AG-A-AG--G-----G---A-AA-T-AAT--GAG--G--C--C-----A-G-G----A---ATG---G 4046
O.US.10.LTNP                 ---------------A-GT-----T--TA----G----------AT----------A-------T-AAACA--A---C------G-------------AG-ACA---G-A---G---A-CA-A-A-T--GA---G--C--C-----A---G----A---ATG---G 3968

N.CM.02.DJO0131              ------T--------A-C---G--A--CACT--G--------A-A----------A-------T-AA--A--A-----G----------T-T---G--AG---------A---G-----G---T-TA-TGAA----C--C-----A-----------------T- 3499
N.CM.02.SJGddd               ------T--------A-C------A--CAGT--G-T------A-A----------A-------T--A--C--A-----G----------T-T---G--AG---A---------G-----G---T-TA-TG-A----C--C-----A---G----A---------T- 3483
N.CM.04.04CM_1015_04         ------T--------A-C---G--A--TAGT--G--------A-A----------A-------T-AA--A--A-----G----------T-T---G--AG-------------G-----G---T-T--TGAAA---C--C-----A-----------------T- 3498
N.CM.06.U14296               ------T--------A-CT--G--A--CAGT--G--------A-A----------A-------T--A--A--A-----G----------T-T---G--AG-------------G-----G--TT-TA-TGYA----C--C-----A---G-------------T- 3496
N.CM.06.U14842               ------T--------A-C---G--A--CTGT--G--------A-A----------A-------T--A--A--A-G---G----------C-T---G--AG---A---------G-----G---T-TA-TG-A----C--C-----A-----------------T- 3500
N.FR.11.N1_FR_2011           ------T-----------C---G--A--C-CT--G--------A-A----------A----------A--A--A-----G----------T-T---G--AG-------------G-----GA--T-TA-TGAAA---C--C-----A---G-------------T- 3394

P.CM.06.U14788               C-----T---C----A-----C------CC----A---------T---------------A----A----C--G----Y----T-------A--AG---G--T--G--G--GA-AA---AA---GAA--R-----T-----A--RG-A--A---G-----G 3464
P.FR.09.RBF168               C-----T---C----A-----C------CC----AT--------T---------------A----A--------------Y----T-------A--AG---G--T--G--R--GA-AA---AA---GAA-GA-----C-----A--AG-A--A---G-----G 4028

CPZ.CD.06.BF1167             -------A-TC-GTTA-C------T--GCC--A------G----AT-----------A-----A--------CT---------G--------C-----G--AG----------A-AGC-GG--AA-AGAAT----A--A--T--T------C--G-A--A-----G--T- 4095
CPZ.CM.05.SIVcpzMT145        -----------------C---G--T----A------------A-A------------------C----A--A--A--AG-----------C-----G--A--C--G--G--------GA--CAT-GT--TGAA--T--C--T-----A--AG-A--A---------G 3535
CPZ.GA.88.GAB1               -------AGT-------C------T---CC-ACCA----T--T-AT-----------A-----A--T-----A--A--ACT--G-----------A--AG-C-A---G-A-------A-CA-T-G-T--GAAA-T--C--T------C-AG-A--A---A-G--TT 4054
CPZ.TZ.06.SIVcpzTAN13        -------A-T--GCTCTCT--C-----TCC--A--------A--T-----------A--T-----A-G---AAG-----------------C--AG-C--G--T---T-T---ACCAGGGAA---GAAA-T--C----------C-AG-A--AC--TGG--TG 3634
CPZ.US.85.US_Marilyn         T-----TA---------C---GG------AG--G------A--T-----G-----A-----CC-AC-A--A--C--GAA------G----------------G---G-T---------A--A-A--TT--GAAA-T--T--T-----A---G-A---C-CAC---GG 4053

GOR.CM.07.SIVgor2139_287     C-----T--A--G--A--G---------CC---------------T----------A--------T--A--A-G---CA-----------------A--AG--AG-T-----G---A-AAGA-AA---GAG-G------T--------AG-A--AC--G-G---G 3530
GOR.CM.07.SIVgorCP2135con    C-----T--A--G--A------------CC---------------T-----C-----A--------T--A--A-GC---CA---------------C--A--AG---G--T-----G---A-AAGG-AA---GAG-G------T--------AG-A--AC--G-G---G 3529
```

```
B.FR.83.HXB2                 TTTATCTAGCTTTGCAGGATTCGGGATTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAACCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAGGAAAAGGTCTATCTGGCATGGGTACCAGCA 4163
Pol                          I__Y__L__A__L__Q__D__S__G__L__E__V__N__I__V__T__D__S__Q__Y__A__L__G__I__I__Q__A__Q__P__D__Q__S__E__S__E__L__V__N__Q__I__I__E__Q__L__I__K__K__E__K__V__Y__L__A__W__V__P__A_
A1.AU.03.PS1044_Day0         -C--------C-----------A----C---------------------------G--------------------G-----------CAG------------A-----T-----G---------A-AC-----GG---A--C--A-----C---T-------------- 3371
A1.CH.03.HIV_CH_BID_V3538    -CC----------A--------A----C---------T-----------T-----G--------------------G-----------CAGG-----------AC--------------------G--C----------------A-----C---T-------------- 3368
A1.CY.08.CY236               -C--------------------A----C---------------------T-----G--------------------G--T--------CA-----------------------------------A--C-----GG---A--C--AA----C---T-------------- 3374
A1.ES.05.X1608_8             -C--------C-----------A----C---------T-----------------G--------------------------G-----CAG------------------T---------------A--C-----G-------C--A-----C---T-------------- 3611
A1.KE.11.DEMA11KE001         -CCTC-----------A-----A----C---------------------------G--------------------GT----------CAGG-----G-----AA--------------------A--C-----GG------C--A-----C---T-------G------ 3502
A1.RU.11.11RU6950            -CC-------------------A----C---------T--------------------------------T-----G-----------CAGG------------A--------A-----------A-AC-----G----A----GA-----C---T-------------G 3687
A1.RW.11.DEMA111RW002        -CCT------C-----------A----C---------T--------------C--G------C-------------G-----------CAGG-----TG-----G--------------------A--C-----G-------C--A-----C--AT-------------- 3514
A1.SN.01.DDI579              --C-------C-----------A----C---------------------------G--------------------------G-----C--G-----C------A----------G--------AA--C-----G----------A--A------T-------------- 3361
A1.UG.11.DEMA110UG001        -C--------C-----A-----A---GC---------T--------T--------G--------G---G-------G-----------CA-G-----C--------------------------AA-AC-----GG------C--A-----C---T----------C--- 3518
A1.ZA.04.04ZASK162B1         -C--------C-----------A---CC---------------------------G--------------------G-----------CAGG------G-----A-------------------AA--C-----G----A----G------C---T-------------- 3604
A2.CM.01.01CM_1445MV         -----T-G-----A--A-----A----C------G--------------T-----G--C-----------------G------------AGG--------------------------------AG-A----------------G-A----C---T-------------- 3362
A2.CY.94.94CY017_41          -C---T-G-----A--------A-----------G--------G-----T-----G--------------------------------AAGG------------A-------------------AA-A------G---------G------C--AT-------------- 3520
B.BR.10.10BR_MG029           -------------------C--A----C------------G--------------------------------------C---------A---------------------G------C------G--C----------------------C------------------ 3625
B.CA.07.502_1191_03          --C-C--------A--------------G------------------------------------------------------------A-G-----G------C------G------------A--------------------------CT----------------- 3590
B.CH.08.M2_0803101_NFLG8     ----------CC----A-----A----C------------------------------------------------T------------A-G--C----------------G---------------------------------A--T--------------------- 3632
B.CN.10.DEMB10CN002          ----------C---------------C----------T-----------------G---------------C-----------------A-G------------------------------------------GGC-----G--------C------------------ 3535
B.ES.10.DEMB10ES002          --C-------------------A--GC-------G------------------------------------C-------G---------AGG--------------------------------------------C-----------------A--------------- 3525
B.FR.11.DEMB11FR001          --C----------------C------C-----------T----------------G--------------T--C---------------A-------------A-----------------------------------------------CT----------------- 3409
B.GB.05.MM45d213_GN1         ----------------------A------A---------------------------------------------------------A---------------------G---------------A-----------------------C------------------ 3564
B.HK.06.HK003                --C----G--------------A-------------------------------------------------------G----------A-------------------T--------------A---------------------------C------------------ 3533
B.HT.05.05HT_129389          -CC-------------------A------------------------------------------------------------------A-G---------------------------------G--------------------A----C------------------ 3364
B.JP.12.DEMB12JP001          -CC--T----------------A--G---------------------------------------------------------------A-G-----------A-------G---------------------------------------C--A--------------- 3550
B.KR.07.HP_18_07JHS10_3909   ------------------------A--C----------------------------------A--------------------------A-------------------------------------------------------------C------------------ 3827
B.PE.07.502_0525_wg5         -------------A--------------------------------------C-----------G--------------C---------A-G-----------T--------C---------------------------A---G------C-------------G--- 3628
B.RU.11.11RU21n              --C-------------------A--------------------------------------------G---------------------A---------------------G---G--------------G--------------A-----C------------------ 3734
B.TH.08.MERLBDTRC10          --A----------A--A-----T------------------------------------------------------------------A-G-------------------G---------------------------------A-----C----------G------ 3396
B.US.11.ES38                 --A-------C-----------A----C------------------------T-----------------------------G------A---------------------G-A---------------------G------G--------C--A--------------- 4098
C.AR.01.ARG4006              --C-G--------A--A-----A--------------------------------G--------------T------------------A-G--------------------------------A--A------C----------------C---T-------------- 3347
C.BR.07.DEMC07BR003          --C-G--------A--------A--C-C---------------------------G--------G-----T------------------A--------------C-------------------A-------------------G------C---T-------------- 3495
C.BW.00.00BW5031_1           -CC-G-----------A--C--A----C---------------------------G-----------------C-----G---------A-G--------------------------------A-----------G--A----G-A----C---T-------------- 3534
C.CN.10.YNFL19               ----------------A-----A----C---------------------T-----G-----------G---------------------A-G-----------------T--C-----------A--A-----------------------C---T-------------- 3516
C.CY.09.CY260                ----C-----------A-----A---CC---------------------T-----G------C------------------------A-G----------GA-------G------------A--A------C---------G----------T-------------- 3386
C.ES.08.X2363_2              --C-GT-----C----A-----A---CC---------------------------G-----------------A--C------------A-G-----------------T--------------AA----------C-------GA-----C--AT-------------- 3594
C.ET.02.02ET_288             --C-G-----AC-A--------A----C---------------------------G-----------G--------------------CA-G--------C---G-------------------AG-A----------------G------C---T-------------- 3365
C.IN.09.T125_2139            ---G------------A-----A----C---------------------T-----G------C----G---------------------A-------G-----------T-GC-----------A--A----------------G------C---T-------------- 4139
C.KE.00.KER2010              --C-G-----------A-----A----C---------------------------G---------------------------------A-G-----------A--------C-----------A--A-------G---------------C---T-------------- 3344
C.MW.09.703010256_CH256.w96  --C-G-----------A--C--A----C---------T-----------------G---------------------------------A-G--------------------------------A---C---------------GAA----C---T-------------- 4143
C.TZ.08.707010457_CH457.w8   --C-G-----C--------C--A----C---------------------T-----G--------------------------T------A-G------------A----T--------------A-----------C-------GA----------T-------------- 4137
C.YE.02.02YE511              --C-------------A-----A--G-C------T--------------------G-----------------C---------------A-G--------------------------------A--A-----------------------C---T-------------- 3338
C.ZA.08.705010534_CH534.w12  --C-G-----------------A----C---------------------------G---------------------------------A-G------------A----T--C-----------A-----------C-------GAAC---C---T-------------- 4153
C.ZA.10.DEMC10ZA001          --C-GA-----C----A-----A---CC---------T-----------------G---------------------------------A-G-----------------T-----G--------AG------------------G------C---T----------C--- 3503
C.ZM.11.DEMC11ZM006          --C-G-----------------A----C---------------------------G---------------------------------A-G-----G-----A-----T--C-----------A----------G---A----G----------T-------------- 3512
D.CM.10.DEMD10CM009          --A-------C---------------CC------G--------------T---------------------------------------A-G-----------A-------G------------A--------------------------------------------- 3538
D.CY.06.CY163                --A----------A--------A---CC------G--------------T---------------------------------------A-G-----------A-------G--T---------A-----G--------------------C------------------ 3371
D.KE.11.DEMD11KE003          --A--A-G-----------------------G--------------------T------------------------------------A-G-------------------G---------------------------------A--T--C--AT-------------- 3513
D.KR.04.04KBH8               --A----------------------------------------------T-----------C-----------C---------------A-G-------------------G------------A--------------------------C-----C------------ 4122
D.SN.90.SE365                --A-------------A----------C---------------C-----T-----------------------------C--G------A-------------AG------G------C------------------------------C------------------ 4180
D.TZ.01.A280                 -AA-C---------------------C--------------A-------------------------G---------------------A---------------------G------------A-----------G-----------T--C--A-----------G--- 3364
D.UG.10.DEMD10UG004          --A------------------------C---------------------------------------------------------------G-----C--------------------------A--------------------------C--AT-------------- 3501
D.UG.11.DEMD11UG003          --A----------------C--A--------------T-----G-----------------------G--T------------------A-G-----------------T-----------------------C-----A-----A-----C--AT-------------- 3518
D.YE.02.02YE516              -CA---G---------------A----C------G--T-----------T---------------------------------------A-G-----------AG------G------------A--------------------A-----C------------------ 3374
D.ZA.90.R1                   --A---G------------------------------------------T--------------------------------G------A---------------------G----------------C---------------G---T--C------------------ 3509
F1.AO.06.AO_06_ANG32         -CC--T-------A-------------C------C--T-----------------G-----C---------------------------A-------------------------G-----------A------C---------G------C---T-------------- 3374
F1.AR.02.ARE933              --C-GT-------A--------A---GC------G--------------------G---------------------------------A-G-----------A-----A------------------------------------------------------------ 3457
F1.BR.10.10BR_RJ015          --C-GT-------A--------A----C------G--------------------G-----------------------G---------A-G-----------A-----------------------A-----------------------C---T-------------- 3636
F1.CY.08.CY222               --C--T-------A--------A----C------G--------------------G---------------------------------A-G-------------------G-------------AG-------C----------------C---T-------------- 3356
F1.ES.02.ES_X845_4           --C--T-------A--------A---CC---------T-----------------G---------------------------------A-------------------------G-----------A------C----AAC--G-A----C--AT-------------- 3615
F1.RO.96.BCI_R07             --C--T-------A--------A----C------G-----------------G------------------------------------A-G-----------------------------------A------C----A----G-A----C---T-------------- 4187
F1.RU.08.D88_845             -----T-------A--------A----C------G-----------------M--G---------------------------------A-G------------------------------------------C---------G------C---T-------------- 3622
F2.CM.02.02CM_0016BBY        --C--T-------------C--A----C---------------------------G--------------------------C------A-G--------------------C--------------A------C----------------C---T-------------- 3353
F2.CM.10.DEMF210CM001        ---G-T----------A-----A----C---------------------------G-----------G--------G-----C-----CA-G--------------------C-----------A--A------C-G--------------CT-AT-------------- 3427
F2.CM.10.DEMF210CM007        --C-GT-------------C--A---CC---------------------------G-------------------------TC------A-G--------------------C--------------A------C----------------C---T-------------- 3510
G.BE.96.DRCBL                --C-G------C-------C--ACA--C---------------------------G--------------------------G------AGG-----------AA-------------------A---C----------------------C---T-------------- 4120
G.CM.10.DEMG10CM008          --A-------------A--C--A--G-----------------------------G---------------------------------A-G-----------AA----------G--------A--AC------G---------------C---T-------------- 3544
G.CN.08.GX_2084_08           --C-G--------------C--A----C------G--------------------G---------------------------------AGG-----------A---------A-W--------A---C-----M-----------A----CT----------------- 3403
G.CU.99.Cu74                 --C-G--------------C--A----C---------------------------G---------------------------------AGG-----------A--------------------A--AC----------------------C---T-------------- 3767
G.ES.09.X2634_2              --C-AA-------------C--A----C---------------------------G--------------T------------------AGG-----------A-----------G--------A---C----------------------C---T-------------- 3646
G.GH.03.03GH175G             -CC-AT-------------C--AAA--C------G--------------------G---------------------------------AGG-----------A-----T--------------A---C----------------------C---T-------------- 4203
G.KE.09.DEMG09KE001          --C-G-----------------A----C---G-----------------------G---------------------------------AGG-----C--------G------A-G--------A---C--------------------------T-------G------ 3538
G.NG.09.09NG_SC62            -CC-GT-------A-----C--A---CC---------------------------G---------------------------------AGG-----------AA-------------------A---C----------------------CT--A-------------- 3355
G.ZA.01.TV546                --C-G--------------C--A----C---G-----------------------G------------------------ATC------AGG-----------A--------------------A---C-------T--------------C---T-------------- 3365
H.BE.93.VI991                ----------------A--G--A--GCC---------------G--------------------------------G--------T---A-G-----------AC----T---------------G-A------------------T----CT--T-------------- 3552
H.CF.90.056                  ----------------A-----A--G--------G--------G-----T-----G------C----------------------C---A-G-----------------T---------------G-A--------G--------------C---T-------------- 3510
H.GB.00.00GBAC4001           ----------C-----A--C--A--GC----------------------T-----G--------------T--------------T---A-G-----------A-----T-----G---------G-A--G-----G--------------C---T-------------- 3681
J.CM.04.04CMU11421           ---TAT-------AM-------A--GK------C---------R-----T--C-----------R--R--T-----------R------A-G-----------A---------------------G-A------M-G--R----GR-----C---T----------R--- 3669
J.SE.93.SE9280_7887          -----T-------A-G------A--GC----------------------T--------------G--G--T--A---------------A-G-----------A---------------------G----------G--------------C---T-G------------ 3477
J.SE.94.SE9173_7022          -----T-------A--------A--GA----------------------T--------------G--G--T------------------A-G-----------A---------------------G----------G--------------C---T-G------------ 3478
K.CD.97.97ZR_EQTB11          -CC--T-------A--------A----C---------------------------G---------------------------------A-------------A-----T-----------------A--------------C-G------C--AT-------------- 3359
K.CM.96.96CM_MP535           -C-G-T-------A--------A----C---------------------------G--------------------------G------A-------------T-----T-----------------A----------------G-A----C--AT-------------- 3359
U.CA.01.TV749                -C---T----C-----------A----C---G-----------------------G------A-------------G-----------CAGG--------T--------T-----G--------AA--C-----G-----------A-----A--T-------R------ 4180
U.CA.99.TV721                -C--------C-----A-----A----C---G-----------------T-----G------A-------------G-----------CAGG------------A--------TGG--------AA--C-----G-----------A--------T-------------- 4179
U.CD.83.83CD003_Z3           -------C--A-----------A----------C---------------T-----G---------------------------------AGC--------------------------------A--------------------A-----C---T-------G------ 3658
U.CD.90.90CD121E12           -------T--A-----------A----------C---------------T-----G---------------------------------AGC--------------------------------A--------------------A-----C---T-------G------ 3362
U.CY.05.CY090                ----CT----A-----------A----C---G--G--------G-----T-----G------C-------T-----G-----C------A-G-----------A-----------------------AC-------------C--------C--AT-C------------ 3362
U.CY.08.CY223                -C---T----C--A--------A----C---------T-----------T------------C-------------G--T--------CAGG-----G-------------G------------A---C-------G------------------T-------------- 3380
U.ES.10.DEURF10DZ001         -AC-------------------A--------------------------------G------A----------------C---------A------------CA-------G------------AA--------C-----A-C--A-----C---T-------------- 3439
U.GR.99.99GR303              -CC-CT-------A--------A----C---------------T-----------G--------------------G-----------CACG-----G-----A---------------------A--C-----GG---A----G---T--C---T-------------- 3434
U.NL.95.U_NL_95_H10986_D1    --C-AT----------------A----CT-C------------------------G--------------------------G------A-C-----------T-----T-----------------A------C-G-----CC-------C---T-------------- 4048
01_AE.AF.07.569M             -CC--T----C-----------A----C---------T--------------------------------------------------CAGG-----------AG-------C------------G--C----------A-----------C---T-------------- 3369
01_AE.CF.90.90CF11697        -CC--T----C-----------A----C---T-----T--------------C---C-------------------G-----------CAGG-----------AA-------C-----------AA--C-----G---------GA-----C---T-------------- 4112
01_AE.CN.10.YNFL03           -CC-AT----C-----------A----C---------T--------------------------------------G-----------CAGG-----------AG-------C------------G--C----------------A-----C---T----------G--- 3541
01_AE.HK.04.HK001            -C---T----C-----A-----A----C---------T--------------------------------------G-----------CAGG-----------AA----T-GC------------G--C----------------A-----C---T-------------- 3525
01_AE.IR.10.10IR.THR48F      -CC--T----C--------C--A----C---------T---------------------------------C----G-----------CAGG-----------AG----A--C------------G--C----------------------C---T-C------------ 3353
01_AE.JP.x.JRC77AE           -CC-GT----CC----A-----A---CC---------T--------------------------------------G-----------CAGG-----------AG-------C------------G--C------G---------A-----C---T-C------------ 4160
01_AE.TH.09.AA090a_WG11      -CC-G-----C-----A---Y-A--------------T--------------------------------------G-----------CAGG-----------AG-------C-----------AG--C-------------C--A-----C---T-------------- 3586
01_AE.TH.90.CM240            -CC--T----C-----------A----C---------T--------------------------------------G-----------CAGG-----------AG-------C------------G--C-------------G--A-----C---T-------------- 3737
01_AE.US.05.306163_FL        -CC-AT----C-----------A----C---------T--------------------------------------G-----------CAGG-----T-----AG----------G---------G-AC----------------------C--AT-------------- 3354
01_AE.VN.98.98VNND15         -CC--T----C-----------A----C---------T--------------------------------------G-----------CAG------------AG-------C------------G--C----------------A----------T-------------- 3418
```

```
B.FR.83.HXB2                  TTTATCTAGCTTTGCAGGATTCGGGATTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAACCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAGGAAAAGGTCTATCTGGCATGGGTACCAGCA 4163
Pol                           I__Y__L__A__L__Q__D__S__G__L__E__V__N__I__V__T__D__S__Q__Y__A__L__G__I__I__Q__A__Q__P__D__Q__S__E__S__E__L__V__N__Q__I__I__E__Q__L__I__K__K__E__K__V__Y__L__A__W__V__P__A_
02_AG.CM.10.DE00210CM013      --C-------C-----------A----C---------T-----------------G--------------------G-----------CAGG--------------------------------AA--C-----G-G--A-----A-----C---T-------------- 3530
02_AG.CY.09.CY256             --C-------C-----------A----C---------T-----------------G--------------------G-----------CAGG-----------A--------------------A---C-----GG------C--A-----C---T-------------- 3359
02_AG.ES.06.P1423             -Y--------C-----------A----C---------T-----------Y-----G--------------------G-----------CAGG-----C----------------T----------G--C-------------C--A-----C---T-------------- 3600
02_AG.GW.05.CC_0048           --C-------C-----------A----C---------T-----------------G--------------T-----G-----------CAGG------------A--------------------G--C------G------T--AA--------T-------------- 3513
02_AG.KR.12.12MHR9            -----T----C-----------A----C------G--T--------------G--G--------------------G------------AGG-----------A----------T----------A--C-----G-------C--A-----C---T-------------- 3822
02_AG.LR.x.POC44951           --C-------C-----------A----C---------T-----------------G--------------T-----G-----------CAGG---------------------------------G--------GG------C------------T-------------- 4158
02_AG.NG.09.09NG_SC61         --C-------C---------G-T----C---------T-----------------G--------------------G-----------CAG----------------------------------G--C-------------C-GA--A--C---T-------------- 3355
02_AG.NG.x.IBNG               --C--T----C--A--------A----C---------T-----------------G--------------------G-----------CAGG---------------------------------A--C-------------C--A-----C---T-------------- 3688
02_AG.SE.94.SE7812            --C--T----C-----------A----C---------T-----------------G--------------------G-----------CAGG------------C------C-------------G--C-------------T-GA-----C---T-------------- 3535
02_AG.SN.98.98SE_MP1211       --CT------C--A-----A--A----C---------T-----------------G--------------------G-----------CAGG---------------------------------A--C-----G-------C--A-----C---T-------------- 3360
03_AB.RU.97.KAL153_2          --C--------------------------------------------------------------------------------------A-G-------------------G------------------G--------------------C------------------ 3389
04_cpx.CY.94.94CY032_3        ----CT----------------A----C---------------------------G------A--------------------------AG------------T-----T------------------------CGG-----C--------C---T-------------- 3529
05_DF.BE.x.VI1310             --C-C-----------A-----A-----------G--T-----------------G--C------------------------------A-G-----------------------------------A------G-----A--------------T-------------- 3545
06_cpx.AU.96.BFP90            --A-------A--A--------A---CC---------------------------G---------------------------------A-------------AC-------C-----------A---C----------------AA----C---T-------------- 4191
07_BC.CN.98.98CN009           ---G-A----------A-----A----C---------------------T-----G-----------G---------------------A-G-----------------T--C-----------A--A----------------GA-----C---T-------------- 3495
08_BC.CN.97.97CNGX_6F         -----A----------A-----A----C---------------------T-----G-----------G---------------------A-G-----------------T--C-----------A--A----------------G------C---T-------------- 3343
09_cpx.GH.96.96GH2911         -AC-------C--A--------A----C---------T--------------------------------------G-----------CAGG---------------------------------A--C-----G----A-----------C---T-------------- 3365
10_CD.TZ.96.96TZ_BF061        --A-------------------A----C---------T-----------------G-----------C--T--------------C---A-G-----------------------------------------------------A-----C--AC-------------- 3542
11_cpx.CM.95.95CM_1816        --C-------------------A---C----------T-----T-----T--------------G--G--T------------------A-G-----------A-------G---------------------------------------C---T-------------- 3527
12_BF.AR.99.ARMA159           ----CT-------A--------A----C---------------------------G--------------T------------------A-G-----G-------------G---------------A-----------------------C---T-------------- 4168
13_cpx.CM.96.96CM_1849        --C-G-----------------A----C---------------------------G---------------------------------AGG--------------------------------A---C-------G--------------C---T-T------------ 3568
14_BG.ES.05.X1870             --C-AA----Y--------C--A----C---------------------------G-----------G--T-----G------------AGG-----------A--------------------A---C----------------------C------------------ 3630
15_01B.TH.99.99TH_MU2079      -CC-GT----C----G------A----C---------T--------------------------------------G-----------CAGG-----------AG-------C------------G--C----------------------C---T-------------- 3562
16_A2D.KR.97.97KR004          -C---T-G-----A--------A--------------------------T-----G--------------------G------------AGG------------A---------T---------AA--------G----A----GA--------AT-------------- 3526
17_BF.AR.99.ARMA038           -CA-GT----C--A--------A----C---------------------------G---------------------------------A-G-----G------A----------------------A-----------------------C---T-------------- 3377
18_cpx.CU.99.CU76             --C-CT----A--------------GCAG--------------G-----T-----G-----------------------------T---A-------------AA--------------------G----------G--------------C---T-------------- 3468
19_cpx.CU.99.CU7              --A----------------------------------------------T--------------------T------------------A-G-------------------G------------------------G-------G---A--C------------------ 3389
20_BG.CU.99.Cu103             --C----------------------------------------G---------------------------------------------A---------------------G------------------------------------T--C--A--------------- 3629
21_A2D.KE.99.KER2003          -C---T-G-----A--------A-----------G--------------T-----G--------------------G-----G------AGG--------------------------------A---------GG---AA---G----------T-------------- 3368
22_01A1.CM.01.01CM_0001BBY    -CC--T----C-----------A----C---------T--------------------------------------G-----------CAGG-----------A----------T---------AA-AC-----G----------A-----C---T-------------- 3371
23_BG.CU.03.CB118             --C----------------C--A----C---------------G---------------------------------------------A---------------------G---------------A--------------------T--C--A--------------- 3606
24_BG.ES.08.X2456_2           --C-------------------A--------------------G--------------C-----------M------------------A-------G------G------G---------------------------------A--T--C--A--------G------ 3614
25_cpx.CM.02.1918LE           -CC-------C-----------A----C---------T-----G-----T-----G--------------T------------------A-G------------A----------------------------------------------C---T-------------- 3371
26_AU.CD.02.02CD_MBTB047      -A--------------A-----A---C----------T-----------------G--------------T-----G-----------CAGG-----------AA----T--------------AA-A------G-G--------A--G--C---T-------------- 4168
27_cpx.FR.04.04CD_FR_KZS      --C-G--------A--------A----C---------------------------G-------------------------GG------A-------------------------------------A--------G--------------C---T-------------- 4154
28_BF.BR.99.BREPM12609        -CC-------------------A--------G-----------------------------------------C-----C---------A---------------------G---------------------------------------------------------- 3529
29_BF.BR.01.BREPM16704        --C-GT-G--C--A-------------C-------------------------------------------------------------A-G------------C----------------------A--G-----C--A-----A-----C---T-------------- 3564
31_BC.BR.04.04BR142           -CC-AT-------A--------A----C---------------------------G--------------T------------------A----------------------------------A-T---------C-----G-G------C---T-------------- 3635
32_06A1.EE.01.EE0369          --A----------A--------A----C---------------------------G---------------------------------A-------------A------------------------C-------------G------------T-------------- 3798
33_01B.ID.07.JKT189_C         -CC--T----C-----------A----C------------------------------------------T-----G-----------CAGG-----------AG-------C------------G--C-------------C--------C---T-------------G 3464
34_01B.TH.99.OUR2478P         -CC--T----C-----------A----C---------T--------------------------G-----------G-----------CA-------------AG-------C------------G-AC----------------A-----C---T-------------- 3365
35_AD.AF.07.169H              -CC-------C-----------A----C---------T--------------------------------------G-----------CA-------G-----A---------A-----------A-AC--G--G-------C--A-----C--AT-------------- 3363
36_cpx.CM.00.00CMNYU830       --C--T----C-----------A----C-A-------T-----------------G--------------------G-----------CAGG-----------------T---------------A--C-----G-------T--A-----C---T-------------- 3368
37_cpx.CM.00.00CMNYU926       --C-------CC----------A-A--C------G--T-----------------G--------------------G-----------CAG------C-----AA-------C--G---------A-AA-T---G-------G-GA-----C---T-------------- 3353
38_BF1.UY.03.UY03_3389        --AGCT-------A--------A----C---G--G--------------------G------C--------------------------A-G-----------------------------------A--------G--------------C---T-------------- 3586
39_BF.BR.04.04BRRJ179         ----------------------A----C---G------------------------------------------------A--------A-G-------------------G---G---------G-A------------------A----C--AA-------------- 3640
40_BF.BR.05.05BRRJ055         --C-------------------------------------------------------------------------------G------A-G-------------------G---------------------------------------C--A-----------G--- 3646
42_BF.LU.06.luBF_18_06        --C-GT-----C----------A--------------------------T-----G-----------G---------------------A-G------------A----------------------A-----------------------C---T-------------- 3717
43_02G.SA.03.J11223           ----C-----C--------C--A----C---------T--------------C--G--------------------G-----------CAG------------A---------------------A--C-----G----A--C--A-----C---T-------------- 3688
44_BF.CL.00.CH80              --C-G------C-A--------A---CC------G--------------------G---------------------------------A-G----------G------------------------A------G---------G------C---T-------------- 3605
45_cpx.FR.04.04FR_AUK         --C-CT-------A--------A----C---G-----------------------G---------------------------------A-------------T-----T-----------------A-------TG-----C--------C---T-G------------ 4149
46_BF.BR.07.07BR_FPS625       --C-GT-------A--------A----C---------------------------G-----------------------------------G-----------C----------------------TA-----------------------C------------------ 3601
47_BF.ES.08.P1942             -CCT---------A--------A--------------------------------G---------------------------------A-G-----------A-----------------------A--------C--------------C---T-------------- 3602
48_01B.MY.07.07MYKT021        -CC--T----C-----------A----C------------------------G--------------G--------G-----------CAGG-----------AG-------C------------G----------------------T--C---T-------------- 3368
49_cpx.GM.03.N26677           -C--------C-----------A----C---------T-----------T-----G---------------------------------A-G-----------A--------------------A--A----------------G------C---T-------------- 3585
50_A1D.GB.10.12792            -CC-------------------AA-------------T-----C-----------G--------------------G----T------CA-G----------GA---------------------A--C-----G----A--C--A-----C-----------G------ 3570
51_01B.SG.11.11SG_HM021       --C-------------------A--------------------------T-----------------------------------------G---------------------A----------------G---C-------------T--CT-A-----------G--- 3370
52_01B.MY.03.03MYKL018_1      -CC--T----A-----------A----C---------T--------------------------------------G-----------CAGG-----------AG-------C------------G-AC----------------------C---T-------------- 3477
53_01B.MY.11.11FIR164         -CC--T---------------------C---------T--------------------------------------G-----------CA-G-----------AG-------C------------G-------------------------C---T-------R------ 3503
54_01B.MY.09.09MYSB023        --C--T-------A--------A--------------T---------------------------------------------------A-G-------------------G------------A---------C---------GA-----C--A--------G------ 3565
55_01B.CN.10.HNCS102056       -----------C-A-------------C------------------------G------------------------------------A--------------C----T------------------A-G----GC-------G------C------------------ 3473
56_cpx.FR.10.URF5_patient_A   ---------------------------C---------T-----------------G--------------------G------------A--------------G----T-G---------------------------------A-----C---T-------------- 3514
57_BC.CN.09.09YNLX19sg        -----A----------A-----A----C---------T-----------T-----------------G---------------------A-G-----------C-----T--C-----------A--A--------G-------G------CT--T-------------- 3347
58_01B.MY.09.09MYPR37         -CC--T----C-----------A----C---------T--------------------------------------G-----------CAGG-----------AG----Y--C------------G--C----------------A-----C---T-------------- 3508
59_01B.CN.09.09LNA423         -CC--T----C-----------A----C---------T-----------------------G--------------G-----------CAGG-----------AG-------C------------G--C-------------C--------C--AT-------------- 3347
60_BC.IT.11.BAV499            --C-G--------A--------A----C------------------------T--G--------------T---------A--------A----------------G--T---------G--------C-------------CG----T--C---T----------G--- 4120
61_BC.CN.10.JL100010          ---G-A----------A-----A---GC---------------------T-----G-----------G---------------------A-G-----------------T--C-----------A--A--------------G--------C---A-------------- 3584
62_BC.CN.10.YNFL13            -----A----------A-----A----C---------------------T-----G-----------G---------------------A-------------------T--C-----------A-------------------GA-----C---T-------------- 3492
63_02A1.RU.10.10RU6637        --C-------------------AA---C---------T-----------------G-----------------C--G-----------CAGG------------A--------A-----------A-AC-----G----A-----A-----C---T-------------- 3704
64_BC.CN.09.YNFL31            -A---A----------A-----A----C---------------------T-----G------C----G---------------------A-------------------T--C-----------A--A-------G------G-G------C---T-------------- 3517
65_cpx.CN.10.YNFL01           --C--T----------------A---G----------------------------------------G---------------------A-G-----------------G--C-----------A--A----------------G------C---T-------------- 3534
67_01B.CN.11.ANHUI_HF115      --C--T----A-----------A--------------------------------------------G-----C--T------------A-G-------------------G---------------------------------A------T----------------- 3506
68_01B.CN.11.ANHUI_WH73       --C--T----A-----------A--------------------------------------------G-----A--T------------AGG-----------------------------------------------------A------T----------------- 3497
70_BF1.BR.10.10BR_PE004       -CC-GT-------A---------------------------------------------------------------------------A-------------A-------G------------------G--------------------C------------------ 3619
71_BF1.BR.10.10BR_PE008       --C-A------C-A--------A----C------------------G------------------------------------------A-------G--------------------------------G-----------G-G-A-A--C--AA-------------- 3827
72_BF1.BR.10.10BR_MG002       -AC-GT-------A--------A----C-AC---G--------------------------------------C---------------A-------------AA--------------------G-A-----------------------C---T-------G------ 3671
O.BE.87.ANT70                 -A-TA-----C--A--------CAA-GA-AC------T-----------T--------------G--CG----CTCCT-C-----TACA--G--------CCCTA----TC-G------------G-AC---C----------C----G-----TA-------T--T--C 4217
O.CM.98.98CMA104              -ACTAT----C--A--------CAAGGA-A-------T-----------T----------T---G--C-----CTCCT-T-----TACA--G---------CCTA----C-A---G--------AG-AC---C---G------C-A--G--C--TA-------T--T--C 3649
O.CM.98.98CMABB141            -A-TAT----C--A--------CAA-GA-A-GA----T-----------T--------------G--C-----CTC-T-------TACA--G--------CCCTA----TC-A--G---------G-AC---C---------------G-----TA-------T--T--T 3648
O.CM.98.98CMABB212            -ACTGT----C--A--------CAA-GAGA------------------------------T---G--C-----ATCCT-C------ACA--------C--CCCTA----TC-G--G--------AG-AC---C----------C-A--G--C--TA-------C--C--T 3653
O.CM.98.98CMU5337             -A-TAT----CC-A--------CAA-GARR-------T-----------T--------------G--C------TCCT-------TACA-----------CCCTA----TC-G--G---------G-AC---C--G-------CG---G-----T--------T--T--T 3650
O.CM.99.99CMU4122             -A-TRT----C--AA-------CAAGGA-AC------------------T-------T------G--C------TCCT-C-----TACG--G-----T--CCCTA----TC-G--G---------G-AC---C----------C----A-----T--------T--T--T 3648
O.FR.92.VAU                   -A-TGT-------A--------TAA-GA-A-------------------T----------T-C-G--C-----ATCTT-C-----TACA--G-----C--CCCTA----TC-G--G--------AG-AC-G-C-----G----C-A--G-----TA-------T--T--T 3729
O.GA.11.11Gab6352             -G-TAT----C--A--------CAA-GA-AC------T-----------T---------A----G--C------TCCT-C-----TAC---G-----G--CCCTA----TC-G--G---------G-AC---C---------C-----G-----TA-------T--T--T 3374
O.SN.99.99SE_MP1299           -A-TAT----C--A--------CAAGGA-AC------T------CA---T--C-------T---G--C-----CTCCT----G--TACA--G--------CTCTC----TC-G--G---------G-AC---C---------C-----G-----TA-------T--T--C 4216
O.US.10.LTNP                  -G-TAT----C--A--A-----CAA-GA-A-G-----T-----G-----T----------T---G--C-----CTCCT-T-----TACA--G--------CCCTA----TC-G--G---------G-AC---C--G------C-----A-----CA-------T--T--T 4138
N.CM.02.DJO0131               -CCT---G--C--A--A-----A---CAG--------------C--T--T--T--G-----TC-G-----A-----CT----C------A-------------A--G--G-GC-----------AG--C-C-------------GA--T-----CT-T--------T--- 3669
N.CM.02.SJGddd                -CCT---G--C--A--A-----A---CAG--------------T--T-----T--G-----TA-G-----A--A--CT-----------A-------------A--G--G-GC-----------AG-AC-C-------------GA--T-----CT-T--------T--- 3653
N.CM.04.04CM_1015_04          -CCT-A-G--C--A--A---K-A---CAG--------------C--T--T--T--G-----TA-G-----A-----CT-----------A-G-----------A--G--G-GC-----------AG-AC-C-------------GA--------CT-T-----R--T--- 3668
N.CM.06.U14296                -CCT-A-G--Y--A--A--G--A-A-AAG--------------C--T-----T--G-----TA-G-----A-----C------------A-------------AG-G--G-GC-----------AG-AC-C-------------G---------YT-T--------T--- 3666
N.CM.06.U14842                -CCT-A-G--C--A--A-----A---CAG--T-----------C--T-----T--G-----TA-G-----A-----CT-T---------A-------------A--G--G-GC-----------AG-AC-C-------------GA--------CT-T--------T--- 3670
N.FR.11.N1_FR_2011            -CCT-A-G--C--A--A--G--A---CGG----C---------C--T-----T--G-----TA-G-----A-----TT-----------A-------------A--G--G-GC-----------AG-AC---------------GA--T-----CT-T--------T--- 3564
P.CM.06.U14788                -G-TGT----C--AA-A--A--A---GA-C-G-CT-----------------C-------TG-----R--T--CTC----ACT--------------C--CCCCY----GC-AA----------AG-AA-G-C---------------A--C--RT-------------- 3634
P.FR.09.RBF168                -GCTGT----C--AA-A--A--A---AA-C-G-CT-----------------C-------TR-----G-----CTC----ACT--------------C--CCCCC----GC-AA----------AG-AA-G-C---R--------A--A--C--AT-------------- 4198
CPZ.CD.06.BF1167              -ACTG--G--A--AA-A---G-AAA-GA-A-G--C--------G-----T--C-------TGC--------T-A----G-ACT--T--C--------C--CCCCC----AG-GG--------C-AG-A-----G--T--A---GCA--T---T-AAAT--------T--T 4265
CPZ.CM.05.SIVcpzMT145         ---TAA-G--C--A-----C--A---CAG-----C--TG-G--C--T-----C-------TGC-G------T-A---T----C-----C--C-----------TA----A-----G--------AG------------------AA--T-----AT-------------- 3705
CPZ.GA.88.GAB1                -GCT---G--C-----------A-ATCA-C-G--T--------G--T-------------TG-----G--T-----GT----G--------C-----------A-----------G--------AG-------T-----A----AA-----C--CT-C------------ 4224
CPZ.TZ.06.SIVcpzTAN13         -AG--T-----C-AA-------A--G-C-C----------------------C--G----TCA-G--GG-AC-G----GGATG--T-----------C--CCCCA----TG-GG----C---C-AA-AC-C--GC-G--AA-TGCA--G-----A-----------T--T 3804
CPZ.US.85.US_Marilyn          -A--CT----A--AA-A-----A-AGAATAC---T--TG----C--T-G---C-------TC-----------C--CT-T--G-----C--------------A--G--------G--------AG-A--------G--A----AAGT----A-CT-------------- 4223
GOR.CM.07.SIVgor2139_287      -CCTCT-G--C--AG-A--A--A---AA-A-T--T--------G--------C--G----T-C----G--T--GC-----AGT--T-----------C--CCCC-----GC-GA-----------G-AA-G-C--CG--------A------T-A--T-----------T 3700
GOR.CM.07.SIVgorCP2135con     -CCTCT-G--C---G-A--A--A--GAA-A-T--T--T-----G-----T--C--G----T-C----G--T--GC-----AGT--T-----------C--CCCC-----GC-GA-----------G-AA-G-C--C---A------------T-A--T-----------T 3699
```

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

Pol p66 RT, Pol p15 Rnase H end Pol p31 Integrase start

```
B.FR.83.HXB2                CACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTTAGATGGAATAGATAAGGCCCAAGAT.GAACATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAGTGATTTTAACCTGCCACCTGTAGTAGCAAAAG 4332
Pol                         _H__K__G__I__G__G__N__E__Q__V__D__K__L__V__S__A__G__I__R__K__V__L__F__L__D__G__I__D__K__A__Q__D_#_E__H__E__K__Y__H__S__N__W__R__A__M__A__S__D__F__N__L__P__P__V__V__A__K__
02_AG.CM.10.DE00210CM013    --------G-------------------------C--------AG-------------G----------------C--------A--------A.--G-------G-----------------------------------T---------A----------- 3699
02_AG.CY.09.CY256           --------------G-------------------------------AG--G--------G----------------C--------A--A-----A.--T----ACG---------C---------------G---------T---------A----------- 3528
02_AG.ES.06.P1423           --------------------------------------------A-------------------------------C--------A--------A.--G----A----------AC-----------------------------T--A------A----------- 3769
02_AG.GW.05.CC_0048         --------G-----------------------------------AG-----------G------------------C--------A--------A.--G----A-G-----T--C--------------------------------------A----------- 3682
02_AG.KR.12.12MHR9          --------G-----------------A-----------------AG------------------------------C--------A--------A.--G----A-G---------C-----------------------------T---------A----------- 3991
02_AG.LR.x.POC44951         --------G---------------G-------------------AA------------G-----------------C--------A--------A.--T----A-G---------C-----------------A-------C--T---------A----------- 4327
02_AG.NG.09.09NG_SC61       --------G-----------------------------------AG------------------------------T--------A--------A.--G----A-G---------C------------------------------T---------R----G------- 3524
02_AG.NG.x.IBNG             --------G-----------------------------------AA------------G-----------------C--------A--------A.--G----A-G---------C-------AG-----------------------T---------A----------- 3857
02_AG.SE.94.SE7812          --------G----------------G--G---------------AG------------G-----------------C-------------A.--G----A-G---------C---------------------C-----T------------------------- 3704
02_AG.SN.98.98SE_MP1211     --------G-----------------------------------AAG-----------------------------C--------A--------A.--G----A-G---------C---------------------A--C--T---------A----------- 3529
03_AB.RU.97.KAL153_2        ---------------------------T-----------------------------G------------------------------A-----A.--------------------G-----------------------------------G---------- 3558
04_cpx.CY.94.94CY032_3      --------G--------------------------------CAA---------A--G--G-----------------------------T----A.---------------------A---C---------------------T------T-A--G---------- 3698
05_DF.BE.x.VI1310           --------G-------------------------------------------T--A-------------G---------------A--T--G---.------------------T--C------------------------------T---------------------- 3714
06_cpx.AU.96.BFP90          --------G-----------------------------------AA------------G-----------------C--------A----G--A.--T----A-G---------C----------C------A---------T---------A----------- 4360
07_BC.CN.98.98CN009         --T-------------------------------A---AG--------------G-----C----------------A--T----A.--G----A-GG---------C----------------------C-----T---------C------------ 3664
08_BC.CN.97.97CNGX_6F       --T-------------------------------A---AA--C-----------G-----C----------------A--T----A.--G----A--G---------C----------------------C-----T--------CA------------ 3512
09_cpx.GH.96.96GH2911       --T-----------------------------------AG--------A-----------------------C--------A--T----A.--------A-G-----T--C---------------------------T------------------------ 3534
10_CD.TZ.96.96TZ_BF061      ---------------------------G------------A--AGG-------A--------G--C----------------A--T----A.--G-----------C---AC-------------------------------------------G---------- 3711
11_cpx.CM.95.95CM_1816      --------------------------------------------CA----T---GG---------C---------------------T----A.--------A-G--C-----C--C-----------------------C--T---------A-------G--G- 3696
12_BF.AR.99.ARMA159         -----G--G-------------------------------------GG----A---A---G---------G---------A--G--G.---------A---------AC-----------------------------T-------------------------- 4337
13_cpx.CM.96.96CM_1849      --------G--------------------------------T---AA-------------G--------G----C-----C--A--------A.--T----A-G---------C---------A------------------T--A------------------- 3737
14_BG.ES.05.X1870           --------G-----------------------------------AG-----------G---T----------T----C--A--------G.--G----A-G---------A---------------------------C--TA-A------------------- 3799
15_01B.TH.99.99TH_MU2079    ------------------------------------------T-A-----------G--G------------G----------T----A.--------A-G---------C---------A------------------TT---------A---------G- 3731
16_A2D.KR.97.97KR004        --------G------A--------G-----C-------------T---C-----------GT--------------G----------T----------.--------A-----C-G-----------------CA---C-----T-----------G---------- 3695
17_BF.AR.99.ARMA038         --------G--------------------------------T-----------A---A---G-----------G----------A--G--G.--------A--G-------AC--------------------------------T------------------------- 3546
18_cpx.CU.99.CU76           --------G-----------------------------------CAG---------A-GG---T------------------A----T----A.--------T---------AC-------------------------------TT--------A-------------- 3637
19_cpx.CU.99.CU7            --------G-----------------------------------T------T----G--G-----C-------G---------T----A.--------A----------C-------------------------------T---------A---------G- 3558
20_BG.CU.99.Cu103           --------G-----------------------------------AG------------------------------C--------A--------A.--------A-G---------C-------------------A---------T---------A----------- 3798
21_A2D.KE.99.KER2003        --------G------------------------C--------T---G--T---------GT--------------------T----A.--------A---------------------------------------------CA---C-----T--A------A----------- 3537
22_01A1.CM.01.01CM_0001BBY  ---------------------------------T---T-A--------A--GA--T-----C-------G------------T----A.--C----ACG---------C--------GA------T-----------------T---------A---------G- 3540
23_BG.CU.03.CB118           --------G-----------------------------------AG------------------------------C--------A--------A.--G----A----------C--------A--------------C-----T------------------------- 3775
24_BG.ES.08.X2456_2         --------G-----------------------------------AG------------------------------C--------A--------A.C-G----A-G---------C---------G--C------------------T--A--T----------------- 3783
25_cpx.CM.02.1918LE         --------G-----------------------------------AA------------G-----------------T----C--A----G--A.--G----A-G--C-----C-----------------------------T---------A----------- 3540
26_AU.CD.02.02CD_MBTB047    --T-----G-----------------A-----------------AG--------A----G-----------------------T----A.--C----A------------------------A------------A-----A-------T---- 4337
27_cpx.FR.04.04CD_FR_KZS    ------------------------------G-----T---AG-----------------T-----------------C--------A--------A.--G----A-G---------C-----------------C-----------T---------A----------- 4323
28_BF.BR.99.BREPM12609      -----G----------------------------------------------A--GA------C------------------------------.-----------------------------T------A------------------A---------A----------- 3698
29_BF.BR.01.BREPM16704      --------G------------------------------------------A---A---G------------G------------A--G--G.--------A---------AC------------------------------T------------------------- 3733
31_BC.BR.04.04BR142         --------------G-------------------A---AA------------G--G---C------------------T----A.--G----A----------C----------------------G-----T--------CA------------ 3804
32_06A1.EE.01.EE0369        --------G-----------------------------------AG-----------G--------C-------C--------A--------A.--G----A-G---------C----------C-------A-----------T---------A----------- 3967
33_01B.ID.07.JKT189_C       -----G--GR----G-------------------------T-A--------A----R---------------G-----Y-----T----A.--------A-G---------C--------A-------------------TT--------CA---------G- 3633
34_01B.TH.99.OUR2478P       --------G-----------------------------------T-A-----------G--G--------------G--------A--T----A.--------A-G---------C---------A-------------------TT--------A---------G- 3534
35_AD.AF.07.169H            --T-----G-----------------------------------CAA-----------GA-----------------------G----T----A.--------ACG----------------------G---------T---------A---------G- 3532
36_cpx.CM.00.00CMNYU830     ---------------------------------------------AAG----------A---------------T-------A----G--A.--T----A-G---------C------------------------------T---------A----G------- 3537
37_cpx.CM.00.00CMNYU926     -----G--T-----------------A-----------------T-A---------G-GGG-------------G----------T----A.--T--A-A-G---------C-------------------A--------T---------A----------- 3522
38_BF1.UY.03.UY03_3389      --------G---------------------------------------A------A---G--------------------A--G--G.--------A--G-------AC------------------------------T------------------------- 3755
39_BF.BR.04.04BRRJ179       ---------------------------------C---------A----T-----------------------------A--------A.--G-------------A---------------------------------A------A----------- 3809
40_BF.BR.05.05BRRJ055       ---------------------------------C---------------------A---G-----------------------------.--G----A--------------------------------------------A------------------------ 3815
42_BF.LU.06.luBF_18_06      --------G-----------------------------------A------T-----A--T----C----------------A--------A.--T----A--G-----------------G--------------------A-A---------G----------- 3886
43_02G.SA.03.J11223         --------G-----------------A-----------------AG--G-----------------------C--------A--------A.--G--Y--A-G---------C-----------------------G-T-----C---------------- 3857
44_BF.CL.00.CH80            --------G----------------------------------------A-------G-------------------------A--G--G.--C----A----------C-------------------------------T---------A--A---------- 3774
45_cpx.FR.04.04FR_AUK       --T--G--G-----------------------------------CAG---------A--G--G--G-------------------------T----A.--------A----------C-------------------------------T------------------G- 4318
46_BF.BR.07.07BR_FPS625     --------G-------G-------------------------------G----A-------G-------------G----------A--G--C.--------A----------C-------------------------------T--A------------------ 3770
47_BF.ES.08.P1942           --------G-------------------------------T----------------------------G--------A--------A.--------A---------A---------G---------C-------------A-C---------G----------- 3771
48_01B.MY.07.07MYKT021      -----G--G---------------------------G-----T-A-------A-GG--G---------------------G--T----A.--C----A----------C------------------------C--TT--------A--A-------G- 3537
49_cpx.GM.03.N26677         --T-------------------G---------------A---AAG-------A--------G--C--------G----------T--G--A.--------A----------C-----------------------G-----T--------CA------------ 3754
50_A1D.GB.10.12792          --T-----------G--C-------------------------T-C--------A--G--G--G----------G--------A--T----A.--C----A-G---------C--------A---------------------T--A----------------G- 3739
51_01B.SG.11.11SG_HM021     ------------------------------------------T----------GG-------------------------------------A.--C----A--------------------------------------------------G--G---------- 3539
52_01B.MY.03.03MYKL018_1    --T--G--G-----------------------------------T-A----------G--G-----C-------G---------T----A.--------A-G---------C---------------------------------TT---------A-------G- 3646
53_01B.MY.11.11FIR164       --T--G--G-----------------------------------AAA----------G--G------------G---------------A.--G----A-G---------C---------A--------------------TT--------A---------G- 3672
54_01B.MY.09.09MYSB023      ----------------G--------A------------T--T------T--A-----T---------------A--------A.--------------------A-------------------------------------A------------------------- 3734
55_01B.CN.10.HNCS102056     --------------------------------------T--T------T--A--G----------------G----C----------A.--C----A-G------------------A------------------------------G-----------G- 3642
56_cpx.FR.10.URF5_patient_A --------G-------------G--A------C-------AG----------G----------------C--------A--------A.--C----A-----C----C---------A---------------------T-------------------------- 3683
57_BC.CN.09.09YNLX19sg      --T-------------C-----------------A---AG------------G-----C----------------A--T----A.--T----A-GG---------C----------------------C----------------CA---------G- 3516
58_01B.MY.09.09MYPR37       -----G--G-----------------------------------T-A----------G--G------------G---------------A.--C----A-G---------C---------A--------------------TT-A------A---------G- 3677
59_01B.CN.09.09LNA423       -----G--G-----------------------------------T-G----------G----------------------T----A.--------A-G---------C-----------G------------------T--------A--A-------G- 3516
60_BC.IT.11.BAV499          --------------G-------------------A---AAG----------G--G--G--C-----------C----T----A.--T----A----------C------------------C-A---G-----T----T---CA------------ 4289
61_BC.CN.10.JL100010        --T-------------------------------A---AG--------------G--G--C----------------A--T----A.--G----A-GG---------C----------------------C-----T--------CA------------ 3753
62_BC.CN.10.YNFL13          --T-------------------R-----------G---CA----------------------------------A--T----A.--G----A-GG---------C---------------C----C-----TA-A------A----R-------- 3661
63_02A1.RU.10.10RU6637      --------G-----------------------------------AG--------A-----------------C--------A--------A.--C----A-G---------C------------------A---------T---------A----------- 3873
64_BC.CN.09.YNFL31          --T-------------------------------A-------G---------G-----C----------------A--T----A.--G----A--G---------C----------------------C-----T--------CA------------ 3686
65_cpx.CN.10.YNFL01         --T-------------------------------A--CAG------------G-----C----------------A--T----A.--G----A--G----T-AC---------------------------C-----T--------CA------------ 3703
67_01B.CN.11.ANHUI_HF115    --T---------------------------------------A-------------------------G--------------------A.------------------------------------T------------------------G-----C---- 3675
68_01B.CN.11.ANHUI_WH73     --T------------------------------------------------------------------------------------A.------------------------------------T------------------------G-----C---- 3666
70_BF1.BR.10.10BR_PE004     --------C-----------------------------------CAA-------------T-----------G----A--A--------A.---------------------------G-------------C----------------------------- 3788
71_BF1.BR.10.10BR_PE008     --T-----------------------------------------CA----------------------------------A--------A.---------------C--T-A---------G---------------------------G---------- 3996
72_BF1.BR.10.10BR_MG002     --------G------------------------------------------T--A---A---G------------------GA--A--G--A.--------A-----------AC------------------------------T------------------------- 3840
O.BE.87.ANT70               --T-----C--A-----------A--A-------------A--CAAG-AT--T--A-G---C--G--CC----A--------CC---A----A.--T-----A---------T-----------A----C-A---------A---GGA--A-----A--G--G--C--G- 4386
O.CM.98.98CMA104            --T--G--T--A--G-----------A-------------A--CAAG-AT--T--A-G---CT----CC----A--G------C---A--G--A.--T-----A---------T--CM-----------T-A---------C---GGA---------A-----G--C---- 3818
O.CM.98.98CMABB141          --T-----C--A-----------A--A-------------G--CAAA-AT--T--A-G---C-----CC----A---------CC---A--G--A.--T-----A---------T--C-----------T-A---------C---GGA--A-----A--G--G--C--G- 3817
O.CM.98.98CMABB212          --T-----C--A-----------A--A----C--------A--CAAA-AC--T--A-G---C--G--CC-G--G---------YC---A-----A.--T-----A---------T-----------------T-A---------C---GGA---------A-------------G- 3822
O.CM.98.98CMU5337           --T-----C--A----G------A--A-------------A--CAAG-AT--T--A-G---C-----CC-R--A---------CC---A-----A.--T-----A---------T--C-----------T-A--------------GGA--A-----A-----G--C---- 3819
O.CM.99.99CMU4122           --T-----C--A-----------A--A-------------A--CAAG-AT--T--A-G---C-----CC----A--K------CC---A-----A.--T-----A--G-----T--C-----------T-A---------C---GGA--A-----A--G-----C--G- 3817
O.FR.92.VAU                 --T-----T--A--G--------A--A-------------A---AAA-AT--T--A-G---C-----CC----A----------C---A-----A.--T-----A---------T-----------------T-A---------C---GGA--A-----A---------C---- 3898
O.GA.11.11Gab6352           --T--G--T--A-----------A--A-------------A--CAAG-AC--T--A-G---C-----CC----A---------CC---A-----A.--T-----A--G-----T--C-----------T-A---------C---GG-A-A---G-A--G--G--C--G- 3543
O.SN.99.99SE_MP1299         --T-----C--A-----------A--A-------------A--CAAG-AT--T--A-G---C-----CC----A---------CC---A-----A.--T-----A--G-----T--C-----------T-A--------------GGAT-A-----A--G--G--C---- 4385
O.US.10.LTNP                --T-----T--A-----------A--A-------------A--CAAG-AT--T--A-G---T--G--CC----A---------CC---A-----A.--C-----A--G-----T----------A----T-A---------A---GGAT-A-----A--G--G--C--G- 4307
N.CM.02.DJO0131             --T-----T-------------G--G-----C--------T--CT-A-----T--A------T----CC-------T-----A--A---------A.--T-----R-G-------------C-----------------C-------------T-A--CT-------G------- 3838
N.CM.02.SJGddd              --T-----T-------------G--GA----C-----------CT-A-----T--A------T----CY-------C-----A--A----------.--T-----A-GG------------------------------C-------------T-A--C--CA---G------- 3822
N.CM.04.04CM_1015_04        --T-----T-------------G--G-----C--------T--CT-A-----T--A------T-----C-------T-----A--A---------A.--G-----A-G-------------C-------C----A---------C-------------T----C--CA---G------- 3837
N.CM.06.U14296              --T-----T-------------G--G-----C--------T--CT-A-----T--A------T----CC-------T-----C--A---------A.--T-----A-G-----------C--C-----C--C---------------C-------------T-A--C--CA------------- 3835
N.CM.06.U14842              --T-----T-------G-----G--G-----C--------T--CT-A-----T--A------T----CC-------T-------RA---------A.--T-----A--------------------------A---------C-------------T-A--C--CA---G------- 3839
N.FR.11.N1_FR_2011          --T-----T-------------G--G-----C-----------CT-A-----T--A------T----C--------T-----------------G.--T-----A-G-------------------------------------C-------------T-A--C--CA---G------- 3733
P.CM.06.U14788              --------C--A--G--T-----RA-YA----Y-----------YAAR-AC--T--A-G---GT----C--G--A--C-----CC-A--A--G--G.--T-----A-----------------------------CC-A--------R---GG---T-----AA----G------- 3803
P.FR.09.RBF168              --------C--A--G--T------G-TA----Y---Y--------AAA-AC--T--A-GR--GT----CC----A--C-----YC-A--A--G--R.--T-----A-------------------------A---CC-A--------C---GG---T-----AA----G------- 4367
CPZ.CD.06.BF1167            -----------A-----G------G------------------A--CAAAAAC--T--A-----T--T--CC----A---------C-A--A-----A.--T-----A-G--TC--T-----------A-ATGT-A---GA---A-A---TT-A--C--CA-------T--G- 4434
CPZ.CM.05.SIVcpzMT145       ----G---T--A-------------G--------C----T---T-A--G-----A-----T--T--C---------------A--A--T-----A.--------A-G---C---TC---------AG----------CAG--C------T-A-----AA----G--T---- 3874
CPZ.GA.88.GAB1              --------T--A----------G-----G---------------A-------------G-----CC-----------C-G---------A.--------A-GG-----T-----------A---T----------------TT-A-----CA------------- 4393
CPZ.TZ.06.SIVcpzTAN13       --T--------A-----C-----GG----------GC--------AAAAAT--T-------C------------A--G---AC-G---A--G--A.--T-----T---------T-----------A-----------GA---A-AC--T--T--C---A-------C--G- 3973
CPZ.US.85.US_Marilyn        --T-----------------------A--------C---------T-A--G-----A-----T--T--CC--------C--------A--A--G--A.--C-----C---------T-----------C---T-----C-----------------------A--G--C--T--G- 4392
GOR.CM.07.SIVgor2139_287    --------T--A--G---------G--A----C------------AAA-AT-----A-G---CT----CC----A--C------C-A--G--G--A.--T-----A---------------------------A--C-A--C-----C---GGT-----G--A--G-----C---- 3869
GOR.CM.07.SIVgorCP2135con   --------T--A--G---------G--A----C------------AAA-AT-----A-G---CT----CC----A--C------C-A--A--G--A.--T-----A------------------------------C-A--C-----C---GGT-----G--A--G-----C---- 3868
```

```
B.FR.83.HXB2                AAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAGGAAAAGTTATCCTGGTAGCAGTTCATGTAGCCAGTGGATATATAGAAGCAGAAGTTATTCCAGCA 4502
Pol                         E__I__V__A__S__C__D__K__C__Q__L__K__G__E__A__M__H__G__Q__V__D__C__S__P__G__I__W__Q__L__D__C__T__H__L__E__G__K__V__I__L__V__A__V__H__V__A__S__G__Y__I__E__A__E__V__I__P__A_
A1.AU.03.PS1044_Day0        -----------T-------------C--AT-------G-----------------------------------G---------T-------C------C-G-----------A--T-----------A--------------C--------------------C------ 3710
A1.CH.03.HIV_CH_BID_V3538   -------G--------------------A--------G-----------------------------------G---------T----------C---C-------------A--TA-A--------C-----G--------T--------------------C------ 3707
A1.CY.08.CY236              -----------------A-------C--A--------------------------------------------G--G------T-G-----C------C-----------A-A--T-----------C--------------C--------------------C------ 3713
A1.ES.05.X1608_8            ----------------------------A--------G-------C---------------------------G--G------T-------C------C-----------A-A--T-----------C--------------C--------------------C------ 3950
A1.KE.11.DEMA11KE001        -------------------------------------G--------------------------C--------G---------T-------C------C-------------A--T--A-----------------------C--------------------C------ 3841
A1.RU.11.11RU6950           ----------------------------A--------G-----T--------C--G-----------------G---------T----C--C------C-------------A---A-A--------C--------------C--------------------C------ 4026
A1.RW.11.DEMA111RW002       ----------------------------A--------G-----------------------------------G---------T-------C------C-------------A--------------C-----G--------C--------------------C------ 3853
A1.SN.01.DDI579             ----------------------------A--------G---------------------------------------------T-------C------C-------------C--TA-A--------C--------------C--------------------C------ 3700
A1.UG.11.DEMA110UG001       -G-----------------C--------A--------G-----------------------T-----C-----G---------T-------C------C-------------A--------------C--------------C--------------G-----C------ 3857
A1.ZA.04.04ZASK162B1        ----------A-----------------A--------G-----------------------------------G---------T-------C-----CC-----------A-A--------------C--------------C--------G-----------C------ 3943
A2.CM.01.01CM_1445MV        ----------T------------------T-------G--------A--------------------------------------------C------C----------------------------C--------------C-----------------C--C------ 3701
A2.CY.94.94CY017_41         ----------T--------------------------G-----------------------------------------------------C------C-T--------------------G-----C-----G--------C-----------------C--C---A-- 3859
B.BR.10.10BR_MG029          -R-----------Y---------------M----------------------------------------------R-----R-----Y--------YY-----------R----Y-----------C--------------------------------Y--------- 3964
B.CA.07.502_1191_03         -G--------------------------A----------------------------------------------------------------G---------------A----------------------------------------------------------- 3929
B.CH.08.M2_0803101_NFLG8    -G--------T-----------G--------------------------------------------C------G-------------------------------G---A----------------C------------------------------A----C------ 3971
B.CN.10.DEMB10CN002         -------------T-----------------------------------------------------------G--------G---------------------------A-------A-----------------T--------------------------------- 3874
B.ES.10.DEMB10ES002         -----------T----------------------------G-----------------------------------------G-----C--------------------G----------------------C---------------------------C--C------ 3864
B.FR.11.DEMB11FR001         -----------------------------T--------------------------------------------------------------------------------A-------------------------T-----------------------C--C------ 3748
B.GB.05.MM45d213_GN1        -------------------------------G--------------------------------------------------------------C--------------------------------C-----G------------------------------------ 3903
B.HK.06.HK003               -------------T-----------C---A-------------T-----------------------C--------------------C---------------------A----------------------------------------------------------- 3872
B.HT.05.05HT_129389         ----------------------------A--------------------------------T--------------------------------------------------------A--------------G-----------------G-----------C------ 3703
B.JP.12.DEMB12JP001         -----------T-------------------------------------------------------------------------------------C-----G------A-------------------------------------------------C--------- 3889
B.KR.07.HP_18_07JHS10_3909  ----------------------------A--G--------------------------------------------------G--------C------------------A----------------------------------------------------------- 4166
B.PE.07.502_0525_wg5        ----------------------------AT----------------------------------------------------------C---------------------A-------------------------------G--------------------------- 3967
B.RU.11.11RU21n             ----------------------------AA-------------------C------------------------------------------------------------A-----T-A--------------------------------------------------- 4073
B.TH.08.MERLBDTRC10         ----------------------------------------------------------------------------------------C-----C---------------A----------G-----A-----G--------C--------------------------- 3735
B.US.11.ES38                -------------------------------C-----------------------------------C--------------G---------------------------A----------------------------------C-----------------C------ 4437
C.AR.01.ARG4006             ----------T--------------------------G-----T-CA------------------------------------T--------------------------A-C--------------C-----------------C-----------G-----C------ 3686
C.BR.07.DEMC07BR003         ----------T--------------------------G-------CA-----------------------T------------T--------------------------A-C--------------C--------------G--C-----------------C------ 3834
C.BW.00.00BW5031_1          -------G--------C--------------------G-----------------------------C-----GG--------T--------------------------A-C--A--A--G-----C--------------C--C-----G--------C--------- 3873
C.CN.10.YNFL19              ----------T---------C----------------G-----------------------------------G---------T--------------------------A-C--------------C--------------C--C--G--------G-----C------ 3855
C.CY.09.CY260               ----------T-----------------A-A------G--------A--------------------C---------------T-------------------G--G---A-C--------------C--------------C--C-----------G------------ 3725
C.ES.08.X2363_2             -----A----T-A-------------------------------T-A------------------------------------T-------C----------------G---C--------------C--------------C--C-----------G-----------G 3933
C.ET.02.02ET_288            ----------T--------------------------G-----------------------T-----------G-----------G-----------C------------A-C--------------C--------------C--C-----------G------------ 3704
C.IN.09.T125_2139           ----------T---------C--------T-------G-----------C---------------------------------T----C---------------------A-C--------------C--------------C--C--G--------G-----C------ 4478
C.KE.00.KER2010             ----------T--------------------------G-------CA-----------------------G--G---------T-G------------------------A-C--------------C-----------------C-----------G-----C------ 3683
C.MW.09.703010256_CH256.w96 -------------------------------------G--------A--------------------------GG----------------------C--------G---A-C--------------C-----------------C--------------C--C---A-- 4482
C.TZ.08.707010457_CH457.w8  ----------T-------------------A------G-----A--A--------------------------G---------T----C---------------------A----------------C--------------C--C-----------G----------A- 4476
C.YE.02.02YE511             -------------T---------------T----G--G--------A--------------------------G---------T----------------------G---A-C--------------C--------------C--C-----------G-----C------ 3677
C.ZA.08.705010534_CH534.w12 ----------T-----------G--C-----------G-----------------------------C-----G---------T----C---------C----G------------T----------C--------------C--C--G--------G-----C------ 4492
C.ZA.10.DEMC10ZA001         ------------------------------A------G-----T-CA--------------------------G---------T--------------------------A-C-----A--------C--------------C--C-----------------C------ 3842
C.ZM.11.DEMC11ZM006         -----------------------------T-------G--------------------------C--------G--------------------------------------C---T----------C--------------C--C-----------G-----C------ 3851
D.CM.10.DEMD10CM009         -----------------------------------------------------------------------------------T--------------C-------G---A-------------------------------C--------------------------- 3877
D.CY.06.CY163               ----------T--------------C--A-----------------------------------G------------------T--------------C-------------------------------------------C------------A-------------- 3710
D.KE.11.DEMD11KE003         ----------T-----------G--C-----------------T---------------------------------------T----------------------G--------------------C--------------C--------------------------- 3852
D.KR.04.04KBH8              ----------TT------------------A----------------------------------------------------T----------C---C-------------------------------------------C--------------------------- 4461
D.SN.90.SE365               ----------T-----------------------------------A------------------------------------T----------------------------------------C--------G--------C--------------------------- 4519
D.TZ.01.A280                -----------T-----------------------------------------------------------------------T-G-----C--------------------------------------------------C--------------------------- 3703
D.UG.10.DEMD10UG004         ----------T---------------------------------T----------------T---------------------T--------------------------A----------------C--------------C-----------------C--------- 3840
D.UG.11.DEMD11UG003         -------G--T--------------------------------TT-------------------------------------------------------------G--------------------C--------------C--------------------------- 3857
D.YE.02.02YE516             -G--------T--------C-----C--A-A----------------------------------------------------T--------------C-----------A-------------------------------C--------------------------- 3713
D.ZA.90.R1                  ----------T------------------------------------------------------------------------T--------------------------------A-A-----------------------C--------------------------- 3848
F1.AO.06.AO_06_ANG32        ----------T--------------------------G-----------------------------------G---------T-G------------------------------A-------T--G--------T-----C-----------------C--------- 3713
F1.AR.02.ARE933             ----------T------------------T-------G-----------------------T---------------------T-------------------------------------------C--------T--------CC----------------C------ 3796
F1.BR.10.10BR_RJ015         -------G--T--T--------G-----A--------G-----------------------------C-----G---------T-------C------------------A----T-----------C--------T-----G--CC-------------C--C------ 3975
F1.CY.08.CY222              ----------T----------G---------------G--G--------------------------------G---------T---------------------A----A-C--------------C--------T-----------------------C--------- 3695
F1.ES.02.ES_X845_4          -----A----T--------------------------G-----------------------------------G---------T--------------------------A-------------T-----------T-----C-----------------C--------- 3954
F1.RO.96.BCI_R07            ----------T--------------------------G--GA-------------------------------G---------T----------------------------C--------------C--------T-----------------------C--------- 4526
F1.RU.08.D88_845            ----------T--------------------------G-----------------------------------G---------T--------------C-----------A-C-----T--------C--------------C--------G-----------C------ 3961
F2.CM.02.02CM_0016BBY       -----------------------------T-------G--------A-----------------C--------G---------T--------------------------A----------------C--------T-----C--------------------------- 3692
F2.CM.10.DEMF210CM001       -------------------------C-----------G--------A-----------------C--------G---------T--------------C-----------A----------------C--------T-----C--------G-----------C------ 3766
F2.CM.10.DEMF210CM007       ----------------------------A--------G-----------A--------------C---------------------------------C-----------A----------------C--------T-----C--------------G-----C------ 3849
G.BE.96.DRCBL               -------G-----------------C---T-------G---------------------------------------------T-G------------------------A----------------C--------------C--------------------C------ 4459
G.CM.10.DEMG10CM008         -------G-----------C-----------------G--------A-----------------------T------------T-G-----------C--------GC--A-----T----------C-----------------------------------C------ 3883
G.CN.08.GX_2084_08          -------G--------------G--------------G--G--------------G--------C------------------T--------------------------A----------------C--------------C--------------------C------ 3742
G.CU.99.Cu74                -G-----G-----------------------------G--------------T------------------------------T-------------------------GA----------------C-----G--------C--------------------------- 4106
G.ES.09.X2634_2             -------G--------------------AT-------G--------------C-----------------------------GT--------------------------A----------------C-----G--------C--------------------C------ 3985
G.GH.03.03GH175G            -------G-----------------------------G------T-A--------------------------G---------T--------------------------A-----A-A--------C--------------C--------------------C------ 4542
G.KE.09.DEMG09KE001         -------G-----------------------------------T--------T------------------------------T--------------------------------A----------A--------------C-----------------C--C------ 3877
G.NG.09.09NG_SC62           -------------------------------------------------------G---------------------------T-G------------------------A----------------C--------------C--------------------C---A-- 3694
G.ZA.01.TV546               -------G-----------------------------G---------------------------------------------T-------------C------------A-----A-A--------C--------------C--------------------C------ 3704
H.BE.93.VI991               ----------T------------------T-------G-----------------------------C-----C---------T-------------C------------A----T-----------C--------------C--------------------C------ 3891
H.CF.90.056                 ----------T--------------------------G-----------------------------C---------------T-------C--------G------C-------T-----------C--------------C-----------------C--C------ 3849
H.GB.00.00GBAC4001          ----------T--------------------------G--------A-----------------C--C---------------T-------------------------------A-----------C--------------C-----------------C--------- 4020
J.CM.04.04CMU11421          ----------------------------------R--G-----------------------Y-----------R---------T----------M---------------A-------------------------------Y--------------------Y------ 4008
J.SE.93.SE9280_7887         -G--------T-----------------------G--G-----------------------------------G---------T--------------C-------------------------------------------C--------------------C------ 3816
J.SE.94.SE9173_7022         ----------T-----------------------G--G-----------------------------------G---------T----------------------------------------------------------C--------------------C------ 3817
K.CD.97.97ZR_EQTB11         ----------T--------------------------G--------A--------------------------G--------GT--------------------------A-------A-----------------T-----C--------------------C------ 3699
K.CM.96.96CM_MP535          ----------T--------C-----------------G--------A------------------------------------T--------------------------A----------------C--------T--C--C--------------------C------ 3698
U.CA.01.TV749               ----------T------------------A----G--G--------------G--------------C-----GG--------T----C---------C-----------A-------A--------C--------------C-----------------G--C------ 4519
U.CA.99.TV721               ----------T------------------T----G--G--------A--------------------C-----GG-------------C---------C----------GA----------------C--------------C--------G-----------C------ 4518
U.CD.83.83CD003_Z3          -------------------------------------G--------------------------C--------G---------T-------------CC----G-----------T-----------C--------------C---C----------------C--G--- 3997
U.CD.90.90CD121E12          ----------T-----------------------------------A-----------------C--------G-----------------------CC----G------A----------------C--------------C--------------------C------ 3701
U.CY.05.CY090               ----------T--------C---------G-------G--------------------------C------------------T-------------C------------------T----------C--------------C--C-----------------CT----- 3701
U.CY.08.CY223               -------G--------------------A---------------------------------------------G--------T----C--C------C-------------A--------------C-----------------C-----------------C------ 3719
U.ES.10.DEURF10DZ001        -------G--T------------------------------------------------------G-----------------T----C---------------------A----------------C--------------C--------------------C------ 3778
U.GR.99.99GR303             -------G----------C-GG-AT--G---------G--------------------------------------------GT-------------------------------T-C---G-----C--------T-A---C--------------------C------ 3773
U.NL.95.U_NL_95_H10986_D1   ----------T--T--------------A--------G-----T-----------------------------G--G------T--------------------------A---------------CC--------T-----C-----------------C--C------ 4387
01_AE.AF.07.569M            ------------A---------------A--------G-----T-----------------Y-----------G---------T-------C------C-----------A-C---A----------C--C--G-----------------------------C------ 3708
01_AE.CF.90.90CF11697       ------------A---------------A--------------T--------------G--------------G---------T-------C--------------------C--------------C--C--G-----------------------------C------ 4451
01_AE.CN.10.YNFL03          ------------A---------------A--------G-----T-----------G-----------------G---------T-------C------C-------------C--------------C-----G-----------------------------C------ 3880
01_AE.HK.04.HK001           ------------A---------------A--------G-----T-----------------T-----------G---------T-------C------C----G--------C--------------C--C--G-----------------------------C------ 3864
01_AE.IR.10.10IR.THR48F     ----------------------------A--------G-----T-----------------------------G---------T----C--C------C-----------A-C--------------C--C--G-----------------------------C------ 3692
01_AE.JP.x.JRC77AE          ------------A---------------A--------G-----T--------------G--------------G---------T-------C------C-G-----------C--------------C--C--G--------------------------C--C------ 4499
01_AE.TH.09.AA090a_WG11     ------------A---------------A--------G-----T-----------------------------G---------T-------C------C-------------C--------------C--C--G-----------------------------C------ 3925
01_AE.TH.90.CM240           --------A---A---------------A--------G-----T-----------------------------G---------T-------C------C-------------C--------------C--C--G-----------------------------C------ 4076
01_AE.US.05.306163_FL       ------------A---------------A--------G-----T-----------------------------G---------T-------C------C-----------A-C---A-A--------C-----G--------G--------------------C------ 3693
01_AE.VN.98.98VNND15        ------------A---------------A--------G--------------------G------------------------T--------------C-------------C--------------C--C--G--------G--------------------C------ 3757
```

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

B.FR.83.HXB2 GAAACAGGGCAGGAAACAGCATATTTTCTTTTAAAATTAGCAGGAAGATGGCCAGTAAAAACAATACATACTGACAATGGCAGCAATTTCACCGGTGCTACGGTTAGGGCCGCCTGTTGGTGGGCGGGAATCAAGCAGGAATTTGGAATTCCCTACAATCCCCAAAGTCA 4672
Pol E T G Q E T A Y F L L K L A G R W P V K T I H T D N G S N F T G A T V R A A C W W A G I K Q E F G I P Y N P Q S Q
A1.AU.03.PS1044_Day0 3880
A1.CH.03.HIV_CH_BID_V3538 3877
A1.CY.08.CY236 3883
A1.ES.05.X1608_8 4120
A1.KE.11.DEMA11KE001 4011
A1.RU.11.11RU6950 4196
A1.RW.11.DEMA111RW002 4023
A1.SN.01.DDI579 3870
A1.UG.11.DEMA110UG001 4027
A1.ZA.04.04ZASK162B1 4113
A2.CM.01.01CM_1445MV 3871
A2.CY.94.94CY017_41 4029
B.BR.10.10BR_MG029 4134
B.CA.07.502_1191_03 4099
B.CH.08.M2_0803101_NFLG8 4141
B.CN.10.DEMB10CN002 4044
B.ES.10.DEMB10ES002 4034
B.FR.11.DEMB11FR001 3918
B.GB.05.MM45d213_GN1 4073
B.HK.06.HK003 4042
B.HT.05.05HT_129389 3873
B.JP.12.DEMB12JP001 4059
B.KR.07.HP_18_07JHS10_3909 4336
B.PE.07.502_0525_wg5 4137
B.RU.11.11RU2ln 4243
B.TH.08.MERLBDTRC10 3905
B.US.11.ES38 4607
C.AR.01.ARG4006 3856
C.BR.07.DEMC07BR003 4004
C.BW.00.00BW5031_1 4043
C.CN.10.YNFL19 4025
C.CY.09.CY260 3895
C.ES.08.X2363_2 4103
C.ET.02.02ET_288 3874
C.IN.09.T125_2139 4648
C.KE.00.KER2010 3853
C.MW.09.703010256_CH256.w96 4652
C.TZ.08.707010457_CH457.w8 4646
C.YE.02.02YE511 3847
C.ZA.08.705010534_CH534.w12 4662
C.ZA.10.DEMC10ZA001 4012
C.ZM.11.DEMC11ZM006 4021
D.CM.10.DEMD10CM009 4047
D.CY.06.CY163 3880
D.KE.11.DEMD11KE003 4022
D.KR.04.04KBH8 4631
D.SN.90.SE365 4689
D.TZ.01.A280 3873
D.UG.10.DEMD10UG004 4010
D.UG.11.DEMD11UG003 4027
D.YE.02.02YE516 3883
D.ZA.90.R1 4018
F1.AO.06.AO_06_ANG32 3883
F1.AR.02.ARE933 3966
F1.BR.10.10BR_RJ015 4145
F1.CY.08.CY222 3865
F1.ES.02.ES_X845_4 4124
F1.RO.96.BCI_R07 4696
F1.RU.08.D88_845 4131
F2.CM.02.02CM_0016BBY 3862
F2.CM.10.DEMF210CM001 3936
F2.CM.10.DEMF210CM007 4019
G.BE.96.DRCBL 4629
G.CM.10.DEMG10CM008 4053
G.CN.08.GX_2084_08 3912
G.CU.99.Cu74 4276
G.ES.09.X2634_2 4155
G.GH.03.03GH175G 4712
G.KE.09.DEMG09KE001 4047
G.NG.09.09NG_SC62 3864
G.ZA.01.TV546 3874
H.BE.93.VI991 4061
H.CF.90.056 4019
H.GB.00.00GBAC4001 4190
J.CM.04.04CMU11421 4178
J.SE.93.SE9280_7887 3986
J.SE.94.SE9173_7022 3987
K.CD.97.97ZR_EQTB11 3869
K.CM.96.96CM_MP535 3868
U.CA.01.TV749 4689
U.CA.99.TV721 4688
U.CD.83.83CD003_Z3 4167
U.CD.90.90CD121E12 3871
U.CY.05.CY090 3871
U.CY.08.CY223 3889
U.ES.10.DEURF10DZ001 3948
U.GR.99.99GR303 3943
U.NL.95.U_NL_95_H10986_D1 4557
01_AE.AF.07.569M 3878
01_AE.CF.90.90CF11697 4621
01_AE.CN.10.YNFL03 4050
01_AE.HK.04.HK001 4034
01_AE.IR.10.10IR.THR48F 3862
01_AE.JP.x.JRC77AE 4669
01_AE.TH.09.AA090a_WG11 4095
01_AE.TH.90.CM240 4246
01_AE.US.05.306163_FL 3863
01_AE.VN.98.98VNND15 3927

```
B.FR.83.HXB2                 GAAACAGGGCAGGAAACAGCATATTTTCTTTTAAAATTAGCAGGAAGATGGCCAGTAAAAACAATACATACTGACAATGGCAGCAATTTCACCGGTGCTACGGTTAGGGCCGCCTGTTGGTGGGCGGGAATCAAGCAGGAATTTGGAATTCCCTACAATCCCCAAAGTCA 4672
Pol                          _E__T__G__Q__E__T__A__Y__F__L__L__K__L__A__G__R__W__P__V__K__T__I__H__T__D__N__G__S__N__F__T__G__A__T__V__R__A__A__C__W__W__A__G__I__K__Q__E__F__G__I__P__Y__N__P__Q__S__Q
02_AG.CM.10.DE00210CM013     --------A-----G-------TC--CA-A--------------------------G---GT------C--A---------------------A-----G-A--A-A---A--A--C--------AAATG---CA--A-----------------------------C-- 4039
02_AG.CY.09.CY256            --CC----A--------------C---A-A--------------------------G----T------C--A---------------------A-----G-AT-A-A---A--A------C----AAATG---CA--A-----------------------------C-- 3868
02_AG.ES.06.P1423            --------A-----G--------C--CA-A--------------------------G----T------C--A-----------------T---A-----G-A--A-A---A--A-----------AAATG---CA--A-----------------------------C-- 4109
02_AG.GW.05.CC_0048          --------A--A--G------------A-A--------------------------G---GT------C--A---------------------A-----G-A--A-A---G--A--C--------AAATG---CA--A-----------------------------C-- 4022
02_AG.KR.12.12MHR9           --------A-----G--------C---A-A------------------------------GTG-----C--A--------T------------A-----G-A--A-AA--A--A-----------AAATG---CA--A-----------------------------C-- 4331
02_AG.LR.x.POC44951          --------A--------------C---A-A--------------------------G---GT------C--A---------------------A-----G-A--A-----A--A-----------AAATG---CA--T-----------------------------C-- 4667
02_AG.NG.09.09NG_SC61        --------A--------------C---A-A------------------------------GT------C--A--T------------------A-----G-A--A-AA--A--A-----------A-ATG---CA--A-----------------T--C----------- 3864
02_AG.NG.x.IBNG              --------A-----G--------C---A-A--------------------------G---GT------C--A---------------------A-----G-A--A-A---G--A-----------AAATG---CA--A-----------------------------C-- 4197
02_AG.SE.94.SE7812           --------A--A--G--------C---A-A--------------------------G---GT------C--A---------------------A-----G-A--A-A---A--A-----------AAATG---CA--A-----------------------------C-- 4044
02_AG.SN.98.98SE_MP1211      --------A-----G-----C--C---A-A--------------G-AGAT-G-CAGT---GT-G----C--A---------------------A-----G-A--A-A---A--A-----------CAA-G-T-CA-CA-G-A----G-A-T------------------- 3869
03_AB.RU.97.KAL153_2         --------A--------------C---G-C-------------------------------T---------A---------------------A--A--G----A-----T--------------A--G-----------------C----------------------- 3898
04_cpx.CY.94.94CY032_3       --------A-----------C--C--CA-AC-------------------------G----TG------G-A-----C---CC----------A-----G----A-----A--------------A-AT-----C--------------------------------C-- 4038
05_DF.BE.x.VI1310            --G-----A-----C-----C--C--CA-A-----GC-G----------------------TG--------A--T------CC----------A---G-G----A-----A--------------A--T---C----A-------------------------------- 4054
06_cpx.AU.96.BFP90           --------A-----G--------C---A-A------------------------------GTG-----C--A--------T------------A-----G-A----AA--A--------------AAAT----CA--A--------------T----------------- 4700
07_BC.CN.98.98CN009          --------A--A--G--------C---A-AC-------------------------C---GT---------A--------T--T---------A--A--G-A--A-----A--------------A--T---C-A----------------------------------- 4004
08_BC.CN.97.97CNGX_6F        --------A--A-----------C---A-AC-------------------------C---GT---------A--------T--T---------A-----G-A--A-----A--------------A--T---C-A----------------------------------- 3852
09_cpx.GH.96.96GH2911        --------A--A-----------C---A-A------------------------------GTG-----C--G-----------T---------A-----G-A--A-A---A--A-----------AAATG---CA--A-----------------------------C-- 3874
10_CD.TZ.96.96TZ_BF061       --------------------C-----C--C--G---------------------------GT-G-------A---------------------A-C---G-A--A--------------------A--C----------------------------------------- 4051
11_cpx.CM.95.95CM_1816       --------A--------------C---A-A------------------------------GT---------A---------------------A------TA--AA----A--------------A-AT---C-A-----------G--C-------------------- 4036
12_BF.AR.99.ARMA159          --------A--A--------C--C--CA-AC----G-----------------------------------A---------CC--------T-A----C-T---A-----A--------------A--T---C-A-----------------------C----------- 4677
13_cpx.CM.96.96CM_1849       --------A--------------C---A-A------------------------------GT------C--A-----------------T---A-----G-A--A-AA--A--A-----------AAAT----CA--A-----------------T-------------- 4077
14_BG.ES.05.X1870            --------A--------------C---A-AC-------G--------G--------G--------------A--T-------C----------A-----G-A--A-A---A--A-----------AAAT----CA--A-----------------------------C-- 4139
15_01B.TH.99.99TH_MU2079     --------A--------------C-----GC-----------------------------GT------C--A-----C--T------------A-C---G-AA-G-AA--A--------------CAATG--CGA---------T-GGA--------------------- 4071
16_A2D.KR.97.97KR004         --------A-----G-----T--C---A-A------C-----------------------GT---------A--------GCC--------T-A----AC-A--A-----A-----------------TG--C-A-----------G----------------------- 4035
17_BF.AR.99.ARMA038          --------A--A--------C--C--CA-AC----G-----------------------------------A---------CC--------G-A--AACG----A-----A--------------A--T---C-------------------------C----------- 3886
18_cpx.CU.99.CU76            --------A-----------C--CC----AC-----------------------------GT------C--A---------CC----------A-----G-A--A-A---A--T-----------AAAT--T-CA--A-----------------------------C-- 3977
19_cpx.CU.99.CU7             --------A-----G--------C-----GC------------C---------------GGT------C--A--------TCC----------A-C---G-A----AA--A--------------AAATG--C-A-----------G-----------------G----- 3898
20_BG.CU.99.Cu103            --------A--------------C---A-AC----------T-----G--------G---GT---------A---------CC--------T-A-----G-A--A-A---A--A-----------AAAT----CA--A-----------------------------C-- 4138
21_A2D.KE.99.KER2003         --------A--------------C---T-A-----GC-----------------------GT---------A--------GCC--------T-A----AG-A--A-----------C--------A--T-----A-----------G-----T----------------- 3877
22_01A1.CM.01.01CM_0001BBY   --------A-----G--------C---A-AC-------------G---------------GT------C--A--------T-----C------A-C---G-A--G-AA--A--------------A-A----C-A-----------------T----------------- 3880
23_BG.CU.03.CB118            --------A--------------C---A-AC----------T-----G--------G---GT---------A--Y------CC--------T-A-----G-A--A-A---A--A-----------A-AT----CA--A-----------------------------C-- 4115
24_BG.ES.08.X2456_2          --------A--------------C--CA-AC----------T-----G------------GTG--------A---------CC--------T-A-----G-A--A-A---A--A-----------A-AT----CA--A-----------------------------C-- 4123
25_cpx.CM.02.1918LE          --------A--A--G--------C-----A------------------------------GT-G-------A---------CC----------A----CG----A-----A--------------A--T--------------------------------------C-- 3880
26_AU.CD.02.02CD_MBTB047     --------A--A--G--------C--CA-AC---------------------------GGGTGG-------A---------------------A-C---G-A--C-A---A--------------AAAT---C-A----------------------------------- 4677
27_cpx.FR.04.04CD_FR_KZS     -----------------------C--C--A------------------------------GT------C--A---------------------A-----G-AA-G-A---A--A-----------A--TG-T--A--A-------------------------------- 4663
28_BF.BR.99.BREPM12609       --G--------A-----------C-----C--------------------------G------G-------A--T------------------A--AA-G------A--------------------AG-----------------C----------------------- 4038
29_BF.BR.01.BREPM16704       --------A--A--------C--C--CA-AC----G-----------------------------------A----------C----------A----C--A----A---A--------------A--T---C--------------G----------C----------- 4073
31_BC.BR.04.04BR142          --------A--A-----------C---A-AC-------------------------T---GT---------A-----------T--------AA--AA---A----AA--A--------------A--C---C-A----------------------------------- 4144
32_06A1.EE.01.EE0369         --------A-----G--------C---A-AC-----------------------------GT------C--A--------T------------A-----G-A----AA--A--------------AAAT--T-CA--A-------------------------------- 4307
33_01B.ID.07.JKT189_C        --------A-----R--------C-----GC----R------------------------GT------C--A-----Y--T------------A-Y--AG-A----AA--A--------------CAATG--C-------------G--C-----------T-----C-- 3973
34_01B.TH.99.OUR2478P        --------A-----G--------C-----GC-----------------------------GT------C--A-----C--T------------A-C---G-A----AA--A--A-----------CAATG--CGA-----------G--C-------------------- 3874
35_AD.AF.07.169H             --------------G--------C-----GC-----------------------------GT-G----C--A---------------------A-----G-AT---AA--A--------------AA-T---C-A----------------------------------- 3872
36_cpx.CM.00.00CMNYU830      --------A-----G--------C-----A---------------C--------------GT------C--A--------T-----C------A-C---G-A--G-AA--A--------------A-AT---C-A--------------------------------C-- 3877
37_cpx.CM.00.00CMNYU926      --------A--------------C-----GC----GC---------------------GGGT------C--A--------T------------A-C---G-A----AA--A--------------AAATG----A--------------------T--C----------- 3862
38_BF1.UY.03.UY03_3389       --------A--A--G-----C--C--CA-AC----G-----------------------------------A---------CC--------T-A----C--A--A-----A--------------A--T---C-------------------------C----------- 4095
39_BF.BR.04.04BRRJ179        --G--------A-----------C---A-C-----------------------------------------A---------------------A--A----T--A-----------------------G-----------------C----------------------- 4149
40_BF.BR.05.05BRRJ055        --G--------A-----------C---A-------------------G-----------------------A---------CC--------T-A----CG----C-A---A--------------A--T---CG------------------------C----------- 4155
42_BF.LU.06.luBF_18_06       --------A--A--------C--C--CA-AC----------------------------------------A-----------------T---A-----GT---A-----------C-----------G-----------------C----------------------- 4226
43_02G.SA.03.J11223          --------A-----C--------C---A-A------------------------------GT------C--A--------------C------A-----G-A--A-----A--A-----------AA-TG---CA--A-----------------------------C-- 4197
44_BF.CL.00.CH80             --------A--A--------C--C--C--AC----G-----------------------------------A---------CC----------A----CG----A-----A--------------A-AT---C----------------C--------C----------- 4114
45_cpx.FR.04.04FR_AUK        --------A-----G--------C---A-AC-----------------------------GT------C--A--------T------------A-----G-A--G-AA--A--------------A-ATG----A----------------------------------- 4658
46_BF.BR.07.07BR_FPS625      --------A--A--------C--C--CA-A-----G-----------------------------------A----------CT---------A----CG------A---A--------------A--T--TC-------------------------C----------- 4110
47_BF.ES.08.P1942            -----------------------C-----C-----------------------------------------A---------------------A-CA----A----A---------------------G-----------------C----------------------- 4111
48_01B.MY.07.07MYKT021       --------A----GG--------C----CGC-----------------------------GT------C--A--------T--------------C---G-A----AA--A--------------CAATG--CG---------------------------------C-- 3877
49_cpx.GM.03.N26677          --------A--A-----------C-A-A-A--------------------------C---GT---------A-----------T---------A-C---G-A----A---A--------------A--T---C-A--A-------------------------------- 4094
50_A1D.GB.10.12792           --G-----A--A--G--------C-----A------------------------------GT-G----C--A---------------------A-----G-A----AA--A--------------AAAT---C-A-----------G----------------------- 4079
51_01B.SG.11.11SG_HM021      -----------------------C---A-A----------------------------C------C-----A---------CC----------A--A--G-A-C--A------------------A--G-----------------C----------------------- 3879
52_01B.MY.03.03MYKL018_1     --------A-----G--------C---A-AC-----------------------------GT------C--A--------T------------A-C---G-A----AA--A--------------CAATG--C-A-----------G--C-------------------- 3986
53_01B.MY.11.11FIR164        --------A-----G--------C-----GC-----------------------------GT------C--A-----C--T------------A-C---G-A----A---A--------------CAATG--C-------------G--C--------------G----- 4012
54_01B.MY.09.09MYSB023       --G--------------------C---A-CC-------------------------G------G-------A-----------A--------AA-CAA-G-A----A---------C-----------GG----------------C--------------T-------- 4074
55_01B.CN.10.HNCS102056      --G--------------------C---A-C-----------------G--------------------C--A-----C--T------------A-C---G-A----AA--A-----C--------CAATG--C-A-----------G--C-------------------- 3982
56_cpx.FR.10.URF5_patient_A  --------A--------------C--CA-C--G-----------G--------------------------A-------------------T-A----C--A----A------------------A--G----------------------------------------- 4023
57_BC.CN.09.09YNLX19sg       --------A--A-----------C---A-AC-------------------------C---GT-G-------A---T----T--T---------A-----G-A----A---A--------------A--T----CA--------------C-----T-------------- 3856
58_01B.MY.09.09MYPR37        --------A-----T--------C-----GC----------------G------------GT------C--A-----C--T------------A-C---G-A-C--AA--A--------------CAATG--CG------------G--C-------------------- 4017
59_01B.CN.09.09LNA423        --------A--A--G------------A-AC-----------------------------GT------C--A-----C--T------------A-C---G-A----AA--A--------------CAATG--C-A-----------G--C-------------------- 3856
60_BC.IT.11.BAV499           --------A--A-----------C---A-AC--------------C----------C---GT---------A-----------T---------A--AA---A----AA--A--------------A--T---C-A-----------G----------------------- 4629
61_BC.CN.10.JL100010         --------A--A--G--------C-----AC-------------G-----------C---GT---------A-----------T---------A-----G-A----A---------------------GG----A----------------------------------- 4093
62_BC.CN.10.YNFL13           --------A--A--G--------C---A-AC----G--------------------C---GT---------A--------T--T---------A--AG-G-A----A---A--------------A--T---C-A----------------------------------- 4001
63_02A1.RU.10.10RU6637       --------A--A--G--------C-----GC---------------------------G-GT-G----C--A--T------CC----------A-CAG-G-A----A---T--------------AAAT---C-A--------------C-------------------- 4213
64_BC.CN.09.YNFL31           --------A--A--G--------C---A-AC-------------------------C---GT---------A--------T--T--------TA-----G-A----A---A--T-----------A-AT--T--------------C----------------------- 4026
65_cpx.CN.10.YNFL01          --------A--A-----------C---A-AC-------------------------C---GT---------A--------T--T---------A--A--G-A----A---A--------------A--T---C-A--------------------------------C-- 4043
67_01B.CN.11.ANHUI_HF115     --G--------A-----------C---A-C-----G--------------------G--G-----------A---------------------A--AG---A----A---------------------G-----------------C----------------------- 4015
68_01B.CN.11.ANHUI_WH73      --G--------A-----------C---A-C-----G--------------------G--G-----------A---------------------A---G-G-A----A---T-----------------G-----------------C----------------------- 4006
70_BF1.BR.10.10BR_PE004      -----------------------C-----C------C----------------------------------A-----C--T-----C------A-CAA-G----A-A---------C--------A--G-----------------T--------------------C-- 4128
71_BF1.BR.10.10BR_PE008      --G--------------------C---A-C-----------------------------------------A---------------------A------------A---------------------------------------C----------------------- 4336
72_BF1.BR.10.10BR_MG002      --------A--A--------C--C--CA-CC------------------------------T---------A---------------------A-------A----A---------C-----------G--------A--------T----------------------- 4180
O.BE.87.ANT70                --------A--A-----T--C--C--C--G------C-G--T-C---------T--T---GT---------A-----C--GCCT-----T--AA--A-A--TA-G-A---T--A-----------CAAC--AC-A--T--G--------A--A--T-----A-------- 4726
O.CM.98.98CMA104             --------A--A-----T--C-----C--G------C------C---------T--T---GT---------A--------GCCT--------AA----A--CA-G-A---T--A-----------CAAC--AW-A--T--G--------A--A--T-----A-------- 4158
O.CM.98.98CMABB141           -----------A-----T--T--C--C--G------C----T-C---------T--T--GGT---------A--------GCCT-----T--AA----A--TA-G-AA--T--A---------A-TAAC--AC-A--T--G--------A--A--T-----A-------- 4157
O.CM.98.98CMABB212           ---T----A--------T--C-----C--G-------------C---------T--C---GT---------A--------GCCT-----T--AA----AG-CA-G-AA--T--A---------A-A--C--AC-A--T--G-----G--A--A--T-----A-------- 4162
O.CM.98.98CMU5337            --RT----A--A--G--T--C--C--C--G-----G-----T-C---------T--T----T---------A-----C--GCCT-----T--AA----A-TCA-G-A---T------------A-CAAC--A--A--T--G--------A--A--T-----A-------- 4159
O.CM.99.99CMU4122            --------A--A-----T--C--C--C--G-----------T-C---------T--T---GT---------A-----C--GCCT--------AA----A--CA-G-A---T--A---------A-CAAC--A--A--T--G--------A--A--T-----A-------- 4157
O.FR.92.VAU                  --G-----A--------T--C--C--C--G-------------C---------TA-T----T-C-------A--------GCCT--C--T--AA----AG-CA-G-AA--T--A-----------AAAC--AC-A--T--G--------A--A--------A-------- 4238
O.GA.11.11Gab6352            --------A--A-----T--T--C--C--G-----------CTC---------T--T---GT---------A--------GCCT-----T--AA--A-A--CA-G-A---T--A---------A-TAAC--A--A--T--G--------A--A--T-----A-------- 3883
O.SN.99.99SE_MP1299          --------A--------T--C--C--C--G-----------T-C---------T--T--GGT---------A--------GCCT-----T--AA----AG-CA-G-A---T--A-----------CAAC--A--A--T--G--------A--A--T-----A-------- 4725
O.US.10.LTNP                 --G--------A-----T--C--C--C--GC-----C----T-C---------T--T----T---------A-----C--ACCT-----T--AA----A--TA-G-A---T--A-----------CAAC--AC-A--T--------G--A--A--T-----A-------- 4647
N.CM.02.DJO0131              --------A------------------A-------------T------------------GTT-----C-----T-----ACC---------TA----C--T--A-AA--A--------------AAAT----CA--------C--G--A-----------T-------- 4178
N.CM.02.SJGddd               -----T--A------------------A-------------T------------------GTT-----C-----T-----ACC---------TA----C--T--A-AA--A--------------AAAC---C-A-----G-----G--A-----------T-------- 4162
N.CM.04.04CM_1015_04         --------A--K---------------A-------G-----T------------------GTT-----C-----T-----ACC---------TA----C--T--A-AA--A--------------AA-T-----A-----------G--A-----------T-------- 4177
N.CM.06.U14296               --------A------G-----------A-------------T------------------GTT-----C-----T-----ARC---------TA----C--T--A-AA--A--------------AAAT-----A--------C--G--A--A--------T-------- 4175
N.CM.06.U14842               --------A--A---------------A-------G-----T-C----------------GTT-----C-----T-----ATC---------TA----C--T--A-AA--A--------------A--T-----A--------C--G--A-----------T-------- 4179
N.FR.11.N1_FR_2011           --------A------------------A-G-----G-----T------------------GTT-----C-----T-----ACC---C-----TA----C--T--A-AA--A--------------AAAC--T--A--------C-----A-----------T-------- 4073
P.CM.06.U14788               --------A--A---G-T-----C---G-GC-C-----G---TC---------T------GT------C--A--------GCCT--C--T--TA-C---G-A--A-AA--A------------CTTAAT--A-CT--T--G-----G--A-------------------- 4143
P.FR.09.RBF168               --------A--R-----T-----Y---G-GC-C-----G----C---------T------GTG-----C--A--------GCCT--C--Y--TA-Y---G-A--A-AR--A------------CTTAAT--A-CT--T--G-----G--A--------------G----- 4707
CPZ.CD.06.BF1167             --------CA----------T-TC--CA--------C------------------------AG--------A--T-----AGC------T--TA-----G-A--A-A---A--A--C--------AAAC---C-A--------------------------A-------- 4774
CPZ.CM.05.SIVcpzMT145        --------A------------------G-CC-T--G-----------------------G--T-----C--A--------GCCT--------TA--AA-G-A--A-A---A--A-----------A--T--T-GA-----------------A-----C--A-------- 4214
CPZ.GA.88.GAB1               --G-----A--A--------T-----C--G--------------T-----------------T--T--C--A--T-----GCCA-----T--AA-----G-A--C-A---T--------------A-AC--------------------A-----T-----A--G----- 4733
CPZ.TZ.06.SIVcpzTAN13        --------AA-A--------T--C--CA-C--T-----G-----G--------T-------G---------A--------ACCA-----T--TA-CA-AG-A--A-A---A--A-----------ACA----C-A--T--------C--A--T--T-------------- 4313
CPZ.US.85.US_Marilyn         --------A--A--G-----C------A-C-----------------------T------GTG--T--------T-----ATCT--C--T--AA--AG---A-----A--A--T--C--------A--C--AC-A--A--G-----------A--------A-------- 4732
GOR.CM.07.SIVgor2139_287     --------------------C--C--C--GC-C--G-------CT--G-------------T------C--A--------GCCT-----T--TA----C--A--C-A---T------------CTCAATG-A-CA--T--G--------------T-------------- 4209
GOR.CM.07.SIVgorCP2135con    --------------------C--C--C--GC-C--G-------CT----------------T---------A-----C--GCCT-----T--TA----C--A--C-A---T------------CTCAATG-A-CA--T--G--------------T-------------- 4208
```

HIV-1 Genomes

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

B.FR.83.HXB2 CAACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAAATCCACTTTGGAAAGGACCAGCAAAGCTCCTCTGGAAAGGTGAAGGGGCAGTAGTAATACAAGATAATAGTGACATAAAAGTAGTGCCA 5012
Pol A T D I Q T K E L Q K Q I T K I Q N F R V Y Y R D S R N P L W K G P A K L L W K G E G A V V I Q D N S D I K V V P
A1.AU.03.PS1044_Day0 4220
A1.CH.03.HIV_CH_BID_V3538 4217
A1.CY.08.CY236 4223
A1.ES.05.X1608_8 4460
A1.KE.11.DEMA11KE001 4351
A1.RU.11.11RU6950 4536
A1.RW.11.DEMA111RW002 4363
A1.SN.01.DDI579 4210
A1.UG.11.DEMA110UG001 4367
A1.ZA.04.04ZASK162B1 4453
A2.CM.01.01CM_1445MV 4211
A2.CY.94.94CY017_41 4369
B.BR.10.10BR_MG029 4474
B.CA.07.502_1191_03 4439
B.CH.08.M2_0803101_NFLG8 4481
B.CN.10.DEMB10CN002 4384
B.ES.10.DEMB10ES002 4374
B.FR.11.DEMB11FR001 4258
B.GB.05.MM45d213_GN1 4413
B.HK.06.HK003 4382
B.HT.05.05HT_129389 4213
B.JP.12.DEMB12JP001 4399
B.KR.07.HP_18_07JHS10_3909 4676
B.PE.07.502_0525_wg5 4477
B.RU.11.11RU2ln 4583
B.TH.08.MERLBDTRC10 4245
B.US.11.ES38 4947
C.AR.01.ARG4006 4196
C.BR.07.DEMC07BR003 4344
C.BW.00.00BW5031_1 4383
C.CN.10.YNFL19 4365
C.CY.09.CY260 4235
C.ES.08.X2363_2 4443
C.ET.02.02ET_288 4214
C.IN.09.T125_2139 4988
C.KE.00.KER2010 4193
C.MW.09.703010256_CH256.w96 4992
C.TZ.08.707010457_CH457.w8 4986
C.YE.02.02YE511 4187
C.ZA.08.705010534_CH534.w12 5002
C.ZA.10.DEMC10ZA001 4352
C.ZM.11.DEMC11ZM006 4361
D.CM.10.DEMD10CM009 4387
D.CY.06.CY163 4220
D.KE.11.DEMD11KE003 4362
D.KR.04.04KBH8 4971
D.SN.90.SE365 5029
D.TZ.01.A280 4213
D.UG.10.DEMD10UG004 4350
D.UG.11.DEMD11UG003 4367
D.YE.02.02YE516 4223
D.ZA.90.R1 4358
F1.AO.06.AO_06_ANG32 4223
F1.AR.02.ARE933 4306
F1.BR.10.10BR_RJ015 4485
F1.CY.08.CY222 4205
F1.ES.02.ES_X845_4 4464
F1.RO.96.BCI_R07 5036
F1.RU.08.D88_845 4471
F2.CM.02.02CM_0016BBY 4202
F2.CM.10.DEMF210CM001 4276
F2.CM.10.DEMF210CM007 4359
G.BE.96.DRCBL 4969
G.CM.10.DEMG10CM008 4393
G.CN.08.GX_2084_08 4252
G.CU.99.Cu74 4616
G.ES.09.X2634_2 4495
G.GH.03.03GH175G 5052
G.KE.09.DEMG09KE001 4387
G.NG.09.09NG_SC62 4204
G.ZA.01.TV546 4214
H.BE.93.VI991 4401
H.CF.90.056 4359
H.GB.00.00GBAC4001 4530
J.CM.04.04CMU11421 4518
J.SE.93.SE9280_7887 4326
J.SE.94.SE9173_7022 4327
K.CD.97.97ZR_EQTB11 4204
K.CM.96.96CM_MP535 4208
U.CA.01.TV749 5029
U.CA.99.TV721 5028
U.CD.83.83CD003_Z3 4507
U.CD.90.90CD121E12 4211
U.CY.05.CY090 4211
U.CY.08.CY223 4229
U.ES.10.DEURF10DZ001 4288
U.GR.99.99GR303 4283
U.NL.95.U_NL_95_H10986_D1 4897
01_AE.AF.07.569M 4218
01_AE.CF.90.90CF11697 4961
01_AE.CN.10.YNFL03 4390
01_AE.HK.04.HK001 4374
01_AE.IR.10.10IR.THR48F 4202
01_AE.JP.x.JRC77AE 5009
01_AE.TH.09.AA090a_WG11 4435
01_AE.TH.90.CM240 4586
01_AE.US.05.306163_FL 4203
01_AE.VN.98.98VNND15 4267

```
B.FR.83.HXB2                 CAACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAAATCCACTTTGGAAAGGACCAGCAAAGCTCCTCTGGAAAGGTGAAGGGGCAGTAGTAATACAAGATAATAGTGACATAAAAGTAGTGCCA 5012
Pol                          A__T__D__I__Q__T__K__E__L__Q__K__Q__I__T__K__I__Q__N__F__R__V__Y__Y__R__D__S__R__N__P__L__W__K__G__P__A__K__L__L__W__K__G__E__G__A__V__V__I__Q__D__N__S__D__I__K__V__V__P_
02_AG.CM.10.DE00210CM013     --T----T-------T-------C-------------------------------------------------T------G-C--CA-------------------A--A--------------------------------G--C--------T-----------A--- 4379
02_AG.CY.09.CY256            --T--------------------C--------------------------C-----------------------------G-C--CA----------------------A--T-----------------------------G--C--------T-----G-----A--- 4208
02_AG.ES.06.P1423            --T----T---------------C---------------T----------------------------------------G-----A-------------------A--A--------------------------------G--C--------T-----G-----A--- 4449
02_AG.GW.05.CC_0048          --T----T---------------C--------------------------------------------------------G-C--CA-------------------A--A--------------------------------G--C--------T-----G-----A--- 4362
02_AG.KR.12.12MHR9           --T----T---------------C----------T----T----------------------------G-----------C-C--CA-------------------A--A--------------------------------G--C--------T-----G-----A--- 4671
02_AG.LR.x.POC44951          --T----T-----------G---C-----------------C--------------------------------------G-C--CA-------------------A--A--------------------------------G--C--------T-----G-----A--- 5007
02_AG.NG.09.09NG_SC61        --T----T---------------C-----------------C--------------------------------------G-C---A-------------------A--A--------------------G--------------C----AC--------G--------- 4204
02_AG.NG.x.IBNG              --T----T---------------C----------G----T----------------------------------------G-C--CA-------------------A--A--------------------------------G--C--------T-----G-----A--- 4537
02_AG.SE.94.SE7812           --T----T---------------C----------G---------------------------------------------G-C--CA-------------------A--A--------------------------------G--C--------T-----G-----A--- 4384
02_AG.SN.98.98SE_MP1211      --T----T---------------C-------------------T--------A---------------------------G-C--CA-------------------A--A-----------------------------------C--------T-----G-----A--- 4209
03_AB.RU.97.KAL153_2         ---------------------------------------T------------------------------A---------G-----A-------------------A--A-----------------------G--------G--C----AC--T-----------A--- 4238
04_cpx.CY.94.94CY032_3       --T----T---------------C--------------------------------------------------------G-A---A-------G-----------A--A--------------------------------G--C--C-----T--C--------A--- 4378
05_DF.BE.x.VI1310            -------------------------------------C-T----------------------------------------G-C---G----------------------A--------------G--A--------C--------C--------A-----G-----A--- 4394
06_cpx.AU.96.BFP90           --T----T---------------C---------------T----------------------------------------G-C---A-------------------A--A-----------------------------------C--------A-----G-----A--- 5040
07_BC.CN.98.98CN009          ----------------------------------G-----------------------------------A---------G-C--CAGC--------------C--A--A--------------------------------------------------G-----A--- 4344
08_BC.CN.97.97CNGX_6F        ------------------G--------------------T------------------------------A---------G-C--CA----------------C--A--A--------------------------------------------------G-----A--- 4192
09_cpx.GH.96.96GH2911        --T----T---------------C-------------C------------------------------------------G-C---A-------------------A--A--------------------------------G--C--A-----T-----G-----A--- 4214
10_CD.TZ.96.96TZ_BF061       -----------------------C---------------T------------------------------A---------GGC--TA----------------C--A--A-----------------------------------A--C-----------G-----A--- 4391
11_cpx.CM.95.95CM_1816       --------T--------------C----------------------------A---------------------------G-C-----------------------A--A--------------------------------G--C--------T---------A-A--- 4376
12_BF.AR.99.ARMA159          ------------------G--------------------T---------------C------------------------G-C---G----------------------A--------------------------C--------CG-------A-----G-----A--C 5017
13_cpx.CM.96.96CM_1849       --------T-------A------C---------------T----------------------------------------G-C---A-------------------A--G-----------------------------------C--------A-----------A--- 4417
14_BG.ES.05.X1870            --T----T--------A------C------------------------------------------T-------------G-C---A-------------------A--A-----------------------------------C----AC--A-----G--------- 4479
15_01B.TH.99.99TH_MU2079     ----------T-C--------------C---C---TCC...---------------------------------------G-C---A-------------------A--A-----------------------------------C--------T-----------A--- 4408
16_A2D.KR.97.97KR004         -------T---------------C------G-------------------------------------------------G-----A-------------------A--------------------------------------C--------T-----G-----A--- 4375
17_BF.AR.99.ARMA038          ------------------G--------------------T---------------C-----C------------------G-C---G----------------------A--------------------------C--------C--------A-----G-----A--- 4226
18_cpx.CU.99.CU76            --T----A-----------------C------------------------------------------------------G-C---A-------------------A--A--------------------------------G--C--C-----A--T--G-----A--- 4317
19_cpx.CU.99.CU7             --T--------------------C-----------------------TC---C---------------------------G-C--CA-------------------A--A-----------------------------------C--------------------T--- 4238
20_BG.CU.99.Cu103            --T----T---------------C--------------------------------------------------------G-C---A-------------------A--A-----------------------------------C----AC--A-----G-----A--- 4478
21_A2D.KE.99.KER2003         -------T----------------------G-----------------------------------T---A---------G-----A-------------------A--------------------------------------C--------T-----G-----A--- 4217
22_01A1.CM.01.01CM_0001BBY   ----------------------------------------G---------------------------------------G-C---A-------------------A--A-----------------------------------C--------T-----------A--- 4220
23_BG.CU.03.CB118            -----------------------C--------------------------------------------------------G-C---A-------------------A--A-----------------------G-----------C-G--AC--A-----------A--- 4455
24_BG.ES.08.X2456_2          --T----T---------------C------------------------------------------T---A---------G-C---A-------------------A--------------------------------------C----AC--A-----G-----A--- 4463
25_cpx.CM.02.1918LE          --T--------------------C---------------T------------C---------------------------G-----A----------------------A--------------------------------G--C--------A-----------A--- 4220
26_AU.CD.02.02CD_MBTB047     ------------------G--------------------T----------------------------------------G-----A-------------G-----A--A--T--------------------------------C--------T-----G-----A--- 5017
27_cpx.FR.04.04CD_FR_KZS     --T----T----------G---------------------------C-----------------------A---------G-C---G----------------T--A--G-----------------------------------C--------A-----G-----A--- 5003
28_BF.BR.99.BREPM12609       ------------------G---------------------------C---------------------------------G-C---G----------------------A--------------------------C--------------------------------- 4378
29_BF.BR.01.BREPM16704       ---------------------------------------T----------------------------------------G-C---G----------------------A--------------------------C--------C--------A-----G-----A--- 4413
31_BC.BR.04.04BR142          -----------------------C-------G--------------------------------------A---------G----TA----------------C--A--A---------------------------C----------C-----------G-----A--- 4484
32_06A1.EE.01.EE0369         --T----T---------------C----------------T---------------------------------------G-C---A----------------T--A--A--T-----------------------------G--C--------A-----G-----A--- 4647
33_01B.ID.07.JKT189_C        --------------------------------------------------------------------------------G-C---A-------------------A--A--------------------------------G--C--------T-----------A--- 4313
34_01B.TH.99.OUR2478P        --------------------------------------------------------------------------------G-C---A-------------------A--A-----------------------------------C--------T-----------A--- 4214
35_AD.AF.07.169H             ----------------------GC--------------------------------------------------------G-------------------------A--A--------------------------------G--C--------T-----G--------T 4212
36_cpx.CM.00.00CMNYU830      --T----T---------------C--------------------------------------------------------G-C---A-------------------A--A-----------------------------------C--------T-----------A--- 4217
37_cpx.CM.00.00CMNYU926      --T----T--------C------C--------------------------------------------------------G-C--CA-------------------A--A--------------------------------------------T-----------A--- 4202
38_BF1.UY.03.UY03_3389       --T---------------G------T-------------T---------------C-------T--C-------------G-C---G----------------------A--------------------------C--------C--------A-----G-----A--- 4435
39_BF.BR.04.04BRRJ179        -----------------------C--------------------------------------------------------G-C--------------------------T--------------------------------------C-----A--------------- 4489
40_BF.BR.05.05BRRJ055        ------------------G-----------------------------------------------T-------------G-----G-------------------A--A-----------------------------C--------C--------A-----G-----A--- 4495
42_BF.LU.06.luBF_18_06       --------------------------------------------------------------------------------G-C---A----------------------A--------------------------C-----------C--------------------- 4566
43_02G.SA.03.J11223          --T----TC--------------C--------------------------------------------------------G-C---A-------------------A--A--T-----------------G-----G--------C--------A-----G-----A--- 4537
44_BF.CL.00.CH80             ------------------G--------------------T---------------C------------------------G-C---G----------------------A--------------------------C--------C--------A-----G-----A--- 4454
45_cpx.FR.04.04FR_AUK        -------------G------------------------------------------------------------------G-C---A----------------C--A--A-----------------------------------C--------T-----G--------- 4998
46_BF.BR.07.07BR_FPS625      ------------------G--------------------T---------------C------------------------G-C---G----------------------A--------------------------C--------C--------A-----G-----A--- 4450
47_BF.ES.08.P1942            --------C-----------------------------------------------------------------------G-C---G-------------------A--A-----------------------G--C--------C--------A--------------- 4451
48_01B.MY.07.07MYKT021       --------------------------------------------------------------------------------G-C---A-------------------A--A--------------------------------------------T-----------A--- 4217
49_cpx.GM.03.N26677          ----------------------------------------------------------------------A---------G-C--TA----------------C--A--A--T-----------------------------------A-----------G-----A--- 4434
50_A1D.GB.10.12792           --T----------------------------------C------------------------------------------G----------------------------A--G-----------------------------G--C--------T-----G---C-A--- 4419
51_01B.SG.11.11SG_HM021      -----------------------C---------------------------------------------------A----G-----A----------------------T--------------------------G-----------------T--------G--A--- 4219
52_01B.MY.03.03MYKL018_1     --------------------------------------------------------------------------------G-C---A-------------------A--A-----------------------------------C--------T-----------A--- 4326
53_01B.MY.11.11FIR164        --------------------------------------------------------------------------------G-C---A-------------------A--A-----------------------------------C--------T-----------A--- 4352
54_01B.MY.09.09MYSB023       ------------------G-------------------------------------------------------------G-------------------------------T--------------------------------C--------T-----------A--- 4414
55_01B.CN.10.HNCS102056      --------------------------------------------------------------------------------G-----A-------------------A--A-----------------------------------C--------T-----------A--- 4322
56_cpx.FR.10.URF5_patient_A  --T----T------T--------C--------------------C-----------------------------------G-C--C--------------------A--A--------------------------------G--C--------T-----G--G--A--- 4363
57_BC.CN.09.09YNLX19sg       --T------------------------------------T----G-------------------------A---------G-C--CA----------------C--A--A--------------------------------------------------G-----A--- 4196
58_01B.MY.09.09MYPR37        --------------------------------------------------------------------------------G-C---A-------------------A--G-----------------------------------C--------T-----G-----A--T 4357
59_01B.CN.09.09LNA423        -------------------------------G------------------------------------------------G-C---A-------------------A--A-----------------------------------C--------T-----------A--- 4196
60_BC.IT.11.BAV499           -------------------G---C---------------T------------------------------A---------G--G-TA----------------C--A--A---------------------------C-C--------C-----T--C--G-----A--- 4969
61_BC.CN.10.JL100010         ----------------------------------------------------------------------A---------G-C--CA----------------C--A--A--------------------------------------------------G-----A--- 4433
62_BC.CN.10.YNFL13           ---------------------------------------T----G-----------------------------------G----CA----------------C--A--A--------------------------------G-----G-----------G-----A--- 4341
63_02A1.RU.10.10RU6637       ----------------------------------------------------------------------A---------G-----A-------------------A--A-----------------------G--------G--C----AC--T-----------A--- 4553
64_BC.CN.09.YNFL31           ------------------G---------------G--------------------C------------------------G-C--CAG---------------C--A--A--------------------------------------------------G-----A--- 4366
65_cpx.CN.10.YNFL01          --------------------------------------------------------------------------------G-C---A-------------------A--A--------------------------G--------C--------T-----------A--- 4383
67_01B.CN.11.ANHUI_HF115     --------------------------------------------------------------------------------G-C---A-------------------A--A-----------------------------------C--------T-----------A--- 4355
68_01B.CN.11.ANHUI_WH73      --------------------------------------------------------------------------------G-C---A-------------------A--A-----------------------------------C--------T-----------A--- 4346
70_BF1.BR.10.10BR_PE004      -----------------------Y--------------------------------------------------------G-C---A-------------------R--A--------------------R-----C--------C--------AC----G-----A--- 4468
71_BF1.BR.10.10BR_PE008      ---------------------------------------T---------------------------------A------G-C---G----------------------A-----------------A--------C--------C--------A-----G--G--A--- 4676
72_BF1.BR.10.10BR_MG002      -----------------------C----------------------C---------------------------------G-C---G----------------------A--------------------------C--------------------------------- 4520
O.BE.87.ANT70                --T--C-A--------A-C-------------------TT----N----C--A--------C--------A---------G-C--TA-C-----------G---C-------G-----------G--A--------C-----------GG-A-----T--G-----A--- 5066
O.CM.98.98CMA104             --T--C-A--------A-C----C--------------TT---------------------C--------A---------G-C--TA-------------G---C-A-----G-----------G--A--------C-----------AG-A-----T--------A--- 4498
O.CM.98.98CMABB141           --T--C----------A-C----C--------------T----------------------C--------A---------G-C--CA-------------G--CC-------G-----------G--A--------C---------C-AG-A-----T--------A--- 4497
O.CM.98.98CMABB212           --T--C-A--------A-C----C--------------TT---------------C--------------A--T------G----TA-------------G---C-A-----G-----------G-----------C-----------AG-A-----T--------A--- 4502
O.CM.98.98CMU5337            --T--C-A--------A-C-------------------TTT-----------A----A---C--------A---------G----TA-------------G---C-A-----G-----------R--A--------C-----------AG-A-----Y--------A--- 4499
O.CM.99.99CMU4122            --T--C-A--------A-C-------------------TTC-----------A----A---C--------A---------G----TAC------------G---C-------G-----G--------A--------C-----------AG-A-----T--G-----A--- 4497
O.FR.92.VAU                  --T--C-A--------A-C----C--------------TT---------------------C--------A---------G-C--TA-------------G---C-------G-----------G--A--------C-----------AG-A-----T--------A--- 4578
O.GA.11.11Gab6352            --T--C----------A-C----C--------------TT---------------------C--------A---------G-C--TA-------------G---C-------G-----------G--A--------C-----------AG-A-----T--------A--- 4223
O.SN.99.99SE_MP1299          --T--C-A--------A-C-------------------TTT-----------A----A---C--------A---------G----TA-------------G---C-------G-----------G--A--------C-----------AG-A-----T--G-----A--- 5065
O.US.10.LTNP                 --T--C-A--------A-C-------------------TT---------------------C--------A---------G-C--TA-------------G---C-------G-----------G--A--------C-----------AG-A-----T--G-----A--- 4987

N.CM.02.DJO0131              ----------------A-C-A---------C-------TT---G--------------------------A---------G----CA----------------C--A--T--G--------A-----------G-----C--------CG-A--T-----------C--- 4518
N.CM.02.SJGddd               -------------G--A-C-A---------C-------TT-C--G-------------------------A---------G----CA----------------C--A--T--G--------A-----------G-----C--------AG-G--T-----------C--- 4502
N.CM.04.04CM_1015_04         -------------G--A-C-A---------C-------TT----G-------------------------A---------G----CA----------------C--A--G--G--------A-----------G-----C--------CG-G--T-----------C--- 4517
N.CM.06.U14296               -------------G--A---A-C-------C-------TT----G-------------------------A---------G----CA----------------C--A--T--G--------A-----------G-----C--------CG-G--T-----------C--- 4515
N.CM.06.U14842               -------------G--A-------------C-------TT----G---------------------T---A---------G-C--TA----------------C--A--T--G--------A-----------G-----C--------CG-G--T-----------C--- 4519
N.FR.11.N1_FR_2011           -------------G--A-C-A---------C-------TT----G-------------------------A---------G-C--CA----------------C--A--T--G--------A-----------G-----C--------CG-G--T-----------C--- 4413

P.CM.06.U14788               -T---C-R--------A-C-------------------TT-----------------A---C--------A---------G-C--TA----------------G-CA-----G-----------------------C--------C--AG-A--T--T--G-----C--T 4483
P.FR.09.RBF168               -T---C-A--------A-C-------------------TT---------------------C--------A---------G-C--TA----------------G-C---T--G-----------------------C--------C--AG-A--T--T--------C--T 5047

CPZ.CD.06.BF1167             --------------------------------------TTT---C----------------C-----T------TCA---G-C---G-G--------------CCGT-----G-----------------------C---A-G--GGGAGAA--A--C--------A--- 5114
CPZ.CM.05.SIVcpzMT145        ----------------A------C--------------TT----------C-A--------C--------A-------A-G-----A----------------G--A-----G--------------------G-----T------C-GGAC--A--------G-----T 4554
CPZ.GA.88.GAB1               ----------------A-GC------------------TT----G-------A-----------------A---------G-C---A-------------G----CCT-A--G-----------------------G--C--G---C-AG-G--AC----G-----A--- 5073
CPZ.TZ.06.SIVcpzTAN13        -C---A-T-----G--AC--C--C--------------TT----G----------------------T------GC----G-C--GA-------G--------GCG---A--G--------------A--------T---A----GGG-GAG-----C--------A--C 4653
CPZ.US.85.US_Marilyn         --T----AC-------AG-CTT----------------TT----G----------------C------------------G-----A----------------C--A--T--G--------------------------CA-G--A--CGAG--GG-T--------A--C 5072

GOR.CM.07.SIVgor2139_287     -C---C-A--------A-C-------------------TT---------------------C--------A---------G-C--TA----------------G-C------G-----------------T-----C--T--G--C--AG-G-----T--------C--T 4549
GOR.CM.07.SIVgorCP2135con    -C---C-A--------A-C-------------------TT---------------------C--------A---------G-C--TA----------------G-C------G-----------------T-----C--T--G--C--AG-G-----T--------T--T 4548
```

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

Vif start
Pol, Gag-Pol, and p31 integrase end
B.FR.83.HXB2 AGAAGAAAAGCAAAGATCATTAGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAGAACATGGAAAAGTTTAGTAAAACACCATATGTATGTTTCAGGGAAAGCTAGGGGATGGTTTTATAGACATCACTATGAAAGCCCTC 5182
Vif M E N R W Q V M I V W Q V D R M R I R T W K S L V K H H M Y V S G K A R G W F Y R H H Y E S P
Pol R R K A K I I R D Y G K Q M A G D D C V A S R Q D E D *
02_AG.CM.10.DE00210CM013 TAGG- 4549
02_AG.CY.09.CY256 TAGG- 4378
02_AG.ES.06.P1423 TAGA- 4619
02_AG.GW.05.CC_0048 CTAGG- 4532
02_AG.KR.12.12MHR9 TAA- 4841
02_AG.LR.x.POC44951 TAGG- 5177
02_AG.NG.09.09NG_SC61 TAGG- 4374
02_AG.NG.x.IBNG TAGG- 4707
02_AG.SE.94.SE7812 AG-- 4554
02_AG.SN.98.98SE_MP1211 TAGG- 4379
03_AB.RU.97.KAL153_2 AGGA 4408
04_cpx.CY.94.94CY032_3 TAGG- 4548
05_DF.BE.x.VI1310 TAGAA 4564
06_cpx.AU.96.BFP90 C--AA-- 5210
07_BC.CN.98.98CN009 AGA- 4514
08_BC.CN.97.97CNGX_6F AGA- 4362
09_cpx.GH.96.96GH2911 T--AA- 4384
10_CD.TZ.96.96TZ_BF061 AGA- 4561
11_cpx.CM.95.95CM_1816 AGG- 4546
12_BF.AR.99.ARMA159 AGG- 5187
13_cpx.CM.96.96CM_1849 C--AA- 4587
14_BG.ES.05.X1870 AGG- 4649
15_01B.TH.99.99TH_MU2079 -AG- 4578
16_A2D.KR.97.97KR004 ATAA 4545
17_BF.AR.99.ARMA038 AGGA 4396
18_cpx.CU.99.CU76 TAGA- 4487
19_cpx.CU.99.CU7 AGGA 4408
20_BG.CU.99.Cu103 AGA- 4648
21_A2D.KE.99.KER2003 AGA- 4387
22_01A1.CM.01.01CM_0001BBY AGG- 4390
23_BG.CU.03.CB118 G--A-A 4625
24_BG.ES.08.X2456_2 AGA- 4633
25_cpx.CM.02.1918LE AGA- 4390
26_AU.CD.02.02CD_MBTB047 AGG- 5187
27_cpx.FR.04.04CD_FR_KZS A-- 5173
28_BF.BR.99.BREPM12609 A--- 4548
29_BF.BR.01.BREPM16704 AGA- 4583
31_BC.BR.04.04BR142 AGAA 4654
32_06A1.EE.01.EE0369 C---AC- 4817
33_01B.ID.07.JKT189_C AG- 4483
34_01B.TH.99.OUR2478P T-AG- 4384
35_AD.AF.07.169H AGGA 4382
36_cpx.CM.00.00CMNYU830 AGA- 4387
37_cpx.CM.00.00CMNYU926 AGGG 4372
38_BF1.UY.03.UY03_3389 AGGA 4605
39_BF.BR.04.04BRRJ179 A--- 4659
40_BF.BR.05.05BRRJ055 A--- 4665
42_BF.LU.06.luBF_18_06 AT-- 4736
43_02G.SA.03.J11223 AGGA 4707
44_BF.CL.00.CH80 AGG- 4624
45_cpx.FR.04.04FR_AUK C-AGG- 5168
46_BF.BR.07.07BR_FPS625 C-AGG- 4620
47_BF.ES.08.P1942 AGG- 4621
48_01B.MY.07.07MYKT021 -AGA 4387
49_cpx.GM.03.N26677 AGA- 4604
50_A1D.GB.10.12792 AGGA 4589
51_01B.SG.11.11SG_HM021 AG- 4389
52_01B.MY.03.03MYKL018_1 AG- 4496
53_01B.MY.11.11FIR164 AGA 4519
54_01B.MY.09.09MYSB023 T-ATG 4584
55_01B.CN.10.HNCS102056 AG- 4492
56_cpx.FR.10.URF5_patient_A TAGG- 4533
57_BC.CN.09.09YNLX19sg AGA- 4366
58_01B.MY.09.09MYPR37 AG- 4527
59_01B.CN.09.09LNA423 AG- 4366
60_BC.IT.11.BAV499 AGA- 5139
61_BC.CN.10.JL100010 AGA- 4603
62_BC.CN.10.YNFL13 AGAA 4511
63_02A1.RU.10.10RU6637 TAGG- 4723
64_BC.CN.09.YNFL31 AGA- 4536
65_cpx.CN.10.YNFL01 AGA- 4553
67_01B.CN.11.ANHUI_HF115 AG- 4525
68_01B.CN.11.ANHUI_WH73 AG- 4516
70_BF1.BR.10.10BR_PE004 AGG- 4638
71_BF1.BR.10.10BR_PE008 AGG- 4846
72_BF1.BR.10.10BR_MG002 TA-- 4690
O.BE.87.ANT70 TC-AGAA 5236
O.CM.98.98CMA104 MTCTAGAA 4668
O.CM.98.98CMABB141 TCTAGAA 4667
O.CM.98.98CMABB212 TT-AGGA 4672
O.CM.98.98CMU5337 KSTAGAA 4669
O.CM.99.99CMU4122 TT-AGAA 4667
O.FR.92.VAU TCTATAA 4748
O.GA.11.11Gab6352 CATAGAA 4393
O.SN.99.99SE_MP1299 TCTAGAA 5235
O.US.10.LTNP CTAGAA 5157
N.CM.02.DJO0131 CAAA-- 4688
N.CM.02.SJGddd CAAGG- 4672
N.CM.04.04CM_1015_04 TCAAGG- 4687
N.CM.06.U14296 TAAA-- 4685
N.CM.06.U14842 CARA-- 4689
N.FR.11.N1_FR_2011 CAAA-- 4583
P.CM.06.U14788 G-CAA-C- 4653
P.FR.09.RBF168 G-CAGTC- 5217
CPZ.CD.06.BF1167 T-TAGAC- 5284
CPZ.CM.05.SIVcpzMT145 TG--A-CA 4724
CPZ.GA.88.GAB1 TCA---A 5243
CPZ.TZ.06.SIVcpzTAN13 G-CAGA-G 4823
CPZ.US.85.US_Marilyn A--TA-AA 5242
GOR.CM.07.SIVgor2139_287 GCAGTCA 4719
GOR.CM.07.SIVgorCP2135con GCAGTCA 4718

```
B.FR.83.HXB2                 ATCCAAGAATAAGTTCAGAAGTACACATCCCACTA...GGGGAT...GCTAGA............TTGGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGCATTTGGGTCAGGGAGTCTCCATAGAATGGAGGAAAAAGAGATATAGCACACAAGTA 5334
Vif                          H__P__R__I__S__S__E__V__H__I__P__L__.__G__D__.__A__R__.__.__.__.__L__V__I__T__T__Y__W__G__L__H__T__G__E__R__D__W__H__L__G__Q__G__V__S__I__E__W__R__K__K__R__Y__S__T__Q__V_
A1.AU.03.PS1044_Day0         ------A-G----C------------------T--...------...----A-............A-A---G---G------------------G----------A---T-----A--------T--G------------------CT---A---------------A-- 4542
A1.CH.03.HIV_CH_BID_V3538    ------A-G--------------------------...-----G...-----G............--A--GG---G------------A----------------A---------C--------T--G------------------C-G--A---------------A-- 4539
A1.CY.08.CY236               ------A-G--------------------------...A-----...----C-............--A---G---A------------G-----G----------A---T-----A--------T--G------------------CT---A---------------A-- 4545
A1.ES.05.X1608_8             ------A-G--------------------------...------...----A-............--A---G-G-T---------------A-------------A---T-----C--------T--G------------------CT---A---------------A-- 4782
A1.KE.11.DEMA11KE001         --------GC-------------------------...------...------............--A---G---G-----------------------------A---A-----C--------T--G----------------A-CT--GA---------------A-- 4673
A1.RU.11.11RU6950            --------G--------------------------...--A---...-----G............C-A---G---A-------------------G---------A---------C--------T--G-----------------AC-GGGA--G------------A-- 4858
A1.RW.11.DEMA111RW002        --------GA-------------------------...------...----C-............A-AA--G---A------------A--A-----------G-A---A-----A--------T--G------------------TT---A--C-----------GA-- 4685
A1.SN.01.DDI579              --------GC-------------------------...------...----TG............--A---G---G------------C-----CC--------CA------------A-----T--G-----------------AC-G--A---------------A-- 4532
A1.UG.11.DEMA110UG001        ------A-G-------------G------------...------...----A-............--A---G---G------------A--A-------------AG--T-----C--------T--G------------------CT---A------------------ 4689
A1.ZA.04.04ZASK162B1         ------A-G--------------------------...------...------............--A---G---G------------C-----GC---------A---------C--------T--G-C---------------AC-G--A--------T------A-- 4775
A2.CM.01.01CM_1445MV         ------A-G----------------------G---...--A--A...------............--AA--G---G------------------CC---------A---A-----C--------T--G------------------C-G--A-AG--C--T------A-- 4533
A2.CY.94.94CY017_41          -------------------------------G---...-----G...------............A-AA--G---G------------------C-T--------A---------C--------T---------------------C----C--G---CAT------A-- 4691

B.BR.10.10BR_MG029           -----------------------------------...------...----A-............--------------------------------------------T------------------------------------------------------------ 4796
B.CA.07.502_1191_03          ------A------------------T---------...------...--G-A-............---------G----------------------------------T--------A-----------T-------------------G------C------------ 4761
B.CH.08.M2_0803101_NFLG8     -----------------------------------...------...----A-............------G-------------------------------------------------C------------------------G---G--A-------------A-- 4803
B.CN.10.DEMB10CN002          -----------------------T-----------...------...----A-............--A--TG------------------------------------C---------------T----C-------------------GG------------------- 4706
B.ES.10.DEMB10ES002          -------GG----C---------------------...------...----A-............--A-------A---------------------------------T----------------------------------------G---C----A---------- 4696
B.FR.11.DEMB11FR001          -------G---------------------------...-----A...----AT............---------------------------------------------------------------------------------G---G------------------- 4580
B.GB.05.MM45d213_GN1         ----G--G--------------------T------...-----A...----A-............----------AG--------------------------------A-------------------------------------T---------C------------ 4735
B.HK.06.HK003                ------AG--------------------T------...------...-AG-C-............-----------------------------G---------------------A-------T---------------------------G-------A--------- 4704
B.HT.05.05HT_129389          ----------------------------A------...------...------............--------------------------------------------T----------------------------------------G------------------- 4535
B.JP.12.DEMB12JP001          -----------------------------------...----GA...----A-............----------T---------------------------------T--------------A-----T--------------A---C-------------------- 4721
B.KR.07.HP_18_07JHS10_3909   -----------------------------------...------...----A-............--A-----------------------------------------A------C----------------T-------------------------A---------- 4998
B.PE.07.502_0525_wg5         -----------------------------------...-----G...-AG-A-TTGGAGGAGAAAC-------------------------------------------------------------------------------A--GG-C-A---------------- 4811
B.RU.11.11RU21n              -------------------G------------T--...------...----A-............----------A-----------------------------A---------------C-------C---------------A---G-------------------- 4905
B.TH.08.MERLBDTRC10          -C--T------------------------------...------...---GA-............------G---T---------------AA----------------G--------A-----A----C---------------A--------G--C------------ 4567
B.US.11.ES38                 -------------C--------T------------...------...--G-AG............----------A---------------------------------------------------------------G----A-----G------------------- 5269

C.AR.01.ARG4006              ------A-G----------------T------T--...--A---...------............--A-------T-------------T----A--------------T--------------T--------------------ACTG-RA-A---------------- 4518
C.BR.07.DEMC07BR003          ------A------------------T--T---G--...-----G...----A-............--A---------------------T----A-------------CT--------------T--------TG----G-----ATTG-GA------------------ 4666
C.BW.00.00BW5031_1           ------A-G-----------------------T--...------...------............--A-------A-------------T-------------------T--------------T--------------------ACTG-GA-------A---------- 4705
C.CN.10.YNFL19               ------A-G-----------------------T--...--A--G...------............--A-------A------------GT----A----------A---T--------------T--------------------ATTG-GA-A-----A-------A-- 4687
C.CY.09.CY260                --------G-----------------------T--...-----G...-----C............--A-------G------------------A--------------A--------------T--------------------ATGG--A--------T--------- 4557
C.ES.08.X2363_2              ------A-------------A-----------T--...------...------............--AA------A-------------T----G-----------G--A-----C--------T--------------------ATTGGGA------------------ 4765
C.ET.02.02ET_288             ---------C----------------------T--...-----A...------............--AA--G---A-------------T-A--A--------------T--------------T--------------------ATTG-GAG------C---------- 4536
C.IN.09.T125_2139            ------A-G-----------------------T--...-----A...----A-............--A-------A------------GT----AC----------G--T--------------T--------------------ACTG-GA--C----A---------- 5310
C.KE.00.KER2010              ------A-G-------------------A---T--...-----G...------............--A-------A-------------T----A--------------T--------------T-----T-------------AATTG--A--G----A---------- 4515
C.MW.09.703010256_CH256.w96  ------A-------------------------T--...------...------............--A-------A-------------T----G-------------CT--------------T--------------------ACTG--A------------------ 5314
C.TZ.08.707010457_CH457.w8   ------A-G----------------T------T--...------...------............--A-------A------------GT----G--------------G--------------T--------------------ACTT-GA------------------ 5308
C.YE.02.02YE511              ------A-G-----------------------T--...-----G...------............--A-------A-------------T-------------------T--------------T--------------------ATTG-GA------------------ 4509
C.ZA.08.705010534_CH534.w12  ------A-G----------------T--T---T--...------...----A-............--A---G---A------------CT----G-------------C---------------T----CT---------------CTG-GAGA---------------- 5324
C.ZA.10.DEMC10ZA001          ------A---C------------T--------T--...------...--A---............--AA------A-------------T----A--------------A--------------T--------------------ATTG-GA-A---------------- 4674
C.ZM.11.DEMC11ZM006          ------A-G-----------------------T--...-----G...------............--A-------A-------------T----A--------------G--------------T--------------------ATTG-GA------------------ 4683

D.CM.10.DEMD10CM009          -C----A---------------G---------T--...--A---...----A-............C-----G---TC-----C--------------------------T------C-A-----A--------------------A----G------------------- 4709
D.CY.06.CY163                -C----A---------------G---------T--...--A---...----A-............C-----G--------------------A-------G--------T------C---------------------------------G------------------- 4542
D.KE.11.DEMD11KE003          -------------------------------G---...--A--A...----A-............C-----G---A---------------------------------A------C---------------------------------G------C--T--------- 4684
D.KR.04.04KBH8               -------------C---------------------...--A---...--A-A-............--A---G-G------------------A--------------A-G------C----------G--------T--------ACTG--A--G--C--A------A-- 5293
D.SN.90.SE365                -C----A---------------G-----T---T--...--A--A...------............C-----G---AG--------------------------------A------C--------------------------------G-------------------- 5351
D.TZ.01.A280                 -C----A---------------------T------...--A---...------............C----GG-------------------------------------A------C----------------------------A----G------------------- 4535
D.UG.10.DEMD10UG004          -C---------------------------------...--A--A...------............C-----G---A-------------------------------A-A--------------------------------------GGG---------T--------- 4672
D.UG.11.DEMD11UG003          ------A------------------T---------...--A--A...-----G............C-----G---G------------------G--------------A------C---------------------------------G------------------- 4689
D.YE.02.02YE516              -C--C-AG--------------G--T------T--...--A--G...----A-............C-----G---T----------------A----------------T------C----------------------------AC---G------------------- 4545
D.ZA.90.R1                   ------A----------------------------...--A--A...-AG---............C-----G---A-------------------------------A-A------C------CT--------------------A----G--------T---------- 4680

F1.AO.06.AO_06_ANG32         --------G-----------------------T--...A--A-G...-A--A-............C-A------------------------A----------------G------C----------------------------ACTGGG--AG--C--T------A-- 4545
F1.AR.02.ARE933              --------G--------------------------...-A---A...----A-............--A-------T------C-----C--------------------T------C----G------------------------C-NGG--AC-----G------A-- 4628
F1.BR.10.10BR_RJ015          ----------------------C------------...-ACA-A...----AC............A-A---G----------C-----G---------------------------------------------------------C--GG---G--C--A------A-- 4807
F1.CY.08.CY222               ------A-------------------------A--...--A--A...----T-............A-A---G-------------------------------------G------C----G-------C----------------C--GGA--G-----T------A-- 4527
F1.ES.02.ES_X845_4           -----------------------------------...-AT---...----A-............G-A-----------------------------------------T------C----------------------------AC-GGG------C--A------A-- 4786
F1.RO.96.BCI_R07             --------G--------------------------...-AA--A...----A-............--A---G-------------------------------------G------C-----------------------------C-GGG---G------------A-- 5358
F1.RU.08.D88_845             -----------------------------------...-A---A...----A-............A-A---G-----------------------C-------------T------C-----------------------------CTGGG---------G------A-- 4793
F2.CM.02.02CM_0016BBY        --------G----------------------T---...-A-A--GATT---A-............--A------GT---C-----------A-------------A---T--------------T---------------------C-G--A--G-----G--------- 4527
F2.CM.10.DEMF210CM001        -C------G--------------------------...-CA---GATT---C-............--A------GT---C-----------A-----------------T--------------A---------------------CGG--A--G-----G--------- 4601
F2.CM.10.DEMF210CM007        --------G----------------T---------...AA----GATT---A-............C-A------GT---C-----------A--C--------G-----T------C-------C---------------------C-G--A-RG-----G--------- 4684

G.BE.96.DRCBL                -C----A-G--------------------------...--A---...----A-............C-A---G---G---------------------------C-A---A-----C-----G--T--G------------------C-GGGA--------T--------- 5291
G.CM.10.DEMG10CM008          ------A-G----------------T---------...--A---...------............--AA--G---G----------------A------------A--C---------------T--G-----------------CCTG-GA---------------A-- 4715
G.CN.08.GX_2084_08           -C------G----------------T---------...--A---...----A-............C-AA--G-G-G-----------------------------A---------A--------T--G-----------------AC-G-GA------CA---------- 4574
G.CU.99.Cu74                 ------A----------------------------...--A---...------............C-AA--G---G-----------------------------A---T-----C--------T--G----------------A-C-G-GA-A-------------A-- 4938
G.ES.09.X2634_2              ------A-G---------------------T-T--...--A---...----A-............C-A---G---G---------------A--------G--------T--------------T--G-----------------AC-GGGA---------------A-- 4817
G.GH.03.03GH175G             ------A-G--------------------------...--A---...--C-CC............C-A---G---A------------C---------------CA--C-------C----CA-T--GACT---------------C-G-GA-------AT------A-- 5374
G.KE.09.DEMG09KE001          ------A-G--------------------------...--A--G...----C-............C-A-------A------C-----C---A------------A---T-----A--------T--G------------------C-G-GA-----C-T-------A-- 4709
G.NG.09.09NG_SC62            -C------G-------------------T------...-AA---...----A-............C-A---G---G------------G----------------A---------A--------T--G------------------C-G-GA--G------------A-- 4526
G.ZA.01.TV546                -C------G--------------------------...--A--G...------............A-A---G---G------------G-------A--------A---A-----G--------T---------------------CTG-GA---------------A-- 4536

H.BE.93.VI991                -C----AG-C----------------------G--...------...-----G............--A--C--------------------------C--------------------------T--------------------AC-GG-A--G------------A-- 4723
H.CF.90.056                  -------G------------------------T--...--A--A...-----G............--A--C-----C-----C---------A----------------A--------A--C-----------------------ACTG--A--G--------------- 4681
H.GB.00.00GBAC4001           ------AG--------------------T---T--...-----A...-----G............--A--C----T------C---------A-----------------------------------------------------C-G-GA--G--------------- 4852

J.CM.04.04CMU11421           ------A------C----T---RT----Y---S--...--A--R...---GA-............--AR-------------------------G--------------Y------Y----------------------------RCTG--A-RG--C--G--------- 4840
J.SE.93.SE9280_7887          ------A----------------------------...--A--G...----T-............--A--------------------------G--------------T-----C------------------------------C-G-GA--G--C--A--------- 4648
J.SE.94.SE9173_7022          ------A----------------------------...--A--G...------............--A---G----------------------A--------------T-----C------------------------------CGG--A--G--C--A--------- 4649

K.CD.97.97ZR_EQTB11          ------A----------------------------...------...--AGA-............--A--GG----------------------------------G--A-----C------------------------------CT-----AG--C--A--------- 4526
K.CM.96.96CM_MP535           ------A----------------------------...------...--AGAG............--A---G---------------------T------------G---------------------------------------CTG--A--G--C--A--------- 4530

U.CA.01.TV749                --------G----------------------G---...-----A...------............--A---G-------------------------------------A------------------------------------CTGGGA--------A------A-- 5351
U.CA.99.TV721                ------A-G-----------A----------G---...-----A...------............C-AA--G-------------------------------------A-----------AAGT------------------GA-CTG-GA--------A------A-- 5350
U.CD.83.83CD003_Z3           -------------------------------T---...---A--...------............A-A---G-----G----------------------------------------A---------------------------C-GGGA--G--C---------A-- 4829
U.CD.90.90CD121E12           -C-----------------------------T---...-A-A-C...------............A-A---------G-------------------------------T--------A---------------------------C-GGGA---------------A-- 4533
U.CY.05.CY090                --------G--------A----------T---T--...AA---A...----A-............--A--G---------------------------------------------------------------------------C-GGGA--G---TA---------- 4533
U.CY.08.CY223                --------G--------------------------...-----A...----AG............A-AA--G---G---------------A-------------A---------A--------T--G-----------------AC----A--C------------A-- 4551
U.ES.10.DEURF10DZ001         ------A------C------------------A--...--A--A...---GT-............--A--------------------GT-AA---------------CA-----------C-----G------------------TTG--A--------A------A-- 4610
U.GR.99.99GR303              -------------C-----G---------------...A-A--A...----A-............--A---G-G-----------------------------------G--------------A--G-------------------TG--A--T--C--A--------- 4605
U.NL.95.U_NL_95_H10986_D1    --------G--------------------------...AA---C...----A-............--A---------------C--------------------G--------T-----------------------------G--C-G-GA-AG--C------------ 5219

01_AE.AF.07.569M             ------AGG----------------G---------...--A--G...-G----............--AA------G------------C-----G----------AG--------A-----C--T-----T--T------------CTG-GA-A-----A-------A-- 4540
01_AE.CF.90.90CF11697        ------A-G-----------------------A--...------...------............--A-------G---------------------G-------A---------C--------T---------------------C-G-GA-A---------------- 5283
01_AE.CN.10.YNFL03           -C----AGG----C-----------T---------...--A--G...----A-............--A-------T--------------C---G----------AG--T-----CC-------T---------------------C-G--A---------------A-- 4712
01_AE.HK.04.HK001            ------AG-----C---------------------...--A--G...------............--A-------A------------------G----------AG--------A--------T--------------------AC-G-GA-AC------------A-- 4696
01_AE.IR.10.10IR.THR48F      ------A-G----------------T--T---A--...--A--G...----A-............--A-------G---------------A--G----------AG--T-----CC----A--T---------------------C-GGGA-A-------------A-- 4524
01_AE.JP.x.JRC77AE           ------AGG----------------T---------...--A--G...------............--A--C----G------------------A----------AG--------C--------T---------------------C-GGGA-A-----A-------A-- 5331
01_AE.TH.09.AA090a_WG11      ------AGG----------------T------T--...--A--G...----A-............A-A-------G-------------T----G----------AG--------A--------T---------------------C-GGGACA-----A---------- 4757
01_AE.TH.90.CM240            ------AGG----C-----------T---------...--A--G...------............--A-------G------------------A----------AG--------C--------T---------------------C-G-GA-A-------------A-- 4908
01_AE.US.05.306163_FL        ------AGG----------------G------A--...--A--G...------............C-A-------G------------------G----------AG--------G--------T---------------------C-G-GAGA-----A-------A-- 4525
01_AE.VN.98.98VNNDI5         ------AGG----------------T---------...--A--G...----T-............--A-------G------------------G----------AG--------A--------T---------------------C-G-GAGA-----A-------A-- 4589
```

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

```
B.FR.83.HXB2                ATCCAAGAATAAGTTCAGAAGTACACATCCCACTA...GGGGAT...GCTAGA............TTGGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGCATTTGGGTCAGGGAGTCTCCATAGAATGGAGGAAAAAGAGATATAGCACACAAGTA 5334
Vif                         H__P__R__I__S__S__E__V__H__I__P__L__.__G__D__.__A__R__.__.__.__.__L__V__I__T__T__Y__W__G__L__H__T__G__E__R__D__W__H__L__G__Q__G__V__S__I__E__W__R__K__K__R__Y__S__T__Q__V_
02_AG.CM.10.DE00210CM013    --------G--------------------------...--A---...----A-............--AA--G---G-------------------------------------A--------T-----T--T-----------A-C-GG-A--T-------------A-- 4701
02_AG.CY.09.CY256           --------G---C-----------------T----...------...------............--A---G---A-------------------------------------C--A-----T--G---------------GA-C-G-GA---------------A-- 4530
02_AG.ES.06.P1423           ------A-----------G--------T-------...-AAAG-...----C-.........................................................C--------T--G----------------A-TG--A---------------A-- 4771
02_AG.GW.05.CC_0048         ------A-G---------------T----------...------...------............--A---G---A---------------A------------------------C--------T-------------------A-C-G--A---------------A-- 4684
02_AG.KR.12.12MHR9          ------A-G-C------------------------...------...----A-............C-A-------G------------------------------------C--------T--G---------------A-C-G--A---------------A-- 4993
02_AG.LR.x.POC44951         ----T-A-G---------------T----------...------...----C-............--A-------G------------------------G--------------------T--G-----------------C-G--A---------------A-- 5329
02_AG.NG.09.09NG_SC61       ------A-GC-------------------------...------...------............--AA--G---G------------------------G-------T----C--------T--G--------------GAAC-G--A--G-----T------A-- 4526
02_AG.NG.x.IBNG             ------A-G--T-----------------------...------...-----G............--A---G---G----------------------------C-------C--------T--G---------------A-C-G--A---------------A-- 4859
02_AG.SE.94.SE7812          -------GG--------------------------...------...------............--A---G---G---G---------------------------------C--------T--G--T------------A-C-G-GA--------------GA-- 4706
02_AG.SN.98.98SE_MP1211     ------A-G---------------T----------...------...----C-............--AA--G---G-----------------GC------------------------T--G---------T----CA-C-G--A---------------A-- 4531
03_AB.RU.97.KAL153_2        -------------------------------T---...------...----A-............----------A--------------------------------------C---------C------C--------------A---G------------------- 4560
04_cpx.CY.94.94CY032_3      -C----A-G---------------T----------...-----G...------............--A---G---G-----------------GC---G---CA--------C--------T---------------------CTC-GA------------------G 4700
05_DF.BE.x.VI1310           -----------------------------------...--A--C...----A-............C-----G---------------A-------------------------------------------------G---------G------------------- 4716
06_cpx.AU.96.BFP90          ------A-----------------A---T------...--AAG-...---GA-............--A-----------------AA-----------A-A-----C-----C-----C---------------------TTG-GA-------A--------- 5362
07_BC.CN.98.98CN009         ------A-G-------------------G------...--A--G...----A-............--A------A----------GT---A-------------T-----------------T------------------ATTG-GA------C-------A-- 4666
08_BC.CN.97.97CNGX_6F       ------A-G-------------------T------...--A--G...------............--A------A----------GT---A-------------T-----------------T----C--------------ATTG-GA--C----A-------A-- 4514
09_cpx.GH.96.96GH2911       ------A-----C---------------G------...--A-C-...---GA-............--A---G-G---------------------G-------A-A-----------------------G-----------------CTG--A--G--C--G---------- 4536
10_CD.TZ.96.96TZ_BF061      ------A----------------------------...--A--A...-----T............C--A------A-----C-----------------------A-----------------T--G-----------------ACTG--A-------T------A-- 4713
11_cpx.CM.95.95CM_1816      --------G---C----------------------...---CC-...--A-C-............--A--GG---T-------------------------G----A-----------------A--G-----------------CTG--A--G--C--A---------- 4698
12_BF.AR.99.ARMA159         --------G--------------------------...-A-A-A...---GA-............--A------------C----G----GC-------------T-----C--------------------------------C-GGG---G----G------A-- 5339
13_cpx.CM.96.96CM_1849      ------A----------------------------...------...----A-............--A---G-G-------------------------A-A-----------------A--G--------T---------CTG--A-----C--A---------- 4739
14_BG.ES.05.X1870           ------A-G-------------------T-G....--A---...----A-............W-A---G---G----------------------A---T-----------------T--G-----T------------C-G-GA---------------A-- 4801
15_01B.TH.99.99TH_MU2079    ------AGG---C---C------T-----------...-AA--G...------............--A------A-----------G--------AG--T-----------------T-----------------------C-G-GAG---------------A-- 4730
16_A2D.KR.97.97KR004        --------G---------G---------G------...----GG...----A-............A-AA--G---A-----------CC-------A--------C-----------G-----------------C-GG-A-----C--T------A-- 4697
17_BF.AR.99.ARMA038         --------G--------------------------...-A---C...----A-............--A-----G-----C----G----CC----------T-----C---A-------------------------------C--GG---G-----G------A-- 4548
18_cpx.CU.99.CU76           ------A-G---C-------------------G..------...----C-............--A---G----------------------A---A-----------------T--G-----------------CTG-GA--G--C--A---------- 4639
19_cpx.CU.99.CU7            ------AGG--------------------------...------...----A-............--A---G---A--------------------A-------C--------T--G-----------------C-G-GA---------------A-- 4560
20_BG.CU.99.Cu103           ------A-G---C-------G--------------...--A--C...----A-............C-AA--G---G--------------------A---T----------Y--T--G--T-------------A-C-G-GAGA-------------A-- 4800
21_A2D.KE.99.KER2003        --------G-G------------T--G--------...-----A...----C-............C-AA--G---G----------------CT-G------AG-------C----G-T--G---------------AC-G--A--------T------A-- 4539
22_01A1.CM.01.01CM_0001BBY  ------A----------------------------...------...------............--A---G---G-----------G--------------A---A-----C--------T--G-----------------C-G--A--------T--------- 4542
23_BG.CU.03.CB118           -------G----C-----------A----------...-----A...----A-............---A-----AG---------------------------T-----------C----------------------------------------------------- 4777
24_BG.ES.08.X2456_2         ------A-G---Y-------G----T---------...--A--C...----AT............--AA--G---G----------------------A---T-----------------T--G---------------A-C-G-GA-A-------------A-- 4785
25_cpx.CM.02.1918LE         ------------C----------------------...--A---...---GA-............--A---G-------------------------------A-------A--------G-----------------ACTG-------C--G---------- 4542
26_AU.CD.02.02CD_MBTB047    ------A-G--------------------------...------...------............C-A----C-G----------G--A--G-----------------A------------T--G-----T----------ACT--GA------------------ 5339
27_cpx.FR.04.04CD_FR_KZS    ------A-----C----------------------...------...---GA-............--A---G-----------------A--------------------C---------G-----------------TTG--A--G--C--A---------- 5325
28_BF.BR.99.BREPMI2609      --------G--------------------------...------...----A-............--A---G---T---T---------------------------------------------------------------------G------------------ 4700
29_BF.BR.01.BREPM16704      --------G---------------T----------...-A---A...-A-GA-TTA.........G-A-----------C----G--------------------T-----C-A--------------------------------GG--AG----G-----A-- 4738
31_BC.BR.04.04BR142         ------A----------------------------...-----A...------............--A------A----------AT-A--A-------------T-----------------T--------C----G-----ACTG-GAG--------------- 4806
32_06A1.EE.01.EE0369        ------A----------------------------...--AAG-...---GA-............--A-------------------AA-------------A-----C-----C-----C---------------------TTG--A-------AA--C---A-- 4969
33_01B.ID.07.JKT189_C       ------AGG---------------T----------...--A--G...----A-............A-A--G---G-------------------G--------AG-------------------T-----------------C-G-GAGAC------------A-- 4635
34_01B.TH.99.OUR2478P       ------A-G--------------------------...--A--G...-C----............--A-----CG-------------------G--------AG--T----AC--------T-----------------C-G-GA-AC------------A-- 4536
35_AD.AF.07.169H            --------G---------------A-----T----...------...------............A-A---G---G-------------------------A---------------------T--G---------------CT-C-A----------------A-- 4534
36_cpx.CM.00.00CMNYU830     ------A-G---------------T----------...------...--A---............----------G---------------------------------------------T--G-----T---------AAC-G-GA-A---C---------A-- 4539
37_cpx.CM.00.00CMNYU926     --------G--------------------------...-----A...---T-G............--A---G---G-----------A-----GC---G----------------C--------T--G--------------------CTG-GA-A------A------A-- 4524
38_BF1.UY.03.UY03_3389      --------G--------------------------...CAA--A...-A--A-............--A-----------C----G------------------T-----C---A-----------------------------AC-GGG---C-----G------A-- 4757
39_BF.BR.04.04BRRJ179       -----------------------------------...-----A...------............--A---G---------------------------------------A--------A-----------------------------G------C--A---------- 4811
40_BF.BR.05.05BRRJ055       ------A----------------------------...-----A...---GC-............--A---G---------------------------------A-----------------A------------A-------------GA-A----------------- 4817
42_BF.LU.06.luBF_18_06      ------A----------------------------...------..AG----............--A------A-----------A-------------------------------------------------------------A---G------------------ 4888
43_02G.SA.03.J11223         ------A-G-------------------M--..R------...----C-............--AR--R---R------------R--G----------G--Y----A----C--T--G----------------R-C-G--A-A----------------- 4859
44_BF.CL.00.CH80            --------G-------------------Y--...-A------...----M-............--A------------C----G---------------------T-----C-------------C----------------C-GGG---G-----A------A-- 4776
45_cpx.FR.04.04FR_AUK       ------A-G--GCAG--------------------...------...----C-............--A---G---G-----------------A----------A-------C--------T--G-----------------ACTG--A------------------ 5320
46_BF.BR.07.07BR_FPS625     --------G--------------------------...-ATA-A...-A--CC............----------------C----A--------------------T-----C--------------------------------AC-GGG---G-----G---------- 4772
47_BF.ES.08.P1942           ------A-----C---A-----------W------...-A---AGAAAA--A-............--A---G-------------------A------G-------------------------------------------C-GGG--AG-----G------A-- 4776
48_01B.MY.07.07MYKT021      ------AGG---------------T-----G----...-AA-GG...----A-............--A------A-------------------G--------AG-------AC--------T----------T---------C-G-GAGAT------------A-- 4539
49_cpx.GM.03.N26677         ------A-G-------------------T------...-----G...------............--A---G---G------------T-A--A-----G-------G---C--------T-------------------ATTG-GA------------------ 4756
50_A1D.GB.10.12792          ------A--C-------------------------...------...----A-............--A---G---A-------------G-------------------A--------------T--G-----------------CT--GA--------------A-- 4741
51_01B.SG.11.11SG_HM021     ------AGG---------------T----------...--AA-A...------............--A-------G--------------------G-------CA--------A--------T-----------------C-G-GAGAC---A-------A-- 4541
52_01B.MY.03.03MYKL018_1    ------AGG---------------T----------...--A--G...----A-............--A---G---G--------------------G---------AG-----------------T-----------------C-GGGACT------------A-- 4648
53_01B.MY.11.11FIR164       ------AGG--------------------------...--A---...----A-............A-A-------G--------------------G--------CAG-------C--------T-----------------C-GGGA-AT---A-------A-- 4671
54_01B.MY.09.09MYSB023      ------A---------------------G------...---G-...----A-............-------------------------A-----------------------------C---------------C-----A-GGG-A--C--------------- 4736
55_01B.CN.10.HNCS102056     ------AGG---------------T------A---...--A--A...------............--A-------G--------------------G--------AG-------A--------T-----------------C-G-GA-C-----A------A-- 4644
56_cpx.FR.10.URF5_patient_A ------A-G-------------------A------...------...------............--AA--G---G--------------------G---------G-------C--------T--G-----------------AC-G--A--G-----------A-- 4685
57_BC.CN.09.09YNLX19sg      ------A-G-------------------G------...-----G...----C-............--AA-----A----------GT---A-------------T-----------------T--------G----------ATTG-GA------C-------A-- 4518
58_01B.MY.09.09MYPR37       ------AGG---------------T----------...--A---...----A-............A-A-------G--------------------G--------AG-------C--------T-----------------C-GGGA-AT---A-------A-- 4679
59_01B.CN.09.09LNA423       ------AG----------------T--T---A---...--A--A...----A-............A-A-------A--------------------A--------AG-------C--------T--------G--------CTG-G-C-----A-------A-- 4518
60_BC.IT.11.BAV499          -------------------------A---T-----...-A--G..A-----............--A---------------A--T-A--A---------------T-------A-----T-------A----G-----ACTG-GA--------A--------- 5291
61_BC.CN.10.JL100010        ------A-G------------------GT------...--A--G...------............--A------A----------AT---A-------------T-----------------T--------G----------ATTG-GA-A-------------A-- 4755
62_BC.CN.10.YNFL13          ------A-G-G-------G-------T---G----...--A--G...------............--A--C---A----------GT---A-------------T-----------------T-----------------A-TG--A---------------A-- 4663
63_02A1.RU.10.10RU6637      -C----A-G---C----------------------...------...------............--A---G---G----------G----A-A--------G-------C--------T--G-----------------CGGG-A--G------------------ 4875
64_BC.CN.09.YNFL31          ------A---------------------GT-----...--A--G...------............--A------A----------GT---AC-------------T-----------------T-----------------ATGG-GA------A-------A-- 4688
65_cpx.CN.10.YNFL01         ------A---------------------T------...-----C...------............--A----G-A----------GT---A-------------A-----------------T-----------------A-TG-GAG-----A-------A-- 4705
67_01B.CN.11.ANHUI_HF115    ------A-G--------------------------...--A--A...---T-............--A------G--------------------G--------AG-------C--------T-----------------C-G--A-AC----A------A-- 4677
68_01B.CN.11.ANHUI_WH73     ------A-G--------------------------...--A--A...---TG............--A---G---G--------------------G--------AG-------A--------T---------G---------C-G--A-AC------------A-- 4668
70_BF1.BR.10.10BR_PE004     ------A-G--------------------------...--AA-A...--GAG............--A--C---T----------G--A-----------------------C-------------------------------C-GGG--------G-----A-- 4790
71_BF1.BR.10.10BR_PE008     -C-----G--G---T--------------------...-A---A...---C-............--A---G----------C----C------------------T-----C-----------A----------------GGGG--------G-----A-- 4998
72_BF1.BR.10.10BR_MG002     ----G-----C------------------------...------...-AGA-............C-----G----------AAA----A---------------------------------------------------------A---G------CA----------- 4842
O.BE.87.ANT70               --------G-T------AGT---T-T--T---G--...---TG...-CCAT............G-A---G---C----------AT--ATGC---G--------TGAA--------A--T--G--TAGT---------C-TT-T----AG----AA-----GA-T 5388
O.CM.98.98CMA104            --------G-C-------GT---T-T--T---G--...--C-G...--T-T............A-A---G-G--C----------AT-AATGC---G--------GGAT--G--A--A--T-----TAGT---------CA-T-C----AG------A-----GA-T 4820
O.CM.98.98CMABB141          ----------C------AGT---T-T--T---T--...--CCG...--T-T............A-A---G-G-GT----------AT--ATGC------------GGAA---A----A--T--G--TAGT---------CA-T-C--A-AG----CA-----GA-T 4819
O.CM.98.98CMABB212          GC--------TG------GT---T-T-----G--...-CA--A...----AG............A-AA--G-C--C----------AT-AATGC---G--------GGAA--CC----A--T--G--TAGT---------CAAT-C----AT----CA-----GA-T 4824
O.CM.98.98CMU5337           --------G-C------AGT---T-T--T---G--...---ATG...-----G............G-AAA-G-G--C-----C-----AT-AATGC---G--------TGAA---A----A--T--G--TAGT---------CA-T-C---MAG----CA-----GA-T 4821
O.CM.99.99CMU4122           --------GCC-------GT---T-T--T---G--...---ATG...---T-T............G-A---G-G--C--C--------AT--ATGC---G-----G--TGAA--------A--T--G--TAGT---------CA-T-C----AC----CA-----GA-T 4819
O.FR.92.VAU                 -------GG-C------AGG---T-T--T---G--...--TCCA...---T-T............A-AA-GG-G--C----------AT-AATGC------------GGA---G----A--T------AGT---------CA-TGC-G--AG----TA-----GA-T 4900
O.GA.11.11Gab6352           -C--------C------AGT---T-T--T---G--...-----A...---T-G............G-A---G-G--C----------AT--ATGC---G--------TGAA--------A--T--G--TAGT---------CA-T-C----AG----AA-----GA-T 4545
O.SN.99.99SE_MP1299         --------G-C-------GT---T-T--T---G--...----GG..C--T-G............A-A---G-G--C----------AT--ATGC---G--------TGAA--------A--T--G--TAGT-----------T-C----AG----AA-----GA-T 5387
O.US.10.LTNP                ----T-----C------AGT---T-T--T---G--...---ATG...--CT-T............G-A---G-G--C----------AT--ATGC---G--------GGAC--G-----A--T--G--TAGT---------CAAT-T----A-----AA-----GA-T 5309

N.CM.02.DJO0131             ------A-----------------T------G--...--TCTG...--A--G............--A--G-----C--T--------A--AACA-----------GTCT------C-A--A--T-----A-------------AACT-GGA-A---C-AG---------T 4840
N.CM.02.SJGddd              ------A-----------------T------G--...---TC-G...--A---............--A--G---GTT--T--------A---ACA---------CAGTCT------C-A--A--T-----A--------------ACT--G--A---C-AG---------T 4824
N.CM.04.04CM_1015_04        ------A--C--------------T------G--...---TC-G...--A-C-............--A--G---GTC--T--------AT-AACA---------CAGCCT------C-A--A--T-----A--------------ACT--GA-A---C-AG-----G--T 4839
N.CM.06.U14296              ---------A--------------T------G-R...---TC-G...--A-A-............--A--GG--TC--T--------A--AACA---------CA-CCT------C-A--A--T-----A--------------ACT-GGA-A---C-AG---------T 4837
N.CM.06.U14842              -Y----AG-A--------------T------G-G...---TC-G...--A-A-............--A--G---TC--T-----------AACA---------CAGCCT------C-A--A--C-----A--------------ACT-GGA-A---C-AG---------T 4841
N.FR.11.N1_FR_2011          ------A----G-C----------T------T--...---TCCG...--A---............--A--G---TC--T--------A---ACA---------CA-CCT------C-A--A--T-----A--------------ACT--GA-A---C-AG---------T 4735

P.CM.06.U14788              ----T--G--T-------C---TT-T--T---G--...--AACA...--A-CC............A-TA-TG-G--T--Y--------G--CATGC-T--G--------AGAA--A--A--A--T----CAAGTG-G--G-----AC--GGT-A---C-C------GA-- 4805
P.FR.09.RBF168              ----T--G--T-------G---TT----T---G--...---A-CA...--A-CC............A-T--TG-G--T--T--------G--TATGC-T--G--------AGAA--A--A--A--T----CAAGTG-G--G-----AC--GGT-A---C-C---C--GA-- 5369

CPZ.CD.06.BF1167            -C--C----AGG-AGG------G--T--A--CA--ACT-A-A-A...--A-A-.............-------------T---------C--A--CTGT--G----A--CT------C-T--A--T--G-CAAG-------------AC--GG------CT-G---G----- 5439
CPZ.CM.05.SIVcpzMT145       -C--T--G--GGC----C-GA--T-T--T---A--...AATC-G...-----AG............C-AA-TG-C---------------A--AATGC-----------------T------C---A--T-----AG-A--------------AG---GAG-T--C--T--C--G--- 4876
CPZ.GA.88.GAB1              ------A-G-TGC---T---A-------A--TT-CAGA-ATT--...T---A-.............--AA--G-G----C------CCT-AAGCC----------G-C----------A--C--T------------TC---------CT-GGATCT---GT------G--G 5398
CPZ.TZ.06.SIVcpzTAN13       TG----A--G-GCAGGG---A-------A--T--TGAA--AA-G...--A---............A-A--TG-GCT-CAT--------AT-AGC-TGT--G--G---CCA-----------A--T-----AGGG-------------AC--GGA--------------GA-- 4978
CPZ.US.85.US_Marilyn        ----C---G---A-----------T-----TA--...-ATA-A...--A-A-............A-A--G--T-T---G--------GT--A-A---------CA--C------C--A--G--T-----A--A-------------AC--GGA-CT--C-AG------A-- 5394

GOR.CM.07.SIVgor2139_287    ----T--GG---------GG---T----A---G--...--CCCC...---GA-............A-TA--G----C-----------A--CATGC---------AG--AGAA--A--A--C--T--G--AGG------------AC-GGGA-A---C-TA------T-- 4871
GOR.CM.07.SIVgorCP2135con   ----T--GG---------GG--GT----A---G--...--CCCC...---GA-............A-TA--G-G--C-----------A--CATGC---------AG--AGAA--A--A--C--T--G--AGG------------AC-GGGA-A---C-TA------T-- 4870
```

```
B.FR.83.HXB2                         GACCCTGAACTAGCAGACCAACTAATTCATCTGTATTACTTTGACTGTTTTTCAGACTCTGCTATAAGAAAGGCCTTATTAGGACACATAGTTAGCCCTAGGTGTGAATATCAAGCAGGACAT...AACAAGGTAGGATCTCTACAATACTTGGCACTAGCAGCATTAAT 5501
Vif                                  _D__P__E__L__A__D__Q__L__I__H__L__Y__Y__F__D__C__F__S__D__S__A__I__R__K__A__L__L__G__H__I__V__S__P__R__C__E__Y__Q__A__G__H__.__N__K__V__G__S__L__Q__Y__L__A__L__A__A__L__I
A1.AU.03.PS1044_Day0                 --T-----CT-G-----------------C---C----T---A-------------------C-----G--A---A----------AG-----------------------------------...-----------------------T--------GAG-------G- 4709
A1.CH.03.HIV_CH_BID_V3538            --T-----C------------T--------A-AC----T-----------------A-----C--------A---A--------G-A--------A---------------------------...-----------------------T---------AA------GG- 4706
A1.CY.08.CY236                       --T-----CA-G--G-----C--------C--------T---A-------------------C-----G--A---A----------AG-------------T-----T--------------C...--T--------------------T--------GAG-------G- 4712
A1.ES.05.X1608_8                     --T-----C--G------------------A--C----T---A-------------------C-----G--A---A----------A--------------T--------------------C...-G---------------------T--------GAA-------G- 4949
A1.KE.11.DEMA11KE001                 --T-----C-AG--------------------------T---A-------------A-----C--------A---A----------AG-------T-----------------TA--------...-----------------------TC-A-----GAA-------G- 4840
A1.RU.11.11RU6950                    --------C--G--G----------------------CT-----------------A-----C-----G--A---A--G------------G---A---------A-C----C---------C...--------------------------------GAA-------G- 5025
A1.RW.11.DEMA111RW002                --T-----CA-G---------------------C----T-----------------------C-----G--A---A-----------G-------------T----------T-T--------...-----------------------TC--------AA-------G- 4852
A1.SN.01.DDI579                      --T-----C-----------------------AC----T-----------------------C-----G-GA---A----------A--------A------------------A-------C...---------------T-------T--------GAG-------G- 4699
A1.UG.11.DEMA110UG001                --T-----C--G---------------------C----T-----------------A-----C--C--G--A---A----------AG-------------T--------C------------...-----------------------T--------GAA----C---- 4856
A1.ZA.04.04ZASK162B1                 --T-----T-----------------------------T-----------------------C-----G--A-T-A----------AG-------A---------------TC----------...--T--------------------T--------GAA-------G- 4942
A2.CM.01.01CM_1445MV                 --------T--G--G-----T-----C-----------T-----------------A-----C-----G--A---A--------G-AT-------T----A----------------------...-----------G-----G-----T------T-GAA-------G- 4700
A2.CY.94.94CY017_41                  --------T--G--------T-----C-----------T-----------------A-----C-----G--A---A--A-----G-A--------T---------------------------...-----------G-----G-----T------T-GAA----G--G- 4858
B.BR.10.10BR_MG029                   -------GC---------------------A----------------------------C-----C-----A---A----------T-G------T---------------------------...--T-----------------G------------------C--C- 4963
B.CA.07.502_1191_03                  ------A-TA-------------------CT-------T-----T--------------------------T--------------T-G------A---------------------------...---C----------------G-------------T--------C 4928
B.CH.08.M2_0803101_NFLG8             -------GC----------------------AA---C-T-----T-----------A--------------T---A--C-----A-T-G------A---------------------------...---G-----------------------------A-------G-- 4970
B.CN.10.DEMB10CN002                  -------GC-----------------------------T-----------------------------G-GT---A--G------GT--------T-----------------------C---...---------------T------------------T-------G- 4873
B.ES.10.DEMB10ES002                  --------C-----------------------------------T-----------G--------------T---A--C-------G--------T--------------C------------...-G---------G--------GC--------T--A---------- 4863
B.FR.11.DEMB11FR001                  ---------AC---------------C----------------------------------------A---T--AA---------G---------T-----T--------C------------...--------------------G----------------------A 4747
B.GB.05.MM45d213_GN1                 ------A-C-----------------C-----------------T-----------A--------------TA--A----------T-G------T---------------------------...--------------------G---------------A------- 4902
B.HK.06.HK003                        -------GC---------A-----------T--C----------T------------------G-----------A----------T-G------T-----T---------------------...---C-------G--------G------------AA--------- 4871
B.HT.05.05HT_129389                  --T----GC-----------------C-----------T-----T-----------A-----C--------A---A--------T-T-G------T---------------------------...--------------A-----G-------------T--------- 4702
B.JP.12.DEMB12JP001                  ------A-T-----------------------A-----T-----T-----------A------------C-A--------------T-G------T-----T----------G----------...-GT-----------C---------------T--A---------- 4888
B.KR.07.HP_18_07JHS10_3909           --------C---------A----------C---C----T-----T------------------------C-T---A---------GTT-------G---GA-----------C----------...----C------G------A-G------------A---------- 5165
B.PE.07.502_0525_wg5                 ------A-C-----------------------------------T-----------A--------------T--AA--G------GT--------------------------C---------...--------------C-----G------------A--------G- 4978
B.RU.11.11RU21n                      ------A-CA----------------------------T-----T-----------A--------------T---A---------GT-G------------T---------------------...--------------------G----------------------G 5072
B.TH.08.MERLBDTRC10                  ------A-TG----------------------------------T-----------A------------------A--------AGT-G------A---------------------------...--------------------G------------A--------C- 4734
B.US.11.ES38                         ------C----------------------C--------------T-----------A------------C-A---A----------T--------T---------------------------...--------------------G---------------------G- 5436
C.AR.01.ARG4006                      -------GC--G------------------A--C----T-----T------A----------C-----G--A---C----------T-G----------------------------------...--T------------T--------C-------GA-------GT- 4685
C.BR.07.DEMC07BR003                  --------C--G--------------C---A-AC----T-----T------G----------C-----G--A--------------T-G------------------C---------------...------------------------C-------GA---------- 4833
C.BW.00.00BW5031_1                   -------GC--G--------G-----C--CA--C----T-----T------G----------C--------A---A---------G--------TT-----------C---------------...--TC-------------------T--------GA-------G-- 4872
C.CN.10.YNFL19                       --A----GC--G--------G---------A-------T-----T------G----------C-----G--A---A------------------TT-----------C---------------AAT--T-----------------------------GA-----C-G-- 4857
C.CY.09.CY260                        -------GC--G--------G---------A--C----T-----T------G----------C------C-A---A----------T-------TT--------------------------C...--G--------------------TC-------GA-------G-- 4724
C.ES.08.X2363_2                      -------GC--G--------G---------A--C----T-----T------A----------C--------A---A----------T-G-----CT-----------T---------------...------------------------C-------GA---T---GG- 4932
C.ET.02.02ET_288                     ------TGC--G------------------A--C----T-----T------G----A-----C-----G--A---A----------G-G------------------C---------------...--T---------------------------T-GA-------GC- 4703
C.IN.09.T125_2139                    -------GC--G--------G---------A--C----T---A-T------G----A-----C-----G--A---A-----------T------TT-----------C--C-C----------...-CT-----------C-----------------GA-----C---- 5477
C.KE.00.KER2010                      -------GC--G--------G---------A--C----T-----T------G----------C--------A---A------------------TT---------------------------...--T---------------------C-------GA-------G-- 4682
C.MW.09.703010256_CH256.w96          --------C--G-----T------------A--C----T-----T------G----------C------C-A---A--C--------G-------------------C---------------...--------------------------------GA-------G-- 5481
C.TZ.08.707010457_CH457.w8           -------GC--G------------------A--C----T-----T------G----------C--------A---A------------------CT-----------C-----------C---...-TC--------------G--------------GA--------G- 5475
C.YE.02.02YE511                      -------GT--G--------G-----C--CA--C-C--T-----T------A----------C--------A---A----------T-------T------------T---------------...-CT-----------------------------GA-------G-- 4676
C.ZA.08.705010534_CH534.w12          --T----GC--G------------------A--C----T-----T------G----------A--------A---C----------T------------------------------------...-----------------------T--------GA---------- 5491
C.ZA.10.DEMC10ZA001                  -------GC--G--------G---------A--C----T-----T------G----------C------C-A---A------------------TT-----------C---------------...--T------------T----------------GA-----C-G-- 4841
C.ZM.11.DEMC11ZM006                  -------GC--G------------------A--C----T-----T------G----------C--------A---A-------------------------------C----C----------...--------------G-----------------GA-------G-- 4850
D.CM.10.DEMD10CM009                  -------GC--G------------------A-------------T-----------A-----C--------A---A----------T-G------T-----------G--C------------...--------------CT----G--T---------A---------- 4876
D.CY.06.CY163                        -------GC--G------------------A-------------T-----------------C--------A---A----------T-G------T-----------G---------------...--------------CT----G--T---------A---------- 4709
D.KE.11.DEMD11KE003                  -------GC--------------G------ACCC----T-----T------G-----------------C-A---A----------T-G-----TTT--------A-T--C------------...---------------T-------TC--------A-------GC- 4851
D.KR.04.04KBH8                       -------GC--G-----T------------A--C----------T-----------------C-----G--A---A----------T-G------------------T---------------...--------------------G--T------T--A---------- 5460
D.SN.90.SE365                        -------GC--G------------------A-A-----T-----T-----------A------------C-A---A----------T--------------------T---------------...---------------T----G--T---------A---------G 5518
D.TZ.01.A280                         -------G---G------------------A-------T-----T------A----------C-----G--A--AA-------T--T-----G--T-----------T---------------...---------------T----G--T---------AT--------- 4702
D.UG.10.DEMD10UG004                  -------GC--G------------------ACA-----T-----T------G----A------------C-A---A----------T-------TT---------A-T---------------...---------------T-------T---------A---------- 4839
D.UG.11.DEMD11UG003                  -------GC--G-----------------CA-A-----T-----T------G----A-----C-----G--A---A----------T------TTT---------A-T---------------...---------------T-------T------T--A--------G- 4856
D.YE.02.02YE516                      ------AGC--G------------------ACA--C---------T-----------A--------------A---A---------TT-G------T-----------T---------------...--------------CT----G--T---------A---------- 4712
D.ZA.90.R1                           -------GC--G------------------A-A-----T-----T-----------A--------------A---A----------T-G------T---------------------------AAC--------------CT----G--T---------AA--------- 4850
F1.AO.06.AO_06_ANG32                 ------AGCA-G-----------G--C-----A-----T-----T-----------G-----C-----G-G----A-------T--T-G---------------CA-C---------------...--------------C--------T---------A---------- 4712
F1.AR.02.ARE933                      -------GC--G-----------G--C---A-A--C-CA-----------------A-----C-----GGGA---A----------T-G-A--------------A-C---------------...----G-T----------G-----T---------A---------- 4795
F1.BR.10.10BR_RJ015                  -----A-GT--G-----------G--C---A-A-----T-----T-----------A-----C-----G--A---A-------T--T-G------------R---A-C---------------AGC-C---------------------T--------GA---------- 4977
F1.CY.08.CY222                       -------GC--G-----------G--C---A-A-----T-----T-----------A-----------G--A---A----------T-A-A--------------A-C---------------...--------------C--------T---------A---------- 4694
F1.ES.02.ES_X845_4                   -------GCAC------------G--C---A-AC----T-----T-----------------C-----G-GA---A-------T--T-A----------------A-C---------------...--------------C-----G------------A-G-------- 4953
F1.RO.96.BCI_R07                     -------GC--G-----------G--C---A-A-----T-----T-----------T-----C-----G--A---A-------GT-T-G----------------A-C---------------...--------------C--------T---------A---------- 5525
F1.RU.08.D88_845                     ------AGCA-G-----------G------A-AC----T-----------------A-----C-----G--A---C----------T-GGA--------------A-C---------------...--------------C--------T---------A--------G- 4960
F2.CM.02.02CM_0016BBY                --T----GCT-G-----------------AAA-----T-----T-----------A-----C-----G--A---A----------T-G------T---------A-C---------------...--------------C--------T---------A-----C--G- 4694
F2.CM.10.DEMF210CM001                -------GC--G---------------------C----T-----T-----------A---A-C-----G--T--TA--C-------T-G------T---------A-C---------------...-----------G--C--------T---------A---------- 4768
F2.CM.10.DEMF210CM007                -------GCT-G-----T---------------C----T-----T-----------A-----C-----G--T--TC-------T--T-G------T---------A-T--------G------...--------------C--------T---------A---------- 4851
G.BE.96.DRCBL                        --T-----C-----------C------------C----T---A-T-----------A-----C-----G--A---A----------A-C------G-----T--------C--------G---...--T------------T-------T------T--AA--T----G- 5458
G.CM.10.DEMG10CM008                  --T-----CT-------------G------TC-C----T-----------------------C-----G--A---A----------A--------A---------------------------...--T--------------------T--------GAA-------G- 4882
G.CN.08.GX_2084_08                   --------C--------------G---------C----T-----T-----------------C-----G--A---A----------AT-------A---AA----------------------...--T-----------C--------T------T-GAAG-TG--GG- 4741
G.CU.99.Cu74                         --T-----C--------------G---------C----T---A-A-----------------C------C-A---A--------G-AG-------A--------------------------C...--G--------------------T--------GAA-------G- 5105
G.ES.09.X2634_2                      --T-----C-----A------------------C-A--T-----A---------A-------A-----G-----AA--------AGA--------T-----------C---------------...-GT--------------------T--A-----GAA-------G- 4984
G.GH.03.03GH175G                     --T-----C-----------C--G---------C----T-----T-----------A-----C-----G--A---A--------G-G-----------------------------------C...--T--------------------T------T-GAA-------G- 5541
G.KE.09.DEMG09KE001                  --T-----CT-------------G---------C----T---A-T--------------C--A-----G--A---A--A-----G-G--------A----A----------------------...--T--------------------T--------GAA-------G- 4876
G.NG.09.09NG_SC62                    --T-----C--------------G-----C---C----T-----T-----------------C-----G--A--------------AG-------A---------------------------...--T------------T----------------GAA-------G- 4693
G.ZA.01.TV546                        --T-----C--------------G--W------C-C--T-----T-----------A-----C-----G--A---A----------A--------------------C---------------...--T--------------------TC-----T--AA-------G- 4703
H.BE.93.VI991                        --------C--G-----T------------T--C----T-----T-----------------C-----G--A--TA----------T-G------T--CTT------C---------------...-GA------------T-------T--------GA---------- 4890
H.CF.90.056                          --G----GC--G------------------A--C----T-----T-----------A-----C-----G--A---A---------GTG-------A---------A-C--C------------...--AC--------A-----------------T--A--------G- 4848
H.GB.00.00GBAC4001                   --------C--G-----T-----------------C--T---A-------------A-----C-----G--A---A----------T-G----TAT-----T------------A--------...---------------T-------T--------GA--------G- 5019
J.CM.04.04CMU11421                   --T----GC--G-----Y------------A--C----------T-----------A-----C-----G--A---W----------TR-------------------C--C-C----------...--YC-------------------TC----T---A---------- 5007
J.SE.93.SE9280_7887                  -------GC--G------------------A--G----------T-----------------C-----G------A----------A--------------------C--C------------...-----------------------TC--------A---------- 4815
J.SE.94.SE9173_7022                  -------GC--G------------------A--C----------T-----------------C-----G------A----------A--------------------C--C------------...-----------------------TC-----T--A---------- 4816
K.CD.97.97ZR_EQTB11                  --T----GC--G------------------A-A-----T-----T-----------A-----------G--A---C----------T-G------T---------------------------...-C-C----------CT----G--T-----T---A---------- 4693
K.CM.96.96CM_MP535                   --A-----C---------------------A-A-----T-----T-----------A-----AG----G------A----------T-G------T-------------G-------------...--------------CT-------T--------CA--------G- 4697
U.CA.01.TV749                        -------GC--G--------------C--CACAC----------T-----------A-----C-----G--A---A----------A------------------------------------...---------------T-------T------T-GAA-------G- 5518
U.CA.99.TV721                        -------GC--G------------------CACAC----------T-----------A-----C-----G--A---A----------A------------------------------------...---------------T-G-----T------T-GAA-------G- 5517
U.CD.83.83CD003_Z3                   --------C--G--------------C--C---C----T-----T-----------------C-----G--A---A----------------------------------C---A--------...-----------------------T--------GA--A-----G- 4996
U.CD.90.90CD121E12                   --------C--G------------------CA--C----T-----T-----------------C-----G--A--AA-------------------------------C--C---A--------...-----------------------T--------GA--------G- 4700
U.CY.05.CY090                        -------GC--G-----------G------A--C----T-----T------A----------C-----G--A-------------TT-G------T-----------T--C----A-------...-TA-----------------G--T------A--A---------- 4700
U.CY.08.CY223                        --T-----C--------TA----G--------AC-A--T-----------------A-----C-----G--A---A--------G-AG-------A--------C-C--TCTC----------...--T--------------------T--------GAA-------G- 4718
U.ES.10.DEURF10DZ001                 -----CAGT--G------------------G--C----------------------------C-----G--A---A--------G--------TA----------------------------...-G---------------------TC-A----------T----G- 4777
U.GR.99.99GR303                      ------AGC--G------------------A--C----------T-----------A-----C-----G--A---A----------A--------------------C---------------...--T------------T-------T---------A---------- 4772
U.NL.95.U_NL_95_H10986_D1            -------GC--G-----------G------A--C----T-----T-----------------C-----G--A---C----------T-G------T---------------------------...--T-----------CT-------TC--------A---T----G- 5386
01_AE.AF.07.569M                     --T-----C--------------G--------AC-C--T-----------------------C-----G--A---A----------AG-------A-A---------C----C-T--------...--------------C--------T---------AA--------- 4707
01_AE.CF.90.90CF11697                --T-----C----------G--AG---------C-A--T-----------------------C-----G--A---A--G-------AG-------A----------------C-A--------...--------------G--------T--------GAA-------GC 5450
01_AE.CN.10.YNFL03                   --T-----C---------AG---G--------AC-A--T-----------------A-----C-----G-GA---A----------AG-------A-A------C-------C-T--------...------------C-C--------T--------GAA--------C 4879
01_AE.HK.04.HK001                    --T------G---------G---G--------AC-A--T------------G----------C-----G-GA---A--C-------AG-------ATA--A-----------C-T--------...---------------------------------AGG-------C 4863
01_AE.IR.10.10IR.THR48F              --T-----T-----T----G---G--------AC-A--T-----------------AA----C-----GGGA---A--------A-AG-------A--------CW------C-T-------C...--------------C--------T--------GAA--------C 4691
01_AE.JP.x.JRC77AE                   --T-----T---------A----G--------AC-A--T-----------------------C-----G--A---A----------AG-------A-A------------C-C-T--------...--------------C--------T--------GAA--------C 5498
01_AE.TH.09.AA090a_WG11              --T-----C----------G---G--------AC-C--T----------------G------C-----G-GT---A----------AG-------A-A---------T--C-C-T--------...--T-----------C--------T--------GAA--------C 4924
01_AE.TH.90.CM240                    --T-----C---------A----G--------AC-A--T----G------------------C-----G--A---A----------AG-------A-G--------------C-T--------...--------------C--------T--------GAA--------C 5075
01_AE.US.05.306163_FL                --T-----C----------G---G--------AC-A--T-----T------G----------C-----G--A---A----------AG-------A-G--------------C-T--------...--T-----------C--------T--------GAA--------C 4692
01_AE.VN.98.98VNND15                 --T-----C----------G---G--------AC-A--T-----A-----------------C-----G-GA---A----------AG-------A-G------------C-C-T--------...--T-----------C--------T--------GAA--------C 4756
```

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546


```
B.FR.83.HXB2                 GACCCTGAACTAGCAGACCAACTAATTCATCTGTATTACTTTGACTGTTTTTCAGACTCTGCTATAAGAAAGGCCTTATTAGGACACATAGTTAGCCCTAGGTGTGAATATCAAGCAGGACAT...AACAAGGTAGGATCTCTACAATACTTGGCACTAGCAGCATTAAT 5501
Vif                          _D__P__E__L__A__D__Q__L__I__H__L__Y__Y__F__D__C__F__S__D__S__A__I__R__K__A__L__L__G__H__I__V__S__P__R__C__E__Y__Q__A__G__H_._N__K__V__G__S__L__Q__Y__L__A__L__A__A__L__I
02_AG.CM.10.DE00210CM013     --T-----C-------------G-----CA--C----T----------------GA-----C-----G--A--A---------AG-----------------------------------------...--T----------G-------T-------GAA-------G- 4868
02_AG.CY.09.CY256            --------C-------------G-----C---C----T----------------------C-----G--A--C---------AG---C--A-----------------------G--C...--G----------G-------T------T--AA-------G- 4697
02_AG.ES.06.P1423            --T-----C--------T----------C---C--C-T---A------------A-----C-----G--A--A---------AG-----A----------------------------------...--TC-----------------T------T--AA-------G- 4938
02_AG.GW.05.CC_0048          --T-----C-------------G-----C---C----T----------------------C------C-A--A---------AG------A----T-------C-------------...--T----------G-------T-------GAA-------G- 4851
02_AG.KR.12.12MHR9           --T-----T-----T-------G-----C---C----T----------------------C-----G--A--A---------AG------A---TT----------------------------...-GT----------G-------T-------GAA--T----G- 5160
02_AG.LR.x.POC44951          --T-----C-------------G---------C----T--A-A-----A---------C-----G--A--A---------TG----------TC-------------A-----------...--T----------G-------T-------GAA-------G- 5496
02_AG.NG.09.09NG_SC61        --T-----C--------T----G-----C---C----T-----------A---------C-----G--A--C---------AGA-----A--------------------------G----...-GA----------G-------T-------GAA-------G- 4693
02_AG.NG.x.IBNG              --------CT------------G-----C--------T--A-T------------------C-------A--A------G-A-------A-----------------------------...--T----------G-------T-------GAAT--T---G- 5026
02_AG.SE.94.SE7812           --T-----C-------------G-----C---C----T----------------------C-----G--A--A---------A-------A----A--------------------...-CT----------G-------T-------GAA-------G- 4873
02_AG.SN.98.98SE_MP1211      --------T-------------G-----C---C----T-----------------A-----C-----G--A------------AG------A---A-----------------------...--T----------G-------T-------GAA-------G- 4698
03_AB.RU.97.KAL153_2         ------A-C-----------------------------T-----T-----------A--------------T--A---------T-G-----------T---------G------------...-------------------G---------------------G 4727
04_cpx.CY.94.94CY032_3       --T-----C--G----------------A--C-----------T-----------A-----C-----G--A--A---------T-G-----T------------------------------...--T----------CT-------C-------------------- 4867
05_DF.BE.x.VI1310            ------AGC--G----------------G-------T-----T-----------A-----C-------T--A---------GT---------T----------G-T--------------...--------------CT---G--T---------A---------- 4883
06_cpx.AU.96.BFP90           ------AGTA-G----------------A-AC-----------T-----------A--------C--G--A--------------T-G------------------T-------------...-----------------------TC---------A---T----- 5529
07_BC.CN.98.98CN009          --A----GC--G-------G---------T-------T-----T-----G-----------------G--A--A-----------------TT-----------C---------------AAT--T-----------------------GA-------G-- 4836
08_BC.CN.97.97CNGX_6F        --A----GC--G-------G---------A-------T-----T-----G---------C-----G--A--A-----------G-----TT-----------C---------------AAT--T-----------------------GA-----C-G-- 4684
09_cpx.GH.96.96GH2911        -----CAGC--G----------------A-AC----------T-----------A-----C-----G--A--A--A----G-T-G-----------------------------------...--T-----------T-------TC---------A---------- 4703
10_CD.TZ.96.96TZ_BF061       -------GC--G----------------A--C----T-----T-----G---A-----------G--A--A---------AG-----CT----------A-T---------------...--AC----------T-------T---------A-T-------- 4880
11_cpx.CM.95.95CM_1816       ------AGC--G----------------CA-------------T-----A---A-----C-----G--A--A---------T-G-----A----------------C-------------...-------------------------------GAA-------G- 4865
12_BF.AR.99.ARMA159          -------GC--G-------G---G--C---A-A-----T-----T-----------A-----C-------A--A---------T-A-A---------------A-C---------------...-------------------------T---------A----T----- 5506
13_cpx.CM.96.96CM_1849       -------GC--G-----------------AC-C-----------T---------------C-----G--A--A---------T-G-----------------C----------------...--T-------------------TC------T--A---------- 4906
14_BG.ES.05.X1870            --T-----CTY--------R-----------A-C-C--T---------------A-----A-----G-----AA--------GGA-------T---A-----------C-----------...TCT-------------------T--------GAA----G--G- 4968
15_01B.TH.99.99TH_MU2079     --T----------------G---G--------AC-A--T---------------A-----C-----G-GA---A---------AG------ATA---T-----T----C-T--------...-------------C-------T--------GAA--------C 4897
16_A2D.KR.97.97KR004         --------T-----G-G---T-----C------C----T----------------------C-----G------A--------A-A--------A--------------------------...--------------T-G------------T-GAA-------G- 4864
17_BF.AR.99.ARMA038          -------GC--G--G-------G--C---A-A-----T-----T-----------A-----C-----G--A--A--------A-T-A-A---------------A-C------------...-------------------------------A----------- 4715
18_cpx.CU.99.CU76            --T-----C--G--------------------------T-----------------A-----C-----G--T--A---------T-G-----T--------------------------...--T------------T-------TC-T---------------- 4806
19_cpx.CU.99.CU7             --T-----C--------T----G---------C----T-----T------------------C-----G--A------------AG------A-----------------------------...--T--------------------T-------GAA-------G- 4727
20_BG.CU.99.Cu103            --T-----C-----------CA-G--------------T--A-T--------------C--------A--A-------G-A-------A-------------T---C------------...--T--------------------T-------GAA-------G- 4967
21_A2D.KE.99.KER2003         --------T--G-------C-----C-----C----T------------------A---C-----G--A--A-------G-AC-------T-------------------------------...--------------------G-----T------T-GAA--T---G- 4706
22_01A1.CM.01.01CM_0001BBY   --T-----C--G----------------A-------T------------------------C-----G--A--A-----------A--------------T-----------A------...--T-----------C-------T------T-GAG-------G- 4709
23_BG.CU.03.CB118            -------GC-------------------AA-----T-----T-----------A-------------GT--A---------T-G-----T-----T-------------------------...--T--------------------T-------GAA-------G- 4944
24_BG.ES.08.X2456_2          --T-----C----------G---G------A-C----T--CA-A-----------------C-------A--C--------G-----T-----------------G----CT---------...--T--------------------T-------GAA-------G- 4952
25_cpx.CM.02.1918LE          -------GT--G-------G---------A--C-----------T-----------A-----C-----G--A--A---------T-----------A-----T------------------...--G--------------------T-----T--GAG-------G- 4709
26_AU.CD.02.02CD_MBTB047     --T----GC--------------------C---C-A--T-----A-----------------C-----G--A--A--G------G--A--TT-A--A-----------------------...--G-------G-----G-----T-------GAA-------G- 5506
27_cpx.FR.04.04CD_FR_KZS     -------GC--G----------------A-AC-----------T-----------A-----C-----G--A--A-------T-T-G-----T------------C---C-----------...--TC------------------TC--------A---GC---- 5492
28_BF.BR.99.BREPM12609       --------C---------------------------T--C--T-----------A-----------G--T--A--------GT--------A------------T-----A---------...-------------------G-------------A--------C 4867
29_BF.BR.01.BREPM16704       -------GC--G-----------G--C---A-AA----T-----T-----------A-----C-----G--A--A----------A-A-A---------------G--------------...-------------------G--------G--------G----- 4905
31_BC.BR.04.04BR142          -------GT--G-------G---------A--C----T-----T-----G---A-----C-----G--A--A----------T-G----------------C----------------...--TC-------------------G--C-------GA-----C-G-- 4973
32_06A1.EE.01.EE0369         ------AGT--G-----------------A--C-----------T-----------A-----C-----G--A--A----------T-------A-A-----------------G------...-------------------G--TC--------AA---C----- 5136
33_01B.ID.07.JKT189_C        --T-----C--------------G--------AA----T----------------------C-----G--A--A---------AG------A-A---------G---C-T------C...-----------------C-------T-------GAA--------C 4802
34_01B.TH.99.OUR2478P        --T-----C----------G---G--------AC-A--T---------------A-----C-----G-GA--A---------AG------A-G---------------CGT---------...-----------------C-------T-------GAA--------C 4703
35_AD.AF.07.169H             --T-----TG-G-------------------C----T--A-T------------------C-----G--A--A------G-A--------------T-----------------------...-----------------------T--A-----GAA---------- 4701
36_cpx.CM.00.00CMNYU830      --T-----C--------------G-----C---C----T----------------------C-----G--A--A---------A------------------T---C-------------...--T--------------------T-----G--GAA-------G- 4706
37_cpx.CM.00.00CMNYU926      --T-----C--------T-----G---------C----T----------------------C-----G-CA--A---------GA-------A-----------------C---------...-GT--------------------T-------GAA-------G- 4691
38_BF1.UY.03.UY03_3389       --------C--G-----------G--C---A-A-----T-----T-----------A-----C-----G--T--A---------T-G-A---------------A-C-------------...-----------------------T--------A---------- 4924
39_BF.BR.04.04BRRJ179        --------T---------------------------T-----T-----G---A------------GC-T--A---------G---------A-------CA-T----------------...-------------------G-----------A---------- 4978
40_BF.BR.05.05BRRJ055        A-------C------------------C-------T--A-T-----------A--------------T--C---G---G--------T-----T--A-T----------------------...------AA--------G-----------A---------- 4984
42_BF.LU.06.luBF_18_06       --------C--------------G-----C-------------T-----------A--------------TA--A---------T-G-----T----------------------------...---C-----------------G-----------A---------- 5055
43_02G.SA.03.J11223          --T-----C--------------G-----C---C----T-----------------A-----C-----G--A--A---------A-------A----------------------------...--T-----------A--G----T-------GAA-------G- 5026
44_BF.CL.00.CH80             -------GC--G-----------G--C---A-AC----T-----T-----A---------C-------A--C-C-----T-T-A---------------A-C---C--------------...-----------------C------C-----T--A-------G- 4943
45_cpx.FR.04.04FR_AUK        --T-----C--------------G------T----CC-T---------------------C-----G-----A-T--------AG------A-------------------A--------...-----------------------T--------AA-------G- 5487
46_BF.BR.07.07BR_FPS625      --------C--G-----------G--C---A-A-----T---------------------C-----G--A--A---------A-A---------------A-C-----------------...-----------------------T-------GA-------G- 4939
47_BF.ES.08.P1942            -------GC--G-----------G------A-A-----T-----T-----------A-----------G--A--A---------T-G---------------A-C----------------...-----------------------T--------A---------- 4943
48_01B.MY.07.07MYKT021       --T-----C--------------G------T-AC--C-T----------------------C-----G--A--A---------AG------A-G--------------C-A---------...--T--------------------T-------GAA--------C 4706
49_cpx.GM.03.N26677          -------GC--G-------G-----C---A--C----T-----T-----A---------C-------A--A---------T-------A---------TC--C-C---T----------...-----------------------TC----T--A----C----- 4923
50_A1D.GB.10.12792           --T------AG--------------------C----T------------------A-----C-----G--A--A---------AG------------T------------------------...--TC-------------------T--A-----AA---C--G- 4908
51_01B.SG.11.11SG_HM021      --T-----C--------------G--------AC-C--T----A-----------T----C-----G----AA---------A-------TA-A---------------CCT----------...-----------------C-------T--A----GAA-A------C 4708
52_01B.MY.03.03MYKL018_1     --T-----C--------------G--------AC-A--T-----------G---------C-----G--A--A---------AG------A-A--A-----------C-T----------...-----------------C-------T-------GAA--------C 4815
53_01B.MY.11.11FIR164        --T-----T--------------G-----C--AC-CC-T----------------------C-----G--TA--A---------A-G-A---AAG-----C-----C-C-T-------C...-----G---------C-------T-------GA---------C 4838
54_01B.MY.09.09MYSB023       --------C--------------------C-------T-----T-----------A--------------T--A-T------GTG-------T-----T---CT--------------...------A-------------G--T-------GAA--------C 4903
55_01B.CN.10.HNCS102056      --T-----C--------T-G---G--------AC-C--T---A------------------C-----G-GA--A---------AG------A-A--A---------C-T----------...-----------------C-------T-----T--GAA--------C 4811
56_cpx.FR.10.URF5_patient_A  --T-----C--------------G-----C---C----T-----------------A-----C-----G--A--A---------AG------A-----T----------------------...TCT-------------G--G-----T-------GAA-------G- 4852
57_BC.CN.09.09YNLX19sg       --A--C--C--G-------T-----------------T-----T-----G---------C-----G--A--A------------------TT-----------T---------------AAG--T------------------C--T-------GA----GC-G-- 4688
58_01B.MY.09.09MYPR37        --T-----C--------------G--------AA----T----------------------C-----G--A--A---------AG------A-G---------------C-T------C...--T-----------C-------T-------GAA--------C 4846
59_01B.CN.09.09LNA423        --T-----C----------G---G--------AC-A--T----------------------C-----G--A--A-T--------AG------A-G---------------C-C-T-------...-----------------C-------T-------GAA--------C 4685
60_BC.IT.11.BAV499           --------C--G-----------------G--C----T-----T-----G-----------C-----G--A--A---------T-G-----------------T------------------...--T-----------G--------C-------GA-------G-- 5458
61_BC.CN.10.JL100010         --A-----C--G-------G---------T-------T-----T-----G-----------C-----G--A--A---------T------TT-----------C---------------AAT--T--------------------T-------GA-----C-G-- 4925
62_BC.CN.10.YNFL13           --A----GCA-G-------G---------T--C----T-----T-----G-----------C-----G--A--A------------------TT-----------C-------------------......-------------------------------GA-------G-- 4827
63_02A1.RU.10.10RU6637       --T-----C--------------G-----C---C----T-----A-----------AA---C-----G--A--A--G-----G-A---------A---A--A-C---TC---------C...-------------------------------GAA-------G- 5042
64_BC.CN.09.YNFL31           --A----GC--G-------G---------A-------T-----T-----G-----------------G--A--A------------------TT-----------C---------------ACT--T--------------------TC-------GA-------G-- 4858
65_cpx.CN.10.YNFL01          --A-----C--G-------G-----------------T-----T-----G---------C-----G-GA--A----------A-----TTG---------A-T----------------...--T--------------------TC-------GA-------G-- 4872
67_01B.CN.11.ANHUI_HF115     --T----GC--------------G--------AC-C--T----------------------C-----G--A--A---------AG------A-G---------------C-C-T----------...-----------------C-------T-------GAA--------C 4844
68_01B.CN.11.ANHUI_WH73      --T-----C--------------G--------AC-C--T----------------------C-----G--A--A---------AG------A-G-----C-----C-C-T----------...-----------------C-------T-------GAA--------C 4835
70_BF1.BR.10.10BR_PE004      -------GC--T-----------G--C---A-A-----T-----T-----------A--------------T--A---------A-----------------------Y----------C...-G------------Y---G-------------A---Y-W---G 4957
71_BF1.BR.10.10BR_PE008      -------GCAC------------G--C---A-A--C--T-----T-----------A-------------A---------T-G-----T-------------------------------...--------------------G---C-A---------G-------C 5165
72_BF1.BR.10.10BR_MG002      ------A-C-----------------------C----------T-----------A------------GT---A---------GT--------A----------------------------...-------------G--C---------------------------- 5009
O.BE.87.ANT70                ---------AC-------AGGA-G--A------C----T---AC-------A---CA--A--AG-C--G--A---A-CC-------G-G--ACTGA-C-A---------C-CTA-----------...-GTC--------GA-A------CTGC-A--T---AG----G--G- 5555
O.CM.98.98CMA104             ---------AC------TAG-A-G--A-----------T--CAC-------A----G--A--A--C--G--A--ACC-----G--G-G--ACTGA-C-A------C--C-CT-A---------...-GT--------GA-A--G---CTAC-A--TT--AG--T-G-TG- 4987
O.CM.98.98CMABB141           ---------AC-------AGGA-G--A--------C------AC-------A----A--A--A--C--GG-AA--A-CC----G--G-G--ACTGA-C-A---------CTCT----------...-GTC--------GA-A------CTAC-A--T---AG---TG--G- 4986
O.CM.98.98CMABB212           --T------AC-------AGGA-G--A----AC---T-------------A-------C-A--C----------A-TC-------G-G--GCTGA---A---------C-CT----------...-GTC--------GA-A-------TT--A--CT--AAG---A--GA 4991
O.CM.98.98CMU5337            ---------AC------TAGGA----A-----------T--AC-------A----G--A--G--C--G--A---A-C-----G--G-G--ACTGA-C-A---------C-CT----------...-GTC--------GA-AT-G---CTAC-A--T---AG--TGG--G- 4988
O.CM.99.99CMU4122            ---------AC-------AGGA----A----AAC----T--AC-------A----G--A--A--C--G--A---A-CA----G--G-G-A-AYTGA-C-A---------CTCT----------...-GTC--------GA-AT-G---CTAC-A--T---AG--T-G--G- 4986
O.FR.92.VAU                  ---------AC-------A-G-----A------C-------AC-------A----G--A--A--T--G-GA---ACTC----G--G-G--GCTA--CCA---------C-CT-A--------...-GTC--------GA-A-----GCT-C-A--CT--AG----G--G- 5067
O.GA.11.11Gab6352            ---------AC-------AGGA-G--A--A-A-----T---AC-------A----A--A--A--C--G--A---A-C-----G--G-G--ACTGGTC-A---------TC-CT----------...-GTC--------GA-AT-----CTAC-A--T---AG--T-C--GA 4712
O.SN.99.99SE_MP1299          ---------AC-------AG-A-G--A--A-A-----T----C-------A----G--A--A--C--G--A---A-CC----G--G-G--ACTGA-C------------C-CT----------...-GTC--------GA-A--G---CTAC-A--T---AG--TGG--G- 5554
O.US.10.LTNP                 ---------AC-------AGGA----A----A------T--CAC-------------A--A--A--C--G-GA---A-CC----G--G-G--ACTGA-C-A---------C-CT----------...-GTC--------GA-A--G---CTAC-A--T---AG----G--G- 5476

N.CM.02.DJO0131              --T------A-G-----A-G-----A-----T-----T---A-T------A---C------C-----GC-A--GG-C-----GAGACC---ATT-------------------CG-----G--C...--TC--------CA--T-G---C-TC-A-----CA----C-GGG- 5007
N.CM.02.SJGddd               --T------A-G-------G-----A----AT-----T---A-T------G---CT-----C-----GC-A--GA-C-----GAGACC---AGTA----A------------C------G--C..C-AC--------CA-CT--------TC-A-----AG---C-GGG- 4991
N.CM.04.04CM_1015_04         --T------A-G-----A-G-----A-----T-----T----T------A----T-----C-----GC-A--GG-C-----GAGACA---ATTT-------------------CT-----G--C...---C--------ATA-CT--------TC-A---T--A----T-GGG- 5006
N.CM.06.U14296               --T------A-G-------GG-----A-----T-----------A-T------A---C-------C-----GC-A--GGCC-----GAGACC---ATTA----A------------C------G--C...--AC--------CA-CT--------TC-A------A----C-GGG- 5004
N.CM.06.U14842               --T-----GA-G-----A-G-----A-----T-----T---A-T------A---C-------C-----GC-A--GG-C-----GA-ACC---ATTA--------------------C------G---..C-AC--------CA-CT--------TC-A-----CA----C-GGG- 5008
N.FR.11.N1_FR_2011           --T------A-G----TA-G-----A-----T-----T---A-T------A---C-------C-----GC-A--GG-C-----GAGACA---A-TA----A------------C------G--C...--AC--------CA-CT--------TA-A------A----C-GGG- 4902

P.CM.06.U14788               --T--A---AC------TAGG-----------CC-C------C-A-----------T--G---G-G--G-G---AA--C-----GG---GG--ATTGAAY-MA---------CTC------------...-GTC--------C--CT-G--G--T--A--CT--AA--TGG--G- 4972
P.FR.09.RBF168               --T--A---AC------TAGG-----------CC-C------C-A-----------T--G---G-G--G-G---AA--C-----GG---GG--ATTGAAC-A---------CTC------------...-GTC--------C--CT-G--G-TT--A--CT--AA--TGG--G- 5536

CPZ.CD.06.BF1167             -----A---AC---T------A-T--A---AGTAGA--T---CCT------A----AAGG--C-------C-A--TA-T--G---G-A-AGA-CCTA-AGCAT--CC-T-TC-CTA-----------...TCT--------GAGC------G-TTC-A--CT-TAA-AA-G-GC- 5606
CPZ.CM.05.SIVcpzMT145        ---------AC------TAGG-----A--C--TC----T-----T------A---CAA----C--C--G--C--TA-C----A-A-GGC---CCT-----A--------TC-C--A---G--C...-G----------G---T---------G---T---AAGCAG--GG- 5043
CPZ.GA.88.GAB1               --T--CTTCAC------TAG-T-G--C--CTCTC-A--T-----T--C---G-----AA-----C--C-----GA--TA-------G--GC-T--AGCG---C-T---------CA---A------C...-GGC-------------CT----G-T------C---AA---T------ 5565
CPZ.TZ.06.SIVcpzTAN13        ---------ACG--------G--T-----CACTA-G-------CA------A-T-CAGGA--AG-GC-GC-A--AA-C--------G-A-GGA-ACTGA-GTTC--CC-T--------A-----------...-GAC--------GA--T--------TC-A--TT-TCT-AA-G--G- 5145
CPZ.US.85.US_Marilyn         --T-----GA-G------A-G-------------------T-----T------A---C-------C--C------A--TG-T-----GAGACC---A-TT--C-AA----------C-CTA---G--C...--T--------C---T-G----T-------TT--AAG--C---G- 5561

GOR.CM.07.SIVgor2139_287     --T--A---AC------TAGGA----C-----T---------C-A------A----A-----AG-G--G--A--AA-TG-G--GG-A-A-C-CCTGTAC-A----A-C--C-CG-----------...-GTC--------GA-AT-G---------A---T--AG---T---G- 5038
GOR.CM.07.SIVgorCP2135con    --T--A---AC------TAGGA----C-----C---------C-A------A----A-----AG-G--G--A--AA-TG-G--GG-A-A-C-CCTGTGC-A----A-C--C-C--------------...-GTC--------GA-AT-G---------A------AG---T---G- 5037
```

Vpr start
Vif end
B.FR.83.HXB2
Vif
Vpr
A1.AU.03.PS1044_Day0
A1.CH.03.HIV_CH_BID_V3538
A1.CY.08.CY236
A1.ES.05.X1608_8
A1.KE.11.DEMA11IKE001
A1.RU.11.11RU6950
A1.RW.11.DEMA111RW002
A1.SN.01.DDI579
A1.UG.11.DEMA110UG001
A1.ZA.04.04ZASK162B1
A2.CM.01.01CM_1445MV
A2.CY.94.94CY017_41
B.BR.10.10BR_MG029
B.CA.07.502_1191_03
B.CH.08.M2_0803101_NFLG8
B.CN.10.DEMB10CN002
B.ES.10.DEMB10ES002
B.FR.11.DEMB11FR001
B.GB.05.MM45d213_GN1
B.HK.06.HK003
B.HT.05.05HT_129389
B.JP.12.DEMB12JP001
B.KR.07.HP_18_07JHS10_3909
B.PE.07.502_0525_wg5
B.RU.11.11RU21n
B.TH.08.MERLBDTRC10
B.US.11.ES38
C.AR.01.ARG4006
C.BR.07.DEMC07BR003
C.BW.00.00BW5031_1
C.CN.10.YNFL19
C.CY.09.CY260
C.ES.08.X2363_2
C.ET.02.02ET_288
C.IN.09.T125_2139
C.KE.00.KER2010
C.MW.09.703010256_CH256.w96
C.TZ.08.707010457_CH457.w8
C.YE.02.02YE511
C.ZA.08.705010534_CH534.w12
C.ZA.10.DEMC10ZA001
C.ZM.11.DEMC11ZM006
D.CM.10.DEMD10CM009
D.CY.06.CY163
D.KE.11.DEMD11KE003
D.KR.04.04KBH8
D.SN.90.SE365
D.TZ.01.A280
D.UG.10.DEMD10UG004
D.UG.11.DEMD11UG003
D.YE.02.02YE516
D.ZA.90.R1
F1.AO.06.AO_06_ANG32
F1.AR.02.ARE933
F1.BR.10.10BR_RJ015
F1.CY.08.CY222
F1.ES.02.ES_X845_4
F1.RO.96.BCI_R07
F1.RU.08.D88_845
F2.CM.02.02CM_0016BBY
F2.CM.10.DEMF210CM001
F2.CM.10.DEMF210CM007
G.BE.96.DRCBL
G.CM.10.DEMG10CM008
G.CN.08.GX_2084_08
G.CU.99.Cu74
G.ES.09.X2634_2
G.GH.03.03GH175G
G.KE.09.DEMG09KE001
G.NG.09.09NG_SC62
G.ZA.01.TV546
H.BE.93.VI991
H.CF.90.056
H.GB.00.00GBAC4001
J.CM.04.04CMU11421
J.SE.93.SE9280_7887
J.SE.94.SE9173_7022
K.CD.97.97ZR_EQTB11
K.CM.96.96CM_MP535
U.CA.01.TV749
U.CA.99.TV721
U.CD.83.83CD003_Z3
U.CD.90.90CD121E12
U.CY.05.CY090
U.CY.08.CY223
U.ES.10.DEURF10DZ001
U.GR.99.99GR303
U.NL.95.U_NL_95_H10986_D1
01_AE.AF.07.569M
01_AE.CF.90.90CF11697
01_AE.CN.10.YNFL03
01_AE.HK.04.HK001
01_AE.IR.10.10IR.THR48F
01_AE.JP.x.JRC77AE
01_AE.TH.09.AA090a_WG11
01_AE.TH.90.CM240
01_AE.US.05.306163_FL
01_AE.VN.98.98VNND15

Vpr start
Vif end
B.FR.83.HXB2
Vif
Vpr
02_AG.CM.10.DE00210CM013
02_AG.CY.09.CY256
02_AG.ES.06.P1423
02_AG.GW.05.CC_0048
02_AG.KR.12.12MHR9
02_AG.LR.x.POC44951
02_AG.NG.09.09NG_SC61
02_AG.NG.x.IBNG
02_AG.SE.94.SE7812
02_AG.SN.98.98SE_MP1211
03_AB.RU.97.KAL153_2
04_cpx.CY.94.94CY032_3
05_DF.BE.x.VI1310
06_cpx.AU.96.BFP90
07_BC.CN.98.98CN009
08_BC.CN.97.97CNGX_6F
09_cpx.GH.96.96GH2911
10_CD.TZ.96.96TZ_BF061
11_cpx.CM.95.95CM_1816
12_BF.AR.99.ARMA159
13_cpx.CM.96.96CM_1849
14_BG.ES.05.X1870
15_01B.TH.99.99TH_MU2079
16_A2D.KR.97.97KR004
17_BF.AR.99.ARMA038
18_cpx.CU.99.CU76
19_cpx.CU.99.CU7
20_BG.CU.99.Cu103
21_A2D.KE.99.KER2003
22_01A1.CM.01.01CM_0001BBY
23_BG.CU.03.CB118
24_BG.ES.08.X2456_2
25_cpx.CM.02.1918LE
26_AU.CD.02.02CD_MBTB047
27_cpx.FR.04.04CD_FR_KZS
28_BF.BR.99.BREPM12609
29_BF.BR.01.BREPM16704
31_BC.BR.04.04BR142
32_06A1.EE.01.EE0369
33_01B.ID.07.JKT189_C
34_01B.TH.99.OUR2478P
35_AD.AF.07.169H
36_cpx.CM.00.00CMNYU830
37_cpx.CM.00.00CMNYU926
38_BF1.UY.03.UY03_3389
39_BF.BR.04.04BRRJ179
40_BF.BR.05.05BRRJ055
42_BF.LU.06.luBF_18_06
43_02G.SA.03.J11223
44_BF.CL.00.CH80
45_cpx.FR.04.04FR_AUK
46_BF.BR.07.07BR_FPS625
47_BF.ES.08.P1942
48_01B.MY.07.07MYKT021
49_cpx.GM.03.N26677
50_A1D.GB.10.12792
51_01B.SG.11.11SG_HM021
52_01B.MY.03.03MYKL018_1
53_01B.MY.11.11FIR164
54_01B.MY.09.09MYSB023
55_01B.CN.10.HNCS102056
56_cpx.FR.10.URF5_patient_A
57_BC.CN.09.09YNLX19sg
58_01B.MY.09.09MYPR37
59_01B.CN.09.09LNA423
60_BC.IT.11.BAV499
61_BC.CN.10.JL100010
62_BC.CN.10.YNFL13
63_02A1.RU.10.10RU6637
64_BC.CN.09.YNFL31
65_cpx.CN.10.YNFL01
67_01B.CN.11.ANHUI_HF115
68_01B.CN.11.ANHUI_WH73
70_BF1.BR.10.10BR_PE004
71_BF1.BR.10.10BR_PE008
72_BF1.BR.10.10BR_MG002
O.BE.87.ANT70
O.CM.98.98CMA104
O.CM.98.98CMABB141
O.CM.98.98CMABB212
O.CM.98.98CMU5337
O.CM.99.99CMU4122
O.FR.92.VAU
O.GA.11.11Gab6352
O.SN.99.99SE_MP1299
O.US.10.LTNP
N.CM.02.DJO0131
N.CM.02.SJGddd
N.CM.04.04CM_1015_04
N.CM.06.U14296
N.CM.06.U14842
N.FR.11.N1_FR_2011
P.CM.06.U14788
P.FR.09.RBF168
CPZ.CD.06.BF1167
CPZ.CM.05.SIVcpzMT145
CPZ.GA.88.GAB1
CPZ.TZ.06.SIVcpzTAN13
CPZ.US.85.US_Marilyn
GOR.CM.07.SIVgor2139_287
GOR.CM.07.SIVgorCP2135con

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

Tat exon 1 start
Vpr end
Rev exon 1 start
B.FR.83.HXB2
Vpr (frameshifted)
Tat exon 1
Rev exon 1
A1.AU.03.PS1044_Day0
A1.CH.03.HIV_CH_BID_V3538
A1.CY.08.CY236
A1.ES.05.X1608_8
A1.KE.11.DEMA11KE001
A1.RU.11.11RU6950
A1.RW.11.DEMA111RW002
A1.SN.01.DDI579
A1.UG.11.DEMA110UG001
A1.ZA.04.04ZASK162B1
A2.CM.01.01CM_1445MV
A2.CY.94.94CY017_41
B.BR.10.10BR_MG029
B.CA.07.502_1191_03
B.CH.08.M2_0803101_NFLG8
B.CN.10.DEMB10CN002
B.ES.10.DEMB10ES002
B.FR.11.DEMB11FR001
B.GB.05.MM45d213_GN1
B.HK.06.HK003
B.HT.05.05HT_129389
B.JP.12.DEMB12JP001
B.KR.07.HP_18_07JHS10_3909
B.PE.07.502_0525_wg5
B.RU.11.11RU2ln
B.TH.08.MERLBDTRC10
B.US.11.ES38
C.AR.01.ARG4006
C.BR.07.DEMC07BR003
C.BW.00.00BW5031_1
C.CN.10.YNFL19
C.CY.09.CY260
C.ES.08.X2363_2
C.ET.02.02ET_288
C.IN.09.T125_2139
C.KE.00.KER2010
C.MW.09.703010256_CH256.w96
C.TZ.08.707010457_CH457.w8
C.YE.02.02YE511
C.ZA.08.705010534_CH534.w12
C.ZA.10.DEMC10ZA001
C.ZM.11.DEMC11ZM006
D.CM.10.DEMD10CM009
D.CY.06.CY163
D.KE.11.DEMD11KE003
D.KR.04.04KBH8
D.SN.90.SE365
D.TZ.01.A280
D.UG.10.DEMD10UG004
D.UG.11.DEMD11UG003
D.YE.02.02YE516
D.ZA.90.R1
F1.AO.06.AO_06_ANG32
F1.AR.02.ARE933
F1.BR.10.10BR_RJ015
F1.CY.08.CY222
F1.ES.02.ES_X845_4
F1.RO.96.BCI_R07
F1.RU.08.D88_845
F2.CM.02.02CM_0016BBY
F2.CM.10.DEMF210CM001
F2.CM.10.DEMF210CM007
G.BE.96.DRCBL
G.CM.10.DEMG10CM008
G.CN.08.GX_2084_08
G.CU.99.Cu74
G.ES.09.X2634_2
G.GH.03.03GH175G
G.KE.09.DEMG09KE001
G.NG.09.09NG_SC62
G.ZA.01.TV546
H.BE.93.VI991
H.CF.90.056
H.GB.00.00GBAC4001
J.CM.04.04CMU11421
J.SE.93.SE9280_7887
J.SE.94.SE9173_7022
K.CD.97.97ZR_EQTB11
K.CM.96.96CM_MP535
U.CA.01.TV749
U.CA.99.TV721
U.CD.83.83CD003_Z3
U.CD.90.90CD121E12
U.CY.05.CY090
U.CY.08.CY223
U.ES.10.DEURF10DZ001
U.GR.99.99GR303
U.NL.95.U_NL_95_H10986_D1
01_AE.AF.07.569M
01_AE.CF.90.90CF11697
01_AE.CN.10.YNFL03
01_AE.HK.04.HK001
01_AE.IR.10.10IR.THR48F
01_AE.JP.x.JRC77AE
01_AE.TH.09.AA090a_WG11
01_AE.TH.90.CM240
01_AE.US.05.306163_FL
01_AE.VN.98.98VNND15

```
                              Tat exon 1 start  Vpr end                                                                                                    Rev exon 1 start
B.FR.83.HXB2                 .CAGAGGAGAGCA....AGAAATGGAGCCAGTAGATCCTAGACTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTGCCAAGTTTGTTTCATAACAAAAGCCTTAGGCATCTCCTATGGCAGGAA 5979
Vpr (frameshifted)           __Q__R__R__A__._#_R__N__G__A__S__R__S__*
Tat exon 1                                     _M__E__P__V__D__P__R__L__E__P__W__K__H__P__G__S__Q__P__K__T__A__C__T__N__C__Y__C__K__K__C__C__F__H__C__Q__V__C__F__I__T__K__A__L__G__I__S__Y__G__R__K
Rev exon 1                                                                                                                                                                     _M__A__G__
02_AG.CM.10.DE00210CM013     .GG---A----GC....--G--------------------C-------------T--C--G-----------GC----------G------T---------TG---C----GG-----T---C-A--C--TT-G-AC--G-G----------------------- 5345
02_AG.CY.09.CY256            .GG--------GC....--G--------TG----------A--------------A--C-----------------C---------C-A--CG--T--C---------A--C----GG---------T-A--C--TC-G-A-----G----------------------- 5174
02_AG.ES.06.P1423            .AG--AWG--A-T....G-G--A...---G-----------C--------------C--C--G-----------------------G------T----------------C----GG---------TA---C--TC-G-AC--G-G----------------------- 5412
02_AG.GW.05.CC_0048          .GGA--A----GC....--G--------------------C--------------C--C--G-----------C-C-----------G---G--T--C--------TA--C----GG---------T-G--C--TC-G-A---G-G----------------------- 5328
02_AG.KR.12.12MHR9           .GG---A----GC....--G--------TG----------C-----A--A-----C--C--G---------A----C-----------G---G--T--C---------A--C----GG---------T-G--C--TC-G-AC--G-G----------------------- 5637
02_AG.LR.x.POC44951          .GG---A----GC....--G------T-------------CA-------A-----T--C--G-----------------C-----G----G---G--T--C------GC---C----GG-----T---T-G--C--TC-G-AC--G-G----------------------- 5973
02_AG.NG.09.09NG_SC61        .GG---A----GC....--G--------TG----------C--------------C-----G-----------C-C-----------A---------C---------A--C----GG---------T-G--C--TC-G-AC--G-G----------------------- 5170
02_AG.NG.x.IBNG              .GG---A----GC....--G--------TG----------C--------------C--C--G-----------------C-----------G---G--T--C-------T---C----.G---------T-G--C--TC-G-AC--G-G----------------------- 5502
02_AG.SE.94.SE7812           .GG---A----GC....------------TG----------C--------------C--C--G-----------------C-----------G---A--T--C------TC--------GG---------T-A--C--TC-G-AC--G-G----------------------- 5350
02_AG.SN.98.98SE_MP1211      .GG---A----GC....--G--------TG----------A--------------C--C-----------------C-----------G---G--T--C------TT---C----GG---------T-A--C--TC-G-AC--G-G----------------------- 5175
03_AB.RU.97.KAL153_2         .-G-------------....---------------------------------------C------------------------------------------------TC---------------------G-A-----G----------------------- 5204
04_cpx.CY.94.94CY032_3       .--A------AG-GGCAG-GG----------------GAC-------------C-----G------------C-----A-----A---G--T-TC----------------GG--------------C--TC-G-A-----G----------------------- 5348
05_DF.BE.x.VI1310            .-----A---AT-....-----------T-----------AC--------------T------------------GG----------A-C-G-------------T--------A-----------AA---C---------G----G----------------------- 5360
06_cpx.AU.96.BFP90           .-----A----GG....-----------------------A-A-------------T--G---------G-----G--------------A---------------------A-------C---G--C--TC---AC--G-G---------T------------ 6006
07_BC.CN.98.98CN009          .-----A---A--....-----------------------A-T-----T-T---------------------G-----------------------------------C---------------------------G-A-----G-----------T------------ 5313
08_BC.CN.97.97CNGX_6F        .-----A-----....------------------------AC--------------C------------------G----------C-AT--C------------CGC---A---AC-----T-T------C--TCAG-A-----G---------T------------ 5161
09_cpx.GH.96.96GH2911        .AG---A----GC....--GG-------TG----------AC--------------C------------------GC----------A---A--T--C-------GA-------A-------------CAC--C---T-G-AC----G----------------------- 5183
10_CD.TZ.96.96TZ_BF061       .-----A-----....-----------A------C-----AC--------------C--C-----------C------------------AT--A---C-------GCA---A---A-----T-T------C--TCAG------AG---------T------------ 5357
11_cpx.CM.95.95CM_1816       .GG---A----GC....--G--------------------ACA-------------T--G---------C--------G-----------A-C-G--T--------------A--C----A-------------CA------TC---AG----G---------T--A---------- 5342
12_BF.AR.99.ARMA159          .-----A----T-....C-----------T----------ACT-------------C--C---------CA-----TC----C-----------A--T---------CG-------------------TT-CTGG--C---G---------G----------------------- 5983
13_cpx.CM.96.96CM_1849       .--A--A----GT....-----------T-----------A-A-------------C--G-----------------------C-----------G--T-------------------------------T---------C-----------G-G----------------------- 5383
14_BG.ES.05.X1870            .-----A----T-....--GG-------T-----------AC-------T------T-------G------------------------------A-----------CG-------GG----T----------C--TC-G-AC---G----------------------- 5445
15_01B.TH.99.99TH_MU2079     .GG---A-----C.AGG--G------C--G-------T--AC--------------T-----G------G----C-----------A--------T--C-------TA------GG----------A-A--C--TC-G-A-----G----------------------- 5374
16_A2D.KR.97.97KR004         .-CA--AG-----....-----------T-----------AC--------------C----------C------GG----------A---G--T-----------G--------A------T-C---G--C---C---AC----G----------------------- 5338
17_BF.AR.99.ARMA038          .-----A----TC....------------AT----------ACT----T--------C-----------C------C----C---------GA--T-T--------G-----------------TT-CTGG--C--T-C---------G------------AA------------ 5192
18_cpx.CU.99.CU76            .--A--A----GG....-----------------------ACA-------------C--G-----------------------------A---------T-------GA-------A------T---C-A--C--TT---AG----G---------T------------ 5283
19_cpx.CU.99.CU7             .-GA--A----T-....--GG-----C--G----------AT--------------T-----G--G------------------------A--CA---------------------------T---C-G--C--T-C--AG-G-G----------------------- 5204
20_BG.CU.99.Cu103            .-G---A----T-....--G------C-------------AC--------------T-----G--G------------------------A---G-----------GT-------GG----T---------C--TC-G-AC--G-G----------------------- 5444
21_A2D.KE.99.KER2003         .-----------....------------T-----------AC--------------C------------------GG----------A---G--TC---------G--------A-------AT-AAG--C---T-G--G---G---------CG------------ 5183
22_01A1.CM.01.01CM_0001BBY   .GG---A----GC....------------TG----------C--------------T-----G------------C---G--------AT-----T-T------GT-------GG---------------C--TC---A-----G----------------------- 5186
23_BG.CU.03.CB118            .-G---A----G-....--G------C--G----------AT-----C---------T-----G--G------------------------A---G------------T-------GG----T---------C--TC-G-A---G-G------A---------------- 5421
24_BG.ES.08.X2456_2          .-G---A---AG-....--GG-----T--G----------AC--------------T-----G--G------------------------A---A----T------GT-------GG----T---------C--TT-G-AC----G----------------------- 5429
25_cpx.CM.02.1918LE          .-----A----GG....-----------C-----------ACA-------------T--------------A---GG----C------A-----------------G---C----------T---------C--T----AG-G-G----------------------- 5186
26_AU.CD.02.02CD_MBTB047     .AGC--A----GC....--G---------G-----C-G--ACA----C--------C-----G-----------C------------A---A-------------C--------AC-------GC------C--T-----C----G---------A------------ 5983
27_cpx.FR.04.04CD_FR_KZS     .-----A----GG....-----------C------C----AC-----A--------T--G------------A-------------A-C-G--T-----------G--------A-----------CA---C--TT---AG----G---------T-T------------ 5969
28_BF.BR.99.BREPMI2609       .-----------....-----------------------------------C-A--C-----G----------GG----------A---A-------------------------C--------------------------G-A-----G----------------------- 5344
29_BF.BR.01.BREPM16704       .-----A-----....-------------------------------------------C-----G----------G-----------A---A-------------------------------------------C-------------G-----------A----------- 5382
31_BC.BR.04.04BR142          .--A--A-----....-----------T-----------A--------------C---------------------C-AT----T--C-------GA--CA---A------T-----G--C--TC-G-------G---------T------------ 5450
32_06A1.EE.01.EE0369         .--A--A----GG....-----------A---------A--A-A----A--------T--G---------G-----G--C-G-----CCA--T-------------------A-------T-CA-G--C---T---AC--G-G---------T-A------------ 5613
33_01B.ID.07.JKT189_C        .GG---A----GC....--G--------G-----------AC--------------T-----G--------------G--A-------G---G--T--C-------TA------GG---------C-A--C--TC-G-A-----G----------------------- 5279
34_01B.TH.99.OUR2478P        .GG---A----GC....--G--------------------AC--------------T-----G------------C-----------G---G--T--C--------T------GG---------C-A--C--TC-G-A-----G----------------------- 5180
35_AD.AF.07.169H             .GG---A----TC....--GG-----T--G-----C----AT-----T-------C--C--------------CC---GC-----G------T--C-------TA---A-T-A------T-A------T---------C--TC-G-AC---G----------------------- 5178
36_cpx.CM.00.00CMNYU830      .GG---A----GC....--G--------TG----------ACA-------------T-----G----C------G----------A---C--T-T-------------C----GG-------------------C--TC-G-AC--G-G----------------------- 5184
37_cpx.CM.00.00CMNYU926      .-GC--A----TG....--G------C-------------AT--------------------G--G------------------------A---C--------C--------C-----------T---------C--T-C--A---G-G------A---------------- 5169
38_BF1.UY.03.UY03_3389       .-----A----T-....-----------AAT---------A-T-------------C----------C------C---C---------A---A-------------CG-----------------TT-CTGG--C--T-C---G--G-G---------A---C---------- 5401
39_BF.BR.04.04BRRJ179        .-------------....--G------T-----------------------------------------------G-----------T--A-----------------C--------------------------------C-------G----------------------- 5455
40_BF.BR.05.05BRRJ055        .-T-------T--....--G------T-----------------------------------------------G-----------T--A--------------G--------------------------------C--------C---G----------------------- 5461
42_BF.LU.06.luBF_18_06       .-G-----------....------------------A--AGT----------G-----------------------C-----------------------G-------------------------------------C-------G-----G----------------------- 5532
43_02G.SA.03.J11223          .GG---A----GC....--GC-------TG----------C--------T----C--C--G-----------C-----------G---G--T--C-------------C----GG---------T-A--C--TC-G-AC--G-G----------------------- 5503
44_BF.CL.00.CH80             .-----A----T-....-----------T-A---------ACT----T--------T-----------C------C---C---------GA-----------------CG-----------------TT-CTGG--C--TG-G--G-G-G----------------------- 5420
45_cpx.FR.04.04FR_AUK        .GG---A----TC....--GG-----C--G----------AC-------T-----C--C--G-----------C-----------A---A--TC-C---------A-------AC-----T-CAG---C--TT-G-AC----G----------------------- 5964
46_BF.BR.07.07BR_FPS625      .-----A----TG....------------TG----------ACT----T--------C-----------C------C---C---------A--T-----------CG-----------------T---TGG--C--TG---G--G-G---------A------------ 5416
47_BF.ES.08.P1942            .-----A---A--....------------------------------------------------C------GG----G---------A--T-------G-GA----------------------------C------------G----------------------- 5420
48_01B.MY.07.07MYKT021       .GG---A----GC....--G---------C----------AC-------------------G------A---A-C-----------G---G--T--C---------A-------GG---------G-C-A--C--TC---A-----G----------------------- 5183
49_cpx.GM.03.N26677          .-----A----GG....--G-------------C---A-A-------------T----------------------G--------C---C-G--T--------------A-------A-------------A-A--C--TC---AG--G-G------A--T-G------------ 5400
50_A1D.GB.10.12792           .GG---A----GC....--G------T-------------AC--------T-----T--C--G-----------CT---C-G-------G--T--C------------------A-------T-C--C---C--TT-G-AC----G----------------------- 5385
51_01B.SG.11.11SG_HM021      .GG---A----GC....--G-------------C-----AT--------------T-----------------GGG---C----C--TGC--------------G----A-----------------C------------A----------------------- 5185
52_01B.MY.03.03MYKL018_1     .GG---A----GC....--G--------------------AC--------------T-----G------------C-----------G-TCG--T--C-------TA------GG---------T-A--C--TC-G-A-----G-------------C----------- 5292
53_01B.MY.11.11FIR164        .GG---A----GC....--G---------------------C--------------T-----G------R-----C-----------G---A--T--C---------A------GG---------C-A--C--TC-G-A---G-G----------------------- 5315
54_01B.MY.09.09MYSB023       .GG---A----GC....--G---------G----------AC------------------------------C-----------G---A--T--C-------GA------GG---------C-A--C--TC-G-A-----G----------------------- 5380
55_01B.CN.10.HNCS102056      .GG---A----GC....--G---------G----------AC--------------T-----G------A------C-----------G---G--T--C---------A------GG----T---C-A--C--TC-G-A---G-G----------------------- 5288
56_cpx.FR.10.URF5_patient_A  .--AG-A----T-....--GG-----T--GA---------AT--------------T-----G--G------------------------A---A-----------GC---C----GG---------C----C--TC-G-AC----G---S----T--Y------------ 5329
57_BC.CN.09.09YNLX19sg       .-----A-----....------------------------AC--------------C-------------------------------AT--C--T---------CGC---A---AC-----T-T------C--TCAG-A-----G---------T------------ 5165
58_01B.MY.09.09MYPR37        .GG---A----GC....--G---------G----------AC--------------T-----G------------C-----A-----G---G--T--C-------TA------GG---------C-A--C--TC-G-A-----G----------------------- 5323
59_01B.CN.09.09LNA423        .GG---A----GC....--G---------G----------AC--------------T-----G-----A-A----C-----------G---G--T--C---------A------GG---------C-A--C--TC-G-A-----G-------------C----------- 5162
60_BC.IT.11.BAV499           .----AT--G--G....-----------TC----AA---AC--------------C--A-----A-----------------C----C-AT-----T---------CGA--CA---A------T---------C--TC-G-------G---------T-A------------ 5935
61_BC.CN.10.JL100010         .-----A-----....------------------------AC-----A--T-----C------------------G---------C-AT--C----TC-------CGC---A---A------T-T------C--TCAG-------G---------T------------ 5402
62_BC.CN.10.YNFL13           .-----A-----....------------------------AC--------------T------------------G---------C-AT--C--------------CGC--CA----C-----T-T------C--TCAG-A-----G------A--T------------ 5304
63_02A1.RU.10.10RU6637       .GG---A----GC....--G------T-------------AC--------------T-----G----------------------G------T--C----------C----GG-------T-G--C--TT-G-AC--G-G----------------------- 5519
64_BC.CN.09.YNFL31           .--A--A-----....------------------------AC--------------C------------------G---------C-AT--C--T-T---------A---A---AC-----T-T------C--TCAT-G-----G-------G-T------------ 5335
65_cpx.CN.10.YNFL01          .GG---A----GC....--G--------------------AC-----A--T-----------------------C-----------A---------T---------A--C---------------R-----------CG-A------AC---------------------- 5343
67_01B.CN.11.ANHUI_HF115     .GG---A----GC....--G---------G----------AC-------------------G-----------CC----A-----G-CCG--T--C---------A---GT--GG---------T-A--C--TC-G-A-----G-------------C----------- 5321
68_01B.CN.11.ANHUI_WH73      .GG---A----GC....--G---------G----------AC-------------------G-----------CT-----------G---G--T--C---------A---GC--GG---------C-A--C--TT-GCA-----G-----A----------------- 5312
70_BF1.BR.10.10BR_PE004      .-GA--------C....-----------A-TG---------A--------------------------------GG----G---------A---------C--------------A-------------------------C---------------G----------------------- 5434
71_BF1.BR.10.10BR_PE008      .-G-------A--....------------G-----------G--------CA---------------C------GG--------------CA--T------------------C--A--------------------AG---C--TC-G-------G----------------------- 5642
72_BF1.BR.10.10BR_MG002      .TC-----------....-----------------------A--------------------------------G----------AT-------T---C--------------A-----T----C---C---------------G----------------------- 5486
O.BE.87.ANT70                AGGA--AG--AG-....----------T-----------GAGG-GCCC--T---C-C--C--T----------CC-G-TCC------AT-----------C---GA-C----A-------T-T---------G---G---G-GT--G--A---------------- 6035
O.CM.98.98CMA104             AGGA--AG--AG-....----------T-----------GAGA-GCCC--T---C-T--C--T---------A-CC-G--CC------AT-----------C---GC-C----A-------T-T-----C-TG-G-G---G-GT--G--A---------------- 5476
O.CM.98.98CMABB141           AGGA--AG--AG-....----------T--------C---GAGA-GCCC--T---C-T--A--T----------CCCG-A-C------AT-----------C----A-C----A-------T-T-----------C--GC--G-GT--G--A---------------- 5475
O.CM.98.98CMABB212           AGGA--AG--AG-....----------T-----------GAGA-GCCT--T---C-T--C--T--G--CA-A--CC-----C------ATTC---------C--CGC-C----A--------TT-----------C--AG--G-GT--G--A------C--------- 5480
O.CM.98.98CMU5337            AGGA--AG--AG-....----------T-----------GAGA--CCC--T---C-T--C--T----------CCCG-ACC------A-----------C----T-C----A-------T-T-----------C--G---G-GT--G--A---------------- 5477
O.CM.99.99CMU4122            AGGA--AG--AG-....----------T-----------GAGA-GCCC--T---C-T--C--T--------A--CC-G--CC------AT--G--------C----A-C----A-------T-T-----------C--GC--G-GT--G--A---------------- 5475
O.FR.92.VAU                  AAGA--AG--AG-....---G-----T--CA--------GA-A-GCCC--T---C-T--C--T-----C--A--CC-G--CC------A-GC----------C---C-A-C----A-------T-TC----------C--AG--G-GT--G--G------C--------- 5556
O.GA.11.11Gab6352            AGGA--AG--AG-....----------T-----------GAGA-GCCC--T---C-T--C--T-----------CC----CC------AT-C----------C----A-C----A-------T-T-----------C--GC--G-GT--G--G---------------- 5201
O.SN.99.99SE_MP1299          AGGA--AG--AG-....----------T-----------GAGA-GCCC--T---C-T--C--T-----------CC-G--CC------AT--G--------C--GCA-C----AC------T-T---------TGC--GC--G-GT--G--A---------------- 6043
O.US.10.LTNP                 AGGA--AG--AG-....----------T-----------GAGG-GCCC--T---C-C--C--T-----------CC-G--CC------AT-C------C-C----A-C----A-------T-T-----------C--AG--G-GT--G--A---------------- 5956
N.CM.02.DJ00131              .-G-G-----.......--G------------------T-------------T----------C--A--C--G--A-----C-ATGGA-----------T-T--C----A---C--TATGTG---C----C--AG---G---------A--A------------ 5481
N.CM.02.SJGddd               .--A------.......--G------------------T-------------T----------C--A--C--G--A-----C-AT----------------GA-------A---C--TATGTG---C----C--AG---G----------CTA------------ 5465
N.CM.04.04CM_1015_04         .--A------.......---------------------T-------------T-------------A--C-----A-----C-AT--G---------------T--C----A-----TGTG-----C----C--AG---G-------------A------------ 5480
N.CM.06.U14296               .--A------.......--G------------------T-----------------------C-GA--C--G--A-----C-AT-------T--------GA-------A---C--TATGTA---C----C--AG---G-------------A------------ 5478
N.CM.06.U14842               .--A------.......--G------------------T---------------C-------C--A--C--G--A-----C-AT--G---------------GA---------WC--TAT------C----C--AG---G-------------A------------ 5482
N.FR.11.N1_FR_2011           .--A------.......--G------------T-----T-------------T--C-------C--A--------A-----C-AT-------------------GA-------A-----TGTGTG---C-----G-AG---G---------A------K-------- 5376
P.CM.06.U14788               .-GA--AG--AGG....---C-----C-----------GAT--GCCT--A---C-A--G-----G--------CTC--GCC-A---A---------C--C--GCC--C----A-------T-T-----------C--AG--G-GT--G--A---------------- 5455
P.FR.09.RBF168               .-GA--AG--AGG....---C-----C-A---------GAC--GCCT--A---C----G-----G--------CTC--GCC-A---A---------C--C--GCC--C----A-------T-T-----------C--AG--G-GT--G--G---------------- 6019
CPZ.CD.06.BF1167             ....G-ACC-AGG....--G-----------------AGAT--T----T--TTA--C---A-CTC-ACA--A-GC--GC-C--C-ATTC--------------GTA--C--T-A---C------C---C--T-CC-AG--G--TC-C---------A------------ 6083
CPZ.CM.05.SIVcpzMT145        .-----A---AGG....---G------A---A----G---GACT----A--A-----C---------G--CA-A----G----A-AA--C-AT--G----T---C------------T-A-----TAT-C----C--T-C--AG---G----------------------- 5520
CPZ.GA.88.GAB1               .--A--A---AG-....TCC------T---A--------GAC--G-----------A--C-----C---------A-G---A-T----AT-----------C--GCT----T-A---C--TAT-TA---C----C--A-----G-----A----------------- 6042
CPZ.TZ.06.SIVcpzTAN13        ..................C---A-T--TA--------CAGG---CC--A----T-----------GC-GCA---G-G--AC---------A-----T--C----------------C-------------CGC----C--T-CG-A------A-----A----------------- 5614
CPZ.US.85.US_Marilyn         T-CTCAAG-GAG-.....--G------T--------------ACA-----------------G--------C---------G--------C-AT---------C---------------------C--TGTG--------T-CG-AG--G-G-----A--T------------ 6050
GOR.CM.07.SIVgor2139_287     .AGA--AG--AG-....----------C---AC-------GAG--GCCA--A---C----A-----G--C-----CCC---CC-A--C-A--CA-----------------C--T-A---C---T-A---C---T-T-----------C--AG--G-GT--G--A---------------- 5521
GOR.CM.07.SIVgorCP2135con    .AGA--AG--AG-....----------C---AC-------GAG--GCCA--A---C----G-----G--C-----CCC---CC-A--C-A--CA-----------------C--T-A---C---T-A---C---T-T-----------C--AG--G-GT--G--A---------------- 5520
```

```
                                                          Tat Rev exon 1 end_intron start                                                          Vpu start (ACG in HXB2)
                                                                                                                                                    Vpu transmembrane domain start
B.FR.83.HXB2                GAAGCGGAGACAG...CGACGAAGAGCTCATCAGAACAGTCAGACTCATCAAGCTTCTCTATCAAAGCAGTAAGT.....................................AGTACA...TGTAACGCAACCTATACCA...............ATAGTAGCAATAG 6089
Tat exon 1                  __K__R__R__Q__.__R__R__R__A__H__Q__N__S__Q__T__H__Q__A__S__L__S__K__Q
Vpu                                                                                                                                                  _T__Q__P__I__P__.__.__.__.__.__I__V__A__I__
Rev exon 1                  R__S__G__D__S__.__D__E__E__L__I__R__T__V__R__L__I__K__L__L__Y__Q__S__
A1.AU.03.PS1044_Day0        -------------...------G--C---C---AGG--A-A---A-T-----AM-C--G--C-------------.....................ATAAGTA........TTT----T-...-----T-ACT---T-GGA-...............--CTGGT------ 5306
A1.CH.03.HIV_CH_BID_V3538   --------AG--C...------G--A------CA-G----A--GA-------AA-C-------------------.....................ATTTAG.......TAATT----T-...-----T-TC----T--GA-...............--TTAT------- 5307
A1.CY.08.CY236              -------------...---------C---C-----G----G--GA-------AA-C--A--C--G----------.....................ATTAGTA........GTT----T-...-----T-ACT---T--GAG...............--CAGT------- 5309
A1.ES.05.X1608_8            -------------...------G--A---C-----G----A--GA-------AA-C--A--C-------------.....................ATTAGTA........GTT--C-T-...-----T-TCT---T-GGA-...............--CTCT------- 5546
A1.KE.11.DEMA11KE001        -------------...-----C--A---C---C-G----A--GA-------AA-C--G--C-------------.....................AGTTAG.......TAATT----T-...-----T-A-TT--T-GA-T...CCTTGGAAT...--CTGT------- 5450
A1.RU.11.11RU6950           -----------C...A---CG--A---C---C-G----A--GA-------AA-C--A----------------.....................ACTA..............-A--A-...-----T-AC----T--GA-...............--CTAT------- 5616
A1.RW.11.DEMA111RW002       --------AG-G-...------G--A---C-----G--A-A--GA-------AA-C--G--C-------------.....................AGTAGTT........ATT----T-...-----T---G-T-T-GGT-..............G-CTGG------A 5449
A1.SN.01.DDI579             --------A----...------G--A---C-----G--A-A--GA-------AA-C--G--C-------------.....................AGTAGTA........ATT----T-...-----T-ACTT--T-GGA-...............--CTGG------A 5299
A1.UG.11.DEMA110UG001       -------------...-----T---A--TC---C-G----A--GA--------A-CT-G--C-------------.....................A.ATAG.................T-T-----T-C--GA-...............--CTGG------- 5439
A1.ZA.04.04ZASK162B1        -----------GA...--T---G--T---C--CA-G----A--GA-------AA-C--A--CA------------.....................ATTA..............----G-...-----T-AC----T-GGA-...............--CTGT------- 5533
A2.CM.01.01CM_1445MV        -----------CC...---------C--TC---A-G----A--GAC------AA-C--G-----G----------.....................AGTAATA........AGT-A--T-...-----T--TG---T--G--...............---T-GT------ 5297
A2.CY.94.94CY017_41         ------AC---CC...-------AGC--TC--CA-G--A-A--GAC------AA-C--A--C-------------.....................AGTAGTA........ATT-A--T-...-----T-TT----T--GT-...............---T-G------- 5455
B.BR.10.10BR_MG029          -------------...---------A---C---AGG----A-------------AC-----C-------------.......................................--C--G...-----T----A--T--AT-................---C-T------ 5550
B.CA.07.502_1191_03         ----A--------...---------T---C---A-G----A---T-G-------GC-------------------.......................................----T-...-----T-------T--G-C................---T--------- 5518
B.CH.08.M2_0803101_NFLG8    -------------...A--------T------CAG----------------------------------------.......................................---G--...-----T-AC----T--TAT................---TGGT------ 5545
B.CN.10.DEMB10CN002         -------------...---------C---C---AG-----A-----------AACC-----C---------G---.......................................-------.-----T-TTTT--T---A-................------------- 5459
B.ES.10.DEMB10ES002         -------------...------------C---AG------------------T----------------------........................AGA.........---G--...-----T-AC-A--T--GTC................---TGTT------ 5450
B.FR.11.DEMB11FR001         -------------...---A-----A---C---AG------C-----------G---------------------.......................................------...-----T---T--T--A--................---T--------- 5334
B.GB.05.MM45d213_GN1        -------------...---------C--TC-G-AGG-----------------A---------------------.......................................-A-----...-----T-----CT--ATT..............G--A--------- 5489
B.HK.06.HK003               -------------...-------------C--CAG-A-------A-------T------C---------------.......................................-A-----...-----T----T-CC--GA-................---T-----G--- 5458
B.HT.05.05HT_129389         ----A--------...---------R--TC---AR-------------------AG-------------G---.......................................----Y-...---R-T----T-YT--YAW..............R--T--------- 5295
B.JP.12.DEMB12JP001         ------------A...---------C---C---AGG-----------------T-C-----C-------------.......................................-------...---G-T-------T---A-................---T--------- 5478
B.KR.07.HP_18_07JHS10_3909  ------------A...---------------C----G---A-A---A-------T---------------------.......................................-A-----...---G-T----T-CT--A-................---T-TA------ 5752
B.PE.07.502_0525_wg5        -------------...---------------C-A--G-----------------G---C--A-------------.......................................---G--...-----T----T-CT--GT-................---T--------- 5565
B.RU.11.11RU21n             -------------...---------T---C---A----A------G--------------C--------------.......................................-------...-----T-----CT--GT-................----C-A------ 5659
B.TH.08.MERLBDTRC10         -------------...------G--C---C---AG----------------------------------------.......................................----A-...---G-T----G--T--TAC................--TT--------- 5321
B.US.11.ES38                -------------...---------------C---AGG-----C--A-A-----T-C-------G----------.......................................----T-...-----T--C-T-CT--TAC................------T------ 6023
C.AR.01.ARG4006             -------------...--------C----C--CT-G-G--G--GA-------AACCT-A-----G----------.....................ATT...........CCAT--C-G-...----TG..............TTAGAMWTAGATTAT-G-T---G----- 5280
C.BR.07.DEMC07BR003         -------------...--------C----C--CA-G----G-AGA--------A-CT-A---------------.....................ATT...........CAAT---G-...-----T--T-A-CT-----GACAGAGTAGACTAT-G-T---G----- 5442
C.BW.00.00BW5031_1          ------------A...--------C----C-TCA-G----G---A-------AA-C--G----------------.....................A...........ATAAGT---G-...-----T-TTTG--T--TTT...GAAGTAGATTAT-G-T--A------ 5463
C.CN.10.YNFL19              -------------...--------C----C---A-G----G--GA-------AA-C--A----------------.....................ATC..........................-----T-GT-GAC...........TTAGATTAT-A---A--G--- 5444
C.CY.09.CY260               -------------...-----C--C----C--GA------G-AGA--------A-CT-A----------------.....................A..........CTAAAT---G-...-----T-ACTGG-T--AT-GAAAAAGTAGATTAT-A-A----G--- 5333
C.ES.08.X2363_2             ------------A...--------C----C-TCT-G----A---GC-------A-CT-A--C-------------..............................................T-...-----T-TTGGACT---T-GCAAAAGTAGATTAT-G-A----G--- 5531
C.ET.02.02ET_288            -------------...-----C--C----C--CA-G----G---GC-------A-C--A----------------.....................ATC..........AAAT-A--G-...-----T-TTTAG-T-TA--GAAAGAGTAGATGCT-G-----G----- 5312
C.IN.09.T125_2139           -------------...--------C----C---A-G----G-AGA-------AA-C--A----------------.....................ATC..........................-----T-GT-GAG...........TTAAACTAT-AGT-----G--- 6064
C.KE.00.KER2010             ----T-------A...-------C-------C--C--GT---G--GA--------A-CT-A----------------.............................................T-...-----T-TT-GA-T---T-GCAAGAGTAGATTAT-G-T---G-G--- 5281
C.MW.09.703010256_CH256.w96 -------------...--------C----C--CA-G----A--GA-------AA-C-----C-------------.............................................T-...-----T-TT-GACT--GT-GCAAGAGTAGATTAT-G-A---G-G--- 6080
C.TZ.08.707010457_CH457.w8  ----------A--...--------C----C--CA-G----G--GA-------AA-C--A----------------.....................A.CCACATAATAACACAT---T-...-----T-TTGGA-T---T-GCAAGAGTAGATTAT-G-A---G-G--- 6094
C.YE.02.02YE511             ------------A...-----C--C----C--CAG-----G-AGA-------AA-CT-G----------------.................................ACAAGT---G-...-----T-TT-GA-T---T-GCAAAAGTAGATTAT-G-A---G----- 5284
C.ZA.08.705010534_CH534.w12 ------------A...--------CA---C-TCA-G----A--GA-------AA-C--------------------.....................A..........CCCAAT---G-...-----T-TT-AAC...........TTAGATTAT-G-----GG---- 6088
C.ZA.10.DEMC10ZA001         -------------...--------T----C--CA-G----G--GA-------AA-C--A----------------.....................ACCAA..........AGT----T-...-----T-TT-AG-T---T-GAAAAAGTAGATTAT-G-A-----G--- 5451
C.ZM.11.DEMC11ZM006         -------------...--------C----C--CAGG----A--GA-------AA-C--A----------------.....................A..........CTTAAT---G-...-----T-AT-AG-T--AT-GAAAGAGTAGATTAT-G-T---G-G--- 5459
D.CM.10.DEMD10CM009         ----A-AG-----...-------A-C---C-A-AG---A-ACA----------T-C--G--C--G-T--------.......................................-------...-----T-------T--GTG.................---A-TT------ 5463
D.CY.06.CY163               ----A-AG-----...---------C---C---AGG--A-A------------G-C--G--C--G----------.......................................-------...-----T---T--T--GAG.................---T--------- 5296
D.KE.11.DEMD11KE003         -------------...---------------C---AGG-G-----GT--------A-C--A--C-------------.......................................-------...-----T---T--T---AG.................----C-------- 5438
D.KR.04.04KBH8              -------------...---------C---C---AGG-GA----G---------T-C-------------------.......................................-AC-T-...---G-T-------G--GTG.................A--------- 6044
D.SN.90.SE365               -------------...---------A---C---AGG--A--------------T---G--C----A--------.......................................-A--T-...-----T---GT--T--GA-.................---T--------- 6105
D.TZ.01.A280                -------------...---------A---C----GG-GA----G----------A-C--A--C-------------.......................................-------...-----T-AC----T--GAG.................---AC-------- 5289
D.UG.10.DEMD10UG004         -------------...------C--C---C---A-G-G-----G----------A-C--A--C-------------.......................................-A--T-...-----T-------T--GTG.................---T--------- 5426
D.UG.11.DEMD11UG003         -------------...------G--A---C---AGG-G-----G---------T-C--G--C--------G---.......................................-A--T-...-----T-A-C-T-T--GTG.................A---G------ 5440
D.YE.02.02YE516             ----A-AG-----...---------T---C---A-G--A--------------T-C--G--C----A--------.......................................----T-...-----T----A-CT--GAG.................---T--T------ 5299
D.ZA.90.R1                  -------------...-------A-C---C---C-G-GA----GA--------T-C--A--C--G----------.......................................---T--...-----T-------T--GTG.................---T--------- 5437
F1.AO.06.AO_06_ANG32        ------------A...---------A---C-----G----A---A-------AA-C--G--C--G----------.....................AGTTGT.......TAAGA-A--T-...-----T-TCTGA-T-GTT-..............GC-A--------- 5310
F1.AR.02.ARE933             -------------...-----AC--A---C---CGG--C-----TA-------A-CT-G--C-------------.....................ATTGTTA........AGT----T-...-----T-TC-TACT-GTT-..............GC-A---GC---A 5392
F1.BR.10.10BR_RJ015         -------------...-----AC--CA--G------T----ACTA-------A-C--G--C-------------.....................ATTGTTG........AAT----T-...-----T-AC-GA-T-G-T-..............G--A--AGC---- 5504
F1.CY.08.CY222              -------------...-----AC--A---C-----G--------AC-------A-C--G----------------.....................ATTGTTT..........................-----T-AC-AA-T-GTT-..............GC-A----G---- 5288
F1.ES.02.ES_X845_4          -------------...---------C---C-----G--A-A---AC------AA-C--G--C--G----------.....................ATTGGTT........AAA----T-...-----T-TCTAA-T-GTT-..............GC-A--T------ 5550
F1.RO.96.BCI_R07            -------------...-----AC--A---C---A-G-----CA-TA------AA-C--G--C-------------.....................ATTGTTA........AGT----T-...-----T-TC--A-G--TT-..............GC-A----G---- 6122
F1.RU.08.D88_845            -------------...------CG--A--YC---A-G------A-TA------AA-C--G--C--G----------.....................ATTGTTA........AAT--C-T-...-----T-AC-GAA---AT-.................---A---G----- 5557
F2.CM.02.02CM_0016BBY       -------------...---------C---TC---C-G----A--GAC-------A-C--G----------------..............................................T-...-----T-TC-TA-T-GAT-.................---T---T----- 5281
F2.CM.10.DEMF210CM001       -------------...---------A---C---A-G----G---TA------AA-C--G----------------..............................................G-...-----T--C-TA-T-GAT-.................A--CTG---- 5352
F2.CM.10.DEMF210CM007       -------------...---------T---C---A-G----A--GTA-------A-C--G--C--------G---.........................................A-T--...---G-T-TC--T-TCTAT-.................------T----- 5437
G.BE.96.DRCBL               --------AG--C...------G--A---C---C-G----A--GA-------A--C--G--C-------------.....................AGAAATA........AAT----G-...-----T------AC--GA-.................---TCT------- 6055
G.CM.10.DEMG10CM008         --------AG--C...------G-CT--TC---CGG----A--GA-------AA-C--G--C--G----------.....................AATAATA........ATT----TT...---G-T---GG-AT--GTC.................----CT------- 5479
G.CN.08.GX_2084_08          --------AG--C...A-----GC-A---C-TCA-G--A-A--GA-------AA-C--G--C-------------.....................................ACT-A----...---G-T---G--AT--GA-.................----CT------- 5334
G.CU.99.Cu74                --------AG--C...------G--A---C----GG----A--GA-------AA-C--G--C-------------.....................ACTGACA........ATT---CT-...-----T-A-GT-AT--GA-.................---TCT------- 5702
G.ES.09.X2634_2             --------A---C...---------A---C------T-A-ACAGAC------AA-C--G--C--G----------.....................AGTAATA........ATT----T-...-----T------AT--GA-.................----CT--G---- 5581
G.GH.03.03GH175G            --------AG--A...------G--A---C----GG----A--GA-------AA-C--G--C-------------.....................AGTAATA........ATT-T--T-...-----T----T-CT--GA-.................---TCT------- 6138
G.KE.09.DEMG09KE001         --------AG--C...-----G------TC--CAGG----A--GA-------AA-C--G--C--G----------.....................ATTAGTAATTA.........---TT-...G----T------AT--GA-.................----CTT------ 5474
G.NG.09.09NG_SC62           --------AC--C...A-------A---C-TCAGG------A-A--------AA-C--G--C-------------.....................AGCTATAATTAGTAT........................-----T----T-AT--GA-.................----C-------- 5289
G.ZA.01.TV546               --------AG---...------G--A---C----G-----A--GA------RAA-C--G--C-------------.....................AGTAATA........ATT----T-...-----T---G--AT--GA-.................---TCT------- 5300
H.BE.93.VI991               ------------A...------G-GA---C-A---G-TTG---GA------GA--CT-A--C-------------.....................AACATAA...........................-T-A-TATAT--GG-.................A---G----- 5471
H.CF.90.056                 ---------C---...-----AC---A---C-GCA-GTTTG--AGA-------AA----A----------------.....................ATTATC..................ATA--T-T-TATAT--GG-.................T---G----- 5432
H.GB.00.00GBAC4001          -------------...------G-GA---C-GCAGGTTTG--AGA-A------A-C--A--C-------------.....................ACCATA..........................--T-T-CATAT--GG-.................A---G----- 5600
J.CM.04.04CMU11421          -------------...------GC-----C--CTGG----A-R-R--------A-CY-G--Y-GG----------...........................AAATCAATT---CTC...-----T-A-CT--T---A-.................----C-T------ 5603
J.SE.93.SE9280_7887         -------------...------TC-----C--CTGG----A------------A-CT-A--C-------------...........................AAATCAAGT--C-T-...-----T-AT----T-G-A-.................----C-------- 5411
J.SE.94.SE9173_7022         ------------A...------TC-----C--CTGG----A---A--------A-CT-A--C--G----------...........................AACTTAATT--C-T-...-----T-GT-T--T---A-.................------------- 5412
K.CD.97.97ZR_EQTB11         ----A--------...---AC--C-A---C-T-CGC----A---A----A---A-C--A--C-------------.....................AGTACCA........AGT-T----...-----T-GT----T--A--.................-----G----A 5287
K.CM.96.96CM_MP535          ----A------CC...-------C-A---C-T-C------G---AC-------A-C-----CG--------G---.....................AGTGCTA........AGT-------...-----T-GTGT-CT--G--.................A-TT------ 5291
U.CA.01.TV749               ------------A...------GCG----C--TC-G----A--GA-------AA-C--G--C-------------.....................ACATTTGTAGAGTAT..........................-----T-T-CA--T---AG.................------------- 6114
U.CA.99.TV721               -----------C-...------GC-----C--TT-G----A--GAC------AA-C--G--C-------------.....................AAATTTGTATAGTAT..........................-----T----A--T---AG.................------T------ 6113
U.CD.83.83CD003_Z3          -------G--A--...---------A---C-----G-G---C--A-------AA-AT-G----------------.....................ACTACC..........................--T-TCCAT-T-G-A-.................------------- 5580
U.CD.90.90CD121E12          -------G--A--...---------A---C-----G-GA--C-GA-------AA-AT-G--C-------------.....................ATTACC..........................--T-TGGAT-T-G---.................----G----- 5281
U.CY.05.CY090               -------------...-----C------TC-----GT-A-----A-------AA-C--A--C-------------.....................ATTATTAAGTAGTGT..........................-----T---GA--C-GTT-.................---C-T-----A 5296
U.CY.08.CY223               -------------...------GC----TC--GC-G----A---A--------A-CA-A--C--G----------.....................A.TAAG.....TAA...TAGTTC...---T-T-A-CA--CC-A--.................---T--AGC---- 5313
U.ES.10.DEURF10DZ001        -----------GC...-----C-G--C---C----GG-C-----GA--------A---G--C-------------.....................AGCAAC..............----T-...-----T-------T--GA-.................----CT------- 5370
U.GR.99.99GR303             -------------...------GC-----C---CG---A-A---A--------A-CT-A--C-------------.....................A.TAAT......TTAAGT-AGGTC...-----T----T--T--GA-.................----C-AGC---- 5370
U.NL.95.U_NL_95_H10986_D1   ----A-----CA...A-----C-A---C-T-CG-----A---A--------A-C--A--C--G----------.....................ATTGTTA........AAT----T-...-----T-A-TAT-C--GT-.................-----G----- 5980
01_AE.AF.07.569M            --------AG--C...------G--A---C-----G----A--GA-------AA-C--A--C-------------.....................AAAA..............-A--T-...-----T-AC----T-GGA-...............--TAGT------- 5295
01_AE.CF.90.90CF11697       --------A---C...------G--C---C--CAGG----A--GA-------AA-C--A--C-------------.....................ATTA..............----T-...-----T-TC-G--T-G-A-...............--CTAG------- 6041
01_AE.CN.10.YNFL03          --------AG--C...------G--A---C-----G----A--GA-------AA-C--A--C-------------.....................ACTA..............----T-...-----T-AC----T-GGA-...............--TAGT------- 5470
01_AE.HK.04.HK001           --------AG--C...------G--A---C--G--G----G--GGC------AA-C--G--C-------------ATTTGTAAA..............ATA..............----A-...-----T-AC----C-T-A-...............--TAGT------- 5462
01_AE.IR.10.10IR.THR48F     --------AG--C...------C--A---C-----G--A-A---A--------A-C--A--C--G----------.....................AATA..............-A--T-...-----T-AC----T-GGA-...............--TAGT------- 5282
01_AE.JP.x.JRC77AE          --------AG--C...------G--A---C-----G-C--A--GA-------AA-C--G--C--G----------.....................AATA..............-A--T-...-----T-AC----T-GGA-...............--TAGT------- 6089
01_AE.TH.09.AA090a_WG11     --------A---C...------G--A---------G--A-A--GA-------AA-C--A--C-------------.......................AAA..............----T-...-----T-AC----T-GGA-...............--TAGT------- 5514
01_AE.TH.90.CM240           --------AG--C...------G--A---C-----G----A--GA-------AA-C--A--C-------------.....................AATA..............----T-...-----T-AC----T-GGA-...............--TAGT------- 5660
01_AE.US.05.306163_FL       --------AG--C...------G--A---C-----G----A--GA-------AA-C--G--C-------------.....................AAAA..............--C-T-...-----T-TC----T-GGA-...............--TAGT------- 5283
01_AE.VN.98.98VNND15        --------AG--C...------G--A---C-----G----A--GA------GAA-C--A--CG-G----------.....................AATA..............----T-...-----T-AC----T-GGA-...............--TTGT------- 5347
```

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

Vpu transmembrane domain end

```
B.FR.83.HXB2                TAGCATTAGTAGTAGCAATAATAATAGCAATAGTTGTGTGGTCCATA...........................GTAATCATAGAATATAGG.....................AAAATATTAAGACAAAGAAAAATAGACAGGTTAATTGATAGACTAATAGAAAGAGCA 6211
Vpu                         V__A__L__V__V__A__I__I__I__A__I__V__V__W__S__I__.__.__.__.__.__.__.__.__.__V__I__I__E__Y__R__.__.__.__.__.__.__.__K__I__L__R__Q__R__K__I__D__R__L__I__D__R__L__I__E__R__A_
A1.AU.03.PS1044_Day0        ---G-C-G-----T--GC----C------------------A-T---...........................---GGT----------A-.....................---T-GC---AG--------------------------G---A---G---------- 5428
A1.CH.03.HIV_CH_BID_V3538   C--GGC--A-------GC----C-----------A------A-T---...........................---GGA----------A-.....................-G-T-GC---AG-------G------------------G---A---G---------- 5429
A1.CY.08.CY236              ---G-C-G-----G-T-TC---C------------------A-T---...........................---GGT------AT--A-.....................---G-GC----G------------------------A-G---A---G---------- 5431
A1.ES.05.X1608_8            ---G-Y-G-----GA-KY----Y------------------A-TC--...........................---GCT-----------A.....................---G-GC---AG--------------A-----------G---A---G---------- 5668
A1.KE.11.DEMA11KE001        ---G-C-GA-------GC-G--CT-----------------A-T---...........................---GCT------------.....................----------A------G--------------------G---A---A--------A- 5572
A1.RU.11.11RU6950           C--------------TGT-T--C------------------A-T---...........................---GGT------------.....................-G-T-GC---A------G--------------------G---A---G---------- 5738
A1.RW.11.DEMA111RW002       ---G-C-GA--------C----TT-G------A--------A-T---...........................---GG-T---------A-.....................---T-G----A------G--------------GC----G---A---G---------- 5571
A1.SN.01.DDI579             ---G-C-GA-------GC----CT----------A------A-----...........................---GGG----------A-.....................---G-GC---AG----A--G------------------A---A---G---------- 5421
A1.UG.11.DEMA110UG001       C--G-C-T--------GC----CC--------A--------A-T---...........................---GGT----------A-.....................-G-T-GCA--------AG--------------------A--GA---G---------- 5561
A1.ZA.04.04ZASK162B1        C--G-C-G--------GC----C-----------A------A-T---...........................---GGT------------.....................---T-G-----G--------------T---------A-G---A---C---------- 5655
A2.CM.01.01CM_1445MV        ---G-C-GA--------T-C--CT-------------A---A-T---...........................---T-----C------AA.....................G--C-TAAG-AG-----G---------T-------CA-A---A---GT--------- 5419
A2.CY.94.94CY017_41         ---G-C-GA-------TT----CT-------------A---A-T---...........................---T------------A-.....................-----TAAG-AG-----G---------T-------CA-A---A---GT--G------ 5577

B.BR.10.10BR_MG029          ---------------T-GC--------------C---------TG--...........................---T-A-----------A.....................------C----------G-----------A------------A---G---------- 5672
B.CA.07.502_1191_03         C----------------GC----C-----------------A-----...........................---GG-------------.....................------------------------A-T--A---C--------A--GC---------- 5640
B.CH.08.M2_0803101_NFLG8    -----------------GC----------------------A-----...........................-C-C--------------.....................-----------------G------------A--C--------A--GG---------- 5667
B.CN.10.DEMB10CN002         -----------------GC----C-----------------A-T---...........................--GTAT---C--------.....................----------A---------G-----T---------------A-----------A-- 5581
B.ES.10.DEMB10ES002         C--------------T-GC-------------A--------A-T---...........................--GT-A------------.....................-------------------G----------A-----------A---G---------- 5572
B.FR.11.DEMB11FR001         ----------------------TT-----------------A-T---...........................---T--------------.....................--------------------------T--T-----------GA---G---------- 5456
B.GB.05.MM45d213_GN1        ---------------T-GG------------------------T---...........................---C--------------.....................--------------------------T--A------------A---G---------- 5611
B.HK.06.HK003               -----------------GC----T--------A--------------...........................---T--------------.....................-G----AG------G-----------T---------------A---G---------- 5580
B.HT.05.05HT_129389         -----------------GY----M-R------A--------W-----...........................---T--------------.....................-R----AAR-------A--G------T--A------------A---G---Y------ 5417
B.JP.12.DEMB12JP001         -------------G---GC----C-----------------A-T---...........................---GG-------------.....................----------A------G------------------A-----A---GT--------- 5600
B.KR.07.HP_18_07JHS10_3909  -----------------GC------------------------T---...........................---T------------A-.....................----------A-------------------------------A---G---------- 5874
B.PE.07.502_0525_wg5        -----------------GC----T-----------------A-----...........................---T--T-----------.....................--------------------G-----T---------------A--GC---------- 5687
B.RU.11.11RU21n             -------T-------T-GGG------------------------C--...........................------G-----------.....................-------------------------------------C----A-------------- 5781
B.TH.08.MERLBDTRC10         ----T------A-----GC------------------A---A-----...........................-----T------------.....................--------------G-A---------T--A------------A---G---------- 5443
B.US.11.ES38                -----------------GC----C--------A----------T---...........................---T--------------.....................------C---A----CA------------A------------A---G---------- 6145

C.AR.01.ARG4006             G------GA--------C----C---------A--------A-----...........................-CTTAT------------.....................---T-GG---------A--G-------T----------A---A-T-GG--------- 5402
C.BR.07.DEMC07BR003         G------GA--------C----C---------A--------A-T---...........................-CTTAT------------.....................---T-GG------------G-------C----------A---A-T-GG--------- 5564
C.BW.00.00BW5031_1          G---C--TA--------T--T-CT----G------------A-----...........................-C-TATT-----------.....................---T-GG------------G-------CA----G--A-A---A-T-GG--------- 5585
C.CN.10.YNFL19              G------TA--------C----CT--------A--------A-----...........................-C-TAT------------.....................---T-GAC--A----------------T-------CA-A---A-TGGG--------- 5566
C.CY.09.CY260               -------TA--A-----T--G-C------------------ATA---...........................---TATT----G------.....................---T-GG-------------------------------A---A-T-G------G--- 5455
C.ES.08.X2363_2             -------CC--A-----T----CT-----------------AT----...........................-C-TAT-----------A.....................---T-GG---A-----C--G------TCA-------A-A---A-T-GG--------- 5653
C.ET.02.02ET_288            G-------A--------C-G--CC--------A--------A-----...........................-C-TAT-----------A.....................---T-G---------C-G--------Y-----------A---A-T-GG--------- 5434
C.IN.09.T125_2139           G---T--GA--------C----CC-----------------AT----...........................---TAT------------.....................---T-GG---A----------------T-----G--A-A---A---GG--------- 6186
C.KE.00.KER2010             G-------A--------C----C---------A--------A-T---...........................---TAT------------.....................---T-G-----------G---------T----------A---A-T-GG--------- 5403
C.MW.09.703010256_CH256.w96 G-------A--------C----C------------------AT-C--...........................-C-TAT----------AA.....................-GGTGG----A-----AG---------T--------A-A---A-T-GG--------- 6202
C.TZ.08.707010457_CH457.w8  C---C--GT--A-----C----C------------------A-----...........................---TAT----------A-.....................---T-G---C-G----A--GG------T--------A-A---A-T-GG--------- 6216
C.YE.02.02YE511             C------T---------C----C---------A--------A-T---...........................-C-TAT------------.....................---T-GG---T----------------T-----G--A-A---A-T-GG--------- 5406
C.ZA.08.705010534_CH534.w12 G------GA--------C---AT-----------A------AT----...........................---TAT------------.....................---C-GG------------G-------T-----G--A-A---A-T-GG--------- 6210
C.ZA.10.DEMC10ZA001         C------TAC-A-----T----CC-----------------A-----...........................-C-TAT-----G------.....................--GG-GG---------C-------------------A-A--GA-T-G------G--- 5573
C.ZM.11.DEMC11ZM006         G------GA--A-----T-G--T------------------A-----...........................-C-TATC-----------.....................---T-GC---------AG---------T----G-----A---G-T-GG--------- 5581

D.CM.10.DEMD10CM009         --------------A-GC-----------------------A-----...........................---TAT------------.....................------CAG-------AG---------T----------C---A---G--------A- 5585
D.CY.06.CY163               C----------------C-----C-----------------A-----...........................---TAT------------.....................------AA---G-----G---------T--C-------C---A--GG---------- 5418
D.KE.11.DEMD11KE003         ----C-----------GC-----C----------G------A-T---...........................---T-------G-G---A.....................-G-C--AG---------.---------T--------------A---G---------- 5559
D.KR.04.04KBH8              K-------------A-GC--------------A--------A-----...........................--TT-------G-GCT-T.....................-G-T--AG--AG--------G---------------A-----A---G---------- 6166
D.SN.90.SE365               C---A-T---------GC-----C----C------------A-----...........................---T--------GT----.....................--G---AG---G-----G---------T----------C---A---A---------- 6227
D.TZ.01.A280                C----------A----TTC----C----------G------A-T---...........................---TA------G------.....................-G----AG--AG----AG---------T--------------A---G---------- 5411
D.UG.10.DEMD10UG004         ----------------GC--T--T-----------------A-T---...........................---T-------G-G---A.....................---T--AG---G----A--G-------T--------------A---G---------- 5548
D.UG.11.DEMD11UG003         -----------------------T-----------------G-T---...........................---T-T-----G-GG--A.....................-G-T--AA---G--G-------------A-A-----A-----A--GC---------- 5562
D.YE.02.02YE516             -----------A----GGC-------------A--------A-----...........................---TA------------A.....................------AGG-G----------------CA-C-----------A---G---------- 5421
D.ZA.90.R1                  -----------------T-----T-----------------A-T---...........................---T--------------.....................-G----AG--AG---------------TAT------------A---G---------- 5559

F1.AO.06.AO_06_ANG32        -----------------C-----------------------A-T---...........................-C-TATC----------A.....................---G-GG---AG----A--G----A-T------TA---A---A---G---------- 5432
F1.AR.02.ARE933             C-------A--------C-----------------------A-T---...........................-C-TAT----------AA.....................---C-GG----G-----G-G----A-T------TA-A-A---A---G---------- 5514
F1.BR.10.10BR_RJ015         C---T---A--------C---------T-------------A-T---...........................---TATC-----------.....................---C-GG----G------------A-T------TA-C-A-------G---------- 5626
F1.CY.08.CY222              C-------A--T-----C-----G-----------------A-T---...........................-C-TAT----------A-.....................---G-GA----G------------A-T------TA-A-A-A-A---G---------- 5410
F1.ES.02.ES_X845_4          -----------------C---------T----A--------A-T---...........................---T-TT-----------.....................--GT-G-----G------------A-T------TA---A-------G---------- 5672
F1.RO.96.BCI_R07            --------A--------C--------------A--------A-T---...........................-C-TAT------------.....................---T-GG----G------------A-T-----GTA---A---A---G---------- 6244
F1.RU.08.D88_845            -------GA-------GC-----G-----------------A-----...........................---T-T------------.....................------------------------A-T------TA---A---A---G---------- 5679
F2.CM.02.02CM_0016BBY       -------TA-------GC-----GC-------A--------A-T---...........................---TAT----------A-.....................---CAG-----G----A--G----A-T------TA---A---A---G---------- 5403
F2.CM.10.DEMF210CM001       -T-----TA---C--TGC--------------A--------A-T---...........................---TAT------------.....................----------AG------------A-T-A----TA-A-A---A---G---------- 5474
F2.CM.10.DEMF210CM007       C------TA------T-C--------------A--------A-T---...........................--G--T----------A-.....................---C-G-----G----A--G----A-T-A----TA---A---A---GG--G------ 5559

G.BE.96.DRCBL               ---GC---A--------TCC---GC---C------------A-T---...........................---T-T-----------A.....................------AG--A---G-A--G------A-A----C-------GA---G---------- 6177
G.CM.10.DEMG10CM008         ---G----A--------T-TT--GC---C---AC-------A-T---...........................---T-T---C--------.....................G-----AG--A---G-AC--------T-A----C--------A---G-------AT- 5601
G.CN.08.GX_2084_08          ---G-C--A--------T-T---GC---C------------A-A---...........................---TAT-----------A.....................------AG------G-A--G------T-A-A--C--------A---GG--------- 5456
G.CU.99.Cu74                ---G----A--A-----GGC---GC---C------------A-----...........................--TT-----C-------A.....................G-----AG--A---G-A---------T---A--C----A---A---G---------- 5824
G.ES.09.X2634_2             ---G-C--A--------GGCT--GG---C---A--------A-----...........................--TT-T---CT-----A-.....................G-----AG--A---GCAG-G----C-A---A--C--------A-GG----------- 5703
G.GH.03.03GH175G            ---GGC--A---------GC---GC---C---A--------AG----...........................---T-T------------.....................-G----AG--A---G-A---G-----A-AA------------A---G---------- 6260
G.KE.09.DEMG09KE001         ---G-C--A-------TT-CT--GC---C-C-A--------A-----...........................---T-T---C--------.....................G-----AG--A---G-----------T------TA-------A---G---------- 5596
G.NG.09.09NG_SC62           ---G-C--A--------T-C---GC---C---A--------A-A---...........................---T-T------------.....................------AG--A---G-A--G------T--T---C--A-----A---G---------- 5411
G.ZA.01.TV546               ---G-C--A-------MT-C---GC---Y------------A-T---...........................---T-AT----------A.....................------AG--A-G-G-----------A-AA---C-------GA---G---------- 5422

H.BE.93.VI991               G----C-G--------TT-T--C-----C------------A-----...........................-C-TAT------------.....................---C-G.---AG--G--G--------------------A---A---G---------- 5592
H.CF.90.056                 G--GC---------A--T-T--C-----CG--A--------A-----...........................---TAT----------AA.....................---T-GG----G----AG------------C-------A---A--GG---------- 5554
H.GB.00.00GBAC4001          G--GC---A--------GGT--CT--TGT---A-A------A-T---...........................---TATT-----------.....................---C-GG--CAG----A--G-------------C----A---A---G---------- 5722

J.CM.04.04CMU11421          -----Y-----------T-TT-CC-T---G--R---C----RYTG--...........................-C-TAT-----------R.....................---C-GG----G-----G-G------T-A----R--------A---G---R----A- 5725
J.SE.93.SE9280_7887         -------TA----------TT-CC-T-------G-A-----A-T---...........................--GTAT----------A-.....................---C-G-----------G--------T-A-------------A---G---------- 5533
J.SE.94.SE9173_7022         --------A--------T-TT-CC-T------TG-------A-T---...........................--GTAT----------A-.....................---C-G-----------G--------T-A-------A-----A---G------G--- 5534

K.CD.97.97ZR_EQTB11         ------------C---GC-----C--------A--------A-----...........................-CCTATC----------A.....................---G-GG---------A--G----A--T-----T--------A---G---------- 5409
K.CM.96.96CM_MP535          -----------------C-----T--------A----A---A-----...........................---TAT------------.....................---T-GG---A-----A--G----A--T-------A------A---G---------- 5413

U.CA.01.TV749               --------A---------CTC-T-----C---TG---A---A-----...........................--CTATG----------A.....................--GC-G-----G----AG--------T---------------A---C---------- 6236
U.CA.99.TV721               --------A--------GCTC-T-----C--CTG---A---A-----...........................--CTATG----------A.....................-G-C-G-----G----AG--------T-A-------------A---G---------- 6235
U.CD.83.83CD003_Z3          -----A--A--------T----CT-----------------A-----...........................---TA-----------A-.....................-G-C-G--------------------TT--------------A--GG---------- 5702
U.CD.90.90CD121E12          -----A--A--------T--GGCT-----------------A-----...........................---TA-C---------A-.....................---C-G--------CA--G-------TT--------------A--GG---------- 5403
U.CY.05.CY090               -----C-----------TCTC------T-----CA------A-----...........................--GT-G----------A-.....................---C--AG------GA--G-----A-T------TA---G---A---G-------A-- 5418
U.CY.08.CY223               --------A--------T-C--T----T-G--A--------A-----...........................---T-T----------AA.....................--GC-GAG--AG--------------T-----G---C-A---A---G---------- 5435
U.ES.10.DEURF10DZ001        ---G-C-----------T-C---GC---C------------AG----...........................---TAT-----------A.....................------AGG-A---G-AG----------A----C--------A---G---------- 5491
U.GR.99.99GR303             ---G-C--A--------T-T---GC---C------------GA----...........................---C------T------A.....................---TA-GA--AC-------G------T----------T----A---G---------- 5492
U.NL.95.U_NL_95_H10986_D1   C---C--------G---T-T--CT---T-G--A--------A-----...........................---TATTGT-------A-.....................--GC-GAA---G--GA---G------T-A-A-----A-----A---G---------- 6102

01_AE.AF.07.569M            ---G---GA-------GC----CTC---------A------A-T---...........................---GCT------GC--A-.....................---T--C----G---------------------G--A-G---A---G---------- 5417
01_AE.CF.90.90CF11697       ---G-C-GA-------GC----CT----------A------A-T---...........................--GT-T----------A-.....................------C----G--------------------G-----A---A---G---------- 6163
01_AE.CN.10.YNFL03          ---G-C-GA-------GT----CT----------A------A-T---...........................--GGCT-------T----.....................-G----C----G------------------------A-G---A---G---------- 5592
01_AE.HK.04.HK001           ---G-C-GA-------GC----CT----------A------A-T---...........................---GGTT-----GT----.....................------C----G--------------T---------A-G-A-A---G--------A- 5584
01_AE.IR.10.10IR.THR48F     C--G-C-G--------GC----CT----------A------A-T---...........................---GCT------------.....................------C---AT-----G-----------A--G---A-G---A---GT--------- 5404
01_AE.JP.x.JRC77AE          C--G---GA-------GC----CT----------A------A-T---...........................---GCT----------A-.....................-----------G----AG---------------C-A------A---G---------- 6211
01_AE.TH.09.AA090a_WG11     ---G-C-GA-------GT----CT----------A------A-T---...........................---GGT------AT--A-.....................-G-T--C----G------------------------A-G---A---GT--------- 5636
01_AE.TH.90.CM240           ---G-C-GA-------GC----CT----------A-----AA-T---...........................---GCT------GT--A-.....................------C----G---------------------G--A-G---A---G---------- 5782
01_AE.US.05.306163_FL       ---G-C-GA--------GC---CTC---------A------A-T---...........................---GCT------GT--A-.....................--------T--G--------G------------G--A-G---A---G---------- 5405
01_AE.VN.98.98VNND15        ---G-C-GA-------GC----CT----------A------A-T---...........................---GGT-------T----.....................----------AG---------------------G--A-G---A---G---------- 5469
```

Vpu transmembrane domain end

```
B.FR.83.HXB2                 TAGCATTAGTAGTAGCAATAATAATAGCAATAGTTGTGTGGTCCATA.........................GTAATCATAGAATATAGG....................AAAATATTAAGACAAAGAAAAATAGACAGGTTAATTGATAGACTAATAGAAAGAGCA 6211
Vpu                          V__A__L__V__V__A__I__I__I__A__I__V__V__W__S__I__.__.__.__.__.__.__.__.__.__V__I__I__E__Y__R__.__.__.__.__.__.__.__K__I__L__R__Q__R__K__I__D__R__L__I__D__R__L__I__E__R__A_
02_AG.CM.10.DE00210CM013     ---G-C-----------T-C---GC---C-----------------A-----.........................---TA-------------.....................------AAG-A---G-AG---------------C--------A---G---------- 5587
02_AG.CY.09.CY256            ---G-C-----------T-C---GC---C-----A------A-----.........................---TAT-------------.....................------AG--A---G-AG---------------C----A---A---G---------- 5406
02_AG.ES.06.P1423            ---G-------------T-C---GC---C-----------------A-----.........................---TAT-------------.....................------AG--A---G-AG----------A----C--------A---GG--------- 5654
02_AG.GW.05.CC_0048          ---G-C-----------T-T---GC---C---------A---A-----.........................---TAT-------------.....................------AGG-A-G-G---------AG--AC---C----C---A---G---------- 5573
02_AG.KR.12.12MHR9           ---G-C--A--------T-C---GC---C-----------------A-----.........................--GT-T-------------.....................------AGG-A-G-G--G------AG--A----C-------GA---G---------- 5879
02_AG.LR.x.POC44951          ---G-C-----------T-C---GC---CG----G------A-----.........................---G-T-----------A.....................------AG--A---G-AG------C-T------C--------A---G---------- 6215
02_AG.NG.09.09NG_SC61        ---GGC--A--------TCC---GC---C-----------------A-----.........................---T-------G------.....................------AGG-A-G-G-----G---AGT------C--------A---G---------- 5412
02_AG.NG.x.IBNG              ---G-C-----------T-C---GC---C-----------------AG----.........................---TAT-------------.....................------AGG-A---G-AG---------------C--------A---G---------- 5744
02_AG.SE.94.SE7812           ---G-C-----------T-C---GC---C-----------------A-----.........................--GTAT------------.....................------AGG-A---G-AG---------------C--------A---G---------- 5592
02_AG.SN.98.98SE_MP1211      ---G-C-----------T-C---T----C-----------------A-----.........................---T-T-------------.....................------AG--A----G-----------A----C-------GA---G---------- 5422
03_AB.RU.97.KAL153_2         ---------------T-GG------------------G----------.........................---T---------------.....................---------------------------------------A---G---------- 5436
04_cpx.CY.94.94CY032_3       ---G-C-G--------GC----T----T------A------A-TT--.........................---T-T-----------A-.....................---T-GAG---G-----G-G---------C--GTACA-----A---G---------- 5584
05_DF.BE.x.VI1310            --------A--------C-----------------------------ATT---.........................---TAT-----------A-.....................---T-GG----G------------A-T------TA-A-A---A---G---------- 5602
06_cpx.AU.96.BFP90           ---G-C-----------T-CT--GC---C-----------------A-T---.........................---T-T---C---------.....................G-----AG--A---G-A---------A-A----C--------A---G---------- 6250
07_BC.CN.98.98CN009          ----C-----------GGC----------------------------A-----.........................---C---------------.....................----------------G-A--G---------------------------A---G---------- 5545
08_BC.CN.97.97CNGX_6F        G------GA--------C----C-----------------------A-----.........................---TAT-------------.....................-G-T-GG---A------G---------T--------A-A---A-T-GG--------- 5387
09_cpx.GH.96.96GH2911        ---G-A--A--------T-C---GC---C---A---C----A-----.........................---T-T-----GT----.....................--G---AG--A-----AG-G------T-----------------A---G---T------- 5425
10_CD.TZ.96.96TZ_BF061       G-------A--------C----C-----------------------AT--G-.........................---TAT---C---------.....................---T--G--GA--CT-CG---------T-------CA-A---A-T-GG--------- 5611
11_cpx.CM.95.95CM_1816       -----------------T-C--T----T-G-----------------A-T---.........................---GGGC----------A.....................--GT--C------------------------------------A---G---G------ 5583
12_BF.AR.99.ARMA159          -----------------GC----T------------C------------.........................---T-A-------------.....................---------------------G------------------------A---G---------- 6215
13_cpx.CM.96.96CM_1849       --------A--------T-T--TT-----G-----------------A-----.........................-C-TAT-----------A.....................--GC-GG--CAT--G-AG---------A-A----T-G------A---GG--G------ 5624
14_BG.ES.05.X1870            ---G-C--A--------GGC---GC---C---RC---R---A-----.........................--KT-T---C------A-.....................G-----AG--A---G-AG------C-AG--A--C---------A---GG--------- 5687
15_01B.TH.99.99TH_MU2079     ---G-C-GA---------------CT---------A------A-T---.........................---GCT------AT----.....................--T-----G-AG---------------------G--A-G---A---G---------A- 5610
16_A2D.KR.97.97KR004         ---GTC-GA--------T-C--CT------G-------A---A-T---.........................---T-----------A-.....................-----TAAG-AG----AC---------T--------A-G---A---GT--------- 5580
17_BF.AR.99.ARMA038          -----------------C------------------------------T---.........................--GC---------------.....................---T-------------------------------------A---A---G---------- 5424
18_cpx.CU.99.CU76            --------A-------TT-T--T------G--TG-------A-----.........................---TAT-----------A.....................--G--T----A-----AG--------TG-----T----G---A---G---C------ 5513
19_cpx.CU.99.CU7             ---G-C--A--------T-C---GC---C-----------------A-----.........................--GTAT-----------A.....................------AG--A-G-G-------------A-A----C--------A---G---------- 5436
20_BG.CU.99.Cu103            ---G----A--------GGC---GC---C---A--------G-----.........................--TT-----CT-----A-.....................G----TAG--A---G-A---------T---A--C----A---A---G---------- 5687
21_A2D.KE.99.KER2003         C-----------------C-------------------A---A-----.........................---T-T-----------A.....................-G----AA---------G-G-------T-T---------------A---G---------- 5415
22_01A1.CM.01.01CM_0001BBY   ---G-C--A-------GC----CT--------------A---A-T---.........................---GCT------AT--A-.....................---G--AG---G----AG---------T------C----A---A-C-G---------- 5423
23_BG.CU.03.CB118            ---G----A--------GGC---GC---C---A--------G-----.........................--TT-T--TCT-----A-.....................C-----AG--A---G-A---------T---A--C----A---A---G---------- 5663
24_BG.ES.08.X2456_2          ---G----A--------TGC---GC---C---A--------A-----.........................--TTGG---C------A-.....................G-----AG--A---G------------T---A--C----A---A---G---------- 5671
25_cpx.CM.02.1918LE          ---GGC--A--------T-C---G-----G--A--------A-T---.........................---TATC-----AT----.....................------AG--A---G-A---G-----TCA----C----A---A---GGAC---G--- 5428
26_AU.CD.02.02CD_MBTB047     ---G-C--A--------C----CC---------------------T---.........................--GT-AG---------A-.....................--G---C---AG-----G--G---------------C--------A---G---C------ 6214
27_cpx.FR.04.04CD_FR_KZS     C-------A--------T-TC-T----T-G-----------------A-----.........................---TAT-----------A.....................--GT-G-----------A--G------T------C----G---A---G---------- 6210
28_BF.BR.99.BREPM12609       ---------------T-GC----------------A------A-----.........................--GT-T-------------.....................------------------G--------------------------A---G---------- 5576
29_BF.BR.01.BREPM16704       ---------------T-GC-----------------------------A-----.........................--GT-T-----------A.....................--------------------GG--------A-----A------A---G---------- 5614
31_BC.BR.04.04BR142          G---C--GT--A-----T----C---------A------A-----.........................--TTAT-------------.....................---T-GG------------G--------T-----G--A-A---A-T-GG--------- 5703
32_06A1.EE.01.EE0369         ---G-C--A--A----TTC-T--GC---C-----------------A-T---.........................---T-A---C---------.....................G-----AG--A---G------------A-AA---C--------A---G---------- 5855
33_01B.ID.07.JKT189_C        C--G-C-GA--A----GC----CT-----------A------A-T---.........................---GG------G-G-CA-.....................------C--CAG-----G-----------------G--A-G---A---G---------- 5515
34_01B.TH.99.OUR2478P        ---GGC--A-------GC----CT-----------A------ATT---.........................---GCT-------------.....................------C-----G------------T------G--A-G---A--CA--------- 5416
35_AD.AF.07.169H             ---G-C-G--------GC----CC-----------------------A-T---.........................---GG-------AT----.....................-G-G-GC-G--G-----G--------T------G--A-G---A---GT------A-- 5420
36_cpx.CM.00.00CMNYU830      C--G-C-----------T-C---GC---C----G--G---A-----.........................---C-T-----------AA.....................------A----G----A------------------C----A---A---A---------- 5429
37_cpx.CM.00.00CMNYU926      ---G-------------T-C---GC---C-----------------A-T---.........................---C-T-----------A-.....................------AG--A---G---------------------A-----A---G---------- 5411
38_BF1.UY.03.UY03_3389       --------A--------C-----------------------------A-T---.........................---TAT-------------.....................-----TG----G-------G---A-T------TA---A---A---G---------- 5643
39_BF.BR.04.04BRRJ179        C-------A--------C-----------------------------A-T---.........................-C-TATC------------.....................---T-TG------------G---A-T------TA---A---A---G---------- 5697
40_BF.BR.05.05BRRJ055        ---G-----------T-GCT--------------------------A-T---.........................--GC--C--------C---.....................---T------A---G-------------------------G---A--GC--------- 5723
42_BF.LU.06.luBF_18_06       -----------------GC-T-----------------A------------.........................---C---------------.....................-------------A---------------T--------A-G---A---GT--------- 5764
43_02G.SA.03.J11223          ---G-C--A--------T-CY--GC---C-----------------A-A---.........................---T-T-------------.....................------AG------G-A--R------T--A---C--R-----A---G---------- 5745
44_BF.CL.00.CH80             C-------A--------GC--------T------------------AT----.........................---TAT-------------.....................---C-GG----G------------A-T------TA---A---A---G---------- 5662
45_cpx.FR.04.04FR_AUK        C---T---A--------GCTC--C--------A-------A-----.........................---TAT-----------A.....................---G-GAA--A----GA--G----A--T-----TA-A-----A---G---------- 6203
46_BF.BR.07.07BR_FPS625      C-------A--------C-----------------------------A-T---.........................-C-TATT---------AA.....................---C-GG----G------------A-T--A---TA-A-G---A---G---------- 5650
47_BF.ES.08.P1942            -----C-----------GC----T--------------------------T---.........................---C---------------.....................--------------G--------G--T-T-AA--------C---A---G---------- 5652
48_01B.MY.07.07MYKT021       ---G-C-G--------GC----CT-----------A------A-----.........................---GGT------T----.....................------C----G---------------------A-A---A---G-------A-- 5419
49_cpx.GM.03.N26677          --------A--------T-TT-CT---TG---A----A---A-----.........................---TAT---C------.....................C-GC-GG----G-------------A---------------A---G---G------ 5639
50_A1D.GB.10.12792           C----------A---TGTC-----------------------------A-T---.........................---T------G---AA.....................-G-T--AG-----------G--------C-----------A---G---------- 5617
51_01B.SG.11.11SG_HM021      -------G---------GC----C--------A-A--------A---.........................---T-------------A.....................---T------A-----AG-------T-----------------A--GC------G--- 5417
52_01B.MY.03.03MYKL018_1     C--G-C-G--------GC----CT-----------A------A-T---.........................---GCT-------T--A-.....................------C----G------G--------------G--A-A---A---G---------- 5528
53_01B.MY.11.11FIR164        ---G-C-G--------GC----CT-----------A------A-T---.........................---G-TT------------.....................------C----G-------------------GG--A-----A------------ 5551
54_01B.MY.09.09MYSB023       C--GGC-GA-------GC----CT-----------A------A-----.........................---C---------------.....................------------A--------------------C-----------A---G---G------ 5616
55_01B.CN.10.HNCS102056      C--G-C-GA-------GC-G--CT-----------A------A-T---.........................---GCT-----------A-.....................------CA--A-G-------------A-A----C----G---A---G---------- 5524
56_cpx.FR.10.URF5_patient_A  ---G-C--A--------T-C---GC---C---A----A---A-T---.........................---TAT-------------.....................------AG--A---GCA--GG-------------C----C-A-A---G---------- 5571
57_BC.CN.09.09YNLX19sg       G-----------------C----C------------------------A-----.........................---TAT---A---------.....................---T-GG-------------------T-T------A-A-A-A-T-GG--------- 5397
58_01B.MY.09.09MYPR37        ---G-C-GA-------GC----CT-----------A------A-T---.........................---GCT------TGCT-.....................------AG--A-----------------A---G---G---A---G---G------ 5559
59_01B.CN.09.09LNA423        ----C--T---A---T-GGG---------------------------A-----.........................---TA-C----------A.....................-----T----------AG---------------------------A--GC--------- 5401
60_BC.IT.11.BAV499           G------GA-----AG-C----------------A-------AT----.........................--TTAT-------------.....................--GT-GG-------------G--------T----G--A-A---A-T-AG--------G 6180
61_BC.CN.10.JL100010         G------G---------GC-G--C----T-----------------A-T---.........................---TA--------------.....................---T-GG---A------G-------------------A-A---A-T-GG--------- 5634
62_BC.CN.10.YNFL13           ----C-----------GC-----------------------------A-----.........................---T---------------.....................------AG------G----------A-T------------------A---G------T-- 5536
63_02A1.RU.10.10RU6637       ---G-C-----------T-C---GC---C-----------------A-----.........................--GTAT-------------.....................------AG--A---G-AG----------A----C--------A---G---------- 5761
64_BC.CN.09.YNFL31           G------TA--------T----C-----------------------A-----.........................---TAT-------------.....................-G-T-GAA--A------G----------T--------A-A---A-T-GG--------- 5567
65_cpx.CN.10.YNFL01          ---G-C--A-------GC-G--CT----------------------A-----.........................---T-A-------------.....................------G---A------------------A-----A-G---A--CA--------- 5578
67_01B.CN.11.ANHUI_HF115     ---G-C-GA-------GC----CT---------R-A------A-T---.........................---GGT------T----.....................--------------G----------------A-----A-G---A---G---------- 5557
68_01B.CN.11.ANHUI_WH73      ---G-C-GA-------GC----CT-----------A------A-T---.........................---GG-------T--A.....................-----T-----G-------G--------A-----A-G---A---G-------A- 5548
70_BF1.BR.10.10BR_PE004      ---------------T-GC-----------------A--------A---------.........................---T-A-------------.....................------------------G-G---A----------------------A---G---------- 5666
71_BF1.BR.10.10BR_PE008      -----C--A--------GG------------------------------------.........................-----------------A.....................---------G-----------------T-----G--------A---G---------- 5874
72_BF1.BR.10.10BR_MG002      -----------------GG----C-----------------------A-----.........................---C---------------.....................------AG--A-G-G-A--G------A----------A-----A--GC--------- 5718
O.BE.87.ANT70                --ATTA-TAGT-CTTTGC-GT-T---AATG--A--T-A---GGGT-T.........................A-TC-T-G-A-----TTAGAA........CAAAAGGAAC--GACAG--AGG-G---G-----CTTGA-AGGT-AAGA---A---G-----TTAGG 6270
O.CM.98.98CMA104             --ATTG-TAGT-CTTTGC-TC-T---AATG--C--T-A---AG-T-T.........................A-TC-T-G-A-----TTAGAA........CAAAAGAAAC--GACAG---GG-----G-----CTTGAAAGGT-AAGA-A-A---G-----TTAAG 5711
O.CM.98.98CMABB141           --ATTA-TAGT-CTTTGC-GC-T-C-AAT---A--C-A---ATGT-T.........................--TC-T-G-A-----TTAGAA........ATAAAGAAAC--GA-AG---GG-----G-----CTTGAAAG-T-AAGA-R-A---G-----TTAG- 5716
O.CM.98.98CMABB212           ----TG-TAGTA-TTTGTGTC-T---TGT---C----A---A-AT-T.........................AACC-T-G-A-G---TTAGAG........CATAGAAAAC--GATAA----G-GC-GG-----CTTGAAAG-T-AAGG---G---GG----TCAGG 5718
O.CM.98.98CMU5337            --ATTA-TAGT-CTTTGC-G-YT---AATG--A--C-A---ATGT-T.........................A-TC-T-G-C----YTTRAAA........RATARGAAAC--GACAG---GG-----G-----CTTGA-AGGT-AAGA---A---G---G-TTAA- 5712
O.CM.99.99CMU4122            --ATTACTAGT-CTTTGC-GC-T---AATG--A--C-A---CTGT-T.........................--TC-T-AGC---G-TTAGAA........CAAAAGAAAC---CTAA----G-----G-----ATT--AAGGT-AAGA---A---G-----TT-AG 5710
O.FR.92.VAU                  ---TCA-TA-T-CTTTGC-TC-T---AATG--C--C-A---ATGT-T.........................AATC-T-G--C----TTAGAA........CAAAAGAAAC--GATAG--C-GC----G-----CTTGAAAG---AAGA---A---G----GTT-AG 5794
O.GA.11.11Gab6352            --ATT-CTAGT-CTTTGC--T-TC--AATG--C-AA-A---A--T-T.........................---C-T-A-C---G-TTAGAT........CATAAGAAAC--GACAG---GG-----G-----CTTGAAAGGG-AAGA-------G---G-TTAAG 5442
O.SN.99.99SE_MP1299          --A-TACTAGT-CTTTGC-GC-T-C-AATG-----C-A---A-AT-T.........................A-TC-T-G-C-----TTAAAA........CAAAAGAAAC--GATAA---GG-----G-----CTTGAAAGGT-AAGA---A---G-C---TT-AG 6281
O.US.10.LTNP                 --ATTG-TAGT-CTTTGC-GC-T---AATG--A----A---ATGTAT.........................--TC-T-G-A-----TTAGAA........CAAAAGAGAC--GACAG----G-----G-----CTTGAAAGG--AAGG-A-A---GG----TTAGG 6185

N.CM.02.DJO0131              C---G--G-G---G-----T---------G--A-AA-C---GTTT--.........................C--TAT-A---------A-.....................---C--AA-TTG--GGA---------ACAAA---GAC-A---A---G---T---A-- 5707
N.CM.02.SJGddd               ---------GGT-----C-T--C-------G--A-AA-T---GTAT--.........................C--TAT-G---------A-.....................G-----AA-TTG---GA--G----A-ACACA---GAC-G---A---G---C--G--- 5691
N.CM.04.04CM_1015_04         C---GC-C-GG--G-----T---------G-CA-AA-C---GTAT--.........................C--TAT-A---------A-.....................G-----AA-CTG--GGA---------ACAAA---AAC-G---G---G---C------ 5706
N.CM.06.U14296               ---------G-T-G-----T---------G----AA-C---GTAT--.........................A-CTAT-G---------A-.....................------AA-TTG--GG-G-G----A-A-CAA--........................ 5683
N.CM.06.U14842               ----GC---GG--------T---------G----GA-C---GTAT--.........................C--TAT-A---------A-.....................------AA-TTG--GGA-G-----A-ACAA----GAC-G---A---G-AGT------ 5708
N.FR.11.N1_FR_2011           -----G---G-------T-T---------G----AA-T---GTAT--.........................T--TAT-G---------A-.....................------AA-GTG---GA-------AGGCACA---GAC-G---A---G---T--G--- 5601

P.CM.06.U14788               --AT-GCT-GG--GTT-T--C-TTGCTGT-----A-YT---GGG-AG.........................--CC-YC--TT-GTGCTA.....................---GA-AG-GA-AGRGAT--GT-T-TRCAAAGGC-A-CA--GTGG-G----G-GCA- 5697
P.FR.09.RBF168               --AT-GCT-GG---TT-C----TTGCATT-----A--T---GGG-AG.........................--CC-GC--TTTGTGCTA.....................---GAGAG-GA-AGGGAT--GT-T-TACAAAGGC-A-TA--GTGG-G------GCA- 6261

CPZ.CD.06.BF1167             -TAGGC-G-GGA..........GGG--T------A-C---ATA---.........................-CGC-AC---TTAT----AAAGCATATTTGTATATAAAAC--G-GCA-GA-G--GAT-G------GCAAGC---AAT-GATAG-TACC-G---AG- 6315
CPZ.CM.05.SIVcpzMT145        --TT---CA---CTTT---GC-TG----T---T---CT---ATAGC-.........................-CTTAT-A---G----AA.....................-GT--CA-CA-GT---GCG------A--AC-TCAG--C---A-T-G-AGT------ 5747
CPZ.GA.88.GAB1               ---TTC-CA--T--.........G---GCT--A---CA---AA---TTGTATATGGGGATACATTATTAAATGG-G-TAT-G-AGG---AA.....................-G-CATAGGCTTG-G-C-G----T--G--AC-T-ACCTA-T-T-G-G---------- 6278
CPZ.TZ.06.SIVcpzTAN13        -CATCA----TT---T-GG-GG-GG----T--A-A-CT---ATT---.........................A---AAGA-AG-A-C---.....................G---G-CAGCAT--G--GCTG-----G--A---G-AAGG--G----GT.......... 5852
CPZ.US.85.US_Marilyn         ---------G-A---A-GGT--TT---T-G-GA-AA-C---GGAC--.........................--TGCT-G-TTG-GG--ACAA.........ATAAAAATA---GA-AAT-CT---CA-G-G--TC-A-ATC-TT-A--A---A-C-G-AT-----A- 6300

GOR.CM.07.SIVgor2139_287     -TAT-GG-A--AC-TTGC--GCTG--A-TG-GA-AA-T---CTA-A-.........................A-CT-TGC-TT---CTTA.....................-G-GATAG---GG-G---CGCT-TTTTGATCGCT-A--A--G---C--AGC-ATAA- 5760
GOR.CM.07.SIVgorCP2135con    -TAT-GG-A--AC-TTGC--GCTGC-A-TG--A-A--T---CTA-A-.........................-CCC-TGC-TT---CTTA.....................-G-GATAG---GG-G---CGCT-TTTT-ATCGCT-A--G--G---C--AGC-CTAA- 5759
```

```
                                           Env start                                                                                                                   Vpu end
                                                                                                                                                           Env signal peptide end          Env gp120 start
B.FR.83.HXB2                GAAGACAGTGGCAATGAGAGTGAAGGAGAAATATCAGCAC...TTG...TGGAGATGG...........GGGTGG...AGATGGGGCACCATG.........CTCCTTGGGATG...TTGATGATCTGT...AGTGCT.........ACA...GAAAAATTGTGGGTCA 6333
Vpu                         _E__D__S__G__N__E__S__E__G__E__I__S__A__.__L__.__V__E__M__.__.__.__.__G__V__.__E__M__G__H__H__.__.__.__A__P__W__D__.__V__D__D__L__.__*
Env                         ____________M__R__V__K__E__K__Y__Q__H__.__L__.__W__R__W__.__.__.__.__G__W__.__R__W__G__T__M__.__.__.__L__L__G__M__.__L__M__I__C__.__S__A__.__.__.__T__.__E__K__L__W__V__
A1.AU.03.PS1044_Day0        --------------------C-----G--T-C-GAT-A--.........CACT-C..............TTA-T-...-A------A--T---.........A--T-G--A---...A-AT-A------...------.........G--...---C--AG------T- 5547
A1.CH.03.HIV_CH_BID_V3538   -----------------------T--G--TGC-GAG-A-T...-AT...CCACAC..............T-----...-AG-----G-T-G--.........---T-CT--T--...--A--A--T---...-A----.........---...---------------- 5548
A1.CY.08.CY236              -----------------------G--G--T-C-GAT-A--...-C-...CATT-C..............TT--T-...---------A--T---.........A--T-G------...A-A--A------...------.........GT-...------AC------T- 5550
A1.ES.05.X1608_8            --------------------------G--T-C-GAT-A--...-C-...CA-C-C..............TTA-T-...---------G--T---.........A--T-G--A---...A-A--A------...------.........G--...-----C--------K- 5787
A1.KE.11.DEMA11KE001        -----------------------T--G--C-CTGAT-A-T...-G-...C--C-C..............TTTGT-...---------A-TT--.........G--T-G------...A-A--A------...------.........G--...C----C--------T- 5688
A1.RU.11.11RU6950           -----------------A--C--G--G--TGA-GAG-A-T...-AT...CAACAC..............TTA---...---------A--TTG.........---T--T-----...A-A--A--G---...-AG---.........G--...---C-T---------- 5857
A1.RW.11.DEMA111RW002       -----------------A-----T--G--T-C-GAG-A-T...-AT...G--CA-..............CT----...---------A--T---.........A--T-G--A---...A-A--A------...------.........---...-----C--------T- 5690
A1.SN.01.DDI579             -----------------------T--G--T-C-GAG-A-T...-AT...CTACCC..............TT----...--G-----G-TG--A.........TC---CT-T---...A-A--A--T---...-AA---.........---...-----C---------- 5540
A1.UG.11.DEMA110UG001       -----------------------T--G--T-C-GAT-ATT...-G-...CAA-AC..............TT----...---------A--T---.........A-TT-G------...--A--A--T---...----T-.........G--...-G---C--------T- 5680
A1.ZA.04.04ZASK162B1        -----------------------T-----C-C-GAG-A-T...-AT...CCAC-C..............T-----...-T------A-TT---.........A--T-G------...A-A--A--TG--...---G--.........T--...--C--CC--------- 5771
A2.CM.01.01CM_1445MV        -----------------A-----T--G--C-C-GAG-A-T...-GT...CAACAC..............TT----...--------A-T-T--.........A-TT--------...C-A--A--T---...-AG---...............AC---T---------- 5535
A2.CY.94.94CY017_41         -----------------------T--G--C-C-GAG-A--...-AT...CA-CAC..............TT----...---G------T-T--.........A-TT-G------...--A--A--G---...-AA------------------G-T---------- 5693
B.BR.10.10BR_MG029          -----------------------------TCAGGA--A-T...-AT...CA---M..............TAC---...----------T-T--.........------------...--A---------...------.........---...--Y-G----------- 5791
B.CA.07.502_1191_03         --------------------------G--TCAGGAG-A-T...-AT...CA-CAC-C-TGGAGAGGGGGCATC--...----------T----.........------------...--A---------...------......GCAGAG...-TC-GGG--------- 5777
B.CH.08.M2_0803101_NFLG8    -----------------A--------G--CCA-GA--A-T...-AT...CA-CAC..............TT----...----------TG---.........------------...------------AGT------.........G--...-TC------------- 5789
B.CN.10.DEMB10CN002         --------------------C-----T--TCAGGA--AGT...-AT...CA-CCC..............T-----...-T--------G----.........--------A---...--A---------...------.........-A-...---------------- 5700
B.ES.10.DEMB10ES002         -----T--------------------G--TCAGGA--A-T...-AT...CAA-AC-T-...........T-----...-T----------T--.........-----------A...------------...---A--.........G--...--G--C---------- 5694
B.FR.11.DEMB11FR001         -----------------A-------A---CCAGG...............AA-CAC..............TTA---...----------T-T--.........-----------A...------------...------.........G--...---C------------ 5566
B.GB.05.MM45d213_GN1        -----------------------------TCATGA--AG-...-G-...CA-CAT..............T-----...-A-G------T-T--.........-----------A...--A-----T---...---AA-.........G--...--G--C---------- 5730
B.HK.06.HK003               -----------------A--------G--TCAGGA--ACT...-GT...CA-C-C..............TTT---...-A------A------.........-----------A...------------...------ACT......-A-...--G------------- 5702
B.HT.05.05HT_129389         --G--------A--------------G--TCAGGA--A-T...-AT...CA-CAC..............T-----...-C--------T----.........------------...--A---------...------.........G--...-----GA--------- 5533
B.JP.12.DEMB12JP001         -----------------A--------G--CCAGGA--A-T...-GT...CA-CAC..............TTA---...-A--------T----.........------------...--A-----A---...------...GTAGAA-A-...--C--T--------T- 5725
B.KR.07.HP_18_07JHS10_3909  --------------------------G--TCAGGA--A-T...-AT...CA-CAC..............T-----...----------T----.........------------...-G----------...------.........---...-TG---A--------- 5990
B.PE.07.502_0525_wg5        --------------------C--G--G--TCATGAG-A-T...-AT...CA-CAC..............TTA---...--------T--A---.........--------A---...--------A---...-A----.........G--...--CC------------ 5806
B.RU.11.11RU21n             -----------------------T-----TCAGG...............AA-CAC..............TTA---...------------TG-.........------------...--------A---...------.........--T...C----T--A------- 5891
B.TH.08.MERLBDTRC10         -----------------A--C-----G--TCAGGA--A-T...-AT...CA-CAC-T-...........T-----...-A--------G----.........------------...------------...-----A...............A--C------------ 5565
B.US.11.ES38                -----------------------T--G--TCAGGA--A-T...-AT...CA-CAC..............T-----...---------------.........------------...------------...------.........---...---------------- 6264
C.AR.01.ARG4006             --------------------C-----G--T-C-GAG-A-T...-G-...AAACA-..............T-----...-T--------T-T-A.........GG-T--T-C---...--M-----T---...-A--TG.........-GG...-GG--C--A------- 5521
C.BR.07.DEMC07BR003         -----------------------T--G--C-C-GAG-A-T...-G-...AAACAC..............T-----...-T------A-T-T-A.........GG-T--T-----...--C---C-----...-A--TG.........-TG...-GG--C---------- 5683
C.BW.00.00BW5031_1          -----------------------T--G--C-C-GAG-A-T...-GT...CAACA-..............T-----...-T--------T-T-A.........GG-T--T-----...--A-----T-A-...-A--TG.........GG-...-GG--C---------- 5704
C.CN.10.YNFL19              -----------------------G-----T-CTGAG-A-T...-AT...CAACA-..............T-----...-T-------GT-T-A.........GG-T--T-----...--A-----T---...----TG.........GG-...-G---CCG-------- 5685
C.CY.09.CY260               -----------------------T--G--TCATGAG-A-T...-GT...CAA---..............T-----...-T------A-T-T-A.........GG-T--T-----...C-C---------...-A--GG.........GGG...--C------------- 5574
C.ES.08.X2363_2             -----------------------G--G--T-CTGAG-A-T...-G-...CAACAT..............T-----...-T--------T-T-A.........GG-T--T-----...--A-----T--W...-A--TG.........TGG...-GG--C---------- 5772
C.ET.02.02ET_288            -----------------------G--G--T-C-GAG-A-T...-GT...CAACA-..............T-----...-T------ACT-T-A.........GG-T--T-----...--C-----T---...-A--TG.........GTG...-GG--T---------- 5553
C.IN.09.T125_2139           -----------------------T-----T-CTGAG-A-T...-AT...CAACA-..............T-----...-T--------T-T-A.........GG-T--T-----...--A-----T---...-A--TG.........GG-...-G---C---------- 6305
C.KE.00.KER2010             -----------------------G--G--T-C-GAG-A-T...-G-...CAACA-..............T-----...-T--------T-T-A.........GG-TC-T-----...--A----CT---...----TG.........GA-...-GG--C---------- 5522
C.MW.09.703010256_CH256.w96 -----------------------G--G--T-C-GAG-A-T...-AT...CAACA-..............T-----...-T----A---T-T--.........GC-T--T-----...--A---G-T......----TG.........-GG...-G---C--------T- 6318
C.TZ.08.707010457_CH457.w8  -----------------A-----T-----T--TGAG-A-G...-G-...AAACAT..............T-----...-T------A-T-T-A.........GG-T--T-----...--AT-A--TA--...-A--TA.........-TG...-GG--C---------- 6335
C.YE.02.02YE511             -----------------------C--G--T-CTGAG-A-T...-AT...CAACA-..............T-----...-T--------T-T-A.........GG-T--T----A...--A-----T---...-A--GG.........GG-...-GG--C---------- 5525
C.ZA.08.705010534_CH534.w12 -----T-----------------T--G--T-C-GA--A--...-GT...CAACA-..............T-----...-T--------T-T-A.........GG-T--T-----...--AT-----T--...-A--TG...............C-G-GC-CT------- 6326
C.ZA.10.DEMC10ZA001         -----------------A-----T--G--TCA-GAG-ACT...-GT...CAACA-..............T-----...-T--------T-T-A.........GG-T--T--.........-----T-A-...-A--TG.........CTG...-G---C---------- 5686
C.ZM.11.DEMC11ZM006         -----------------------G--G--T--TGAG-A--...-AT...CCACA-..............T-----...-T--------T-T-A.........GG-T--T-----...G-A----GT---...----TA.........-GG...-G---T---------- 5700
D.CM.10.DEMD10CM009         --------------------C--G--G--T-A-GAG-A-T...-AT...CAACAC..............TT----...-C--------T----.........--------AT--TTGC------C--A-...-----A.........G--...--T--T---------- 5707
D.CY.06.CY163               -----------------------G--G--T-A-GAG-A--...-AT...CAC-AC..............TTA---...-A-------------.........--------A---TTGC------CA-A-...-----A.........G--...--T--G---------- 5540
D.KE.11.DEMD11KE003         -----T--------------C-----G--TGA-GAG-A-T...-AT...CA-CAT..............TT----...G-------T-T----.........-----------A...--A----CA-A-...----T-.........G--...---C-G--------T- 5678
D.KR.04.04KBH8              -----T-----------------G--G--TGA-GAG-A-T...-AT...CAACAC..............TTA---...-A-------------.........--------A---...-G------T---...------.........G--...-G---C---------- 6285
D.SN.90.SE365               -----T--------------C--G-----TGA-AAG-A-T...-AT...CA-CAC..............TT----...-T-------------.........--------A---...-------C----...----T-.........G--...---------------- 6346
D.TZ.01.A280                -----T-----------------T-----C-C-GAG-A-T...-AT...CAACAC..............TT----...----------T----.........--------A---...-G-----CA-A-...----T-.........G--...---C-G---------- 5530
D.UG.10.DEMD10UG004         -----T-----------------G-----TGA-GA--A-T...-AT...CAA-AT..............TCA---...----------T----.........--------A--A...-------CA---...----T-.........---...-G---C---------- 5667
D.UG.11.DEMD11UG003         -----T-----------------G--G--T-G-GA--A-T...-AT...C--CAC..............TT----...G----------T---.........-----------A...---G---CA-A-...----TG.........G--...---C------------ 5681
D.YE.02.02YE516             -----------------------G--G--TGA-GAG-A-T...-AT...CAACAC..............TTA---...-A---------GC--.........--------A---TTGC------C--A-...-----A.........G--...--T--G---------- 5543
D.ZA.90.R1                  --------------------C--G--G--TGG-GAG-A-T...-GT...CAA-AC..............TT----...-A------T-TG---.........--------A---...------------...------.........G--...-----T---------- 5678
F1.AO.06.AO_06_ANG32        --G--------A-----------T--G--TGC-GAG-A-T...-G-...CA--AC..............TT-G--...-AG------CTTT-A.........T----G---G-A...--A--A------...-A----.........G--...-----C---------- 5551
F1.AR.02.ARE933             -----T-----------------G--G--TGC-GAG-A-T...-G-...CA-CAC..............TT-G--...-A-------C-TT-A.........T----A-----A...-----A------...-A----.........---...-----C----T---T- 5633
F1.BR.10.10BR_RJ015         -----------------A-----G--G--TGC-GAG-A-T...-G-...CA-CAC..............TT-G--...-A-------CTTT-C.........T----G-----A...--A--A------...-A--A-.........GA-...-----G---------- 5745
F1.CY.08.CY222              --------C--------------G--T--CGC-GAG-A--...-G-...CA-CAC..............TT-G--...--G------CTTT-A.........T----G-----A...--A--A------...------.........G--...--C--C--A------- 5529
F1.ES.02.ES_X845_4          --G--------------------G--G--TGC-GAG-A-T...-G-...CA-CAC..............TT-G--...-AG------CTTT-A.........T----G--A--A...--A--A------...-A----.........G--...--C--C--A------- 5791
F1.RO.96.BCI_R07            --------------C--A-----G--T--TGC-GA--A--...-G-...CA-CAC..............T--G--...-A-------CTTT-A.........T----G-----A...--A--A------...------.........---...--T-----------T- 6363
F1.RU.08.D88_845            -----------------------T--G--TGC-GAG-A-T...-G-...CA-CAC..............T--G--...-A------CTTT---.........T----G------...--A--A------...-A----.........---...A-C--C---------- 5798
F2.CM.02.02CM_0016BBY       -----------------------G--G--TGC-GAG-A-T...-G-...CA-CAC..............TT-G--...-AG------TTTT-A.........T----G--A--A...--A--A------...-A----.........G--...--T--C---------- 5522
F2.CM.10.DEMF210CM001       -------------------A---G-----TGC-GAG-A-T...-G-...CA-CAC..............TT-G--...--G------CTTT-A.........T----G-----A...--A--A------...-A----.........---...--G--C--------T- 5593
F2.CM.10.DEMF210CM007       -----------------------G--G--TGC-GAG-A-T...-G-...CA-CAC..............TT-G--...-A-------CTTT-A.........T----G--A--A...--A--A------...-A----.........G--...--T--C---------- 5678
G.BE.96.DRCBL               -----------A--------------G--T-C-GAG-A-T...-G-...CAACAC..............TTA---...-AC-----G-TTT--.........A--------T--...G---A--T----...------.........G--...---------------- 6293
G.CM.10.DEMG10CM008         -----------A--------------G--T-C-GAG-A-T...-G-...CA-CAC..............TT----...-CG----G--TT---.........A--T----AT--...G-A--A--T---...-----C.........T--...-----C---------- 5720
G.CN.08.GX_2084_08          -----T-----A-----------G--G--T-C-GAG-A-T...-G-...CAACAC..............TT----...-T----T-G-ATT--.........A--------T--...G---A--A----...---A-C.........T--...--G--T---------- 5575
G.CU.99.Cu74                -----------A-----------G--G--T-C-GAG-A-T...-GT...CA-CAC..............TT----...-A-----G--TT---.........A--A-----T--...GC--A--T----...-----C.........T--...ART--C---------- 5943
G.ES.09.X2634_2             -----------A-----A--------G--C-C-GA--AGT...-G-...CAA-AC..............TTA---...-A-----G--TT---.........A--------T--...G---A--T----...-----C.........T--...A-C--C---------- 5822
G.GH.03.03GH175G            -----------A--------------G--T-C-GAC-A-T...-G-...CA-CAC..............TT----...-A-----G--TT---.........A--------T--...G---A--T----...-----C.........T--...--C------------- 6379
G.KE.09.DEMG09KE001         -----------A--------C--------TGA-GAG-A-T...-G-...CAAC-C..............TT----...GA-G---GCTTT---.........A--------T--...G---A--T----...-----C.........T--...ATT--C-------C-- 5715
G.NG.09.09NG_SC62           -----------A-----A--------G--T-C-GA--A-T...-G-...CCACCC..............TT----...-C---T-G--TT---.........A--------T--...G-A--A--T---...-A..............G--...--T--C---------- 5527
G.ZA.01.TV546               -----------A-----------G--G--T-C-GAG-A-T...-G-...CAACAC..............TTA---...-A------R--TT--.........A--------T--...G---AC-T----...-A---C.........T--...A-T--T---------- 5541
H.BE.93.VI991               --------------C--------T-----C-C-GAG-A-T...-AT...CCA--C..............TT----...------------T-A.........A--T-G--T---...---C--------...----T-.........GT-...-G---C---------- 5711
H.CF.90.056                 --------------C--------T-----C-C-GAG-A-T...-AT...CCA--C..............TTA---...------------T-A.........A--T-G--T---...---C--------...------.........G--...C----C--------T- 5673
H.GB.00.00GBAC4001          --------------C--------T-----C-C-GAG-A--...-AT...CCACAC..............TT--T-...--CG------T-T-A.........A--T-G--T---...---C---G----...---A--.........GG-...-----C---------- 5841
J.CM.04.04CMU11421          --G--------A-----C-----T-----C-C-GAG-AGT...-G-...CAA-AC..............TT----...-A------CT-----.........A--T---A----...--------T---...-AA---.........G--...--C---A--------- 5844
J.SE.93.SE9280_7887         -----------T--C--------T-----C-C-GA--A-T...-G-...CA--CC..............TT----...---G----CT-----.........A--T---A----...--A-----T---...-AG--A.........-A-...---G-C---------- 5652
J.SE.94.SE9173_7022         -----------T--C--------T-----C-C-GAC-AGT...-G-...CT---C..............TT----...--------CT-----.........A--T---A----...--A-----T---...------.........-G-...-----C---------- 5653
K.CD.97.97ZR_EQTB11         --------------------C--G-----T-C-GAG-A-T...-G-...CA-CAC..............TT-G--...-A-C------T-T-A.........T--T-G-----A...--A--A------...--C---.........G--...A-C--C---------- 5528
K.CM.96.96CM_MP535          -----------------------G--G--TGC-GAG-A-T...-G-...CA--C-..............TT-G--...-AC-------T-T-A.........T--T-G--A--A...-----A------...--CAA-.........G--...--C------------- 5532
U.CA.01.TV749               -----T-----------------T--G--C-C-GAG-A-T...-AT...CAA--C..............TT----...--------CT-----.........A--T---A----...--A-----T---...-A-A--AAA......GGG...--G--C-------C-- 6358
U.CA.99.TV721               -----T-----------------T-----CGA-GAG-A-T...-AT...CA-C-C..............TT----...--------CT-----.........A--T---A----...--A-----T--C...-AGAT-...GAAGGG-A-...--C--C---------- 6360
U.CD.83.83CD003_Z3          -----------------------------T-C-GAG-A-T...-AT...CAACAC..............TT----...-A----A--CTG--A.........A-TT-A--A---...A-A-----A---...-AA---.........-T-...-------C-------- 5821
U.CD.90.90CD121E12          --------------------C--------T-C-GAG-A-T...-AT...CACTAC..............TT----...-A----A--CTG--A.........A-TT-A--A---...A-A--A--A---...-AA-G-.........GA-...---C---C------T- 5522
U.CY.05.CY090               --------------------C--------C-C-GA--A-T...-G-...CA-CAC..............TT-G--...-A--------T-T--.........T----G-----A...--C--A------...---A--GCAGAAAAAG--...-----T---------- 5546
U.CY.08.CY223               -----------------------T--G--C-C-GAG-A-T...-GT...CTACTC..............TC--T-...--G-----G--A---......TGCA--T--T--T-A...A----G-A----...-A----...............A-C------------- 5554
U.ES.10.DEURF10DZ001        -----------A--------C--G--G--T-C-GAC-A-T...-G-...CAA-AC..............TT----...-A-----G--TT---.........A--------T--...G---A--T----...-----C.........---...A-T--CC--------- 5610
U.GR.99.99GR303             -------------------T------G--T-C-GAG-A-T...-GT...CAA-AC..............TTA---...-------C-CTG---.........A-TT---A----...--A-----T---...-AG---.........---...--C--C---------- 5611
U.NL.95.U_NL_95_H10986_D1   -----------------------G--G--TGC-GAG-A-T...-G-...C--CAC..............TT-G--...--------T-T-T--.........T--T-A-----A...------------...------.........G--...--C------------- 6221
01_AE.AF.07.569M            -----------A--------------G--C-C-GAG-A-T...GG-...CCA-AC..............TT----...-A-----G--TT---.........A--------T--...G---A--T----...-----C.........T--...--C--C--------T- 5536
01_AE.CF.90.90CF11697       -----------A---G----------G--C-C-AAT-A-T...-G-...CCACAC..............TT----...-AG----G--TT---.........A--------T--...G---A--T----...-----C.........T--...--C-CC--------T- 6282
01_AE.CN.10.YNFL03          -----T-----A-----------------C-C-GAT-A-T...-G-...CCA--C..............T-----...-A-----G--TT---.........A--------T--...G------T----...-A---C.........T--...A-C--C--------T- 5711
01_AE.HK.04.HK001           -----------A-----------------C-C-GAT-A-T...-G-...CCA-AC..............TT----...-------G--TT---.........A--------T--...G---A--T----...-----C.........T--...--C--C--------T- 5703
01_AE.IR.10.10IR.THR48F     R----------A-----A-----------C-C-GAG-A-T...-G-...CTA-AC..............T-----...-A-----G--TT---.........A--------T--...G--CA-------...-----C.........T--...--C--C--------T- 5523
01_AE.JP.x.JRC77AE          -----------A--------------G--T-C-GAT-A-T...-G-...CCA-AC..............TT----...-------G--TT---.........A--------T--...GG--A--T----...-----C.........T--...--C-----------G- 6330
01_AE.TH.09.AA090a_WG11     -----------A-----------G--G--C-C-GAT-ACT...-G-...CCA-AC..............TTT---...-------G-TAT---.........A--------T--...G---A--T----...-----C.........T--...--C--C--------T- 5755
01_AE.TH.90.CM240           -----------A-----------------C-C-GAT-A-T...-G-...CCA-AC..............TT----...-A-----G--TT---.........A--------T--...G---A--T----...-----C.........T--...--C--C--------T- 5901
01_AE.US.05.306163_FL       -----------A--------------G--C-C-GAT-A-T...-G-...CCA-AC..............TT----...-A-----G--TT---.........A--------T--...GC--A--T----...-----C.........T--...A-C--C--------T- 5524
01_AE.VN.98.98VNND15        -----------A-----------------C-C-GAT-A-T...-G-...CCA-AC..............TT----...-A-----G--TT---.........A--------T--...G---A--T----...-----C.........T--...--C-----------T- 5588
```

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546


```
                                                                                                                                          Vpu end
                                          Env start                                                                          Env signal peptide end        Env gp120 start
B.FR.83.HXB2                 GAAGACAGTGGCAATGAGAGTGAAGGAGAAATATCAGCAC...TTG...TGGAGATGG...........GGGTGG...AGATGGGGCACCATG.........CTCCTTGGGATG...TTGATGATCTGT...AGTGCT.........ACA...GAAAAATTGTGGGTCA 6333
Vpu                          _E__D__S__G__N__E__S__E__G__E__I__S__A__.__L__.__V__E__M__.__.__.__.__G__V__.__E__M__G__H__H__.__.__.__A__P__W__D__.__V__D__D__L__.__*
Env                                       _M__R__V__K__E__K__Y__Q__H__.__L__.__W__R__W__.__.__.__.__G__W__.__R__W__G__T__M__.__.__.__L__L__G__M__.__L__M__I__C__.__S__A__.__.__.__T__.__E__K__L__W__V__
02_AG.CM.10.DE00210CM013     -----T-----------A-----T--G--TGA-GA--A-T...-AT...CCACTC...............TTA---...----T---T-TG--A.........A--T--T----A...A-A--A--T-----...-A---....................-----T---------- 5703
02_AG.CY.09.CY256            -----T-----------C-----C--G--C-C-GA--AGT...-AT...CCAC-C...............TTA---...----T---T-TG--A.........A-AT-CT----A...A-A--A--T-----...-A---.................A--C-G---------- 5522
02_AG.ES.06.P1423            -----T-----------------G--G--T-C-GAG-A-T...-AT...CCACCC...............TTA-T-...--G-----G-TT--A.........A----CT----A...A-A--AT-T-----...-A---.................C----T--------T- 5770
02_AG.GW.05.CC_0048          -----T-----------------T--G--T-C-GA--A-T...-G-...CCCTTC...............TTA---...-A------G-TA--A.........A-AT--T----A...A-A--A--------...GAA--A.........-AC...AG---G--------T- 5692
02_AG.KR.12.12MHR9           -----T-----------------G--G--T-C-GA--A-T...-AT...CCACTC...............TTA---...--------G--T--A.........A-AT--T----A...A----A--T-----...-A---....................-----T---------- 5995
02_AG.LR.x.POC44951          -----T-----------------T--G--T-C-GA--A-T...-AT...CCACTC...............TTA---...----T---T--G...........A-AT--T----A...A-A--A--T-----...-A---.................A-C--TA--------- 6328
02_AG.NG.09.09NG_SC61        -----T--------C--------T-----C-C-GAG-A-T...-G-...CTACTC...............TTA---...---G----GCTG--A.........A-GT--T----A...A----A--T-----...-A---.................C--CC----------- 5528
02_AG.NG.x.IBNG              -----T-----------------T--G--T-C-GA--A-T...-AT...CCACTC...............TTA---...--------T--G-AT.........A-AT--T----A...A----A--T-----...-A---....................-C------------ 5860
02_AG.SE.94.SE7812           -----T-----------------T--G--T-C-GA--A-T...-AT...CCACTC...............TTA---...--------T-TG--A.........A-AT--T----A...A---CA--T---...-------..................-G---T---------- 5708
02_AG.SN.98.98SE_MP1211      -----T-----------------T--G--T-CTGA--AGT...-GT...CCACCC...............TTC---...--------T-TG--A.....ATG--G---T----A...-----A--T-----...-A---....................-----T---------- 5541
03_AB.RU.97.KAL153_2         -----------------------------TCAGG...............AA-CAC...............TTA---...-----------T--.........T------------...C---------------...------...................-----T--A-------- 5546
04_cpx.CY.94.94CY032_3       -----------------------T--G--TGC-GAG-A-T...-AT...CCACAC...............TT-----...GA------A--TT--.........A--------T--...G----A--------...------C.........T--...A-C--C---------- 5703
05_DF.BE.x.VI1310            -----------------------G--G--TGC-GAG-A-T...-G-...CC-CAC...............TT-G--...-AG------CTTT-A.........T----G-----A...--A--A---------...------...................-C-----T-------- 5721
06_cpx.AU.96.BFP90           -----------A-------------G--T-C-GAC-AGT...-G-...CAACAC...............TTA---...-A------G--TT--.........A--------T--...G----A--T-----...-----C.........T-T...A-G--CA-------T- 6369
07_BC.CN.98.98CN009          -----------------------C--G--TCA-GA--A-T...-AT...C--CAT...............TTA---...-----------------.........------------...----------------...------.........GT-....-G---C---------- 5664
08_BC.CN.97.97CNGX_6F        -----------------------G--G--C-CGGAG-A-T...-AT...CAACA-...............T-------...-T------GT-T-A.........GG-T--T--------...--A-----T-----...-A--TG.........GA-....-G---C---------- 5506
09_cpx.GH.96.96GH2911        -----------------------------T-C-GAG-A-T...-G-...CAA-AC...............TT----...---------CT---.........A-T----A--------...----------T-----...-A---.........G--....-----C---------- 5544
10_CD.TZ.96.96TZ_BF061       ------C-----C----C-----C--G--T-C-GAG-A-T...-GT...C-ACA-...............T-------...-T------A-T-T-A.........GG-T--T--------...--A-----T-----...-A---A.........-GG....-G--G---------- 5730
11_cpx.CM.95.95CM_1816       --G--------------------------C-C-GAG-AGT...-G-...CAC-AC...............TT----...---------CT---.........A--T---A--------...----------------...-A---....................-C--C---------- 5702
12_BF.AR.99.ARMA159          -----------------------G--G--TGC-GAG-A-T...-G-...CA-CAT...............TT-G--...-AC------TT-T-A.........T----G--A--A...-----A-----...-A---....................-----T---------- 6334
13_cpx.CM.96.96CM_1849       -----------A-------C--G--G--T-C-GAT-ACT...-G-...CAA-AC...............TT-----...-A------A--TT--.........A--------T--...G----A--T-----...------C.........T--...A-T--C---------- 5743
14_BG.ES.05.X1870            -----------A-----A--C-----G--C-C-GAG-A--...-G-...CA-TCC...............TTA---...--------GG-TT--.........A--------T--...G----A--TA--...G----C.........T--...A-TG-C---------- 5806
15_01B.TH.99.99TH_MU2079     -----------A-----------------C-C-GAT-A-T...-G-...CCA-AC...............TT-----...-A------G--T--.........A--------T--...G----A--T-----...-------.........G--....-----C---------- 5729
16_A2D.KR.97.97KR004         -----T-----------------G--G--T-C-GAG-A-T...-AT...CAACAC...............TTA---...-A------A-T-T--.........A-TT-G-------...C-A-----TA--...-AA---.........-------G-TC---------- 5699
17_BF.AR.99.ARMA038          -----T-----------------G--G--TGC-GAG-A-A...-CT...CA-CAC...............TTA---...-A------------------.........----------A...--A----------...-A-A-A.........G--...-G--CT---------- 5543
18_cpx.CU.99.CU76            ------------T----------G--G--C-C-GAG-A-T...-AT...CCA--C...............TTA---...-A------A----A.........A---G--T----...---C---G----...------.........------TC-C---------- 5632
19_cpx.CU.99.CU7             -----------------------T--G--C-CGGAG-AGT...-AT...CCACCC...............TTA---...-AT-----G--TT--.........A--------T--...G----A--G-----...------C.........TTT...--C--C---------- 5555
20_BG.CU.99.Cu103            -----------A-------------G--T-C-GAC-A-T...-G-...CA-T-C...............TT-----...-C------G--TT--.........A--A-----T--...GG---A--T-----...------C.........T--...A-T--T--------T- 5806
21_A2D.KE.99.KER2003         -----------------------G-----T-G-GAG-A-T...-GT...CAA-AC...............TT-----...-A--------TT---.........--------A--A...-------C------...--------.........G--....-GC--TC---------- 5534
22_01A1.CM.01.01CM_0001BBY   -----------A-----------G--G--C-C-GAT-ACT...-G-...CAA-AC...............TT----...---------A--TT--.........A--------T--...G-A--A--T-----...------C..........C----T---------- 5539
23_BG.CU.03.CB118            -----------A-------------G--T-C-GA--A-T...-G-...CA-C-C...............TT-----...-C------G--TT--.........A--A-----T--...GG---A--T-----...------C.........T--...--T--C---------- 5782
24_BG.ES.08.X2456_2          -----------A-------C--G--G--T-C-AAT-A-T...-G-...CAACTC...............TT-----...-C------G--TT--.........A--A-----T--...GG---A--T-----...------C.........T--...--T--C--A-------- 5790
25_cpx.CM.02.1918LE          -----------------A-----G--G--T-C-GA--A-T...-AT...CCACTC...............TTA---...---------G-TTT--.........A-G------T--...G----A--T-----...------C.........T--...--C------------- 5547
26_AU.CD.02.02CD_MBTB047     -----------------------T--G--C-CGGA--A--...-GT...CA-CAC...............TT--T-...-A------G-----A.........A-------G-T...A-A---T-T-----...-A---.........G--...--C-----A-------- 6333
27_cpx.FR.04.04CD_FR_KZS     -----------------A-----T--G--T-C-GAG-A-T...-G-...CAAG-C...............TT----...----------CT----.........A--T-G--A---...A-A----T-----...------T-.........CA-...-G---C---------- 6329
28_BF.BR.99.BREPM12609       -----------------------G--G--TGC-GAG-A-T...-G-...AA-CAC...............TT-G--...-C-------CT-T--.........T--T-G--A--A...--A--A--T-----...-A---....................A-CC---CA-------- 5695
29_BF.BR.01.BREPM16704       -------------------C--------CCATGA--ACT...-AT...---CAC...............------...TCG-------TG---.........---------------...--A-------------...------....................-GC------------ 5727
31_BC.BR.04.04BR142          -------------------------G--C-C-GAG-A-T...-G-...AAACA-...............T-------...-T------A-T-T-A.........GG-T--T-------...--CC----T-A-...-A--TA.........-GG...-G---C---------C 5822
32_06A1.EE.01.EE0369         -----------A-----------G--G--C-C-GAT-A-T...-G-...-CACAC...............TTA---...-A------A--TT--.........A--------T--...G----A--T-----...-A---C.........T--...A-GG-CC---------- 5974
33_01B.ID.07.JKT189_C        -----------A-----------------C-C-GAG-A-T...-GC...-CA-AT...............TT-----...-A------A---T--.........A--------T--...G----A--T-----...------C.........T--...A-C--C--------T- 5634
34_01B.TH.99.OUR2478P        -----------A-----------------C-C-GAT-A-T...-G-...CTA-AC...............TT-----...-A------A--TT--.........A-------AT--...G----A--T-----...------C.........T--...--C--C--------T- 5535
35_AD.AF.07.169H             -----T-----------------T--G--C-C-GAC-A-T...-G-...CAA-AC...............TT-----...--------A--T--.........A--T-G-------...A-A--A--T-----...------.........GT-....-----T--------T- 5539
36_cpx.CM.00.00CMNYU830      A----T-----A-------C--G--GA-C-C-GAG-A-T...-GT...CTACAC...............TT-----...-A------G--TT--.........A--A-----T--...G----A--T-----...------C.........T--...--CTTG--------T- 5548
37_cpx.CM.00.00CMNYU926      -----------------------T--G--T-C-GAG-A-T...-AT...CCACCC...............TTA---...-AT-----G--TT--.........A--------T--...G----A--T-----...------A.........-AC...A-CCTC---------- 5530
38_BF1.UY.03.UY03_3389       -----T-----------------G--G--TGC-GAG-A-T...-G-...CA-CAC...............TT-GA-...-AG--CA--TTGT-A.........T----G--A--A...--A--A--T-----...-A---....................-----G---------- 5759
39_BF.BR.04.04BRRJ179        -----------------------G--G--TGC-GAG-A-T...-G-...CA-CAC...............TT-G--...--G------CTTT-A.........T--T-G-----A...--A--A---------...------.........G--...--C--T---------- 5816
40_BF.BR.05.05BRRJ055        -------------------------G--TCAGGA--A-T...-GT...CAACCC...............TTA---...-----------T----...............---------------...-A---C.........--G...A-C--T-G-------- 5842
42_BF.LU.06.luBF_18_06       -----------------------T--G--TCATGA--A-T...-AT...CA-CAC...............TT-----...-A--------T----...........----------A...-----------A--...------.........G--....---G----------- 5883
43_02G.SA.03.J11223          -----------A-----------------T-C-GAG-A-T...-G-...CAATAC...............TT-----...-C-AT-T-G-SYT--.........A--------T--...GK---A--T-----...------C.........T--...A-T--C---------- 5864
44_BF.CL.00.CH80             -----------------------G--G--TGC-GAG-A-T...-G-...CA-CAC...............TT-G--...-AG------CTTT-A.........T----G--A--A...-----A---------...-A---....................------------------ 5781
45_cpx.FR.04.04FR_AUK        -----------------------G-----C-C-GAG-A-T...-G-...CA-CAC...............TT-G--...-----------T-T-A.........T--T-G-----A...---C-A---------...------.........G--...--T--C---------- 6322
46_BF.BR.07.07BR_FPS625      -----------------------G--G--TGC-GAG-A-T...-G-...CA-CAC...............T--G--...-A-------CTTT-A.........T----G--A--A...--A-CA----------...---A--....................-----C--------- 5766
47_BF.ES.08.P1942            -------------------C--G-----TCA-GA--A-T...-AT...CA-CA-...............T-------...-A-G------T-T--.........----------A...--A------------...-A----GAA...AAGT-T.......-----G---------- 5777
48_01B.MY.07.07MYKT021       -----------A-----------------C-C-GGG-A-T...-G-...CTA-AC...............T-------...-A------G--TT--.........A--------T--...G--G-A--T-----...------C.........T--...--T--C--------T- 5538
49_cpx.GM.03.N26677          -----------------------T-----C--GGAT-ATT...-GC....AA-AT...............T-------...---G----TCT-T--.........A-TT-A--A---...C-A-----T-----...-----G-.........G--...-G---C---------- 5757
50_A1D.GB.10.12792           -----T-----------------G-----T-G-GAG-A-T...-AT...CA-CA-...............T-T---...-----------TG---.........---------------...-------CA-----...----TA.........G--...----C-G--A------- 5736
51_01B.SG.11.11SG_HM021      -------------------------G--TCAGGA--A-T...-G-...CA-CAC...............TT-----...---G-----T-T--.........T----G-----A...---------T------...-A-........G--...------------------ 5533
52_01B.MY.03.03MYKL018_1     -----------A-----------------C-C-GAT-A-T...-G-...CCACAC...............TT-----...-A------G--TT--.........A--------T--...G----A--T-----...------C.........T--...------C--------T- 5647
53_01B.MY.11.11FIR164        -----------A-----------------C-C-GAT-A-T...-G-...CCA-AC...............TT-----...-A------A--TT--.........A--------T--...G----M--T-----...------C.........T--...--C-----------T- 5670
54_01B.MY.09.09MYSB023       -----------A-------------G--C-C-GAT-ACT...-G-...CCA--C...............TT-----...-C------G-TTT--.........A-GA-----T--...G-A--A--T-----...------C.........T--...A-C--T--------T- 5735
55_01B.CN.10.HNCS102056      -----------A-----A-----G--T--C-C-GAT-AGT...-G-...CTC-AC...............TT-----...-T------G--TT--.........A--AC----T--...G----A--G-----...----A-C.........T--...--C--C--------T- 5643
56_cpx.FR.10.URF5_patient_A  -----------A-----------G--G--T-A-GAG-A-T...-G-...CCACAC...............TTA---...-C----T-G--TT--.........A-T------T--...G----A--T-----...----A-C.........T--...ACT--T---------- 5690
57_BC.CN.09.09YNLX19sg       -----------------------G--G--C--TGAG-A-T...-GT...CAACA-...............T-------...-T----A---T-T-A.........GG-T--T--------...--A----GT-----...-A--TG.........GG-...-G---C---------- 5516
58_01B.MY.09.09MYPR37        -----------A-----------------C-C-GAT-A-T...-G-...CCA-AC...............TT-----...-A------A--TT--.........A--------T--...G--GCA--T-----...------.........TT-...--T--C--------T- 5678
59_01B.CN.09.09LNA423        -------------------C--T--G--TCAGGA--A-T...-GT...CA-CAC...............TCA---...-A---------G--.........--------A--...------------T-----...---A--.........G--...---C------------- 5520
60_BC.IT.11.BAV499           ---A-------------------AG--TGC-GAG-A-T...-G-...ATA-C-.........................T-G-----T-T-T-A.........GG-T--T-------...--A-----G-A-...-A---....................------------------ 6293
61_BC.CN.10.JL100010         -----------------------G--G--T-CTGA--A-T...-AT...CCACA-...............T-------...-T------A-T-T-A.........GG-T-CT-------...--A-----TG---...-A--GG.........GG-...A--GGC---------- 5753
62_BC.CN.10.YNFL13           -----------------------T--G--CCAGGA--ACT...-GT...CA-CAC...............TTA---...-----------------.........--------A--A...--A--A----------...------.........G--...-----C---------- 5655
63_02A1.RU.10.10RU6637       -----T-----------------T--G--T-GTGA--A-T...-AT...CAACAC...............TTA---...--G-----A-TT---.........---T--T-------...A-A--A--G-----...-AG---.........G--...-----C---------- 5880
64_BC.CN.09.YNFL31           -----------------------T--G--T-CTGAG-A-T...-GT...CAACA-...............T-------...-TC-----A-T-T-A.........GG-T--T-------...C-A--A--T-----...-A--TG.........GG-...-G---C---------- 5686
65_cpx.CN.10.YNFL01          -----------A-----------G--G--C-C-GAT-A-T...-G-...CCA-AC...............TT-----...---------G--TT--.........A--------T--...G----A--T-----...----A-C.........T--...--C--C--------T- 5697
67_01B.CN.11.ANHUI_HF115     -----------A-------------G--C-C-GAT-AMT...-G-...CTA-AC...............TT-----...-A------G--TT--.........A--A-----T--...G----A-CT-----...------C.........T--...--C--C--------T- 5676
68_01B.CN.11.ANHUI_WH73      -----------A-------------G--C-CRGAT-A-T...-G-...CTA-AC...............TT-----...-A------G-TTT--.........A--A-----T--...G----A-CT-----...---A--C.........TT-GACA-C--C--------T- 5670
70_BF1.BR.10.10BR_PE004      -----------------------G--G--TGC-GAG-A-T...-G-...CA-CAC...............T--G--...-C----A--TT-T-A.........T----G-----A...---G-A--T-----...-A-A--.........G--...--T--C---------- 5785
71_BF1.BR.10.10BR_PE008      -----------------------T--G--TGC-GAG-A-T...-G-CAGCA-CA-...............TT-G--...-A-------CT-TG-.........T--T-GA---CA...--AT-A---------...-------C.........------...-----C---------- 5996
72_BF1.BR.10.10BR_MG002      -----------------A--C--T--G--TCA-GAG-A-T...-AT...CA-CAC...............T-------...-A-G------T-T--.........----------A...----------A--...------.........------...--C--G---------- 5837
O.BE.87.ANT70                --T--T----A-T----A--CA-T------GAGGA-CA-G.............-AG...............TTA---...-CC-T-TA-TTA...........GC-A-G-CTT--...A-A-CCCCA----...TTGAGC.........CTT...AG-C-GC-A-AT-CA- 6380
O.CM.98.98CMA104             --T-------A-T----A--CA-T-A----GA-GA-CA-G.............--G...............TTA---...-TC-TATA--TA...........G--A-G--TT--...A-A-CCCCA----...TTGAGC.........T-T...AGCC-G--T-AT-CA- 5821
O.CM.98.98CMABB141           --T--T----A-T----A---A-T-A----GA-GA-CA-G.............-AG...............TTAG--ATC-TC-T-T---TA...........--TT-G-CTT--...C-A--CCCA----...TTGAGA.........T-T...A-GG-CC---AT-CA- 5829
O.CM.98.98CMABB212           --T-------A-T--------R-T------GA-GA-CA-G.............-AG...............TYA---...-TC-TATC--TA...........GC-A-G-CTT--...C-A--CCC-----...TTGAG-.........TATAGT--GC-----AT-C-- 5831
O.CM.98.98CMU5337            --T-------A-T----A---A-T-A----GARGA-CAGG.............-AG...............TTAGA-...-TC-T-TA--TA...........G--A-G--TT--...A-A--CCCA----...TTGAGC.........T-T...A-T--GA---AT-CA- 5822
O.CM.99.99CMU4122            --T--T----A-T-------CA-T------GAGGA-CAGA.............CAG...............TTAG--...-TC-TATA--TA...........G--A-G-CTT--...A-A--CCCA----...TTGAGC.........TAT...--CC-GC---AT-CA- 5820
O.FR.92.VAU                  --T-----C-ATT----A--CA-T--G---GAGGA-CAGG.............-AG...............TTAG--...-TC---T--TTG...........A-TT-G-CTT--...A-A--CCCA----...TTGAGC.........TGT...A-CC--C-A-AT-C-- 5904
O.GA.11.11Gab6352            --T-------A-T----A--CA-T------GAGGA-CAGG.............-AG...............TTA---...-TC---T---TA...........G--A-G--TT-T...A-A-CCCCA----...TTGAGC.........T-TAACC--C-GC---AT-CA- 5555
O.SN.99.99SE_MP1299          --T-------A-T----A-----T---AC-GAGGA-CAGG.............-AG...............TTAG--...-TC-T-T---TA...........G-TA-G-CTT--...A-A-CCCCA----...TTGAGC.........TAT...A-CC-GCAC-AT-CA- 6391
O.US.10.LTNP                 --C--T----A-T----A--CA-T-----GGA-GA-CAGG.............--G...............TCA---...-TC-T-T---TA...........G--A-G-TTT--...A-A--CCCA----...TTGA-C.........T-T...A--C--C---AT-CA- 6295

N.CM.02.DJ00131              -----T-------G-----A---T--G--TGC-GATTGGT...-G-...CCATTC...............TTC--T...CTCAT-ATA-GTT--.........ACAA--..........................-GG.........T-T...A------AT----CA- 5805
N.CM.02.SJGddd               -----T-----------------T--G--T-C-GAGCT-T...-G-...-CACTC...............-TC--T...CTC-TA-TA-G-T--.........A--AAG-T-.......................-T--GG.........T-T...---C--CAT----CA- 5795
N.CM.04.04CM_1015_04         -----T-----------------T--G--TGC-GAGTGGT...-G-...CCATTC...............TTC--T...CTCCT-ATA-G-TC-.........A-AA--..........................-GG.........T-T...---C--CAT-----A- 5804
N.CM.06.U14296               ......----------G--A-A-T--GA-TGA-GAGTGGT...-G-...CCATTC...............TTC--T...CTCCT-ATA-G-T--.........A-AA----TAT...GAACAA.........-T-AGG.........T-T...---CC-CAT-----A- 5790
N.CM.06.U14842               -----T----A------------T-AG--GGC-GAGTGGG...-G-...CTACTC...............T-C--T...TTCCTACTA-G-T--.........ACAA--..........................-GG.........TTT...---CC-AAA-----A- 5806
N.FR.11.N1_FR_2011           -----T-----------------T--G--TGA-GAGTT-T...-G-...-CACTC...............TTC-AC...CTCCT-AYA-G-T--.........A--AAG-TA.......................-T--GG.........T-T...---C---AT-----A- 5705

P.CM.06.U14788               -----TGAG--GT----A---A-T-A----GA-GA--A--...............--C...............TTAG--...-AC-T--AA-T-...........TG-T-G-CTT--...A-C---TA--T-...-A----ATTAGCTAT--TAGT-G--C-CGT-AT--T- 5819
P.FR.09.RBF168               -----TG-G--GT-------CA-T-A----GA-GA--A--...............--C...............TTAG--...-AC-T---A-TT...........TG-T-G-CTT--...G-A--ATA--T-...-A----ATTAGCTGTG-TAGT-G--T-CAC-AT--T- 6383

CPZ.CD.06.BF1167             -GTAGT-C--AT-G--GA-TA----A---TGA-AAT-GC-...--A...-TA-TT...............--A---...-TCCT-ACA-TT--T.........T--TG-T-T-A-...C-AGGA.........................-TT...-G--C--AT-ACAC-- 6422
CPZ.CM.05.SIVcpzMT145        --G--------------A-----T--G--TGA-AT-TT-T...-G-TGGA---AC...............TCA--C...--G-CCAT--G---CAAGTTAATC--GA---ATG-...A-CGCATC------...TT--G-.........GA-...-----T-G-------- 5878
CPZ.GA.88.GAB1               -----T-----A-----A---A-T------GA-GA--AGAGAC-G-...AAC--C...............TTA-CC...-T-ATTACA-T---A.........ACAA-CATTT--...C-A-CCCCA----...TTGA-C...............TCTG-G--A-----A- 6397
CPZ.TZ.06.SIVcpzTAN13        ......-T--AT-G--GA-T-----A---TGA-GA-TTC-.........A-TG-C-TA...........ATT-T-...-CCCT-ATA----CA.........A----A-----.......................-T--GA.........TTT...AGT-C--AC-AT-CT- 5950
CPZ.US.85.US_Marilyn         -----T-----A-----A-----T------GA-GA--AGACTC-G-...CTA--T...............TAT--TCTTCTC-C-A--TTG--A.........A---CA--AT--.................................TCT-GC--A----CT- 6404

GOR.CM.07.SIVgor2139_287     -----TGAG--AT-------CA-T-AG---GA-G.........CA-...CT--AC...............TTA---...-A-----TA-TG...........AG-T-G--TT--...A-TT-A-CT--C...-T--AG..............-CT-CGCAT-AT-CT- 5867
GOR.CM.07.SIVgorCP2135con    -----TGAG--AT-------CA-T-AG---GA-G.........CA-...CT--AC...............TTA---...-A-G---TA-TG...........AG-T-G--TT--...A-TT-A-CT--C...-T--AG..............-CT-CGCAT-AT-CT- 5866
```

```
B.FR.83.HXB2                 CAGTCTATTATGGGGTACCTGTGTGGAAGGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTG...GTAAATGTGACAGAAAAT 6500
Env                          T__V__Y__Y__G__V__P__V__W__K__E__A__T__T__T__L__F__C__A__S__D__A__K__A__Y__D__T__E__V__H__N__V__W__A__T__H__A__C__V__P__T__D__P__N__P__Q__E__V__V__L__.__V__N__V__T__E__N_
A1.AU.03.PS1044_Day0         -T-----C--------------A-----A--T---GAG-----CT------------------G-----------C-----A--G--------C-----T------------------T------------------A--AAT---...RA------------G-G 5714
A1.CH.03.HIV_CH_BID_V3538    -G-----C--------G-----------A--C---GTG-----C-------------------------------------A--G--C------C-----T-----------------------------------------C---A--CCT--A...-CT------------G-G 5715
A1.CY.08.CY236               -T-----C--------------------A--T---GAG-----CT----C--------G-----A---------ACA-----A-CG-----A-C-----T------------------T--G--------------G---A--AAT---...-A------------G-G 5717
A1.ES.05.X1608_8             -T--T--C--------G-----------A--T---GAG----C-----C--------------G----------------AAAG--------C-----Y--G-------------------------------------------------A--CGT---...-AT------------G-G 5954
A1.KE.11.DEMA11KE001         -T---C-C--------------------A--T---GA-----TCT------------------------------A-----AAAG--------C-----T---------C--------------------------------GT------TAT---...-A-------------G-- 5855
A1.RU.11.11RU6950            -G--A--C-----A-------------GA--T---GAG-----C------------------------------A---A-----C-----C-----T----------------------------------TG----------A--A-G---...-A-------------- 6024
A1.RW.11.DEMA111RW002        -T-----C---------------------A-C---GAG-----CT-----------------------------A------G-------C-----T---------------------------------------------G---TAT--A...-A------------G-G 5857
A1.SN.01.DDI579              -G-----C-----------G--------A--C---GAG-----C----------------------------CA---A--G-------------T----C-----------------------------------------GT---AT---...-A------------G-G 5707
A1.UG.11.DEMA110UG001        -T-----C-------------------GA--T---GAG-----CT------------------A-----------------A--G--------C----------------------TG-------------------------A---AG---...AA---------------G 5847
A1.ZA.04.04ZASK162B1         -G-----------A--------------A--T---GAT------------------------------AG--A---A----------C-----T----------------------------------A-----------A-GAAT---...-A------A--------G 5938
A2.CM.01.01CM_1445MV         ----A--C-----A---------A----GA--T---GAG-----C------------------A-------------G---A-----------C-----------------------------T-----------------------ACC--...AA------A--------- 5702
A2.CY.94.94CY017_41          ----A--C-----A--------------A--T---GAT----TC-------------------------------------A--G--------A-----------------------------------------------A--AACC--...-A------A--------- 5860
B.BR.10.10BR_MG029           -------C--------------------A-------T----------------------------T----------A--G--C---C-------------------G--------------------------------------AG----...-A---------------C 5958
B.CA.07.502_1191_03          -------------A--------------A--G---AA-------G--C-----C-------------G-------------C---------------------C-----------------------G------------------A---...-A------------G-A 5944
B.CH.08.M2_0803101_NFLG8     ----A-----------------------A--------T-----C-------------------------------------------------------------------------------------------------------------------...-G---------------- 5956
B.CN.10.DEMB10CN002          -------------A--G-----------A-------T----C-----------------------------CC--T----C--------------------------G----------G---------T--T-----------G----AT---...-A------A------G-- 5867
B.ES.10.DEMB10ES002          ----TC----------------------A-----G-G-------T--------------------------C------------------------------------------------------T-G--------------C------...-A---------------- 5861
B.FR.11.DEMB11FR001          -------------A--------------A-------A-----------------------------------------------------------------------------------------------------G--------------...-G------------G-- 5733
B.GB.05.MM45d213_GN1         ----------------------------A---------------------------------G---------------C-------------------C--T------------------------------------------A-G-------...-A------------G-A 5897
B.HK.06.HK003                -----------------A---------A-------A--------------------------C-------------------C-------------------------------------------------------------------------...-G------------G-A 5869
B.HT.05.05HT_129389          ----------------------------A-------------------------------------------A---------------C------T--------------------------------------------------------A--A----...AA----------------- 5700
B.JP.12.DEMB12JP001          ----------------------------A---------------------T---------------------C------C--------A-------------------------------------------------------CGG--A...-A---------------C 5892
B.KR.07.HP_18_07JHS10_3909   ----------------------------A---------------------C---------------------C-------------------------------------T-----------------------------G-A----...----------------- 6157
B.PE.07.502_0525_wg5         -------C--------------------A---A--G----------------------------------CC-----------------C------------------------------------------------A--C-------...-A------------G-- 5973
B.RU.11.11RU21n              ----------------------------A--------T---------------------------AG--AG--A-C---------------T--T----------------G-------G------G---------A--CC-----...AAG---------------- 6058
B.TH.08.MERLBDTRC10          ----------------------------A--T---T------------------------------------------C----------------G-----------------------------T----------A---C-----...-G---------------G 5732
B.US.11.ES38                 ----------------------------A-------AT--------------------G----A-G-A-----C------CA------T----------------------------------G----------A--AA-----...-A---------------- 6431
C.AR.01.ARG4006              ----A-----------------------A-------AA--T---T----------------------G------------A--G--------C-----T-----------------------------------T--------T-G--T---...AA------A--------- 5688
C.BR.07.DEMC07BR003          ----A-----------------------A--T----AA--T---------------C-----C-----G--------C-----A--G--------C-------------------G-----------------T-----T---A-GA-T---...-A------A--------- 5850
C.BW.00.00BW5031_1           ----------------------A-------------AA--T---T----------------------------G-A---A--G--------C-----T---------------------------------G-----------A----T---...-A------A-------C- 5871
C.CN.10.YNFL19               ----------------G-----------A--T----AA--T-------C------------------G------G-A---A--G--------C-----T-------A---------------------------------------A-G--T---...-A------A--------- 5852
C.CY.09.CY260                ----------------------------A-------AGG-T---T---------------G----G---------G-G---A---------C-----T-----------------------------------------T-G--T---...-A------A--------- 5741
C.ES.08.X2363_2              ----------------------------A-------AA--T--G-----C-------------------W---A-C--------C-----T-----------------------------G----------C----T---...-AC----CA------T-- 5939
C.ET.02.02ET_288             -----------------A---------A--C----G-C-T---------------------------A--G--C-----C-----T--T-----------------------------------------------T-G--TA--...-A---------------- 5720
C.IN.09.T125_2139            ----------------------------A-------AG--T-------C-------------C------C---G----A--G--------C-----T------------------------------A--------A----T---...-G------A--------- 6472
C.KE.00.KER2010              -------C--------------------A-------AAG-TC------C--------------------G-A---A--G--------C-----T--------------------------------G---------A----T---...-A---------------- 5689
C.MW.09.703010256_CH256.w96  ----------------------------A-------AA--T-------C--------------------C-AG-A--G--------C-----T-----------------------------------------A--AA-----...-A------A--------- 6485
C.TZ.08.707010457_CH457.w8   ---------------------------GA-------AA--T---T---C--------------------G-G--A--G--------C-----T-------------------------------------G---A-GAGG--A...-A------A--------- 6502
C.YE.02.02YE511              ----T-----------------------A-------AA--T------------------------G-GC-A--G--------C-----T-------------------------------------------A----T---...-A------A------G- 5692
C.ZA.08.705010534_CH534.w12  -------C----------------------------AA--T---T-----------------G-----------GAG--A----------C-----T---------------------------------G----------A-GT-T---...-AG-----A------G-- 6493
C.ZA.10.DEMC10ZA001          ----------------------------A-------AA--T-----------------------G-GG-----------------T----------------------------------G-----------A---C-----...-G------A------G--- 5853
C.ZM.11.DEMC11ZM006          ----------------------------A-------AA--T-------C--------------------G-A---A--G--------C-----T-----------------------------------------A--TGG--A...-A------A--------- 5867
D.CM.10.DEMD10CM009          ----A-----------G-----A-----A-------AA-------------------------------AC-A-AG------C-------A-C-----T--------------------G-------GT--C------A--A-----...-A------C------CC- 5874
D.CY.06.CY163                ----G-----------G-----------A-------AAG----------------------------T-AC-A-AG------C-------A-C-----T--------------------G--------T--C------A--------GGG-A------C--------C 5710
D.KE.11.DEMD11KE003          ----G-----------------------A------------T--------------------T-----A-AGA---A-C-------A-C-----T------------------A-----------------------C----A...-A------T--------- 5845
D.KR.04.04KBH8               ----T-----------------------A------------------------------G-------CT--CT-A-AG------C-------A-C--------------T----------G--T----G------G--G------C--...-C------C------G--C 6452
D.SN.90.SE365                ----T--------------G--------A-------------------------------T-------A------AA------CA-C-----T--------------------G-------------------A--A-C--...-C------C--------C 6513
D.TZ.01.A280                 --A-T-----------------T----GA-------A-----------------------------T---T---C-------C-------A-C-----T-----------C-----A--------------------A-G-ATC--...----------TT-------C 5697
D.UG.10.DEMD10UG004          ----T-----------G-----------A-------AA--T-----------------------------CA-AC---A-C-------A-C-----T------------------A--------------------A--TAC--A...-A------C--------C 5834
D.UG.11.DEMD11UG003          ----T----------------------GA------------T-------C-----------------T------A-------G--------C------------A-------A--------------------A--CATC-A...-A------C--------C 5848
D.YE.02.02YE516              ----G----------------A-----A-------AAG----------------------------T--GT-A-A------C-------A-C-----TT--------------------G--------T--C------A--AGC--...AA------C--------C 5710
D.ZA.90.R1                   ----T-----------G-----A------------------T--------------------------G----A-A-A------C-------A-C-----T--------------------G-----------------------G-C--...-A---------------C 5845
F1.AO.06.AO_06_ANG32         ----A-----------------------A------------T--T-----------------------G-------A-G---A----------C-----T-------------------------------------G---A--T-GC--...AA------A--------- 5718
F1.AR.02.ARE933              ---------------------------GA------------T--T----C-------------------T-------A-G-------------C-----T-------------------------T---G-T----G---------CC-N...-A------A--------- 5800
F1.BR.10.10BR_RJ015          -G--T-----------G-----------A-------A--T-------C--------------------A-A-A-----C---------C-----T----------------------------------------GCA---TC--...-G------A---------A 5912
F1.CY.08.CY222               ----------------------------A------------T-----------------------G-----A-A-----G--C-----C-------------------C------------------------------G---A--CATC-A...-A---A-A------G-- 5696
F1.ES.02.ES_X845_4           ----T-----------------------------------------T--T----------A--------G-----A-G-------------C-----T-------------------------------G-------G---A--TGGC--...-A------A--------- 5958
F1.RO.96.BCI_R07             ----T----------------------------G-----T-----------C-----------------G-----A-A-------------C-----T----------------------------------T-----G---A--AT---...----------A--------- 6530
F1.RU.08.D88_845             ---------------------------GA--------A--T-------------------------G----------GA-----C---C------C-----T------------------------------------------T-TC--...-A------A------G-- 5965
F2.CM.02.02CM_0016BBY        -------------A-----A--------A------------T---------------G--------------------A-A-----C---------C-----T-----------------T-------------TG----------A--T-TC--...-AC-----A--------- 5689
F2.CM.10.DEMF210CM001        -------------A-----A-------GA-------A---AGT-----------------------G------A-G----------------C-----T-----------------T-------------------A-G--TC-A...-C------A-------G- 5760
F2.CM.10.DEMF210CM007        -------------A--------T-----A-----------TGT-----C-----------A-----------------A-A------G-------CC-----T---T-----------------T-------------------T----TC--...-A------A--------- 5845
G.BE.96.DRCBL                -------------A-----------G-A--T----ATG--C---------------T-----------C--AG---T--AAGT------A-C---------------T-------------------T----G---------A--AATA--...AG------A--------- 6460
G.CM.10.DEMG10CM008          ----T-----------G--------G-A--GA--GAT------------------T---------------AG---T--AAG-------------T--------------------------------G---------GC--T-TC--...-A------A--------- 5887
G.CN.08.GX_2084_08           -----------------A-----G-A-----GAT---C---------------T------G------AG---T--AAA---------C-----------------------------------------------------A--AAGA--...-AC-----A------CC- 5742
G.CU.99.Cu74                 --A---------------------G-A--C---GATG--C----------------T---------------AG---T--AAGC---------C-----T-------A-------------------------------GA--CCTC--...-A------A--------G 6110
G.ES.09.X2634_2              -------------------------G----T---GAT------------------T------G--C---AG---T--AAGC---------C-----T--------------------------------------------A-CCCTC--...----------AT-----G-A 5989
G.GH.03.03GH175G             -------------------------G-A--T---GAT---------C-------T--------T---AG-T-T--AAA---------C--------------------------------T---------A----TA--...AA------A--T---T-- 6546
G.KE.09.DEMG09KE001          -------------------------G-A--T---GAT----C-----------T-----------------AG---T--AAA------A-C-----T--------------------T-------T-----GT--CCCC-A...AC------A--------- 5882
G.NG.09.09NG_SC62            -----------------A-----G-A--T---GAG---C---G--C-------T------------C---AG---C--AAA---------C-----------------------------------------------GA--GTC--...-A------A------T-- 5694
G.ZA.01.TV546                ---------------C--------G-A--C---GATG--C-Y--------------T---------T----AG---T--AAG-------A-A-----T--------T--------Y-------------T-----GA--A-TC-T...AA------A--------- 5708
H.BE.93.VI991                ----A-----------G-----------A-------AG--T-------C------------------------C-----AAGG--------A-----T--------------------------------------------A-G--C----...-A------A-------CC 5878
H.CF.90.056                  ----A-----------------------A--G---AA---------C---------------G-------G----AAAG-------C-----T-------A-----------------------------------GA-G--CA--...-AG-----A-----G-GC 5840
H.GB.00.00GBAC4001           ----A-----------G-----------A-------AA--T-------C---------------G------AGC-----AAA---------A-----T-----------------------T---G---------GA-G--CC--...-A------A-----G--- 6008
J.CM.04.04CMU11421           -------C--------------------A--G---AG-----T--------------T---------T-----AG---T--AAGT--------C-----T-------T----------------T--A-----------GA-GA-TA-A...-C------A--------- 6011
J.SE.93.SE9280_7887          ----------------------------A--T---AG-----T------C-----T-----------------AG---T--AAA---------C-----T-------T----------------T--A-G---------GA-GAATC--...CC------A--------C 5819
J.SE.94.SE9173_7022          ---------------------A----GA--T----AG-----T------C-----T-----------------AG---T--AAA---------C-----T-------T----------------T--A-----------GA-GAGTC--...CC------A--------- 5820
K.CD.97.97ZR_EQTB11          ----------------------------A--------A----------C-------------------------A-------------C-----T--------------------G-----------------------T---...-A------A--------- 5695
K.CM.96.96CM_MP535           ----------------------------A--------AC---------C-----------------------G-AG-------------C-----T--C------------------G----------------------AGA--...-A------A--------C 5699
U.CA.01.TV749                ----T----------------------GA--G----AGG----T-------C----T-----C--T-----AG---T--AAA--Y-----C-----T-------T----G---------G------------------G---A---AKC-A...AA------A--------- 6525
U.CA.99.TV721                ----------------------------A--G----AG---------------------T--C---C--T-----AG---T--AAA---------C-------------T-------------------G-------G--GA---ACC-A...AA------A--------- 6527
U.CD.83.83CD003_Z3           ----------------------------A--T---GAA------------------------G--------G-A---ATCG--------C-----T---------------------------------------G---------GC----T-----...-G------A--------C 5988
U.CD.90.90CD121E12           ----------------------------A--T---GA---------------------------G------A-A---A-C---------C-----T---------------------------------------G---------T----T-----...--------A--------C 5689
U.CY.05.CY090                ----------------G-----------A-------AA--T-------C-------------------------AGCC------C---------C-----T------------------T--------T---------------------A----T--A...-A------A------T-- 5713
U.CY.08.CY223                -G-----C--------------------A--T----AG-----C---------------A------C----T---A-------AAA------C--C-----T--------------------------------------GA-GC-TC--...A--------A--------- 5721
U.ES.10.DEURF10DZ001         -------C-----A--G-----------G---T---GAG-----C-----C------------------G--------CCA---A------------C--------------T----------------------G----A--CCCC--...AA------A--------- 5777
U.GR.99.99GR303              ----------------------------A--T---GAG-----C---------C-----------C------C--AG-G---AAA----------C-T---T----------C-----------------G----------A-GA-T----...-G------A--------- 5778
U.NL.95.U_NL_95_H10986_D1    --A-----------------------C-A--T----AT------T--------------T-----------T-----AG--AT--AAG--------------T--------------------------------------T--CCCC-A...----------A--------- 6388
01_AE.AF.07.569M             ----------------T----------GA--T---GAT-----C--------------------C------C---G-----A--G--C-----C-------------------------------------G---------A--TACC--...-A------A--------- 5703
01_AE.CF.90.90CF11697        ----T-----------T----------GA--T---GAT-----C--------------------C--C---C-A--G-----A--G--C-----C-----T------T----------------------------------A--CACC--...-A------A--------- 6449
01_AE.CN.10.YNFL03           ----T-----------T----------GA--T----AT-----C--------------------C------C---G-----A--G--C-----C-------T-T-----------------------A-----------A--CATC--...----------A--------- 5878
01_AE.HK.04.HK001            ----T-----------C-----------A--T---GAT-----C--------------------C------C---G-----A--G--C-----C-------G-------------------------------------A--AA-C--...-G-G---A--------- 5870
01_AE.IR.10.10IR.THR48F      ----T-----------T----------GA--C---GAA-----C--------------------C------C-G--G-----A-C---C-----C----------------T----------------------------A--TAC---...-A------A--------- 5690
01_AE.JP.x.JRC77AE           ----T-----------T----------GA--T---GAT-----C--------------------C------C---G-----A--G--C-----C----------------------------------------------A--CACC--...-A------A--------- 6497
01_AE.TH.09.AA090a_WG11      ----T---------A--T-----------A--T---GAT-----CT---------------------------C--ATGGA---A-CG--C---A-C-------------------------------------------A--CAC---...-A------A--------- 5922
01_AE.TH.90.CM240            ----T-----------G----------GA--T---GAT-----C--------------------C--G---C---G-----A--G--C-----C-----------------------------------------------A--CACC--...-A------A--------- 6068
01_AE.US.05.306163_FL        ----T-----------T----------GA--T---GAA-----C--G------------------C------C---G-----A-----C-----C---------T----------------T--------T-----------------A-GCACC--...-A------A--------- 5691
01_AE.VN.98.98VNND15         ----T-----------T----------GA--T---GAG-----C--------------------C------C---TGC----A--G--C-----C-----------------------------------------------A--CCCC--...-A------A--------- 5755
```

```
B.FR.83.HXB2                  CAGTCTATTATGGGGTACCTGTGTGGAAGGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTG...GTAAATGTGACAGAAAAT 6500
Env                           T__V__Y__Y__G__V__P__V__W__K__E__A__T__T__T__L__F__C__A__S__D__A__K__A__Y__D__T__E__V__H__N__V__W__A__T__H__A__C__V__P__T__D__P__N__P__Q__E__V__V__L__.__V__N__V__T__E__N_
02_AG.CM.10.DE00210CM013      -G-----C------------------------A--T----AG-----C------------------------------------------A-----------A-----T----------------------------T-----------A---AT---...AA------A---------C 5870
02_AG.CY.09.CY256             -G-----C-----------------------GA--C---GAG-----C--C------------------C-G---------------A-C---------------T--C-----------------------T-C----------A--TAT---...-A------A--------- 5689
02_AG.ES.06.P1423             -------C-----------G-----------GA--C---GAG-----------------------------------C-------A-C---------------T----------------------------------------A-GAAT--A...AA------A--------- 5937
02_AG.GW.05.CC_0048           -G--A--C------------------------A--T----AG-----C------------------------------G----A-----------C-----T-----------------------T--T------------T--CCT---...-A------A------G-- 5859
02_AG.KR.12.12MHR9            -G-----C--------G----------GA--C---GAG-----C-----------------------------------G--A-----------C-----T-----------------------T--T------------A--CAGC--...-A------A------G-- 6162
02_AG.LR.x.POC44951           -G-----C-----------------------GA--C---GAG--A--C-----------C---A--------A-------------A-CT--------C-----T--------------------------T---------------C-T---...-A------A--------- 6495
02_AG.NG.09.09NG_SC61         -GA----C-----------------------GA-----GAG-----C-----------C-----------G-------TG-A---ATC---------C-----T-----------------------T--T------------A--TAT---...AA------A--------- 5695
02_AG.NG.x.IBNG               -G-----C---------------------ACG---GAG-----C---------------------------------------A-----------C-----T--------------------------T------------A--CAT---...-A------A--------G 6027
02_AG.SE.94.SE7812            -G-------------------------GA--C---GAG-----C---------------------------------------A-----------C-----T-----------C--------------T------------A--CAT---...-G------A------G-- 5875
02_AG.SN.98.98SE_MP1211       -G-----C---------------C---GA--T---GAG-----C--------------------C-----------G-----A-----------C-----T--------------------------T-G----------A-GCAT---...-A------A--------C 5708
03_AB.RU.97.KAL153_2          ----T-----------------------A----------------------------------------------AG--AG------------------------T----------------G-------G----------A--CC----...AA---------------- 5713
04_cpx.CY.94.94CY032_3        ----T----------------------GA--C---GAG-----C-----------------A--------------G-A---A---------A-C-----T-----------------------------------------CTC--...A-------A-----G--C 5870
05_DF.BE.x.VI1310             ----------------------A----GA-------AA--T--C--G----------------------G------C-A---------------C-----T--------T-----------------------------------TC--...-G------A--------- 5888
06_cpx.AU.96.BFP90            ----------------G----C----G-A--T---GAT----T---------------T---------------AG-C-T--TAAG--------C-----A--------T------------------------------A--TCTC--...AA------A--------- 6536
07_BC.CN.98.98CN009           ----------------------A-----A--G------------T-------------------------------------------------------T---------------------------------------A-GT-T---...-A------A--------- 5831
08_BC.CN.97.97CNGX_6F         ----------------------------A-------AA--T--------C-----G--------------------G-----A--G--------C-----T-----------------------------G----------A----TA--...-A------A--------- 5673
09_cpx.GH.96.96GH2911         ----------------------------A--G---GAG-----CT--------------------C-G-----CTC--A---AAA---------------T--------------G-----G------------------A---GTC--...-A------A--------- 5711
10_CD.TZ.96.96TZ_BF061        ----A-----------------------A---A-------T-----------------------------C--CA-AG------C-------A-C-----T-----------C-----T----------------G----A----T---...-A------A--------- 5897
11_cpx.CM.95.95CM_1816        ---------------------------GA--T---GAT---C----------------TA--------------AG---T--AAA---------C-----T--------------------------------------GA--CATC--...-A------A--------- 5869
12_BF.AR.99.ARMA159           ----------------------------A-----------T--------C------------------T-------A-GG--------------C-----T--C-------------------T--A--T---------------TC--...-A------A--------- 6501
13_cpx.CM.96.96CM_1849        ----T-----------T----------GA--T---GAG-----C-----------------------G-------TC-----AAA------C--C-----T-----------------------------------------A--AAT--A...-GC---------------C 5910
14_BG.ES.05.X1870             ----------------G-----------A-----------------T--C--------------------------------A---------------------------------------------------------------AAC---...-----------------C 5973
15_01B.TH.99.99TH_MU2079      ----------------------A-----A--C---T------------------------------------------------C--------G------T--------------------G-------G-----------A-A-----...-G----------------C 5896
16_A2D.KR.97.97KR004          ----A--C-----A-------------GA--T---GAG--T--C---------------------------------------A-C-------------------------------------------------------A--AACC--...-A------A--------- 5866
17_BF.AR.99.ARMA038           ----------------------------A-----------T-T------C-------------------G------G---------------------------------------------T-----T---------A-GCCGC--...-A------A--------- 5710
18_cpx.CU.99.CU76             ----A-----------G-----------A--T---GTA--------------------T---------T-----AG---T--AAGC--------C-----T---------------------------G-----------A--ACTC--...-G------A--------- 5799
19_cpx.CU.99.CU7              ----T-----------T----------GA--T---GAA-----C-----------------------------------------G--------C-----T-----------------------------G----------A--AAT---...-A---------------G-G 5722
20_BG.CU.99.Cu103             --------------------------G-A--C----AT---C----------------T---------------AG---T---AGC--------C-----T--------A--------T--------------------GA--CCTC--...-A------A--------- 5973
21_A2D.KE.99.KER2003          ----T-----------G-----A--------------A---T---------------------------------TCA-------AAG------A-C-------------------------G------------------A-G-A-C--...-G----------------C 5701
22_01A1.CM.01.01CM_0001BBY    ----T-----------C----------GA--G---GAG-----C--G--C--------------C------------C----A--G--------C-----T--------------------------T-G----------A---AG---...-CC-----A--------- 5706
23_BG.CU.03.CB118             --------------------------G-A--T---TAT---C----------------G---------------AG---T---AGC--------C-----T--------A--------T--------------------GA--CCTC--...RA------A--------- 5949
24_BG.ES.08.X2456_2           --A-----------------------G-A--T---GATG--C-C--------------T---------T-----AG---T---AGC--------C-----T--------------G-----------------------GA--TCTC--...CA------A--------- 5957
25_cpx.CM.02.1918LE           -------------A--------------A--T----AG--------G-----------T---------------A-A--T--A-CC------A-C---------------------------------------------A---ATA-A...-A------A--------- 5714
26_AU.CD.02.02CD_MBTB047      ---------------------------GA--T---GAA-----C-----C-----------------G--------------A--G--------C-----T--T-----------------------------------GA-GAACC--...ACT-----A------T-- 6500
27_cpx.FR.04.04CD_FR_KZS      ----A------------T---------GA------GAT-----C---------------------------C----G-----A-CG--------C-----T-------------------------------------GA--C-T---...-AC-----A--------- 6496
28_BF.BR.99.BREPMI2609        ----------------------------A--------------------------------------C---GT-A-A-------C-------A-----------------------------------T-T------G------------...AA------C--------- 5862
29_BF.BR.01.BREPM16704        ----A--C--------G-----------A------T-------C---------------------------GT---C-------C------------------------------------------------------G---T--------...-A---------------- 5894
31_BC.BR.04.04BR142           ----A----------------------GA--T----AA--T------------------------------------G-----A--G--------C-----T-----------------------T-----T---------A----T---...CA------A--------- 5989
32_06A1.EE.01.EE0369          ----T-----------G----C----G-A--T---GAT--T-T-T-------------T------------C--AG---T--AAGG--------C-----T--------T------------------------------A--CAT---...AA------A-T------- 6141
33_01B.ID.07.JKT189_C         ----T-----------T----------G---T---GAG-----C--------------------C------C----G-G---A--G--C-----C---------------------------------------------A-GTACA--...-A------A--------- 5801
34_01B.TH.99.OUR2478P         ----T-----------T----------GA--T---GAT-----C--------------------C------C----G-----A--G--C-----C---------T-----------------------------------A--CACC-A...-A------A--------- 5702
35_AD.AF.07.169H              -T--------------------------A------GAG-----CT-------------------------------G-----A--G--------C-----T-----------------------------------G----A--CAT---...CA-------------G-G 5706
36_cpx.CM.00.00CMNYU830       ----T-----------T----------GA--T----AT-----CT-------------------C-----------------A--G--------C-----T-------------------------T-------------A---AT---...--C-----A-----G--- 5715
37_cpx.CM.00.00CMNYU926       ----T--C--------T-----------A--T---GAT-----C----------------------C--------T-C----A--G--------C-----T---TTGC-TG..----------------.-------GA--A--AATG--...-GT----------A--G-- 5694
38_BF1.UY.03.UY03_3389        ----------------------------A------------------------------------------------G-----------------------------------------------T-----T--------G---T-TC--...AA------A------G-A 5926
39_BF.BR.04.04BRRJ179         ----------------------------A--------T--T--------C--------------------------A-GG--------------C-----T-----------------------------G------C----A-T-TC--...-A------A--------- 5983
40_BF.BR.05.05BRRJ055         ----------------------------A------------------------------------------------C-------------------------------------------------T-------------------------...-G---------------- 6009
42_BF.LU.06.luBF_18_06        ----T-----------------------A---------------T------------------------G------C--------------------------------------------------T-----------------AG----...------------------ 6050
43_02G.SA.03.J11223           -------------------A------G-A--T---RAY---C----------------TA--------------AG---T--AAA---------Y----------------------------------------------A-GACTC--...-A------A------CC- 6031
44_BF.CL.00.CH80              ----------------------------A-------------------------------------------------------------------------------------------------T-----T---------A---ATC--...CC------A--------- 5948
45_cpx.FR.04.04FR_AUK         -------C--------------------A--C--CCAG-----C------------------------------------T--AAAG--------C-----T---------------------------------------A---AT---...--C-----A-------GG 6489
46_BF.BR.07.07BR_FPS625       -G-------------------------GA-----------T--------C--------------------------C-A-----C---------C-----T-----------------------T-----T---------A----TA--...-G------A--------C 5933
47_BF.ES.08.P1942             ----------------------------A--------------------------------------G----------------C-------------------------------------------------------A-GT-----...AA---------------- 5944
48_01B.MY.07.07MYKT021        ----T-----------T----------GA--T---GAA---G-C--------------------C------C-A-TGGA---A--G--C-----C---------------------------------------------A--CACC--...-A------A--------- 5705
49_cpx.GM.03.N26677           ----------------------------A-------AG------T-------------T---------------AGC-AG--A-CC--------C-----T--------T-----------------AC------G--------AA-C--...-A------A--------- 5924
50_A1D.GB.10.12792            ----T-----------------------A-----------T--------------------------GT------A-AG------CT------A-C-----T-----------------A-----T---------------A---ATC-A...-A------C--------C 5903
51_01B.SG.11.11SG_HM021       -------------A--------------A--------T-----------------C----------G-------------------------------------------------------------------------T--CCCC-A...-A------A--------- 5700
52_01B.MY.03.03MYKL018_1      ----T-----------T-----------A--T---GAT-----C--------------------C------C----G-----A-CG--C-----C---------------------------------------------A--CACC--...-A------A--------C 5814
53_01B.MY.11.11FIR164         ----T-----------T--G-------GA--T---GAT-----C--------------------Y------C----G-----A-----C---A-C---------------------------------------------A--CACC--...-A------A--------- 5837
54_01B.MY.09.09MYSB023        ----G-----------C----------GA--T---GAG-----C--C-----------------C-----GC----G-----AA-G--C---A-C---------------------------------------------A--C-CC--...-A------A--------- 5902
55_01B.CN.10.HNCS102056       ----T-----------T----------GA--T---GAG---------------T----------C------C---TG-----AA-G--C-------------------------------T--------G------A--TATC--...-A------A--------- 5810
56_cpx.FR.10.URF5_patient_A   -G--------------------A-----A--T---GAG-----C--------C-------------G---------------------------C-----------------------A----------GT--------G---AA-C--...-A---------------- 5857
57_BC.CN.09.09YNLX19sg        ----G----------------------GA-------AA--T-----------------------G-----------G---------------A-C-----T-----------------------------------G----A--TGG---...AA---G--A--------- 5683
58_01B.MY.09.09MYPR37         ----T-----------T----------GA--T---GAT-----C-----------------C---------------GA---------------C-----------------------------------G-----G----------A--...-A---------------C 5845
59_01B.CN.09.09LNA423         ----A----------------------GA-------A-------T----C--------------------------G------AA---------C--------------------------------T--------------------------...-C---------------- 5687
60_BC.IT.11.BAV499            ---------------------------GA-------AA--T-----------------T----------------------A-----------C-----T-----C--A--------------------T---------C-GC-T---...-A------A--T------ 6460
61_BC.CN.10.JL100010          ----A-----------------A----GA--T----AA--T--------C---------------------------G-----A--G--------C-----T---------------------------------------A--A-T---...-A------A--------- 5920
62_BC.CN.10.YNFL13            ----T-----------------A-----A-------G---------G--------C---------------------------------------C-----T---------------------------G----------C-GT-T---...-A------A--------- 5822
63_02A1.RU.10.10RU6637        ----A--C-----------A--------A--T----AG-----CT----------------A--------CC----C-----AA-G------A-C-----T--------------------------T-------------AC-CAT---...-C------A--------C 6047
64_BC.CN.09.YNFL31            -G--------------G-----------A--T----AA--T--------C---------------------C--A-G-----A--G--------C-----T---------------------------------------A-G--T---...-A------A--------C 5853
65_cpx.CN.10.YNFL01           ----T--C-----A--T-----------A--T----GT-----C-----------------G----------------------------------C-----T-----------------------------------G----A--AAT---...-A------A--------- 5864
67_01B.CN.11.ANHUI_HF115      ----T-----------T----------GA--T---GAT-----C--------------------C------C----G-----A--G--C-----C-----T--------------------------------R------A--CCC---...--G-----A--------- 5843
68_01B.CN.11.ANHUI_WH73       --A-T-----------T----------GA--T---GAA-----C--------------------C-G----C----G-----A--G--C-----C--------------------------------------G------A--CCC---...--------A--------- 5837
70_BF1.BR.10.10BR_PE004       ---------------------------RA-------G---------------------R----------------MK-R-------------------------------------------------------------------------...-G------A--------- 5952
71_BF1.BR.10.10BR_PE008       ----------------------------A--------T---------------------------------------G------C-----------------------------------------------G------------GAA----...AA---------------- 6163
72_BF1.BR.10.10BR_MG002       ----------------------------A---------------------------------------------A-------------------------------------------------------------------A-G-AG---...AA---------------- 6004
O.BE.87.ANT70                 --------GC------G-----A---G-A--T-----AC-AGTA-----C-----T-----------CCT-ACAAGC--T--AAAG------A-------AT----A--------T--T----------CT---T-T---TATCC----...CAC-----------TG-C 6547
O.CM.98.98CMA104              --------GC---A--G-----A---G-AA-T----A-C-AGTAT----C-----T-----------CCT-ACAAGCC-T---CA-------A-------AG----A--------T------------------T-----T---TATCC----...CAC-----------T-GA 5988
O.CM.98.98CMABB141            --------GC------G-----A---G-A-------A-C-AGTAT-G--C-----C-----------CCT-ACAAGC--T---CAG------A-------AT----A--------T-----------------CT-----T---TATCC---A...CAC-----------T--A 5996
O.CM.98.98CMABB212            ---------C------------A---G-A--G----AAC-AGTA-----C-----T-----------CTT-ACAAGC--T---AAG------A-------AT----A-----C--T------------------T---A-T--GTATACG......AC-------------T-GC 5995
O.CM.98.98CMU5337             --------GC----------------G-A--T------C-A--A-----C--C--T--T----T---CCT-ACAAGC--T---CA-------A-------AT----A--------T-----------------CT-----T---TAYCC---A...CAC-----------T--C 5989
O.CM.99.99CMU4122             ---------C------------A---G-A--G----AAC-A--RT----C-----T-------T---CCT-ACAAGC--T---AAG------A-------AT----A--------T--------T----CT-----T---TATCC---A...AA------------T--A 5987
O.FR.92.VAU                   ---------C------------A---G-A--T----AAC-A--AT-G--C-----T-----------CTTGACAAGC--T--ACAG------A-------A-----A--------T-----------------GT---A-T---TAT-AGC-A...AA------------T--A 6071
O.GA.11.11Gab6352             --------GC---A--G-----A---G-A--T----G-C-AGTAT----------T-----------CCT-ACAAG-C-T---AAG--C----AA-----AG----AT-------T--------T--T-CT--------GTACCC---A...AA------------T--A 5722
O.SN.99.99SE_MP1299           --------GC------------A---G-A--G------C-AGTAT----C-----T-----------CCT-ACAAGC--T---CAG------A-------AT----A--------T----------------TCT---T-T---TATCCT--A...ACC--G--------T--- 6558
O.US.10.LTNP                  --A-----GC------G-----A---G-A--T---G-AC-AGTAT----------T-----------CCT-ACAAGCCAG--ACAG--------------AG----A--A-----T--T----------GT---A-T---TATAC----...CAC-----------T--A 6462
N.CM.02.DJO0131               ----G--C-----A-----A--A----GA--T-T-GAA--AGT---T--C--C--C--------------CC---AG--------CT--C--CA-C-----------A--A-----C-----T--T-----T------------CC-C-A...-AC---------T---CC- 5972
N.CM.02.SJGddd                ----G--C-----A-----A--A----GA--C---GAG--A-T---T--C--C--T--------------CC---AG--------CT--C--CA-C-----------A--A-----T--T--T--T----------------GC----A...CCC-----A--T------ 5962
N.CM.04.04CM_1015_04          ----G--C--------G--A--A----GA--CA--GAG--AGT---T--C-----T--------------CC---AG--------CT--C--CA-C-----------A--A-----T-----T--T-----T-----------GCC---A...A-C---------T---C-- 5971
N.CM.06.U14296                ----G--C-----A-----A--A----GA--C---GAG--A-TC--T--C--C--C--------------CC---AG--------CT--C--CA-C-----------A--A-----T-----T--T-----T-----------G-A---A...CCC-----A--T---T-- 5957
N.CM.06.U14842                ----G--C-----A-----A--A----GA--C---GAG--AGT---C--C--C--T--------------CC---AG--------CC-----CA-C-----------A--A-----C-----T--T-----T-----------GCC---A...A-------A--T---C-- 5973
N.FR.11.N1_FR_2011            ----G--C-----------A--A-----A--C---GAG--A-T---T--C--C--T---K----------CC---AG--------CC--C--CA-C-----------A--A-----C--T--T-----T--------------GC-C--A...CCC-----A--T---G-- 5872
P.CM.06.U14788                -T--A-----------T--A--T----GA--T--T-AA-TT--CT----C------G-------ATCGCT-AC-AGC-RW---CAG------A----------C--G--------G-----------T-GA---AT---G--CAGGA-A...-AT-----A-----GTC- 5986
P.FR.09.RBF168                -T--A-----------T--A--T----GAA-T--TGAAGTT--CT-G---------G-------ATCGCT-AC-AGC-A----CAG------A----------C--G--------G-----C-----T-CA---AT---GA-CCCG......A-------A-----GTC- 6547
CPZ.CD.06.BF1167              ----------------G--A--A-----A--G-----AC-A--AT-G--C-----T--CTCA--A--G-T-GCAAGC---C-ACC-------A--------------AAT-----G--TCTT--T--TC-G---T-T---A--CCTA-T...AAC-T-ACT--T---... 6586
CPZ.CM.05.SIVcpzMT145         -T--T-----------C--A--A----G---G----AA-----C--T--C-----C--C--------GAGT---AG------A-CT-----CA-C--------C--A--T--------T--T--T----CT--T-----G--C-C---A...CCT-----C--T---G-G 6045
CPZ.GA.88.GAB1                ----A--------A--------T---C-T--T--TGA-C-GGTA--C--------C-----C-----G---C--AG--------CT------A-------------G--A--------T-----T----GT--T--G------T-TC-T...CC------A-T----TCA 6564
CPZ.TZ.06.SIVcpzTAN13         ------T------A--------T-----A--T--T--AC-AC-A--G--------T--T-----AG-T-TTGC-AG----C---C---C--CA-C-----A------AAT-----T---CT---T--TC-T--TT-T-----GCCT--A...CC------AT-CAT-G-A 6117
CPZ.US.85.US_Marilyn          ----A-----------------A----GA--T-T-GAG--A--CT----C-----C--T-----A--G-----CA-GCAG----CC--C--CA--------------G--A-----C--T--T--------T------------CAT---...CC---------T-----G 6571
GOR.CM.07.SIVgor2139_287      -T--G--------A--G--A--C---CGA--G----ATGTA---T----C------G----------TTATGTGAGC-A---AC-A--C---A-------A--G--G--A-----G--T--C-----A-G-G---AC---GT-TCC----...-A------A-------CC 6034
GOR.CM.07.SIVgorCP2135con     -T--G--------A--G--A--C---C-A--G----ATGTA---T----C------G----------TTATGTGAGCCAG--AC-A--C---A-A-----A--G--G--A-----G--T--C-----A-G----AT---GT-TCC----...-A--R---A-----G-CC 6033
```

```
                                                                                                                                            _V1 loop start
B.FR.83.HXB2                TTTAACATGTGGAAAAATGACATGGTAGAACAGATGCATGAGGATATA.ATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACC.CCACTCTGTGTTAGTTTAAAGTGCACT...GATTTG.........AAGAATGATACTAATACC............ 6644
Env                         _F__N__M__W__K__N__D__M__V__E__Q__M__H__E__D__I_#_I__S__L__W__D__Q__S__L__K__P__C__V__K__L__T_#_P__L__C__V__S__L__K__C__T__.__D__L__.__.__.__K__N__D__T__N__T__.__.__.__.__._
A1.AU.03.PS1044_Day0        ------------------A----------G--------G--A------.------C-------C-----T-----------------G------.--T----------C------C--T-GC...C--GA-.........GTC-TCTT--A--GC---CTTAATAATAGC 5870
A1.CH.03.HIV_CH_BID_V3538   ------------------A----------G--A------ACA--C---.------C-------C-----------------------G---G--.--T--------A-C----G----T-A-...---GCC.........--C-GAACA-AC-TC-AT............ 5859
A1.CY.08.CY236              ---------------CT-A----------G---------ACA------.------C--------------------------------------.--T----------C----G-T--T-GA...C--AAT.........GTC-C-ATACA-----ATGTCACTAATAAT 5873
A1.ES.05.X1608_8            ------------------A----------G---------ACA------.------C----------------------------GG-G------.--T----------C------T--T------...--GGCC.........T-C-----C--CG--GT-............ 6098
A1.KE.11.DEMA11KE001        -----T------------AG---------G---------ACA------.------C-------C-----------------------G------.-------------C------T--T-GC...A--GCC.........--TGTCAG-GTC-G--ATGTCAGTGTCACT 6011
A1.RU.11.11RU6950           ---G--------------A-------------A-----AACA------.------C-C-----C-----T---------------C-G------.--T-----C----C----C-T--TG--...-GA-CC.........--C-GCAC--GCTC--A-AACAGTACTAGC 6180
A1.RW.11.DEMA111RW002       --C---------------A----------G--A------ACA------.------C-------C------T--------------C-G------.--T-----C----C------T--T---...A--AAC.........-TC-CC-TG-AC.................. 5995
A1.SN.01.DDI579             ------------------A-------G--G---------ACA--C---.------C-------C------T--------------C-G------.--T-----C----C-C----C--T--C...--CA-C.........CCC--GA-CTT--CC-AT............ 5851
A1.UG.11.DEMA110UG001       ------------------A----------G----------CA--C---.------C-------C--------------------G--G------.--T-----C----C-C-------T--A...--CC-A.........G-CCT-A--GGC-C-CAGAGTCTTAATGGC 6003
A1.ZA.04.04ZASK162B1        -------------------G------------A------ACA------.--------------C----------CA-----------G------.--T----------C------T--T---...A-CAGC.........--CTTCAC--GG----G-AACCTCACTGAT 6094
A2.CM.01.01CM_1445MV        ------------------A----------G--A--------A------.----A-C------------------T-------------C-----.--T-----C--C-C-C----T--T-GC...--AG-T.........--AG--A-C--C.................. 5840
A2.CY.94.94CY017_41         -----T------------A----------G--------A--A------.-----C---------------------------------------.--G-----C--C-T------T--T-GC...A--GCC.........--T-CCAGC--CC---G-AATAGC...... 6010
B.BR.10.10BR_MG029          -----TG-----------A----------C------------------.--------------C-----------------------------T.-------------C------T------...---C-A.........T-TGC-A---A--CCG-TAATGGCCAGAAT 6114
B.CA.07.502_1191_03         -------------------------C----------------------.-G--A----------------T-----------------------.-------------C------T------...A--GCA.........--TGT-AG-GAG-C---TTAT......... 6091
B.CH.08.M2_0803101_NFLG8    ------------------A-------G---------------------.--T-----------------T------------------------.-------------C-C----T------...---GAC.........TT---AA---A-----ATAGTAGTAGTAAT 6112
B.CN.10.DEMB10CN002         ------------------A-----------------------------.---------------G----------------------------A.-------------C------T------...----G-.........---CC-A---A--G---TACTAATAATAAT 6023
B.ES.10.DEMB10ES002         --------------G---A--------C----A--------T------.---------------G-------------------------G---.-------------C------T------...A-AG-A.........-GCTGCAC--A--G---TTCTTCTAATAAT 6017
B.FR.11.DEMB11FR001         ------------------A-------------A---------------.--------------C-------------T----------C-----.-----------C-CC-----T------...-GG---.........--C----C---C----T-............ 5877
B.GB.05.MM45d213_GN1        ---G--------------A-------G-----------G---------.-CA------------------------------------------.-------------C------T------...---GCA.........G-TGT-AC-----CC-ATAGTACT...... 6047
B.HK.06.HK003               ---G----A---------A----------G------------------.----A---------------------------------------A.-------------C------T---T--...A--AAT.........-TT-C-ATC----C--GTACCAATCGC... 6022
B.HT.05.05HT_129389         ------------------A-------G---------------------.-----CC--------G----------------------------T.-------------C------C--T---...A-GCCT.........GTT--CTGC-----GGAGGAGAATGGTACC 5856
B.JP.12.DEMB12JP001         -----T--A---G-----A----------------------A---G--.--------------A------T----A--------G---------.-------------C------C------...---GCT.........GGA---AC----G-AG-TGGGAATACT... 6045
B.KR.07.HP_18_07JHS10_3909  ------------------A-----------------------------.--------------------------------------------T.-------------C-C-G--T--T-G-...A-----.........-G------G-------TAAAGAATGCT... 6310
B.PE.07.502_0525_wg5        ------------------A-----------------------------.--T------------------------------------------.-------------C------T------...---AA-.........TT---A........................ 6105
B.RU.11.11RU21n             ---------G--------A----------C------------------.-------------------------------------GG------.-------------T------T------...A--G-T.........GCT-G-........................ 6190
B.TH.08.MERLBDTRC10         ------------G-----A----------T------------------.-------------------------------------G-------.-----------C-C------T------...---G--.........TG--GAA-----C--AT............. 5876
B.US.11.ES38                --------------------------------A--------A------.-------------------------------------GTC-----.-------------C------A------...A--GCC.........-CT-C-AGC------GT-ACT......... 6578
C.AR.01.ARG4006             ------------G----------------T---------C--------.--------------------------------------G--G---.-------------CC-----T-----G...---GCT.........--C-GAAC--T--GCRGAACTATTAGCAAT 5844
C.BR.07.DEMC07BR003         -----T--------------T-----G--T--------GC--------.-----------------G--------------------G--G---.-------------C-C----C-----G...A--GCT.........-CC-G-ACCC-ACTCTGTGTTACTCCAAAC 6006
C.BW.00.00BW5031_1          --------------------------G--T------------------.--T-----------------------------C-----G--G---.-----------C-CC-----C--T-G-...-C-GCC.........--T---........................ 6003
C.CN.10.YNFL19              --------------------------GA-T-----------A---G--.--------------------------------------G--G---.-----------C-C------C--T--A...A--G-T.........-CA----T--AC-G--ATAGTACCAACAAC 6008
C.CY.09.CY260               -----T--------------------G--T------------------.---------------G----------------------G--G---.-----------C-CC-----T--T-G-...----AT.........---GG-ACCCAC----ATACAACCAGCAAT 5897
C.ES.08.X2363_2             ------------------------S----T--.----T------C--CA-----------------------------M--------G--G---.-----------Y-C----C-C--T--A...A-C-ATTTAATACCACCA-C-......-T--ATAKTACCATAACA 6098
C.ET.02.02ET_288            -----T--------------------G------------C--------.---------------G----------------------G--G---.-------------C----C-C--T-G-TCTA--G-T.........-CC-G-...-A-TC--ATGTTACCAGTAAT 5876
C.IN.09.T125_2139           --------------------------G--T---------------G--.--------------------------------------G--G---.-----------C-C----G-A--TGGA...A--G-T.........-GTGT-AGC-A-GT--GTGTTAGCAATACT 6628
C.KE.00.KER2010             --------------------T-----G--T--A---------------.--------------------------------------G--G---.-----------C-C------T--T-GC...A--G-T.........-CC-G-A--G---CC-G-AATAGCACTAAT 5845
C.MW.09.703010256_CH256.w96 -----T--------------T-----G--T------------------.-----------------------------------G--G--G---.-----T-----C-C------C--T-G-...C--AAT.........-TT-CG-T--A-GG----ATGGGAAATGGT 6641
C.TZ.08.707010457_CH457.w8  --------------------------G--T------------------.--------------------------------------G--G---.-----------C-C-----GA--TGAG...A--GCT.........-CCT-CA--GA--CC-A-............ 6646
C.YE.02.02YE511             --------------------T-----G--T--------G---------.--A------------------T----------------G--G---.-----------C-C------C----A-...A--AAT.........-TT-CCA---A-GT----AATGAT...... 5842
C.ZA.08.705010534_CH534.w12 -----------------------------T------------------.--------------------------------------G--G---.-----------C-C----G-------A...A--GCT.........-CATTCA----................... 6631
C.ZA.10.DEMC10ZA001         ------------G-------------G--T------------------.--T---C-------------T-----A-----------G--G---.-----------C-C------A--TGGA...A-CG-T.........-GT-GAA-CGT--C--A-ACAAATAGTAGT 6009
C.ZM.11.DEMC11ZM006         --------------------------G--T------A-----------.---------------G----------------------G--G---.-----------C-C-----GC--T--C...A----T.........-CT-CCA-C-G--CC-TTATTACTATCAAC 6023
D.CM.10.DEMD10CM009         --------------------------G---------------------.--------------------------A------------------.-----------C-C------C------...--AACT.........C-CCTGA---GC-CCCAGTCGGTTAAC... 6027
D.CY.06.CY163               ---G----------------------G---------------------.---------------G----------A------------------.-----------C-C----G-C----T-...---G-T.........G---CCA-C-A---G--T............ 5854
D.KE.11.DEMD11KE003         ------------------A-------G---------------------.--------------C-----------A-----------G------.-----------C-C------C------...--A-G-.........---G--A---GC-CG-A-AGTAATAGCACA 6001
D.KR.04.04KBH8              --A-CATGT---------A-------G--C------------------.--------------G-----------A------------------.-----------C-C------C--T---...---A-C.........CG---A-G---TC-A-AAC........... 6599
D.SN.90.SE365               ------------------A-------G--G-----------A------.--------------------------A------------------.-----------C-C------C----GG...---A-T.........TCT-G---CG-C-C--GTAACACC...... 6663
D.TZ.01.A280                ------------------A-------G---------------------.-----C--------GG----------A------------------.-----------C-C----C-C---T--...--CGC-.........--C-CGAC--AC-G-GGGAATGGG...... 5847
D.UG.10.DEMD10UG004         ------------------A-------G---------------------.--T-----------------------A--------T--GC-----.---T-A-----C-C---G--C------...A-A-G-.........-C--G-A-C----C--GTAACAGTACC... 5987
D.UG.11.DEMD11UG003         ------------------A----------G------------------.--------------------T-----A------------------.-----------C-C------C------...--G-G-.........G--CCAA--G-C-C--ATGCCAATAATGCC 6004
D.YE.02.02YE516             --------------------------G--G------------------.------C-------------------A------------------.--G--------C-C-----GC--T---...---A-T.........-------T---AGTAGA-AACAAC...... 5860
D.ZA.90.R1                  -----T-------G-------------A-C------------------.--------------------------A------------------.--G--------C-C------C------...---GCC.........-TCTT-A-CT-C-CCT--TTCAGC...... 5995
F1.AO.06.AO_06_ANG32        -----T------C-----A-------------A------ACA------.---------------------------------------------.-------------C----C-C--T-AG...---GCC.........--C-CCAG-T-C-C-GAA............ 5862
F1.AR.02.ARE933             ---G-T------------A-------------A------ACA------.-----------------------------------G--G------.--------C-N--C-----NT--T---...---GCC.........-TC-GG........................ 5932
F1.BR.10.10BR_RJ015         ---G-T--------C-----------------A------ACA------.--T------------G-------------------G---------.-----------C-C------T--T---...A--GCT.........GCTCGG........................ 6044
F1.CY.08.CY222              -----T------------A----------G--A--------A------.--T------------------------------------------.--------C----CC-----C--T---...-C-AAA.........--CTCCACGGAA.................. 5834
F1.ES.02.ES_X845_4          ---G-T------------A-------------A------ACA------.G-------------------------------------G------.-------------C------C--T-G-...---GCC.........CGTG---CC-A-TC-CAA............ 6102
F1.RO.96.BCI_R07            -----T------------A-------------A--------A------.------C--------G----------------------G------.-------------C-C----T--T---...A--GCC.........--TGGCAC-CTGTC-G-TAACCAAAGCAGC 6686
F1.RU.08.D88_845            ---G-T--------C---AG---------------------A------.---------------------T----------------G------.-----T-------C------C--T---...A--GCT.........-CT-TCCCC-A-GGC-ATAATAAC...... 6115
F2.CM.02.02CM_0016BBY       ------------------A----------C-----------A------.--------------------------A-----------GA-----.-------------CC---C-C--T-G-...---G-T.........-T-T-ACGG------ATACC.......... 5836
F2.CM.10.DEMF210CM001       ------------G-----A--------A-C-----------A--C---.--------------------T---C----------G--G-----T.-------------C------C--T---...A--G-T.........-CT-C-AC--AC-----TACCTATGCTAAG 5916
F2.CM.10.DEMF210CM007       -----------------------------T------A----A------.-C---------------------------------G--GC-----.-------------C-----GC--T---...---GCT.........--TG--ACA-G-GC-GTTAACAGC...... 5995
G.BE.96.DRCBL               ------------------A-----------------------------.---------------G----------------------GC-----.--T----------C------C--T---...--GA-A.........--C---A---G--C-CGGAATATC...... 6610
G.CM.10.DEMG10CM008         -----TGCA---------A-----------------------------.-C-------------G----------------------GC-----.--T--T---A---C----C-C--T-A-...A--GGT.........--TTTAA-CTA--C-GATGCCAATGGGGAA 6043
G.CN.08.GX_2084_08          ------------------A-----------------------------.---------------G-G--------------------GC-----.--T----------CC-----C--T---...-C-G-C.........-CCTG--C--A-GCC---CACGGTACT... 5895
G.CU.99.Cu74                ------------------A-----------------------------.--T------------G----T-----------------GC-----.--T----------C------C--TGT-...A--G-A.........-GTTGCACC-AC-GC--TGGG......... 6257
G.ES.09.X2634_2             ------------------A-----------------------------.---------------G----------------------GC-----.--T----------C---G--C--T---...A--G-A.........-CC-CCAG--GC-C--GTACAAAC...... 6139
G.GH.03.03GH175G            ------------------A----------------------A------.------C--------G----------------------GC-----.--T----------C-C----C--T---...--CA-T.........GTCT--TC--A-TG---TAGAAAGCATCCC 6702
G.KE.09.DEMG09KE001         ------------------A-T---------------------------.---------------G----------------------GC-----.--T------A---C------C--T---...A--G-T.........-CC-GCACAGGA-T-GAAAATAATACAGAG 6038
G.NG.09.09NG_SC62           ------------------A-----------------------------.---------------G----------------------G------.--T----------CC-----C--TGT-...ACC--A.........--CTG-AC-GA-GCC--TTATATTGGGAAC 5850
G.ZA.01.TV546               -----T------------CCT---------------------------.-----YC--------G----------------------G------.--T----------C------A--T---...--C-AT.........--TGGGAC-GTG-C--A-AAAACC...... 5858
H.BE.93.VI991               -----T------GT-------------------------ACA------.-----------------G---T----------------G------.-------------C-C--G-C--T-G-...AG-G-A.........--TGCAACA-A-GTA--AAAAAGT...... 6028
H.CF.90.056                 -----T------G-----A-------G--G---------ACA--C---.---------------------T-G--A------------------.-------------C-C----C--T---...A--G-C.........-GA--CA----CTC--A-............ 5984
H.GB.00.00GBAC4001          -----T------G-----A----------C-----------A------.--T--------------G---T----------------G------.-------------C-C--G-A--TGG-...--AG-A.........GGA---A-G----GC-ATAGTAAT...... 6158
J.CM.04.04CMU11421          -----------------------------T-----------A------.------G--------G----TT----------------GA-G---.--T--------C-C------T--T-G-...C-CAAC.........-TC-G-ACC-A--G---AAACAGCAGGACC 6167
J.SE.93.SE9280_7887         -----------------------------C--------G--A------.------G--------G-----T----------------GA-----.--T--A-----C-C------T--T-G-...A-CA-C.........-CT-G-A-C-G------AACTAGT...... 5969
J.SE.94.SE9173_7022         -----------------------------C--------G--A------.------G--------G-----T----------------GA-----.--T--------C-C------T--T-GC...--CG-C.........--T-G-A-C-A--G---AGATAGTAAT... 5973
K.CD.97.97ZR_EQTB11         ------------------A----------G---------ACA------.---------------G----------A-----------G------.--T--------C-C-C---CT--T--C...A--G-A.........-CT---A-CCG--CC-ATGCAAATAAGAAC 5851
K.CM.96.96CM_MP535          ------------------A----------G---------ACA------.---------------G----------A--T------G-G------.--T--------C-C------T--T---...----AC.........--AGGAACC-A--GC---AATAAT...... 5849
U.CA.01.TV749               -----TG----------------------------------A------.---------------G----T-----A--C---------------.--T----------C------T--T---...---A-A.........-CT-G-AG--A-GCA--TATTAGTACT... 6678
U.CA.99.TV721               -----TGC---------------------C-----------A------.---------------G----T-----A-YT---------------.--T----------C-----CC--T--C...AG-G-A.........GCT---A---T--C---AACTAATGTAACT 6683
U.CD.83.83CD003_Z3          --------------G---A-A--------G------------------.-----------------------T--------------G------.TTT--T-----C-C------C--T-T-...---G-A.........------AG------C-ATAAC......... 6135
U.CD.90.90CD121E12          ------------------A-------G--G------------------.--T--------------------T--------------G---T-T.T-T--T-----C-C------C--T---...---G-A.........GG--C-A-C-AC-CAG-T............ 5833
U.CY.05.CY090               --------------------T--------G--A------A-----G--.--T------------------T-----------------------.-------------C----G-C--TGTC...A--A-A.........-CTGC-A---TA-C--GA............ 5857
U.CY.08.CY223               -----T--------------------G--T------------------.---------------G----T-----A--------G---C-----.--T--------C-C------T--T---...A--GCA.........-GAGCAA---AGGTA--AGTAAATGCAACA 5877
U.ES.10.DEURF10DZ001        -----TGCA---------A----------G-----------A--C---.--T---C-------C-----------------------G------.--T----------CG-----T--T-AA...CCC--C.........--C-GCA-C--C-GCGG-............ 5921
U.GR.99.99GR303             -----T------------A----------G---------ACA------.------C-------C-----------A-----------G------.--T-----C--C-C------T--T-AA...A-CCCC.........TCTTG--T-G--TTA-AT............ 5922
U.NL.95.U_NL_95_H10986_D1   ------------------A--------------G--------------.---------------G-------G--A--------T--GC-----.--T--------C-C------T--T---...--CG-A.........-CA---AGA-------GT............ 6532
01_AE.AF.07.569M            ------------------A----------G--------G------G--.--------------------T-----------------------T.--T----------C------T--T-GC...A--GCT.........--CTTAACC-AC-T--A-AACACGATCACT 5859
01_AE.CF.90.90CF11697       ---G--------------A----------G--------A------G--.--------------------T-----------------G-----T.--T--T--C--C-C----C-T--T--C...A-GGCT.........--ACTGA--GAC-C-TATAACGGAACAGCT 6605
01_AE.CN.10.YNFL03          ------------------A----------G--A-----G---------.---------------G----T-----------------G-----T.--T-----C----C-----CT--T-TC...A--GCT.........--TTTGACC-AGCTC--AGCCAATGCCTCT 6034
01_AE.HK.04.HK001           ------------R-----A-A----C---G--------G------G--.--------------------------------------G-----T.--T----------C------C--T--C...C-GGCT.........--TTGGA-AT------A-ACAACCCAGAAT 6026
01_AE.IR.10.10IR.THR48F     ------------------CC-----C---G--------A------G--.--------------------T-----------------G-----T.--T--------C-C------T--T--C...A-GGCT.........--CTTGAC--AATTGG--AATGCCTCTAAC 5846
01_AE.JP.x.JRC77AE          ------------------A----------G--------G------G--.--------------------T-----------------G-----T.--T-----C----C------T--T--C...--GGCC.........--CTGGACC-A-GTC-GTAGTACAAGTAAT 6653
01_AE.TH.09.AA090a_WG11     ------------------A----------G--A-----G------G--.-----------------G--T--------------T---------.--T-----C----C------T--T--C...A--GCT.........G-TTTGAGATA-GCCGATAATTCGACCAAT 6078
01_AE.TH.90.CM240           ------------------A----------G--------G------G--.--------------------T-----------------G-----T.--T-----C----C------T--T--C...A--GCT.........--TTTGACC-A-GGC-GTAGCAAA...... 6218
01_AE.US.05.306163_FL       ------------------A----------G--------G------G--.--------------------T-----------------G--G--T.--T-----C----C------T--T--C...-----C.........---GTCCC-..................... 5826
01_AE.VN.98.98VNND15        ------------------A----------G--------G------G--.--------------------T-----------------G-----T.--T-----C----C------T--T--C...AG-GCT.........-C-TTGCCC-A--G--ATGCTAATTTG... 5908
```

```
                                                                                                                                                          V1 loop start
B.FR.83.HXB2                  TTTAACATGTGGAAAAATGACATGGTAGAACAGATGCATGAGGATATA.ATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACC.CCACTCTGTGTTAGTTTAAAGTGCACT...GATTTG.........AAGAATGATACTAATACC............ 6644
Env                           _F__N__M__W__K__N__D__M__V__E__Q__M__H__E__D__I_#_I__S__L__W__D__Q__S__L__K__P__C__V__K__L__T_#_P__L__C__V__S__L__K__C__T__.__D__L__.__.__.__K__N__D__T__N__T__.__.__.__.__.
02_AG.CM.10.DE00210CM013      --------------------------G--A--------A------.--T---C-------C--------------------C-G-----.--T---------C------T--T-G-...--CA-T.........-TCTT-AGA-GC-GC-GTATCAATAACAGT 6026
02_AG.CY.09.CY256             ------------------A-------------A------A-A------.--T---C-------C-----T---C------------G--G---.--T---------C------C--T-GC.....--GCC..........--CGTCA-C-GC-CCTATATA......... 5836
02_AG.ES.06.P1423             ------------------A-----------------------------.--T-A-C-------C-----T----------------G-------.--T---------C-----GA--TCA-.........-A..........--C---AC--AC-TA--T............ 6081
02_AG.GW.05.CC_0048           ------------------A---------------------A-------.--T-----------C-----T----------------G-------.--T-----C----C-C--G----TT--...A-CG-C.........--C-GAAC--ACGTC-A-GGCACCATACAG 6015
02_AG.KR.12.12MHR9            ------------------A-T-------------------TA------.--T---C-------C-----------A----------G-------.--T-----C----C----G-T--TCA-...A-C--C.........--C-GCAGC-A--G--G-ACC......... 6309
02_AG.LR.x.POC44951           ------------------A----------T----------A-------.--T---C-------C------------------------------.--T-----C----C---GG-T--TCA-...A-CA-C.........--CTC-AGC....................... 6630
02_AG.NG.09.09NG_SC61         -----T------------A----------T----------A-------.--T---C-------CG-----------------------------.--T----------C----G----T-A-...A----C.........--T-G-AG--A---C-G-ACCTCTGCC... 5848
02_AG.NG.x.IBNG               ------------------A---------------------A-------.--T---C-------C----------------------G-------.--T-----C----C----G-T--TCA-...A----C.........--C--CAGCTAC-GC-A-AGT......... 6174
02_AG.SE.94.SE7812            ------------------AG---------G----------A-------.--T---C--------------------------C-G-------.--T----------C----C-T--TCAG......-AAC.........CTC-C-AGC-GC................. 6013
02_AG.SN.98.98SE_MP1211       ------------------A-------------A-------TA------.--T--CC-------------T----------------G-------.--T----------C----G----TCA-...A-C-AT.........--CT--ACC-GA--C-A-AGT......... 5855
03_AB.RU.97.KAL153_2          ---------G--------A-----------------------------.------------------------------------G-------.-------------C------T------.......------..........-----G--GGT--C--GT............ 5857
04_cpx.CY.94.94CY032_3        ------------------------------------------------.------------A--G--G------A------C---GC-----.T-T-----------C---T-CA--T-T-...A--GCA.........-CT-C-AC--A--G----AAT......... 6017
05_DF.BE.x.VI1310             -----T------------------------A------ACA------.--T-----------------T-----------------G-------.---------------C-C----C--T---.........-----C..........---GCCA--T-C-C-G--AATTCC...... 6038
06_cpx.AU.96.BFP90            -----T------------A----------C------------------.------------G---T-----------------GC-----.--T-----C----C-----CA--T----...A--GCT.........-CATTA-G--AC--A--ATTAGGTAACAAT 6692
07_BC.CN.98.98CN009           --------------------A------A-T----------A--G--.------------------------------------G--G---.-----------C-C----G----T-AA...A--G-T.........-GC-G-A---G----GGTACCTGC...... 5981
08_BC.CN.97.97CNGX_6F         --------------T-----T-----GA-T--------------G--.-C----------------------------------G--G---.-----------C-C----G-A--T-GA...A--G-T.........-GC-G-A--GG----GGTACCTAC...... 5823
09_cpx.GH.96.96GH2911         ------------------A----------G---------A-A--C---.------C-----------T----------------G-------.--T-----C--C-CG---G-T---CA-...AGAG-C.........--C--CAGC-AACCC-A-............ 5855
10_CD.TZ.96.96TZ_BF061        -----T-------------G------G--T------------------.-------------------G-------A------A----------.------------C-C------C----A-....-GCA-T.........--TGCCAC--A-G-C--T............ 6041
11_cpx.CM.95.95CM_1816        ------------G-----A---------------A-------------.-------------G---------A------A---G-------.--T---------C---G--C--T----...A--G-A.........-----GA-CTG--C--ATGTAGAGAAGAAT 6025
12_BF.AR.99.ARMA159           ---G-T------------A---------------A------ACA------.------------------------------G--G-------.-------------C------T--T-G-...A--GCT.........--TGCCAC-G-C----ATGCCACTGCCAAT 6657
13_cpx.CM.96.96CM_1849        ------------G---CA-T---------G---------ACA------.--T---A----------------------------GT------.--T-----C----CC-----T--T-A-...A-CA-C.........-CT-G-----GC-GC-T-............ 6054
14_BG.ES.05.X1870             ------------G-----A----------C---------G---------.--------------------------A------------C----.-------------C------A------.......------T..........--TGGGAC--A--C--ATAATACTGATGTT 6129
15_01B.TH.99.99TH_MU2079      ------------G-------T--------------G--A---------.------C-------------------------------------.------------------------T------.......-------..........-G-------------T-............ 6040
16_A2D.KR.97.97KR004          -------------------G--------------------A-------.------C-----------------------------------.--T---------C-C------T--T-GC...AGGG-C.........--A---ACC-TC-G--G-ACCCAGAGC... 6019
17_BF.AR.99.ARMA038           ---G-T--------C---A-------------A------ACA--G--.--------------------------------G---------.-------------C------T--T-A-.......---G-C.........--TGTCAC-G-C-CC-AT............ 5854
18_cpx.CU.99.CU76             -----T------------A----------------------A--C---.------------G---------A-------GC-----.--T-----------C------C--T----...A--G-A.........-MTTR-AG-GGA--C-A-ACTTRTGCTAGC 5955
19_cpx.CU.99.CU7              ------------------A-------------A------CA------.------C-------C----------------------G-------.--T-----C----C----G-T---CAC...................................... 5842
20_BG.CU.99.Cu103             ------------------A-----------------A-------C---.--T-----------G---T------------------GC------.--T----------C------C--T----...A----A.........TCAGTAA---GC-C-GT-............ 6117
21_A2D.KE.99.KER2003          ------------------A-------G-----------------------.--------------------------A--------------------.-----------C-C-------C----A-.......---CAA.........GT--GGA-C-GCG-AGGG............ 5845
22_01A1.CM.01.01CM_0001BBY    ------------------A-A--------G--------G---------.--T-----------------------------------G-----T.--T-----C----C----------TGAA.......---GCT.........-G-TTCA---T--CC-AT............ 5850
23_BG.CU.03.CB118             ------------------C-------------------------C---.--T-----------G---T----------------GC-----.--T-----------C-C----C--T----...A-GG-A.........G-CTG---G--C-GA-ATTGCACTACT... 6102
24_BG.ES.08.X2456_2           ------------------A----------G--------------C---.--T-----------G---T----A-------------GC-----.--T-----------C-----GC--T----...A--G-A.........--C-G-ACC-GA----ATCATGTAAACAGT 6113
25_cpx.CM.02.1918LE           --------------------T-----------------------------.-------------G---------A--------C-GC-----.--T--T-------C------C--T-G-...A--A-A.........--C-TCA---GC-CAGGAAAT......... 5861
26_AU.CD.02.02CD_MBTB047      ------------------A----------G---------CA-------.----C-------------T------------------G-------.--T---------C-C------C--T----......--CA-C.........-C--C-A---GA-C--G-............ 6644
27_cpx.FR.04.04CD_FR_KZS      ------------G---A-----C-------------------------.-------------G---T-------------------GG------.---T-T-------C-C----C--T----......---G-A.........--T-CAACA-GG-T-CATGACAGT...... 6646
28_BF.BR.99.BREPMI2609        ------------------A-------------A---------------.-------------G-G--A------------------------------.----------------C------T------...A--AAT.........GT--C--T--------TGAATGCTACTAAT 6018
29_BF.BR.01.BREPM16704        --------------------------------A---------------.-------------G--------------------------------------.----------------C----------T-...A--C--.........--TGC-AC--A-TTG-ATGATGCTAGCACC 6050
31_BC.BR.04.04BR142           ------------G-----------G--T--------C-----------.------------------------------------G--G---.-------------CA---C-C--T-G-......---GCT.........--C-GA......................... 6121
32_06A1.EE.01.EE0369          ------------------C---------------------------------.-AA-----------G-----------------------GC-----.--T-----------C---C-C--T-AG...A--GCA.........-CCTT-A-G-A-GCA---ATTAATGGTACT 6297
33_01B.ID.07.JKT189_C         ------------------A-A---C---T--------G-----G--.--------------------T---C-------------G-------.--T-----------C------T--TG-C...T--GCT.........--TTTGACC-AA-GC-ATAACACAGTCAAT 5957
34_01B.TH.99.OUR2478P         -----------------CA---------G--------G-----G-G.--------------------T------------------G-----T.--T-----C----C------T--T--C...A--GCT.........GCTTTGACC-A-GTA--TATA......... 5849
35_AD.AF.07.169H              -----------------CA-A-------G----------A------.------C-------C------------------------G-------.--T--T--C----C-----CT--T--C...A-CGCC.........--TGTYA-C--A----ATGACAGCAAC... 5859
36_cpx.CM.00.00CMNYU830       ------------------A----------G--------G---------.--T---------------------------------------G-------.--T-----C--A-C-C--G----TCAG...A-C--T.........--CGGCAGC-AC--CCA-............ 5859
37_cpx.CM.00.00CMNYU926       ------------------A-----------------GATA------.------C------GA-C--AGT-----------A---G------C--T-----C--C-C---T--A-AGTAG...CCA-AAGGTCA....-CCCCCA-G--C--CGGG............ 5844
38_BF1.UY.03.UY03_3389        ---G-T------------A-------------A------ACA------.------------------------------C-----G--G-----.-------------C-C----T--T-G-......---GCC.........-GTGCCCC--AA-CC--T............ 6070
39_BF.BR.04.04BRRJ179         -----T------------A----------G--A--------C------.--------------------------------------G-------.-------------C------T--T-----......---GC-.........GT--GA-CC------GG-............ 6127
40_BF.BR.05.05BRRJ055         ------------------A---------------------------------.------------------------------G--G-------.-----T-------CG---C-T--TG--....-CC--C.........--TGGCAC-CAA--C-G-ACCCAGGGGGGA 6165
42_BF.LU.06.luBF_18_06        ------------------A-------------A---------------.------------------------------G--G-------.-------------CC--------TG--......---GCC.........-CTCCCC---ACGCC-A-AGTAACACCACT 6206
43_02G.SA.03.J11223           --C---------------A-------------------T-----------.------------G---T-------------TGC-----.--T-----------CC----CC--T----...A--GCC.........-CTTG-A---AG-CC-AT............ 6175
44_BF.CL.00.CH80              -----T------------A-------------A------ACA--G--.G-----------------------A-------G--G-------.-------------C----G-A--T----...A--GCC.........-CTGCCACC-A-C-CT-TGAT......... 6095
45_cpx.FR.04.04FR_AUK         ------------------A----------G---------ACA------.------C------C-----T-------------C-G-------.--T---------C-C------T--T-A-......--CA-C.........-CA--CAGA-ACGGC.............. 6630
46_BF.BR.07.07BR_FPS625       ---G-T-----------CC-------------A------ACA------.------C-------------T--------------------G-------.-------------C-C--G----T-A-....-C-GCC.........--A-CCA--GA-................ 6071
47_BF.ES.08.P1942             -----------------CA---------T---------A-------.--T------------------------------------A.-----T-------CC-----T----A-...A--ACT.........GTC-CCAC--A-GC--G-ATG......... 6091
48_01B.MY.07.07MYKT021        ------------------AG---------G--------G--A--G--.------------------TT------------------G-------.--T-----C----C------T--T-G-.....--AAAA.........-TA-CAAT--G-GGA-ATGGCAAC...... 5855
49_cpx.GM.03.N26677           -----TG-A----------T--------G-------------------.--T--G-------G---T----------------------------.--T---------C------C--T----...A--G-T.........--A-CCA---ACTCC--AGCCAATAACACC 6080
50_A1D.GB.10.12792            --C---------------------G-----------------------.------C-------C---------------------------------.--T-----C----C----C-A---GC...A--A-C.........--T----CC--A--C-A-............ 6047
51_01B.SG.11.11SG_HM021       ------------------T-----C----------G-----G--.------------------T-----------------GC---T.--T-----C----C------T--T--C...AC-GCT.........--ATTC--G--A----G-ACAACC...... 5850
52_01B.MY.03.03MYKL018_1      ------------------A-A---C---G---------G-----G--.------------------T-----------------GC---T.--T-----------C------T--T--C...A--GCT.........-CTGTC-C-GACC-C-ATAAAAATAAA... 5967
53_01B.MY.11.11FIR164         ---R-T------------A-A---C---G--A-----A-----G--.---GA--------------T--T-----------G-------.--T-----------C-----CT--T--C...A--GC-..........--TCTAACC-A--TA-A-ACAACCATG... 5990
54_01B.MY.09.09MYSB023        ------------------A-A---C---C-------G-----G--.------------------T-----------------G----T.--T-----C----C---GC-T--TG--...TCGGCC.........--TGTCAC-GGC--G-A-ATAACC...... 6052
55_01B.CN.10.HNCS102056       -----G...---------A----------T-----------G--.-------------G---T-----------------G----T.--T-----C----C---------T--C...A--GCT.........---TTCA-C-A--CA--TGGAGGA...... 5957
56_cpx.FR.10.URF5_patient_A   ------------------A----------G--------A------.--T---C-T-----------------------------------------.-----------------C-C------T------...A--GCT.........--TGT-AC--G--T--GTAATAGT...... 6007
57_BC.CN.09.09YNLX19sg        ----------------------G-T-------G-----G--.G---------------------------------GGA-G---.-AT-------C-CC---G-A--T-AA...A--G-T.........-CCT-CTG--AA---GATACCACCAATATC 5839
58_01B.MY.09.09MYPR37         ------------------A-----------------A-----------.--------------G---T-----------------G----T.--T-----------C------T--T--C...A-GGCT.........--TTT--TC-AGG-C-ATACCACAACCAAT 6001
59_01B.CN.09.09LNA423         ------------G-------T-----G----------G-----------.--------------------G------------------G------------.----------------CG--G--T------...A---AT.........--T---ACC-AG---GTG............ 5831
60_BC.IT.11.BAV499            -----T------G-------------G--T--------C-----G--.----------------------------A-----------G--G---.-----------------CA------C----G-...A--G-T.........--C-GAAC-GAA----ATGCTACT...... 6610
61_BC.CN.10.JL100010          ----------------------GA-T-------G-----G--.--T-----------------------------------G--G---.-----------C-C------C--T--A......---G--.........-------........................ 6052
62_BC.CN.10.YNFL13            -----T----------------GA-G--A--------------G--.------------------G------------A-----------G--G---.-----------C-C------A--T-G-......---G-T.........-GC---....................... 5954
63_02A1.RU.10.10RU6637        ------------------A-T--------G------------A------.--T---C-G-----C----------------------G-------.--T-----------C----G-T--TTA-......-A-GC.........-GA-G-A--GTCTC--A-AGT......... 6194
64_BC.CN.09.YNFL31            ------------G-----------G--T--------------G--.------A-----------------A-----------G-----.-----------C-C----G-C--TCAA........---AT.........--TGG-ACAG--GG--ATTATACCAGCACG 6009
65_cpx.CN.10.YNFL01           ----------------------GA-T--------A--A--G--.--T-----------G--------------------G--G---.-----------C-C------A--TGGA...A--G-C.........-CTGTCAGA-AGT-C-ATGAGAGTGTAAAT 6020
67_01B.CN.11.ANHUI_HF115      ------------------AG----C---G--A-----G--A--G--.--------------------T---MM-----------R-----G.--T--T--C----C------T--TT-C...A--GCT.........M--TTGA--GRC-CA---AAT......... 5990
68_01B.CN.11.ANHUI_WH73       ------------------AG----C---G--A-----G--A--G--.--------------------T-----------------G-----G.--T--T--C----C----C-T--TT-C...A--GCT.........--ATTGA--GTC-CAG--AAC......... 5984
70_BF1.BR.10.10BR_PE004       ------------------A------------------------------.--T-----------------------A-------G-----------.-----------------C----C-T----A-...A--R--.........-CT--GTCGGG--C--ATAGC......... 6099
71_BF1.BR.10.10BR_PE008       ----------------------A-------------T------.--T-----------------TT-----------------------------.-----------------C------T------.......----G-.........GG---CAC-----GC-A-AAT......... 6310
72_BF1.BR.10.10BR_MG002       --------------------T-----------A------CA------.------C-----------------------------G--G------.-----------------C------T--T-G-...A--CAT.........GGC----CC---C-A-G-............ 6148
O.BE.87.ANT70                 -----T--A--------T-------------A-----G--A--C--T.--T----------C--G--T--T--A--T-----TC--A-G--T.TTC--G-----ACAAA-GG----T--A...A-CA-A.........GCTGGA....................... 6679
O.CM.98.98CMA104              ---G-T--A---G-----T-------G-----------A-----C--T.-CTGA---------G--G--TT----A--T-----G---A-G--T.TTCA-G-----ACAAA-GG-A--T-A-......--CA-A.........--T-CA........................ 6120
O.CM.98.98CMABB141            -----T--A--------T---------C--A-----G--A--C--T.--T--C--------A--G--TT----A--T-----TC--A-G--T.TTCA-G--A-ACAAA-G--T--T--A......--C--T.........--A-CA........................ 6128
O.CM.98.98CMABB212            -----T--A--------T-------G-C--A-----------A--C--T.--T----------------G--TT--------T-----GC--A-G--T.TTCT-A-----ACARA-G--C--T--A......---CC-..........--A-T-----G--CAGGGAATAATAGTACA 6151
O.CM.98.98CMU5337             --C--T--A-------CT-T-----------A-----G--A--C--T.--T----------A--G--TT--------T-----TC--A-G--T.TTC--G---A-ACAAA-G--T--T--C......--CGAA.........-CCG-A-GC--A.................. 6127
O.CM.99.99CMU4122             --C--T--A---G-----T-------------A-----G--A--C--T.--T----------A--G--TT----A--T-----TC--A-G--T.TTC--G-----ACAAA-G--T-----A...A-CG-A.........--CG-C........................ 6119
O.FR.92.VAU                   --C--T--A---G-----T-T---------C--A-----G--A--C--T.--AGA---G-----C--G--TT----A--T-----TC--A-G--T.TTCT-G-----ACAAA-G--T--T--A......---A-C.........--A---AG--T--C----............ 6215
O.GA.11.11Gab6352             ---G-T--A--------T-T-------------A-----------C--T.--TGA---------A-----TT----A--T---AC-C--C-GT-T.TTC--G-----AC-AA-G--T--T--A......--CCCA.........--T---A----C-GC-AA............ 5866
O.SN.99.99SE_MP1299           --C--T--A--------T---------------A-----G--A--C--T.--T----------A--G---T----A--T-----TC--A-G--T.TTC--G-----ACAAA-G--T--T--A...A-C-AT.........GTAC-A-GG-A-T----AAACAAT...... 6708
O.US.10.LTNP                  ---G-T--A--------T---------------A-----G--A-----T.--T--C--------A--G--T--T--A--T-----TC--A-G--T.TTCA-G-----ACAAA-G--T--T--A......--CCAT.........--T-CA........................ 6594

N.CM.02.DJO0131               -----T------G-----A-A---C------AG----A-----C--T.--T-----------A--G---T-------C-----T---------.---T-A-----A-C-A-G--T--T-GC...A--AGC.........--TGGAA--CGG-CA--AGATGAGAAAGAA 6128
N.CM.02.SJGddd                -----T------------A-----C------A-----A-----C--C.--------------A--G---T-------C-----TC--------.---T-A-----A-C-A-GTTT--T-AC...A--AGC.........--TGGA--GGAGCCAGTTGATGACCTCACC 6118
N.CM.04.04CM_1015_04          ---G-T------------A-----C------A-----A-----C--T.--------------A--G---T-------C-----TM--------.---T-A-----A-C-A-GC-T--TWAC...A--AGC.........--TGGGKT--A-GRG--AGTTGACAAT... 6124
N.CM.06.U14296                -----T------G-----A-A---C--C--A-----A--------T.G-------------A--G---T-------C-----T---------.---T-A-----A-C-A-G--T--T--A......---AGC.........--TGGGA--GTG-CA-AAAATGAGAAAAAG 6113
N.CM.06.U14842                -----T------------A-AG---C--C--A-----A-----C--T.--------------A-----T-------C-----T---------.---T-A-----A-CCA-GCTT--T--C...A--AGC.........--TGGAA--GTG-CAGAGGGTAGTGGAACT 6129
N.FR.11.N1_FR_2011            -----T------G----CA-T----C------A-----A-----C--T.--------------A--G---T-------C-----T---------.---T-A-----A-C-A-G--T--T--A......---AGC.........G-TGGG---GGG-CA-ATAGCACAATAAAT 6028

P.CM.06.U14788                -----T--T---G-C---T-T-----GAC---A-----G--A--C--C.--T--C-----------G-----T--A--T-----------G--A.GTT--A-------C-A-GR-T--T-GC......--C-GC.........-GC-CG-T-GACTG---TAATAACTCCTCC 6142
P.FR.09.RBF168                -----T--T---------T-T-----GAC---A-----G--A--C--C.--T--C-----------G-----T--A--T-----------G--A.ATC--A-----C-C-A-G--T--T-G-......--A-GC.........C-C-GG--GG---GCTGTAAT......... 6694

CPZ.CD.06.BF1167              -----T---GAA-CT---T-T---------G-----A-AA-A---T--.---TC-C-G-TTC-A--G--TT-G--A--T--------G--G--A.--CT-T-----A-AAA-G--C----G-......--A--C.........-CA-GAA---G--CA--AGGAAATAATGCA 6742
CPZ.CM.05.SIVcpzMT145         ------------G-----T-----------T--A-----G--A--C---.--A---C-C-----A-----T--T--A-----------------A.--CT-A-----G-CCC-T-CT--T-A-...A--CCC.........-CT---AC-T-CTGC--T............ 6189
CPZ.GA.88.GAB1                ------------------A-T-----G--C--A-----------A-C--T.-----C-C--------G-T--C-----T----------G--G--.--TT-A-----A-CC--GC------G-...A-GGCT.........--CTT-AGCCAGGCA-AAAACCTAACAAAC 6720
CPZ.TZ.06.SIVcpzTAN13         ---G-T----CT-CT---T-T---------G-----A-AACA---T-G.---TCAC-T-TTC-G-----TT-T--A------------------A.--CT-T---A-A--AA-G-CA--TG-A...A--G-AACAGAA...GCA-CAACA--A-CAC-ACCAACAACG... 6276
CPZ.US.85.US_Marilyn          ---G---------G-----A-T----C--------A-----G--A-----T.--T--CC--------C--G---T----A--T---A----G-------.---T-G-------C-A-G-CT--TCT-...A-CCCC.........G-T-G-A---G--G-G-TGTA......... 6718

GOR.CM.07.SIVgor2139_287      -----T--C---------T----------T--A-----G--T-------.G-T--C-----------G---T-------T---------C----A.GTTA-G-----C-C-C-G--T--T--A...C-AC-A.........--T-CAACA-A--CA--A............ 6178
GOR.CM.07.SIVgorCP2135con     ---G-T--C-----C---A----------T--A-----G--T-------.--T--C-----------G---T-------T---------C----A.GTTA-G-----C-C---G--T--T--A...A-GACA.........-CAG-GACG...-CA--RCAGACAAATRYA 6186
```

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

```
                                                                                                                                                                  V1 loop end˩    ˫V2 loop start
B.FR.83.HXB2                  ...................................................AATAGTAGT......AGC......................GGGAGA........ATGATAATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAG.. 6728
Env                           _.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__._N__S__S__.__._S__.__.__.__.__.__.__.__._G__R__.__.__.__M__I__M__E__K__G__E__I__K__N__C__S__F__N__I__S__T__S__I__R__G__K__.
02_AG.CM.10.DE00210CM013      ACCAGCAGTACCCCTACC.................................-C--CC-T-......---..........AACTACACCAGA-ATG-G........CAAGA-GAAAT----A----A--GCG------T---------GCG-C-----TGC------A----.. 6140
02_AG.CY.09.CY256             ..................................................................-AT......................--C-CT........-GC-CC-AA-GT--TAT---A--G-------------------G-C----GAAC------A----.. 5911
02_AG.ES.06.P1423             ...........................................................................................TT........GAAGGTGACAT-GGGAAG--A---------------A---GG-G-C----GAAC------A----.. 6150
02_AG.GW.05.CC_0048           .........................................................................................................GAA-AT-GCACCCT-CAT--A-----------------T-----G-C----GAAC---T--A----.. 6081
02_AG.KR.12.12MHR9            ...................................................TT--TC-A-......G--......................A-C-T-........TCT-GTGACAT-CGG-AG--A------------A----------G-C----GAAG------AC---.. 6393
02_AG.LR.x.POC44951           .........................................................................................................C-T-ATGACAC-C---AG--A----G----------------G-G-C----GAAT------A----.. 6696
02_AG.NG.09.09NG_SC61         ..................................................................-A-......................AAC--C........--ATCTGAT-GC-TGCA---A-------------G--------G-C----GTAC------A----.. 5923
02_AG.NG.x.IBNG               ..................................................................---......................AACCT-........-CT-GTGACAT---T--G--A-----------------------A-C----GAAG------A----.. 6249
02_AG.SE.94.SE7812            ..................................................................G--......................AAC-T-........TC-GA--ACAT-C-------A-----------------------G-C----GAAC------A----.. 6088
02_AG.SN.98.98SE_MP1211       .........................................................................................................-A---T-AT--AGTGCA---A--G-------------------G-C----GTAT---A--A---A.. 5921
03_AB.RU.97.KAL153_2          ...................................................-CC-A--C-.......--TAGC..........................................--AA-A-TGATG--A--G-------------------C----C----GA-C-G----A---A.. 5932
04_cpx.CY.94.94CY032_3        .........................................................................................................GGC-CTG--ATT----A--GA------------------G----A-C----GAA-----G-A----.. 6083
05_DF.BE.x.VI1310             ...................................................-CC-CC-AC......TC-......................ACT-CC........C---A-GAA----C---G-CAG--C-------T-----------G-CT---GAAG---AT-AC--A.. 6122
06_cpx.AU.96.BFP90            AGT.........ACTAATAGTACATTA........................GG--AC-A-......--T......................ACT-T-........G-AGATGATATC-GCAAG--A-----------------------A-C----GAA-------A----.. 6794
07_BC.CN.98.98CN009           ..................................................................-AT......................-A--CC........TGCCGTGA-AGC-TGAAG--A--G-----T-----------GCA-C----GTAG------A----.. 6056
08_BC.CN.97.97CNGX_6F         ..................................................................-AT......................-A--CC........TAC-ATGA-AGCGTGAAG-----------T-----------GCA-C----TTAT------A--G-.. 5898
09_cpx.GH.96.96GH2911         ...........................................................................................ACC-CG........-GTGACGG-AGTG---T-AGA--GC-------------------A-C----GAAT------A---A.. 5927
10_CD.TZ.96.96TZ_BF061        ...............................................................................................................G-T-GTG-CCCCAGA--G-C-----------------CA-C----GAAG---A--A----.. 6101
11_cpx.CM.95.95CM_1816        AATACTACTTTAGGGAATAGC..............................--CT-C-C-......---......................AAT--C........-CTG-GGA--GCC---C---A--G-----T-----------------A-C----GAA-------A----.. 6130
12_BF.AR.99.ARMA159           ...................................................GGC-C-CAA......-A-......................AAC-CC........C--GAGGAA---CC---G-CA---C-------T-----------G-C----GAA-------A---A.. 6741
13_cpx.CM.96.96CM_1849        ..................................................................-AT......................A-C--C........-CC---CA--G--G--A---A-----------------AT-----G-C----G-AG--G-C-A--G-.. 6129
14_BG.ES.05.X1870             AAT.........AAGACTAAT..............................-C--A--A-......GAT......................ATA-T-........GGAGGG-------G------A------------------------------GC-C------A---A.. 6225
15_01B.TH.99.99TH_MU2079      .............................................................................................-AT........-G--C---A--AGG------A-----------T-----------A-C---------------AC----.. 6109
16_A2D.KR.97.97KR004          ...............................................................................................................CCC--T-GCAATA-C-CC-TG-----------TG-G-CT-C----GAAC------A----.. 6079
17_BF.AR.99.ARMA038           ...................................................-G--C-CAA......-A-......................A-C-CC........C---AGGAA---GC---G-CA---C-------T-----------G-C----GAAG----G-A----.. 5938
18_cpx.CU.99.CU76             AATA.....................................................................................................G-AGC-TAA-CT-C--A---A--G---------------TT---A-C----GAAG------AW---.. 6025
19_cpx.CU.99.CU7              .............................................................................................................................--G-------------------G-C----GAAC------AC---.. 5887
20_BG.CU.99.Cu103             ...............................................................................................................-C-TCT-GT-A---C-----C-T---------T----A-C----GAG----AG-A--GA.. 6177
21_A2D.KE.99.KER2003          ..................................................................--T......................ACCCA-........-AC--T-ACC---C--TG--T--G-------------------A-C----GTGG------A----.. 5920
22_01A1.CM.01.01CM_0001BBY    ...........................................................................................A-TTCT........G-TGA-GAATTA---ACG--AG-G-C------T-----T-----G-C----GAAC------A---A.. 5922
23_BG.CU.03.CB118             ..................................................................-CT......................AATG-C........-CTG---ACA-CTCT-C---T-------------------------A-C----GAG-----G-A---A.. 6177
24_BG.ES.08.X2456_2           ..............GCCGGA...............................G--G-M-C-......GTA......................A-A-TT........GAC-GC-AT-T--C--AT--T--G-CWW-----------M---G-C---RGA-T--CAG-A---A.. 6203
25_cpx.CM.02.1918LE           ...................................................-CA-CA-AC......-A-......................AAT--C........-CTG---ATA-C-G--A---A-----------T---------G-A-C----GAAT------A----.. 5945
26_AU.CD.02.02CD_MBTB047      ...........................................................................................A-A-AC........CAT-CTGACATA--CA-G--A--G-G------------------G-C----GTAC------A----.. 6716
27_cpx.FR.04.04CD_FR_KZS      ..................................................................-C-......................AAT-CC........-GT--CCACAC--C--AG--AC-------T--T---------G-A-CT---GCA----A--A----.. 6721
28_BF.BR.99.BREPM12609        ...................................................-CC-C-G--......-AT......................A-T--T........G--GA-GC-AT-G-G-----A--G------------------G---C---------G---A----.. 6102
29_BF.BR.01.BREPM16704        CAA.........GCTCCTACT..............................-CC-AG---......-AT......................A-T--T........TG-----AT---G-G--------G------------------G-T-C-------G--A--A----.. 6146
31_BC.BR.04.04BR142           ...........................................................................................ACACAC........-ATGGT-CCATAGTG-----A--------T--TA----------G-C----GAAC------A----.. 6193
32_06A1.EE.01.EE0369          AGT................................................G-GG-A-CC......-AG......................AAT-TC........-CTG-G-ATA-C----A---A--G-C------------------G-A-C----GAA---GA--A----.. 6384
33_01B.ID.07.JKT189_C         GTC................................................TC----GAA......-TA......................-CA-AT........T-AGG--ATATA-C--AG--AG---G------T-----T-----A-C----GAAC-G----A----.. 6044
34_01B.TH.99.OUR2478P         ...........................................................................................ACC-AT........GGCCC--ATATA-C--A---A---G------T-----T-----G-C----GAA-------A----.. 5921
35_AD.AF.07.169H              ...................................................-CA-A--A-......GA-......................A-C-A-........--TGAGGACAT-GT------A-------------------------G-C----GAAC----G-A---A.. 5943
36_cpx.CM.00.00CMNYU830       ...........................................................................................AACGTG........TCT-AGGACAT-C-G--G--A-----------------------CA-C----GT-C------A----.. 5931
37_cpx.CM.00.00CMNYU926       ...............................................................AGC-AG......................A-CTT-........GATGG-GGA---G--AT---A-----------------A------G-C----GAAC------AC---.. 5922
38_BF1.UY.03.UY03_3389        ..................................................................GAT......................A-C-CCCTG..........GAA--------G-CA-----------TA----------G-C----GAA-------A----.. 6142
39_BF.BR.04.04BRRJ179         ......................................................-C-CA-......-A-......................A-C-CC........T---AGGAA--TCC-----CA---C-------T------TT---G-C----GAAG------A----.. 6208
40_BF.BR.05.05BRRJ055         ...................................................G--CAA-AC......---......................ACTGTT........C---AGGAA---CC-----CA---C-------T-----------G-C----GAAG------A----.. 6249
42_BF.LU.06.luBF_18_06        CTC...AATATCACCAGCAGT..............................-C--C----......-CT......................CCTG-G........-A-C-C-ACAGT----A---A---------------------G-T-C-TT-G-AG---A-A-----.. 6308
43_02G.SA.03.J11223           ..............................................................................................GTT........-AATGTGAA--TGTG---A-CT---------------------G-A-C----GAAT------A----.. 6244
44_BF.CL.00.CH80              ...................................................GCC-C-GCC......GC-......................A-C-CC........C---A-GAA---CC---G-CC---C-------T-----------G-C----GAAG------A--G-.. 6179
45_cpx.FR.04.04FR_AUK         ...............................................................................................................-GT-CTGTT-A------G-G-----------------G-CT---GAAC------A---A.. 6690
46_BF.BR.07.07BR_FPS625       ......................................................-C-CAA......-AA......................--C-TC........C---GGGAA---C----G-CA---C-------T-----------G-C----GAAG------A---A.. 6152
47_BF.ES.08.P1942             ...............................................................................................................-ATA-T-GT-----A--G-----T-----------------C---CCAA-CG-----G..... 6148
48_01B.MY.07.07MYKT021        ...................................................-TAGAC-CA......-CA......................--C-A-........--AGG--ATAT--C--AT--AG---G------T-----T-----G-C----GT-C------A----.. 5939
49_cpx.GM.03.N26677           AAT................................................G-C-CC-CA...CCC-CTAGC...................CCA-C-........GCC-CT-GCTCTCCTAT------G--------T-----------A-CT---......G---A----.. 6167
50_A1D.GB.10.12792            ...........................................................................................AAC-C-........GAAT--GA---C--G-A--GA-----------------AT-----A-C----GATC----G-A----.. 6119
51_01B.SG.11.11SG_HM021       ...................................................T----C-A-......GT-......................TCTG-C........--AGG--ATATA-C--AT--AG---G-------A---T----CG-C----GAA-------A----.. 5934
52_01B.MY.03.03MYKL018_1      ...................................................-C--T--CC......TCT......................AAA-C-........--AGG--ATCTA-C--AT--AG---G------T-----T-G---G-C----GAAC--CAG-A--G-.. 6051
53_01B.MY.11.11FIR164         ...................................................----AC-CA......-CT......................-CT--C........--AGG--ATATA-C--AT--AG----------TA---T-----G-C----GCAT---A--A----.. 6074
54_01B.MY.09.09MYSB023        ...................................................-C-GAC-CA......-C-......................AAAGTC........--AGG--ATATA-C--AT--AG-------T--TA---T-----G-C----GTGC--CA--A----.. 6136
55_01B.CN.10.HNCS102056       ...........................................................................................CCTGAC........-CACC--ATATA-C--AT--AG-G-G------TA---T--C--G-C----GAG-------A----.. 6029
56_cpx.FR.10.URF5_patient_A   .........................................................................................................-CTG-TGATAGTGGTACT--T--------------------------C-C---CAG-G----A---A.. 6073
57_BC.CN.09.09YNLX19sg        TGC............................................................AATCA-......................ATC-AC........GGT-GG-CCCTCG-T-AGA-TG---G---T-------------CA-C----CTAG---A--A----.. 5920
58_01B.MY.09.09MYPR37         AGC................................................CC--AC-A-.......-C-ACAACCGCTAGCCCCTTAGGAAATATA-CT........T-AGG--ATATA-C--AT--AG----------T-----T-----A-C----GAA-------A----.. 6112
59_01B.CN.09.09LNA423         ...................................................-TG-AGGA-.......-C-.....................AT--AT........----C---------------A-----G----------------A-C-T---------A--A----.. 5915
60_BC.IT.11.BAV499            ...........................................................................................-CCTAT........GCTCCT-GTAT-C-G-----A---------T---------------G-C----GAAC------A----.. 6682
61_BC.CN.10.JL100010          .........................................................................................................:..-GC-GCA--GGTCAG--A--G---G-G-------------CA-C----GTAT-G----A----.. 6115
62_BC.CN.10.YNFL13            ...............................................................AAC-C-......................AATGTG........-CCTACGAA----GGCA---A---------A-------------CA-C----GAA-----G-A----.. 6032
63_02A1.RU.10.10RU6637        .....................................................................................................................ACY---C--CT------T-T-----------G-C----GTA-------A----.. 6245
64_BC.CN.09.YNFL31            ..................................................................GAT......................AATG-T........-CC--GGA---T--T-CTA-CG-G--G-GT--------------A-C----GTAG------A----.. 6084
65_cpx.CN.10.YNFL01           GTC............AGTAAT..............................G-G---GTA......GC-......................AA-GTC........-GT-ATGA-AGTGT--A-C-T--G-----T-----------G-A-C-----TAG------A--G-.. 6113
67_01B.CN.11.ANHUI_HF115      ...................................................GYM-AY-AW......-CA......................TWTWAC........--AGG--ATATA-C--AT--AG---G---T--TA----T--C--G-C----GAG-------A----.. 6074
68_01B.CN.11.ANHUI_WH73       ...........................................................................................AAC-AT........--CGAT-ACATA-C--AT--AG---G---T--TA----T--C--G-C----GAG-------A----.. 6056
70_BF1.BR.10.10BR_PE004       ...................................................-CC----A-........--G....................ACCTTG........GGA--G------G-------AC-------T-----------------C-----A----G---A----.. 6183
71_BF1.BR.10.10BR_PE008       ..................................................................--T......................A-TGA-........GGA-A--------G------A--G-----T-----------G--CA---CGA---G---AA-----.. 6385
72_BF1.BR.10.10BR_MG002       ..................................................................-A-......................AAC-CC........C--GGGGA----CC---G-CA-----------T-----------G-C----GAAG------A----.. 6223
O.BE.87.ANT70                 ...........................................................................................................-C--CAA-TG--AACCTT--G--G--G--TGAG--T---G-A-C---TGTT--C-A--AC--A.. 6742
O.CM.98.98CMA104              ..........................................ACAAAC...-CC-AC-CA......-CA......................ACACAC.............-ACCCATC-CC---C--G---C-G--TAA---T---G-A-C----GTT--C-A--AC--A.. 6204
O.CM.98.98CMABB141            ..........................................AACGAC...CCCTC--CC......TC-......................CCACCTACCTCGCTCCCT-GCGGTTCAG-GAACCCT--G-----A--TGAG--T---G-A-C----GTTC-C-A--AC--A.. 6227
O.CM.98.98CMABB212            GGG.........AAGAATAGTACA..........GGGAAT...-------CA......GTA......................-AT--C........-C--C-GAACCAGG-AC-CCT--G-------TAG---T-----A-CT---TT-C-C-A--AC--A.. 6256
O.CM.98.98CMU5337             ..........................................AAGAGT...CC--AG-C-......-CA......................-A---C........CCT-GG-ACATAG-GAACCCTG-G-----G--TGAG--T----CA-C----GTTC-T-A--AC---.. 6217
O.CM.99.99CMU4122             ...................................................-CA-CA-AC......GCA.........................................-C-GC--CC-GCA-C-CAG-G-----G--TGAG--T---G-A-C----GTTC-C-A--AC--A.. 6194
O.FR.92.VAU                   ..................................................................-CA......................AAC--TCCCTTA...-ACTC--ACA-T-C-AAG---G-G---C-G--TGAC--T---G-A-CT---GTGC-C-A--AC--A.. 6296
O.GA.11.11Gab6352             ...................................................-GC-A--TA......-CA......................-A--AC..............GAT-GTG--ACCCCT--G---GTG--TGAA--T---G-A-C----GTTC-C-A--AG--AAA 5946
O.SN.99.99SE_MP1299           ...................................................-GCTCC-TA......-AT......................AATGAT........-CC-GC-GTCCAG-GAACCTTG-G---C-G--TGAA--T---G-A-C----GTTG-C-A--AC--A.. 6792
O.US.10.LTNP                  .....................................................................................................................AC---C--G--------TGAG--T---GCA-C---TGTAG-C-AG-AC-GA.. 6645
N.CM.02.DJO0131               AAACCA............................................................G-A......................AATG-G........-CAGACC-A--AGC-A--C-C--G-----T-----A--------A-C---TGAA---CAT-AC--A.. 6209
N.CM.02.SJGddd                CAG...............................................................-AA......................AATGT-........-CAGAC--A-GAG-TAAGAGA--G-----T-----A--T---GCA-C---TGATC------A---A.. 6196
N.CM.04.04CM_1015_04          ...........................................................................................---GA-........T-AGAT--A-GAT-CAA-C-A--G-----T--T--A------G-A-C---TGAG-G--A--A---A.. 6196
N.CM.06.U14296                ACAGCT.......................AACACGACAACTTCT-CAG-G-AC......-TA......................A-TGC-........GAAGACC-A--AG--AA-C-C--G-----T-----A------GCA-C---TGAGG-G-C--A---A.. 6218
N.CM.06.U14842                AACAGT.......................AGCAGTGAAAATGTGGCA--G-A-......--TGAGACAGGGACAACA........AAT--C........C-AGACC-A--A--CA--C-A--G-----T--T--A------GCA-C---TGAGC---A--A---A.. 6249
N.FR.11.N1_FR_2011            AACACA.............................................-G--TA...............................................--AC-AGTTAA-C-A--G-----T-----A--T--GGCA-C---TGAG-G--AG-A---A.. 6100
P.CM.06.U14788                ......AGGGTTAATAACAGC..............................-CCGAC--C......-C-......................AAC-CC........-GC-A--CCA-TCC-TT---AT--TTTC-G---AG---T----CA-C-----TAG---A--A---A.. 6241
P.FR.09.RBF168                ............AATAACAGT..............................-GCTCCTC-......GTG......................AAT-KC........-CCGCCGACA-TCC-TT---TC--TTT--G---AG---T----CA-C----GTAT-G-A--A---A.. 6787
CPZ.CD.06.BF1167              ACAAAA.............................................----A--CAAGT...CTAGAA...................-AT-A-ATAAAA..-GC-ATC-TACA-GTCT---A--GT-TG-A----T---T---CAA-CT---GAAT-T----AC--A.. 6844
CPZ.CM.05.SIVcpzMT145         ...........................................................................................AACTCT........-CTGACGATCGCTTG-----C--G-G---T--TAGC--T---G---CA--TGAAC------A----.. 6261
CPZ.GA.88.GAB1                CAG........................................................................................ACATCT..............TCTCC-CCTCTC--A--G--------T--A--T---G-G-CA---GAAT------A----.. 6789
CPZ.TZ.06.SIVcpzTAN13         ...............................................TCA-C---C-CAACA...GAACCACCAACATCAAAGACCACAGTGAAT-CG........TCAGAG--A--AG-TTATA--G-GCTT--T---AGC--T---CAA-CA--TGAGT-T--G-A----.. 6390
CPZ.US.85.US_Marilyn          .........................................................................................................:.....-ATACT-CT-ATATA--G-G------T-----T-----A-CT--TGAAT------AC--A.. 6778
GOR.CM.07.SIVgor2139_287      ......CAGACTGAGATGACAAATACAACT.........ATA...CCACCACA-...ACA-CT......................ATACC-.........CCACAC--CCCAC-GTT---C--TT-C-----TAGC--T---G-T-C---GGTTC---A--A----.. 6292
GOR.CM.07.SIVgorCP2135con     ACAAAT.........................................................ACA-CT......................ATACC-.........CCATAC--CCCAC-GTT---C--TT-C-----TAGC--T---G-T-C---GGTTC---A--A----.. 6270
```

```
                                                                                                                                         V2 loop end
B.FR.83.HXB2                .GTGCAGAAAGAATATGCATTTTTTTATAAACTTGATATAATACCAATAGATAATGAT...................................................ACTACCAGCTATAAGTTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGG 6846
Env                         __V__Q__K__E__Y__A__F__F__Y__K__L__D__I__I__P__I__D__N__D__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__T__T__S__Y__K__L__T__S__C__N__T__S__V__I__T__Q__A__C__P__K__
A1.AU.03.PS1044_Day0        .AAA------AT----T--C---------G-------G--G-------T-G----A--................................................AGTGA--GTGAG----TA--A-T--A------T------AC---------A--T---------- 6087
A1.CH.03.HIV_CH_BID_V3538   .AAA------ATGC-----C---------G-AC----G--G---A---T-GA-G-A--..........................................AGTAGTGAA-GATAT-AG----GAC-A-T--A------T-------CT--C--------T---------- 6052
A1.CY.08.CY236              .AGA-----C-T----T--C---------G----------G---A---T----G-A--.............................................AGTAAT-A--GTCAT----GA--A-T--A------T------AC---------A--T--------A- 6084
A1.ES.05.X1608_8            .CAA-------T-C--T--C---------GC------C--GC--A---G-G----AG-AGTTATAAGGGGCAAATACAAACG...AGCAATAATAACCTGAGTGTTACCTA-CAT-C-----GA--A-T-CA------T------AC-----G---A--T--------A- 6342
A1.KE.11.DEMA11KE001        .AAAA----G-T----T--C---------G-TC----G--G--T----GA---GCA--................................................GAG-G--GTGAG----GA----T--A------T------AC---------A-----C------- 6228
A1.RU.11.11RU6950           .AA-A-A----T----T--C---------G----------G--T---CT-G--G-A--................................................GAT-G-GAGCAG----GAC-A-TT-A------T--A----C-G-G--------------C---- 6391
A1.RW.11.DEMA111RW002       .AAC------ATG---T--C----------G---------G---A---TA--G-A---AAGAATAATAGTAGT........................AAGAATAATAGT-G-T-TCTGA---G---A-T--A------T-------C------------T---------- 6215
A1.SN.01.DDI579             .AAA------AT----T--C---------G-------G--G-------T--G----G-AGTAACAGT.................................ACTGAGAAC-G---TGAA----GA--A-T--A--------------C------------T---------- 6062
A1.UG.11.DEMA110UG001       .A-AAG-------------C---------GG---------G---A---TA--GGAAG-................................................GAT-AC-GT-CA----GA--A-T--A------T-------C-----A------T---------- 6199
A1.ZA.04.04ZASK162B1        .AG---A----C-------C-------------A---G--G---A---T------A--...................................................-G--GTGAG----TAC-A-T--A------T-------C------------T---------- 6302
A2.CM.01.01CM_1445MV        .ACA---G---TC---T--C-G--------G---------G---A---TA---G-ACGGACAATGTTACGCAG........................CAAAACAATGAT--GCAGCCG----GAC-A-T-C-------T-----------C-----A-----------A- 6039
A2.CY.94.94CY017_41         .ACA--A----TT---T--C-G-------G-------G--G---A-C-T---G-AAG-GAAAATAAGAAT...........................ACATCAGGTAGT-A---TCTG----GAC-A-T--A------T------AC---C-----A--G---------- 6227

B.BR.10.10BR_MG029          .---A---------------A---C-----C------G--G---A----AC-G-GA--.............................................GCTAAT-A--AT-------G----G----C------------------------------------A 6337
B.CA.07.502_1191_03         .--T-G-------C-G---C-----C---GC---------G--------A-GG---GGATAAT.....................................AACAATACC-A-G---CA----T----G-----------------T--------T--------------A 6308
B.CH.08.M2_0803101_NFLG8    .A-A---------------C-----------------G--G----------AG-----AACAAGACAACTAATGGT...............AGCTATAATAATACAAAGTA--GT-A-----GA--A-T----------------A------------------------ 6368
B.CN.10.DEMB10CN002         .--AG------G-C-G---C-----------------G--GA--A----A-----A--GCTAAT.......................................ACAAGTTA-G-A-AT----GT--A---CA--------------------A---A------------- 6234
B.ES.10.DEMB10ES002         .---------AC------TC-----A---CG------G----G-A-------GG----...................................................-A---A-------CA--A-T--A--------------------GG---A------------ 6246
B.FR.11.DEMB11FR001         .TG--------T-------C-------------A---C--G--AA-------G----AATAAT................................................---T-------GA----T---------T------------------------------- 6058
B.GB.05.MM45d213_GN1        .------------A-----A---C------G-----G-GG--GA-------G--AG-AACAGC...........................................AGCTA--GG-------G-----T---------------------G-----C------------- 6258
B.HK.06.HK003               .--AA-A------C-----C-----C----C---------G---A---GA----..........................................................----C-----G-----T---------T------------------------------A 6218
B.HT.05.05HT_129389         .---------A-C-----C-----C----C-------T--G-------A-G--G--GAAT....................................................-AT-------GA--------------T----------C------A--T--------A- 6082
B.JP.12.DEMB12JP001         .-C--G---G---AGC---C---------G-------G--G----------G-----AATAATAGT...............................GATAGTAGTACT-G-G-T--T----G---A-T--A-------------------------A------------ 6268
B.KR.07.HP_18_07JHS10_3909  .-----------------TC-----------------G--G---A---------......................................................------A-------CA----T--A-------T-----A-----------------------A 6506
B.PE.07.502_0525_wg5        GCA-A-AG-----------C-C---------------G--G-------------AG-GATAAT.................................................--T-------G---A-T--A-------------------------------------- 6292
B.RU.11.11RU21n             .---A-A------------C-C--------T--G---G---A---------A--GATACT....................................................-GT-------G-----T---------T----------G------A------------C 6386
B.TH.08.MERLBDTRC10         ....A----CAC-------C-----------------G---C------G---------AGTATCAAGAAATAATATTACA.................AATGATACCAATTA--GT--G----G----TT--A-------------------------------------- 6108
B.US.11.ES38                .--A------A-G-------C---------C----------G--A-----AG-----AATACTTCT..................................ACTAATAATTGG-AT-------G-----T---------------------C------------------A 6789

C.AR.01.ARG4006             .CAAA---G-T-C------C---------G-------R--G-----C-TA-------G.............................................AATGAT-G--GTGAT----GA----T--A------T-------C---A-----A------------- 6043
C.BR.07.DEMC07BR003         .AAAG---G-T-A------C--------------------G-----C-TA---T-AG-TCGGGTAATAATAATAGC............TCTGATGATAATAACAGTTCTGG--AATAT----G---A-T--A------T-------C-G-A-----------------A- 6271
C.BW.00.00BW5031_1          .AGA--A--G----T----C-----A---T----------G-----C-TA------GAATAACAAT..................................ACCAAAAATTTC-GTGAT----GA--A-T---------T------AC---A-----A------------- 6208
C.CN.10.YNFL19              .AG-A-------G-----CC--------------------G------T--GGGG--GAATTCTAGTGAG...............................GGGAAGAATTT--GTGAG----GA--A-T--A------T-------C---A-----A------------- 6231
C.CY.09.CY260               .-AAA--------------C---------GG---------G--GA-C-T-GA-GCA--...................................................-G--GTGAG----GA--A-T--AC-----T-------C---A-----A-----------A- 6102
C.ES.08.X2363_2             .AAAA-T----TG------C---------G----------G-----C-TAG-------AACAGC.....................................TCTAGTAACT--GTGAT----GA--A-T--A-------------AC---A-----A--T---------- 6324
C.ET.02.02ET_288            .AACA------C-------C----C---CG----------G-----C-TAC-------.............................................AACAGCT--GATGAG----GA--A-T--A------T------AC---A-----A------------- 6093
C.IN.09.T125_2139           .-A--------TG-----GC---------G----------G---A-C-TA--G--AG-AGTAACTCTAGTAGT..................AGTAATAACTCTAGTAAGTA--GT-AG----GA--A-T--A----G-T------AC-G-A-----A------------- 6869
C.KE.00.KER2010             .CA--------T----T--C--------------------G--T--C-T--G----G-.............................................AATAAT-G--GT-CG----GA--A-T--AC--C--T-------C---A-----A------------A 6056
C.MW.09.703010256_CH256.w96 .AAAA--CT----------C-----------------G--G--A--C-T--GG-AA-C...................................................T-C-GTGAG--C-GA--A-T--A------T--------AG-------A------------- 6846
C.TZ.08.707010457_CH457.w8  .AGAAGA-------C---C-C---C----G-G--------G-----C-T--C--CA-C...................................................-G-G-TGAG----GA--A-T--A------A------AG-C-A-----A------------A 6833
C.YE.02.02YE511             .AGAA--C-G----C----C---------GG---------G-G---C-T---GGAA-A..........................................GAAGAGAACC-C-GTGAG----GA--A-T--A--------------C-G-A-----A------------- 6044
C.ZA.08.705010534_CH534.w12 .AAAAG-------A-----C--------------------G-----C-TA--GGAA--TCTAGTGGA.................................AATGGAAATGT--GTGAG----GA--A-T--A------T------AC---A-----A------------- 6827
C.ZA.10.DEMC10ZA001         .ACA-----T--G------C----------AC----G---G-----C-TA--G-GA--....................................................G--GTGAG----TA--A-T--AC------T-----ACT--A-----A------------- 6202
C.ZM.11.DEMC11ZM006         .AGTA--C-----C-----C---------G-GA--------C---GC-T-GAGGGA--................................................GAC-G---T--T----GA--A-T--A------T-------C---A-A---A-----------A- 6219

D.CM.10.DEMD10CM009         .AA-------A--------C-C---------A-G---C-GG--GA-----G---GA--...................................................-G---T-AG----G---A-T--A------T-------C---------A--T---------- 6205
D.CY.06.CY163               .AA-A--C---T-C----CC---------G------CG--G--AA--------G-C-GAATAATAGG..............................GCTAATCAGAGT-A-GGT-CT----G---A-T--A------T--T----C------------T--C------- 6065
D.KE.11.DEMD11KE003         .AAAA-----TTC------C----------G------G-GG---A-------GG-AG-AATAGT.......................................ACTGAT-A---G-------GA--A-T--A------T-------C-------T----A---------A 6215
D.KR.04.04KBH8              .AA-A--C-----------C---------G----------G------G--------.....................................................-A------T----GA--A-T--AC-----T------AC-G----------T---------A 6777
D.SN.90.SE365               .AGA-----T-T----T--C-----------------T--G---------GC---AG-AATAAGAAT.................................AGCACCAATTA------T----GA--A-T-CA------T------AC---------A--T---------- 6883
D.TZ.01.A280                .AGAA--C---T-C-A---C----------------G-GG-------------AA-GAATAAT.................................................-AT--T----CA--A-TGCA---C--T-------C-----A---A--T---------- 6052
D.UG.10.DEMD10UG004         .AA-A--C---T-C-A---C---------GC-A----G--G---A----AGG---A--.........................................................-C-----GA--A-T---------T-------C---------A--A---------A 6159
D.UG.11.DEMD11UG003         .AA-A------T-C-----C-------C---------G--G--AA-------G--A--.................................................GAT---GTGC-----GA--A-T--A------T-------C-C----------G---------- 6200
D.YE.02.02YE516             .CAAA--C----------G--G------------------G--G---G-----AG-...........................................................-CT----GA--A-T--A------T------AC------------T---------- 6032
D.ZA.90.R1                  .AGAA--G---C--C---GC---------G--C-------G----------CC-G---AATAAA....................................AATGGTAAT-AG----CT----G---A-T--A------T------TC------------T---------- 6212

F1.AO.06.AO_06_ANG32        .AG-TT-----T-A-----C----C----G----------G-------CA---G-AG-GAT..........................................AGTAAC-G--ATGAA----G---A-T--A------T------AC-----A------T--------A- 6064
F1.AR.02.ARE933             .AA--T-----T-C-C---A---------G----------G-------CAG-GGCAG-AAT..........................................GACAAT-G--GT-CA--C-G-C-A-T--A-------------AC------------T---------- 6143
F1.BR.10.10BR_RJ015         .CA--TA----T-C-A---C--------------------G---A---CA----CA--AGTAGTAATAAC...........................AGTAGTAGTAATTTC-GTGA---C-G-C-A-T--A------T-G----A------A------T---------- 6270
F1.CY.08.CY222              .AA--TA----T-C-----C---------G----------G-------CA---G-AG-GAT..........................................AACGAG-GC-GTGAG--C-G-C-A-T--A------T------AC------------T---------- 6045
F1.ES.02.ES_X845_4          .AA--T-----T-A-----C--------------------G-----C-CAC--G-A-G................................................GGC-G--GTGAA----G-C-A-T--A------T------AC-----A------T---------- 6298
F1.RO.96.BCI_R07            .AA--TA-G--C-C-G---C--------------------G-----C-CAGC---AC-AATGGGCCA.............................CTCAACAATACT-G-GGGGAA----G-C-A-T-CA------T------AC-----A------T---------- 6924
F1.RU.08.D88_845            .ACAA--C--ATGCG-A--C---------GC---------GA--T---CA-C--AC-AAGTAAARAC.................................AATAGAACT-R--AAGAG----TA--A-G---------T------AC--G---------T---------A 6335
F2.CM.02.02CM_0016BBY       .AG-A-A--T----T----AC-----------A-------G------T--A--AA--....................................................--G--ATC-----G-C-A-----------T------AC-G----------T-----G---- 6038
F2.CM.10.DEMF210CM001       .CA-A--------------C---------G-------G--G----G--T-G-G--A-CAATAAT.......................................GCTGAT-G-GGTGAA----G-C-A-T--A------T------AC-G----------T---------- 6133
F2.CM.10.DEMF210CM007       .AA-A-------------GC---------G-------G--G-----C-CA--GCCA--.............................................GCCAAT-G---TGAA----G---A-T--A------T--T---AC------------T---------- 6194

G.BE.96.DRCBL               .AA-A--GC------C--GC----C----G-AC----G--G-------TA--G-GATGAATAATGAAAAT...........................AATGGTACAAAT-G---AT-G----G-C------A------TGT----AC-----A------T---------- 6812
G.CM.10.DEMG10CM008         .AA-G---C-A--------C----C----G-------G--G-------TA--GG-A--...................................................-A--GT-CT----GAC-A-T--A------TGT----AC-G--AG------T---------- 6236
G.CN.08.GX_2084_08          .CA-A--------------C----C------------G-GG------T------A--ACAAAT.....................................AATAATAGT--A-GT-AT----G---A-T--A------TGT----AC-----A---A--T---------- 6103
G.CU.99.Cu74                .AA-A-AC----------GC----C----G----------G-G-----TA--G--A--AGAGATAATAGTACAGTT.....................AGTAATAGTAGCC-AGAA-A-----G-C-A-T--A------TGT-----C-----A------T---------- 6480
G.ES.09.X2634_2             .AA-A-A-----------GC----C-----------G-GT------T--G-G-A--GCAAAT...................................AGTAGTAGTAATTA--GT-AT----G-C-A-T--A------TGT----ACT----AG--A--T---------- 6359
G.GH.03.03GH175G            .AGTA-------G--C--GC-C--C----GG-C----G--G-------T---GG-A--GATAAT.......................................AGTACTTA--GTGAT----G-C-A-T--A------TGT----TAT--C-AG-----T---------- 6928
G.KE.09.DEMG09KE001         .CA-A----T-----C--GC----C----GGGA----G---C--A---TAGGG---GGGTAAT.....................................AGCACTAAC-G---ACTG----G-C-A-T--A------TGT----AC-----A------T---------- 6246
G.NG.09.09NG_SC62           .AGAA--------C----GC---------G-------G-G-------T------A-C.................................................AAT-G--GTGAG----G-C-A-T--A------TGT----AC---C-A---A--T---------- 6055
G.ZA.01.TV546               .MGAAG--C------C--GC----C----G-------G-GG-------TA--G-G---CAGAAT..........................................GCA----GTGA-----G-C-A-T--A------TGT----AC-----A------T---------- 6051

H.BE.93.VI991               .AAC-----G-T-C-----C-C-------G-GC---C---G---AG--T---G-A-G-GAAAGA....................................AATAAAAGTGA--ATCAT----G-C-A-T--A------T-----G-------A---A--T--C------- 6245
H.CF.90.056                 .CA--------T-C-----C-C-------G-------G--G-------T-----CA--...................................................-G---TCAG----G-C-A-T--A------T--------------------T--C------- 6177
H.GB.00.00GBAC4001          .AA-A------T-C-G---C-C-------G-A-----G--G-------TA--G---AGATTATAATGGGACT.........................AATGGGACTAAC-GG--TCTG----G-C-A-T--A------T--------------------T--C------- 6378

J.CM.04.04CMU11421          .ACAAG--G----------C----C----G--AA------G-------TAG-GGCAGA................................................CCT-A-GAT--T----TT--A-T--AC-----T-------------A---A--T---------- 6384
J.SE.93.SE9280_7887         .AG-A--C----G--C--GC----C----G--AA--G---G-------T----G-A-C...................................................-A--AA-AT----TTC-A-T--AC-----T-------------A---A--T---------- 6159
J.SE.94.SE9173_7022         .AG-A--C-------C--GC----C----G--AA--G---G-------TA---G---C...................................................-A--AA--T----TTC-A-T--AC-----T--------T----A------T---------- 6166

K.CD.97.97ZR_EQTB11         .AAAA---G--TG-C----C----C-----G---------G-G-A---A-GC-ATC-GAAATT.....................................AATCAATCTGAAT-TGAAG---G-C-A-T--A------T------AC-G--------------------- 6068
K.CM.96.96CM_MP535          .AA-A-A--G----C----C----C----G-------G--C-G---C-TA--GGC--GGGTAAT.......................................AACAGT-G---TGAA--C-G-C-A-T--A------T------AC----------A-T--------A- 6054

U.CA.01.TV749               .AATA-AC----------GC-----------------C--G---A-C-----G-AG-.....................................................----TGAT---GTTC-A-T--A------TGTT---AC-----AG-----T---------- 6871
U.CA.99.TV721               .AA-A-AC----------GC-----------GC----C--G----C-------A--......................................................----TGAT---GTTC-A-T--A------TGTTA--AC-----AG-----T---------- 6897
U.CD.83.83CD003_Z3          .AC-G---C--T-C--A--C-------------G---G--G-G---C-TA--GTGACAAATAAT....................................TCTAGTATA-G--GT-C-----G---A-T--A------T------AC------------T---------- 6328
U.CD.90.90CD121E12          .AA-A----T-T-C-----C-------------G---G--G-G---C-TA--G-GACA..........................................AATAGTAGCTA-GGG-C-----G-C-A-T--A------T------AC---C--------T---------- 6014
U.CY.05.CY090               .AAA-------T-C-----C----C----G-------G--G---A---TAG-G--A-G................................................AGT-A---ACTG----GAC-A-T--AC-----T-------C------------T--------A- 6056
U.CY.08.CY223               .CA-A--C----------GC----C-------C------GA--A---T---GG---....................................................-A--GT--T----G-C-A-T--A------TGTT---AGT----A------T---------- 6133
U.ES.10.DEURF10DZ001        .AAAA-----AT-C-----C-C---------A--------G--AA---TA--G-CAG-.............................................GAAAGT-AC-GT--T----T---A-T--A------T------ACT-----------T---------- 6105
U.GR.99.99GR303             .AAA-------T----T--C---------G---G---G--G---A---G-G--G----AATAGC....................................AAGAGCAATTA--GTGAG----GA--A-T--A-----GT-------C------------T--------A- 6139
U.NL.95.U_NL_95_H10986_D1   .AA-AGA-----------GC---------G-T-----G--G-----C-TA-C-G---AATAAT.....................................AGCAGCAATTA--GTGAA----G---A-T--A------T--T---AC-----A------T---------- 6743

01_AE.AF.07.569M            .AA-A----G-CT------C----------G---------G--T----T--A--GA--...................................................--A-GTGAG----G---A-T--A------TT-T----------AG-----T---------A 6064
01_AE.CF.90.90CF11697       .AA----G-G-TCC-----C--------GTG-C-------G---G---T-GAG-AA-GAATAAAAAT..............................AGTAGTGGAAAT-G--GTGAG----TA--A-T-CA---------T----------AG--A--T---------- 6816
01_AE.CN.10.YNFL03          .AA------G-TC------C----------G-C-------G--GA---T----G-A--...................................................-A--GTGAG----GA--A-T--A------T--T----------AG-----T---------- 6242
01_AE.HK.04.HK001           .CA------G-TC------C----------G---------G--GA---TA----..........................................................-GT-CG----G---A-TT-A------T--T----------AG--A--T---------- 6222
01_AE.IR.10.10IR.THR48F     .AA-----G--TCC-----C----C-----G------T--G---A---GA--GG-AG-...................................................-A--AT-CT----GA--A-T--A------T-------------AG-----T---------A 6051
01_AE.JP.x.JRC77AE          .AA----C-G-TCC----TC---------GG---------G-------C------.........................................................-GT-CG----G---A-T--A------T-------------AG-----T---------- 6852
01_AE.TH.09.AA090a_WG11     .CA------TTTC------C---------G----------G---A---TAG-G--AC-................................................GAT-A--GT-CA----G---A-T--A------T--T----------AG-----T---------A 6283
01_AE.TH.90.CM240           .AA------G-TCC-----C-C--------G---------G---A---T--AG--A-G.............................................AAGACT-G--GTGAG----G---A-T--A------T--T----------AG-----T---------A 6426
01_AE.US.05.306163_FL       .AA------G-TCC-----C----------G---------G---A---TA----GAG-......................................................-GT-AT----G---A-T--A------T--T----------AG-----T--------AA 6001
01_AE.VN.98.98VNND15        .CA------G-TCC-----C----------G---------G-------TC-----AG-..........................................AATAATAGT-G--GTGAG----G---A-T--A------T--T----------AG-----T---------A 6119
```


Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

```
                                                                                                                          V2 loop end
B.FR.83.HXB2                 .GTGCAGAAAGAATATGCATTTTTTATAAACTTGATATAATACCAATAGATAATGAT.................................................ACTACCAGCTATAAGTTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGG 6846
Env                          __V__Q__K__E__Y__A__F__F__Y__K__L__D__I__I__P__I__D__N__D__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__T__T__S__Y__K__L__T__S__C__N__T__S__V__I__T__Q__A__C__P__K__
02_AG.CM.10.DE00210CM013     .-AAA---G---G-T----C---------G-AC-------G---A----A--G-AAC-AAGAGTAACAATAGG.......................TGCAGTACC-A--GTT-T----G---A-T--A------TT-----AC-----------T--------- 6282
02_AG.CY.09.CY256            .AAAA--C----G------C---------G--AA---G--G---A---T-----CAG-.........................AATCAGT-TCAG----T---A-T--A-------------AC-----------T--------- 6032
02_AG.ES.06.P1423            .-AA-----TATGAC-T--C---------G-GC----T--G--GA-G-TA--G-AA--..........................AG-GTGAG----TAC-A-T--A------T-------C----G-------T---------A 6268
02_AG.GW.05.CC_0048          .AAAA------TG-C----C-G------GT-TC----G--G---A---G---G-AA--AGCAGTAGTAATAATTCTCAT.............GAAGATAGCAGTAGT-T--GTTCT----G---A-T--A------T-------C-G--G-------T--------- 6235
02_AG.KR.12.12MHR9           .AAA-------TG-C----C---------G-TA----G--G---A---TA--G-AA--AAGAAT....................AGC-G--GTCAG----G---A-T--A------T-------C-----------T--------- 6520
02_AG.LR.x.POC44951          .AAA-------TG------C--------GG------G--G---AG--CA-----AG-AAAGGT..................AGTAAAAATGA--GTCAG----G-C-AGT--A------T-------C-----------T--------- 6829
02_AG.NG.09.09NG_SC61        .AGA------AT-GC---GC---------G--AA------G-------G---GGGA--.......................AGCAGT-G--GTCAG----G---A-T--A------T-------CA----------T--------- 6047
02_AG.NG.x.IBNG              .AAAA----GATGC-----C---------G------G--G---A---TA--G-AA--........................AATGG--GTCAG----G---A-T--A------T-------C-----------T--------- 6370
02_AG.SE.94.SE7812           .AAA-------TG------C---------G-TA----G--G---A---TA--G-AAC-GGG.....................GAT-AC-TTCAA----G---A-T--A------T-------C-----------T--------- 6212
02_AG.SN.98.98SE_MP1211      .AAAA----GATGGCG---C-----------A-------G-------T-----AA--.........................G----TTAT----G---A-T--A------T-------C-----------T--------- 6039
03_AB.RU.97.KAL153_2         .---A-A-----------C-C---------------G--G---A-------------..........................---------G-----T---------T---------G-------A-----------A 6044
04_cpx.CY.94.94CY032_3       .AA-A------------GC----C----G-A--------G-G-----TA--GC-AGAGTGCCAATTAATGGTAGTAAT..................AGGAATAATAGT--AGAAGAG----T---A-T--A-------G-----AC-----A------T--C------- 6234
05_DF.BE.x.VI1310            .AA--T-----T-C-----C-C-------G---------G-------CAG--GC---GACAGT.....................AGC-A--G---T----G-C-A-T--A------T------AC-----A------T--------- 6249
06_cpx.AU.96.BFP90           .ACAA------------GC----C----G--C-------G-------T-GGG----...................AGTAAT-A--GTGAT----GA--A-T--A------TGT----AC-----A------T--------- 6918
07_BC.CN.98.98CN009          .AA-----C--TG------C---------G---------G-----T-TAC---GA-GAACTCTAGT.................GGGAACTCT-G-GAGCAT----GA--A-T--A------T-------C---A-----A------------- 6192
08_BC.CN.97.97CNGX_6F        .AA-A---C--TG------C---------G---------G---CC-TA--G----GAACTCTGGT.................GAGAACTCT-G-GAATAT----GA--A-T--A------T-------C---A-----A------------- 6034
09_cpx.GH.96.96GH2911        .CAC-------T----T--C---------G-TC-------G--T----CA--GCAAG-.........................A--ATGAG----GAC-A-T--A------T-------C-----------T--------- 6045
10_CD.TZ.96.96TZ_BF061       .AA-A--C-----------C-----------------G-GG---A-----G-G--AG-.........................A-----------GA--A-T--A------T-------C-----------A--------- 6219
11_cpx.CM.95.95CM_1816       .CA-A--C----------GC----C-------------G-------G-----C-CA--G--A--......................GAT-A---A--T----G-C-A-T--A----R-TGT----AC-G---A------T--C-----A- 6251
12_BF.AR.99.ARMA159          .CA--T-----T-C-----C---------G---------G-G-----C-G-----GC.........................-G-GAA-CA--C-G-C-A-T--A------T------AC-------T----T--------- 6859
13_cpx.CM.96.96CM_1849       .AGA-------T---CT--C----C----G-G--------GA------T--GGGAA--AGTAATAGT..................GGTAATAGTAGT-G--GTGAA----GA--A-T---------T-------------A---A--T--------- 6268
14_BG.ES.05.X1870            .A-AA--------------C---------GC----------C--A---G--C------.........................................--------G------------------------------------------A- 6337
15_01B.TH.99.99TH_MU2079     ....GTA-----C------C-------C-GC------G--GA-G------------.........................----------G-----G--AG-----------A----------A--T--------- 6218
16_A2D.KR.97.97KR004         .AAA-------TCC-A---C-A------G----------G---A-C-TA---G-AG-GACAGTAAT.................GATACACTG-A--GGCAG----GAC-A-T-CA----G-T------AC---C--------T--C------- 6215
17_BF.AR.99.ARMA038          .CA--TA----T-C-----C---------G---------G-------CA--GG-AG-AAT.......................GGCAGTGGAGGAGAA----G-C-A-T--A------T------AC-----------T--------- 6065
18_cpx.CU.99.CU76            .CA-A--C----------GC----C---WWW--A-----G--T------TGGGGATR..................GTAATAGTAATTA--GTGAT----G---A-T--A------TGT----AC-----A---------C-----A- 6154
19_cpx.CU.99.CU7             .AAA--AC--AT----T--C----C-----G---------G-------T-G-G--A-G.........................................A-A-T--A------T-------C-----------T--------- 5990
20_BG.CU.99.Cu103            .AA-A--C----G--C--GC----C----G-------G--G-G-----TA--G--A-GAATAGTAAT..................AATAGTAATAACTA--GTGA-----G-C-A-T--A------TGT-----C-----A------T--------- 6316
21_A2D.KE.99.KER2003         .AG-A--C---T-------C-------------------G--------A-C------GATAAT.....................ACC----GT-C-----G---A-T--A------T-------C-----------T--------- 6047
22_01A1.CM.01.01CM_0001BBY   .AGAAG-G---T----T--C-C-------GG---------G--GG---TA--GGAA--.........................-AT--T-T--GAC-A-T--A------T-------C---------A--T--------- 6037
23_BG.CU.03.CB118            .AGTA----------C--GC----C-------------------G-----T---G--A-GAATAATGGT..................AGTAATAAAAACTA--GTGA-----G-C-A-T--A-----TGT----AC-----A------T--------- 6316
24_BG.ES.08.X2456_2          .CARA----------CK-GC----C-----------------R-G--W--TAG-G--A--AGTAAT.................AATAATACCWA--GT-A-----GAC-A-T-RA---Y--TGT---CAC-----R------T--C--WW--- 6336
25_cpx.CM.02.1918LE          .AA-A-A--G-----C--GC----C-----------------G-----C-T---G-----......................-A--GTTCA----G---ACT---C-----T------AC-----A------T--------- 6063
26_AU.CD.02.02CD_MBTB047     .CAAG-------T-C----TC---------G-T-A------G---T---GA--GG-A--GATAGT.................AAGAATAGT-G-GGGGAA----GA----T--A------T-------C---A---------T--------- 6849
27_cpx.FR.04.04CD_FR_KZS     .AA-------TT-C-A---C-----G-G-GC------G--G-------GA----AAG-.........................-G--AT-CT----G-C-A-T--A------------------A------T---------A 6839
28_BF.BR.99.BREPM12609       .A-----------------C-----------------G--G-------C---------.........................-A-G-A-------C-C---T--A-----------------------A------------- 6220
29_BF.BR.01.BREPM16704       .A--A-A-----------GC-----C---C----------G--T---C--GGG-A-GGGATAAT...................-A-------TAC---T--C-----------------------A----------A 6267
31_BC.BR.04.04BR142          .-GAAG---G-T-C-----C---------G---------G-----C-T--A-G---G.....................AGTTA-GAAGAT----GA--A-T--A------T-------C---A-----A------------- 6314
32_06A1.EE.01.EE0369         .AAAA---C------C--GC----C------------G---------T--GGG-A-AAATAAT.................AATGGTAAT-G--GTGCT----TA--A-T--A------TG--A--AA----GA------T--------- 6517
33_01B.ID.07.JKT189_C        .AA-G----G-CC------C----------G--------G---AGC-T-GA---AG-.......................AAT-G--GT-AA--C-G---A-T--A------T--T----------AG-----T--------- 6165
34_01B.TH.99.OUR2478P        .AA------G-TC------C-----------------C--G---A---TA--GG-AG-.........................GG-GGTGAA----G-----T--A------T--T----------AG-----T---------A 6039
35_AD.AF.07.169H             .AAA-----G-T-C-GT--C-----------------G--G-------T---G-C---.........................-A--GT--T----GA--A-T--A------T-------C---------A--T--------- 6061
36_cpx.CM.00.00CMNYU830      .A-A-------TGC--T--C---------GG---------G-------T---GCAAG-AATAAT.....................GAT-G--GT--G----GAC-A-T--A------TG------C---------A--T--------- 6058
37_cpx.CM.00.00CMNYU926      .AAA--A---AT----T--C----------G------G--G-------T-G-G--A--.........................-A-GAG-CT----GA--A-T--A---C--T-------------------T--------- 6040
38_BF1.UY.03.UY03_3389       .CA--TA----T-C-AT--C-C----------------G-G-----T---GG---...................AATAGC-GC-GATAT--C-G-C-A-T--A------T------AC-----A------T--------- 6266
39_BF.BR.04.04BRRJ179        .AA---A----T-C----TC---------GG---------G-T-----CAGC---AGCAGCAAT...................AGCAGT-G-GGAGAA--C-G-C-A-T--A-----GT-------C------------------------ 6338
40_BF.BR.05.05BRRJ055        .CA-A------T-C-----C-C-------G-----------C------CA--GG---GGGTAAT.....................GAT-A-GGA-CA--C-G-C-A-T--A------T-------C------------------------ 6376
42_BF.LU.06.luBF_18_06       .C-----------C-----C-G---------TC----G--G--------------A--.....................AGTAGT-G---A-AT----T-----T------------T-----TC----------------T--A- 6432
43_02G.SA.03.J11223          .AA-A------------GA----C------------G--G-------T---GGGA--AGTAAT....................AATGAAT-AGGTGAT----G-C-C-T--A------TGT----AC---------A--T--------- 6374
44_BF.CL.00.CH80             .CG--T-----T-A-----C-A-------G---------G-------CAGC---A--........................AGC-G--GAGAA--C-G-C-A-T--A------T------AC-----A------T--------- 6300
45_cpx.FR.04.04FR_AUK        .CAA---G--AG----T--C-----------------G--G---A---TAG-G-AA--........................AATGA--GT-------GAC-A-T--A------T------AC-----A------TGTC------- 6811
46_BF.BR.07.07BR_FPS625      .CA--T-----T-C-----C---------GG---------G-------C-G-G--A--.........................-A--GT-CA--C-G-C-A-T--A------T-G---AC-------------T--------- 6270
47_BF.ES.08.P1942            .AGAA---G----------C---------C------G--G-G-A-T----ATCA--C....................AAAACT-G---T-------G-----T---------------------------T---------A 6272
48_01B.MY.07.07MYKT021       .CA------G-TCC-----C----------G---------G-------T--A---A--AGCAATAGA...........AATAATAGCAATAGTAGTAATAGTAGTGAGTAC-GTGAG--C-G---A-T--AG-----T--T----------AG-----T---------A 6096
49_cpx.GM.03.N26677         .AA-A----------C--G-----C------ACA-------G-----T------A--.........................T----TGAT----TTC-AGT-CA------T---A----------A------G---------- 6285
50_A1D.GB.10.12792           .AAA-----GAT----T--C--------------------C------TA--G-AA--CATAGT.....................AGCGAT-G--GTGAG-----A--A-T----------T------AC-------------T--C------- 6249
51_01B.SG.11.11SG_HM021      .CA-A----GAGCC-----C---------GG---------G-------T-GA-C-C-G.......................AGTAAT---GTGAG----G---A-T--A------T--T----------AG-----T---------A 6058
52_01B.MY.03.03MYKL018_1     .AA------GATCC-G---C----------G---------G---A---GA---G-A--AGTAATAGT...................AATAGTAGTAATAGTAGTAAGTA--GGGAG----G---A-T--A------T--T----------A------T---------A 6199
53_01B.MY.11.11FIR164        .AM------G-CC------C----------G---------G--GA---TA----GA--AGTAATAGT.................AGTAATAGCAGT-G--GTRAA--C-T---A-K--A----R-T--T----------AG-----T--------- 6213
54_01B.MY.09.09MYSB023       .AA------G-TCC-G---C----------GA---------G-------T-----CA--AATAAT.....................AATAAT-G--GTGA-----GA--A-T--A------T--T----------AG-----T--------- 6266
55_01B.CN.10.HNCS102056      .AA----C-G-TCC-----C---------GG---------G--T----T-GG-----A.......................AAT-A--GTGAG----G---A-T--A------T--T----------AG-----T--------- 6150
56_cpx.FR.10.URF5_patient_A  .AG-A-------------------A----C---------G---A---A--G---GG.........................GAT-A--GT-------G---GT------------------------------------------ 6194
57_BC.CN.09.09YNLX19sg       .AA---------G------C---------G---------G-G---C-T---GGGA-GGACTCTAGTGAT....................-G-TCT----GA--A-T--A------T-G---AC---A-----A------------- 6047
58_01B.MY.09.09MYPR37        .AA------G-TC------C---------GGG---------------T--A---A--.......................AATAAG-G--GTGAG----T---A-T--AC-----T--T----------AG-----T--C------A 6236
59_01B.CN.09.09LNA423        .A-AAG-------------C-----------G-------TG-G--------------...........................-----------C-----T---------T---------------------------------- 6033
60_BC.IT.11.BAV499           .AAAA----G-T-C-----C---------G-------G-------C-T--G-----GAGTAGTGACAATGGG...........AATAATAATACTTT-GGGGAG----GA--A-T--A------T-------C---G-----A------------- 6827
61_BC.CN.10.JL100010         .AA-----C--GG------C---------G-A-C---C--G----TC-T-G-G-GA-CTCTAGCGAGTCTTAC...........AGATTGAACTCT-GC-AGTTT--C-GA----T--A------T-------C---A-----A------------- 6260
62_BC.CN.10.YNFL13           .CA-A-A------------C---------G-----------G-----C-TA--G--A-GAATTCTAGT.................GATAAGAATT---GTGAG----GA--A-T--A------T-------C---A-----A------------- 6168
63_02A1.RU.10.10RU6637       .AGA----G-ATG-C----C---------G-------G--G----C--TAC-G-G-GACAAAAC.....................AATAGTCAG-AC-GT-------G---A-T--A------T--T----C------------------------ 6378
64_BC.CN.09.YNFL31           .AA--------TGA-----C--------------------G-----C-TA--G-GAGCAGCAGTAAG....................-GGGAG----GA--A-T--A------T-------C---A-----A------------- 6208
65_cpx.CN.10.YNFL01          .CA-----C--CT-T----C-C-------G-AG-------G-G---C-T--GGGGA-CTCTAATAAC.............TCTAGTAACTCT-G-GGGGCT----GA--A-T--A---C--T------AC---A-----A-----C------- 6252
67_01B.CN.11.ANHUI_HF115     .AA----M-G-TCC-----C----------G---------G--GA---G--A--GA--..........................-AG-GTGAA----G---A-T--A------T--T-----T----AG-----T---------A 6192
68_01B.CN.11.ANHUI_WH73      .CA------G-TCC-----C----------G---------G---A---G--AGGGA--..........................-A--GTGAA----G-C-A-T--A------T--T----------AG-----T---------A 6174
70_BF1.BR.10.10BR_PE004      .--TA-----------G-A---------C-------------G--------------...........................--T-----C-G--R-T---------------R--C-------R--T--------R- 6295
71_BF1.BR.10.10BR_PE008      .-AA----------------G--------------G---C--A----A-AG-A---AATAAT.....................GCTAGC-A--GT-------G---A-T---------------------G-------------C------- 6515
72_BF1.BR.10.10BR_MG002      .CA--G-----T-C-----C-----------------------G-------TA-----A--GAGAACAACAATAGT............GATAATAACAGC-G-GGAGAA----G---A-T--A------T------ACA---------A-T--------AA 6368
O.BE.87.ANT70                .AA-G-----A--C-G--TC-A--C---GT-TCA---T-G--GGA-C-GA--G-GACAAGC..........AGCACAAATAAGACAAAC-GC-AA-TG----CA--A--T-A-------T--A--AC---C--G--A-------------- 6881
O.CM.98.98CMA104             .AAAG-----A--C-G--TC-A------GT-TCA---T-G--CAGGT-TA--G-GACAGAA...................AATAACAAGACA-GAGAA-TA----CA--A-TT-A-------T-TA-TACA-CC--G--A-------------- 6337
O.CM.98.98CMABB141           .AA-G-----A--C-G--TC-A--C----T-TCA---T-GG-TAA-G-GA--G--AGCGAAACAAGTAAAGGGAATAATAGCGAAACAAGTGGCAATAACACAACACCAT-AGGA-TA----CA--A-TT-A-------T--A--AC---C-AG--A-------------- 6396
O.CM.98.98CMABB212           .AA-G--C--A--C-G--TC-A--C----T-TCA---C-G-CTAGGG-TA--GGAAC-...................AATGGAACAAATGAC--A-TA----CA--A-TT-A-------T--A--ACT--C-AG--A-------------- 6386
O.CM.98.98CMU5337            .AAAG-----A--C-G--TC-A--C---TT-TCA---T---GTAA-C-GA---C-ACGGAT.....................AACACAT-A--T-TG----CA--A-TT-A-------T--A--AC-----AG--A-------------- 6344
O.CM.99.99CMU4122            .AA-G-----C--C-G--TC-A--C----G-TCA---T---ATAA------G--CAGC.......................ACAACC-GC--A-TG----CA--A-TT-A---C--TT--A--AC-----AG--A-------------- 6318
O.FR.92.VAU                  .CA-G-----A--C-G--TC-A--C---GTGACA---T-GG-TAAG--TA-CGCCACA.......................TCGAATGAA--A-TG----GA--A-TT-A-------T--A--AC---C-GG----------------- 6420
O.GA.11.11Gab6352            A-A-G-A--GA--C-G--TC-A--C---GT-TCA--CT--G-TAA-T-G--C---A--...................GGCACCAACCAA--A-TA----CAA-A-TT-A------TT--A--AC-----AG--A-------------- 6074
O.SN.99.99SE_MP1299          .AA-G-----A--C-G--TC-A--C---GT-TCA---T-G--GAAG---A--G-A-CAAAT.......................GACACA-AGGA--TG----CA--A-TT-A-------T--A--AC-----AG--A-------------- 6919
O.US.10.LTNP                 .AA-G-----A-GC-G--TC-G--C---GTGTCA---T-G-C----T-G---G-CA-AAAG..................AATAGCTCAAAT-GC--A-TG----CA--A-TT-A-------T--A--AC------G--A-------------- 6778

N.CM.02.DJO0131              .AA-A--C---CT--CT-TC-G-----GT-GAA---G--G-G---C-CA--G---GGAATAAT.........................-GT-CA--C-G-C-A-T--A------T---A---CTG-A-----A--T-----T---A 6330
N.CM.02.SJGddd               .AA-A-A-C--TT--CT-TC-G-----G--GAG---G--G---A---CAG---........................-AT-CA----G-C-A-T--A------T---A---TG-G-A---A--T-----T---A 6308
N.CM.04.04CM_1015_04         .AA-A--CT--CT--CT-TC-G-----GC-GAA---G--G-G-A-C-CA--G-GAG-GACAGT.....................ACG-A-CAA-CA----G-C-A-T---------A---A--TCTG-A-----A---------T---A 6323
N.CM.06.U14296               .CA-A------TT--CT-TC-G-----GC-GAA---G--------C-CAG----CAACTGATAAC.......................-AAGAT----G-C-A-T--------RT---A----TG-G-----A--T-----T---A 6342
N.CM.06.U14842               .AA-A--.......................-AA---K--S-MY-WR-CA-----C--GCCAATAAT.......................-AA-CA----G-C-A-T--A------T---A---CTG-G---------------T---A 6349
N.FR.11.N1_FR_2011           .AA-A---C--TT--CT-TC-G--------TGAA------G---A----A--G-CACAGTCAATAAT..................AGAACAGAA--AGTA----G-C-A-T--AG---G-T---A----TG-G-----A-T------T---A 6233

P.CM.06.U14788               .CA-----C-C-GC-A---C-C-------G-GCA--CC---C-AA-T-G--------C.......................-A--GT-CA----GA--A-TT-A---C------A-TAC------------A--C-----A- 6359
P.FR.09.RBF168               .AA----GG-C-GC-A---C-C--C----G-ACA--CC-----G--C-GA--G-GACA.....................GCAAAT-A---ACTA----GA--A-TT-A---C------A-TAC---------A--A--C-----A- 6911

CPZ.CD.06.BF1167             .AA-A--C-GAT----T----G--C------GAG--------GAA-GACA-A--C--C.......................AGTACT-A----TCT---T-T--A-T--A---C--T---A---C---A--T----AA---GAG---T 6968
CPZ.CM.05.SIVcpzMT145        .AAAAGAC---TG---T-C--A--C---GTGGAG-------C-G------GA---A--...........................-GT-CT--C-G-----TT-A------T---A---C---C--T--A--T---------A 6376
CPZ.GA.88.GAB1               .AA-A-AC-G-TC---T--C-A------GTGGAG---G-GG--AACC---GG-----G.........................-AC-A--CA----G-A-A-TT-A---C--T--TA---C---A-----A--A--------AA 6907
CPZ.TZ.06.SIVcpzTAN13        .AAAA----CATT--CT-T--A--C----GGGAA---G-T--GAA-GA-AGCC-GA--..............AAAACAAACAAT-GCGAATAT---T-T---CAT-A---C--T---A---C---A--T--AAGT---AGT---T 6520
CPZ.US.85.US_Marilyn         .AA-A-AC---TG---T-C--A------GT-GA----C--GCT-AT--CA-----..........................-AT-C-----GAC---T--A---C------A-C-CT--C---------T-----T---A 6890

GOR.CM.07.SIVgor2139_287     .ACAACA--GC-GC-A--T--A-------G--AA---------GA--C--G-G-AACAGGT...................ACAAATGGG-TC-AAG-G----GA--A-TC-A------T-----TACT--AT-G--A--T-----C--A- 6422
GOR.CM.07.SIVgorCP2135con    .AAAACA--GC-GC-A--C--A--------G-AA---------AA--C----G-GACGAATGAG..................ACCAAAGAGTA--GTGAG----G---A-TC-A------T-----TACT--AT-G--A--T-----C--A- 6403
```

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

HIV-1 Genomes

V3 loop start
B.FR.83.HXB2 CTAGCAGAA...GAAGAGGTAGTAATTAGATCTGTCAATTTCACGGACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTA.....GAAATTAATTGTACAAGA.....CCCAACAACAAT...ACAAGAAAAAGAATCCGTATCCAG...AGAGGACCAGG 7165
Env L A E E E V V I R S V N F T D N A K T I I V Q L N E T S V E I N C T R P N N N T R K R I R I Q R G P G
A1.AU.03.PS1044_Day0 6400
A1.CH.03.HIV_CH_BID_V3538 6365
A1.CY.08.CY236 6397
A1.ES.05.X1608_8 6655
A1.KE.11.DEMA11KE001 6544
A1.RU.11.11RU6950 6704
A1.RW.11.DEMA111RW002 6528
A1.SN.01.DDI579 6375
A1.UG.11.DEMA110UG001 6512
A1.ZA.04.04ZASK162B1 6615
A2.CM.01.01CM_1445MV 6352
A2.CY.94.94CY017_41 6543
B.BR.10.10BR_MG029 6659
B.CA.07.502_1191_03 6621
B.CH.08.M2_0803101_NFLG8 6681
B.CN.10.DEMB10CN002 6547
B.ES.10.DEMB10ES002 6559
B.FR.11.DEMB11FR001 6371
B.GB.05.MM45d213_GN1 6571
B.HK.06.HK003 6531
B.HT.05.05HT_129389 6395
B.JP.12.DEMB12JP001 6581
B.KR.07.HP_18_07JHS10_3909 6819
B.PE.07.502_0525_wg5 6605
B.RU.11.11RU2ln 6699
B.TH.08.MERLBDTRC10 6421
B.US.11.ES38 7102
C.AR.01.ARG4006 6356
C.BR.07.DEMC07BR003 6584
C.BW.00.00BW5031_1 6521
C.CN.10.YNFL19 6544
C.CY.09.CY260 6415
C.ES.08.X2363_2 6637
C.ET.02.02ET_288 6406
C.IN.09.T125_2139 7182
C.KE.00.KER2010 6369
C.MW.09.703010256_CH256.w96 7159
C.TZ.08.707010457_CH457.w8 7146
C.YE.02.02YE511 6357
C.ZA.08.705010534_CH534.w12 AG 7140
C.ZA.10.DEMC10ZA001 6515
C.ZM.11.DEMC11ZM006 6532
D.CM.10.DEMD10CM009 6518
D.CY.06.CY163 6378
D.KE.11.DEMD11KE003 6528
D.KR.04.04KBH8 7090
D.SN.90.SE365 7196
D.TZ.01.A280 6365
D.UG.10.DEMD10UG004 6472
D.UG.11.DEMD11UG003 6513
D.YE.02.02YE516 6345
D.ZA.90.R1 6525
F1.AO.06.AO_06_ANG32 6377
F1.AR.02.ARE933 6456
F1.BR.10.10BR_RJ015 6583
F1.CY.08.CY222 6358
F1.ES.02.ES_X845_4 6611
F1.RO.96.BCI_R07 7237
F1.RU.08.D88_845 6648
F2.CM.02.02CM_0016BBY 6351
F2.CM.10.DEMF210CM001 6446
F2.CM.10.DEMF210CM007 6507
G.BE.96.DRCBL 7125
G.CM.10.DEMG10CM008 6549
G.CN.08.GX_2084_08 6416
G.CU.99.Cu74 6793
G.ES.09.X2634_2 6672
G.GH.03.03GH175G 7241
G.KE.09.DEMG09KE001 6559
G.NG.09.09NG_SC62 6368
G.ZA.01.TV546 6364
H.BE.93.VI991 6561
H.CF.90.056 6490
H.GB.00.00GBAC4001 6691
J.CM.04.04CMU11421 6697
J.SE.93.SE9280_7887 6472
J.SE.94.SE9173_7022 6479
K.CD.97.97ZR_EQTB11 6381
K.CM.96.96CM_MP535 6367
U.CA.01.TV749 7184
U.CA.99.TV721 7210
U.CD.83.83CD003_Z3 6647
U.CD.90.90CD121E12 6327
U.CY.05.CY090 6369
U.CY.08.CY223 6446
U.ES.10.DEURF10DZ001 6418
U.GR.99.99GR303 6452
U.NL.95.U_NL_95_H10986_D1 7056
01_AE.AF.07.569M 6377
01_AE.CF.90.90CF11697 7129
01_AE.CN.10.YNFL03 6555
01_AE.HK.04.HK001 6535
01_AE.IR.10.10IR.THR48F 6364
01_AE.JP.x.JRC77AE 7165
01_AE.TH.09.AA090a_WG11 6596
01_AE.TH.90.CM240 6739
01_AE.US.05.306163_FL 6314
01_AE.VN.98.98VNND15 6432

```
                                                                                                              V3 loop start
B.FR.83.HXB2                 CTAGCAGAA...GAAGAGGTAGTAATTAGATCTGTCAATTTCACGGACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTA......GAAATTAATTGTACAAGA......CCCAACAACAAT...ACAAGAAAAAGAATCCGTATCCAG...AGAGGACCAGG 7165
Env                          _L__A__E__.__E__E__V__V__I__R__S__V__N__F__T__D__N__A__K__T__I__I__V__Q__L__N__T__S__V__.__.__E__I__N__C__T__R__.__.__P__N__N__N__.__T__R__K__R__I__R__I__Q__.__R__G__P__G
02_AG.CM.10.DE00210CM013     ---------...------A--A-G----------AA---A----AA-T-----C---------C--------T----GGAGC-----......A-------------C---.......--T---------...---------G-TG-G-A-.........CT--------- 6595
02_AG.CY.09.CY256            AC--------...-G-A-CA--AC----------AA---A-G--AA------C----T------------T-AGCTGATC----G......-------------C---.......--T-G--------...-------G----G-G----GGGA...-T--------- 6351
02_AG.ES.06.P1423            ------A-G...--CC--A--A-G----------AA---A--T-A-------C-----T-------------G-T-GTC-----......A-T----C-----TC--G.......--T----------...------G---TG-A---.........-T--------- 6581
02_AG.GW.05.CC_0048          ----------..A-----A---A-----------AA---A----AA------C----T-----------AT-A--AGAGC-----......A-C----CC----GC---.......--T--------C...------------TG-A---.........-T--------- 6548
02_AG.KR.12.12MHR9           ----------...-G-A------CG--------C-AA---A----AA-----TC----A-------------T-A--T-AGC-----......C-------------C---.......--T----------...------------TG-A--C.........-T--------- 6833
02_AG.LR.x.POC44951          ----------...A--A--A---------------AA---A----AA------C----------------T--G-T-AGC-----......AG------------C---.......--T----------...------------TG-A--A.........-T--------- 7142
02_AG.NG.09.09NG_SC61        ----------...A----T-----G--C-------AA---A--T-AA------C----------------T-A--G-GTC-----......A-------------C---.......--T----------...---------G-TG-A-A-.........-T--------- 6360
02_AG.NG.x.IBNG              ----------...-G---A-----G--C-------AA---A----AA------C----------------T--GCT-ATC-----......A-------------C---.......--T----------...---------G-TG-A-A-.........-T--------- 6683
02_AG.SE.94.SE7812           ----------...------A----G----------AA----T--AA------C----T-------------T--C-TGA---------......A-------------C---.......--TGG--------...------------TG-A---.........-T--------- 6525
02_AG.SN.98.98SE_MP1211      ----------...-G---TA----G----------AA---A--T-AA------C-----------------T----T-AGC-----......TGG-----------C---.......--T-G--------...------------TG-A---.........-T--------- 6352
03_AB.RU.97.KAL153_2         ----------...-------------------------------------A-------------------------AGA-C----G......---------------------.......--------------...---------G-T--T-A-.........-T--------- 6357
04_cpx.CY.94.94CY032_3       T-----ACG...-------------------AAA---A----A------A-C----AT---------------TGCA-AGG-----......A-------------C---.......--TGG--------...------------TG-A-A-.........-T---G----- 6547
05_DF.BE.x.VI1310            ------A--...----GTA--A----C------CAA---A--T-A--T-----A----A-------------C--T--TGA--------......C-T---------------G.........--------------...------------T--A-A-.........TT--------- 6562
06_cpx.AU.96.BFP90           T---------...-G-A-CA--AC-----A-A---AA--CA-----A------A-C----AT--------G-----T--TCAGC------......--------GG-----C---.......---GGT-------...------------T--ATCC.........TTT-------- 7231
07_BC.CN.98.98CN009          ----------...------AA--A------------AA---C-G--AA------TC-----A------------T--T--TCA--------......------GTA-----------.......---GG---T---...------------T--AA-G.........-T--------- 6505
08_BC.CN.97.97CNGX_6F        ----------..AG----A--A------------AA---C-G--AA------TC-----A------------T--T--TCA--------......------GTA-----------.......--------T---...------------T--AA-G.........-T--------- 6347
09_cpx.GH.96.96GH2911        T---------...-----CA--A------------AG---A----AA------TC---GT-------------T--GCTGATC-------......ACG---G-----GTC---.......A-TGG--------...--------C---TG-A---.........-T--------- 6358
10_CD.TZ.96.96TZ_BF061       ----------...------A--A------------AA---C----A--T----------AT--------G-----T--TGAG-------......AC---------CGTG---.......--------T---...--------G---T--AA-G.........-T--------- 6532
11_cpx.CM.95.95CM_1816       ----------...-G-------AG----------AAA--CA----A------------------------AT-TG---GTC-------......C-C-------------C---.......--T----------...------------T--A--G.........-T--------- 6564
12_BF.AR.99.ARMA159          ----------...------TA--A----------CAA--CA--T-A--T---A-A----------------T--T--T--TGA--------......C-G-----C---------.......--------T---...------------T--A-AG.........TT--------- 7172
13_cpx.CM.96.96CM_1849       ----------...-G---AA--AAG------G--AAA--C---T-A-------T-GG--A---------------TGCT--GC-------......AG-------------C---.......--T----------...------------T--A--G.........-T--------- 6581
14_BG.ES.05.X1870            ----------...------AA----------G---AAA--C-------------A------------------------AGACC-----......------------C--------.......--------------...------------G--AAC-.........-TG-------- 6650
15_01B.TH.99.99TH_MU2079     T---------...----------A----C------AG---C------A-------------T-------------A--TGA--------......---------------------.......--------------...------------T--A-A-.........-T---T----- 6531
16_A2D.KR.97.97KR004         ---------GGAA-TA-AAC-A------------CA---A-T--A-----A------A-----------------T-T-CT-AGC------......C-------------C---.......---G--GT-GG-...CA----GG--TG-A--C.........-TT-------- 6531
17_BF.AR.99.ARMA038          ----------...-G---TA--A----C-----AAA---A--T-A--T---A-A----------------C--T--TGA---C----......C-G---------------------.......--------T---...------------T--A-AG.........TT--------- 6378
18_cpx.CU.99.CU76            ------AC-...----A-----A-G--C-------AA---A----A------A-C--G-A-------------T-CG--GC-------......A-C----CC-----C---.......----G---T--C...--G----G---T--A-A-.........-T--------- 6467
19_cpx.CU.99.CU7             ----------...-G-A----GAC-----------AA---A----A-------C----------------T-T-CTRAGC-------......A--------------C--.......--T----------...--------C---TG-A---.........-T--------- 6303
20_BG.CU.99.Cu103            T---------...A-G--AA--A-G----------AA--CC----A-------C---GT--------G-----T--T-A-AG-A--......----------------C---.......--------------...----------G--T--AG-A.........TTT-------- 6629
21_A2D.KE.99.KER2003         T---------...------A-CA-------------AA---C-T--A-------------------------T--TGA--------......-T---------------G.......---GGT-------...--------C--G-T--A-A-.........-T--------- 6360
22_01A1.CM.01.01CM_0001BBY   ----------...A-----A--A-G----------AA-------AA----C--C----T-------------TY-GT--AGC------......ACT----------GC---.......--TGG--------...---------C--T--A---.........-T--------- 6350
23_BG.CU.03.CB118            T---------...A-G--AA--A-G----------AA--CC----AA-------C---GTT--------G-----T--T-GCAG-A--......AC-------------C---.......--------------...---------G--T--AACA.........TTT-------- 6629
24_BG.ES.08.X2456_2          T---------...A-GA-A---A-G----------AA--C---G-AA-------C---GT--------G-----T--TGA-AG-A--......-T-------------C---.......---GG--------...---------GG--TR-AG-A.........-TT-------- 6649
25_cpx.CM.02.1918LE          T---------...------A--------------AA---A-T--AA-------------A---C---G--T--T--TGA--------......ATG------------C---.......--T----------...------------T--A-A-.........-T--------- 6376
26_AU.CD.02.02CD_MBTB047     ----------...-G--GA---AAGG----G----AA----------AA-------C----A--------------T-TG-T-AGG-------......-T------C----GC---.......--------------...------------T--A-A-.........CT--------- 7162
27_cpx.FR.04.04CD_FR_KZS     ----------...------A--A------A----AA---C----A-----A-C-----------------C--T--TGA--------......A-------------C---.......---TT--------...-----G-----TG-AA--.........CT--------- 7152
28_BF.BR.99.BREPM12609       ----------...------A--A------------AA---C----A----------------T--------G-------TGA-A------......-----------------G.......--------------...-G--AC---C-T--A-A-.........CT-----GG-- 6533
29_BF.BR.01.BREPM16704       ----------...-----CA--------A-----AA----T------------------------G--A------GA-A--A--......---------------------.......--------------...------------T--A-A-.........-T--------- 6580
31_BC.BR.04.04BR142          ----------...------A---------------AAA---C-AT-A-------TC-----A-------------TC-G---C-----......A-T----C-----TG---.......--------T---...-----GG----T--AA-G.........-T--------- 6627
32_06A1.EE.01.EE0369         T------G-...---AGTA--A------------AA--CA----A-----A-C-------------G-----T--TCA--------......--C----TA-----C---.......---T-T--T---...--------G---T--A-AC.........-TT-------- 6830
33_01B.ID.07.JKT189_C        ----------...------A--A----C-------AA---C----A--------C-------------G--C--T--T-A---C---......-T---C----------------.......--TTC--------...--------C---T--A-C-.........-T--------- 6478
34_01B.TH.99.OUR2478P        ----------...------A--A----C-------AA---A----AA------A-------------G--C--T--T-A--------......-------C------C---.......---TC--------...--------C---TG-A-A-.........-T--------- 6352
35_AD.AF.07.169H             ----------...-G----A---------------AA---A--T-T---------------------------T--TGAGA-----......AC-------------C---.......--T----------...------------T--A---.........-T--------- 6374
36_cpx.CM.00.00CMNYU830      ------A--...A-T--AAC-A----C--------AA---A----AA------C--G---------A-----T--GT--AGC------......A-T----C----GTC--G......--T----------...---------G-T-CA-C-.........-T--------- 6371
37_cpx.CM.00.00CMNYU926      ----------...-G-A-----A-G----------AA---A----AA----C--C-----------------AT-T--TGAGA-----......----------------C--G......TTT--------C...------------TG-A--C.........-T--------- 6353
38_BF1.UY.03.UY03_3389       ----------...------A--A----C------CAA---A--T-A--T-----A-----------------CT-T--TGA--------......C-R-----C---------.......--------T---...------------T--A-A-.........TT--------- 6579
39_BF.BR.04.04BRRJ179        T---------...----A--A--A----C------CAA---A--T-A--T-----A-----------------C--T--TGA--------......C-G------------------.......--T---------...-----G-----T--A-A-.........TT---C----- 6651
40_BF.BR.05.05BRRJ055        ----------...----A-TA-------C--G---CAA---A--T-A--T-----A---GT-------------T--T--TGA--------......C-G------------------.......--------------...-------C---T--AAA-.........-T--------- 6689
42_BF.LU.06.luBF_18_06       ----------...-----TA--A----C------CAA---A--T-A--------A-----------------CT-T--TGA--------......C-G--------CT-G---.......--------T---...------------T--A-C-.........-T--------- 6745
43_02G.SA.03.J11223          T---------...-----AA--------------AAA--C---T-A-------CC--GT-C--------G-----T--TRGC---A--......------------C--C---.......--------------...-----R-----T--AG-A.........TTC--T--T-- 6687
44_BF.CL.00.CH80             ----------...----A--A--A----C------CAA---A----A--T---A-A-----------------C--T--TGA--------......C-G------------------.......--------T---...------------T--A-A-.........-T--------- 6613
45_cpx.FR.04.04FR_AUK        ----------...----A--A--A-G----------AA---C----AA------TC--G-----------G-----TGTT-GGC------......C-----------T-C---.......--T-CA-------...--------C---TG-G---.........-T--------- 7124
46_BF.BR.07.07BR_FPS625      ----------...-----TA--A----C------CAA---A----A--T-----A----T-------------C--T--TGA------G......C-G------------------.......-----T-------...------------T--A-AA.........-T--------- 6583
47_BF.ES.08.P1942            ----------...-----TA--A----C--G---CAA---A--T-A--T-----A----G-------------T----TGA---A--......AGC----------T---.......---GG---T---...------------T--A-A-.........-TG-------- 6585
48_01B.MY.07.07MYKT021       ----------...------A--A----C-------AA---C----AA---------G-A----------G--C--T--T-A-----------......------C--------C---.......---TC--C----...--------------G--GAC-.........-TG-------- 6409
49_cpx.GM.03.N26677          A---------...A---GAA--A------------AA---A--T-A--T---A----G-A------------A--TG-T-A-C------......AC-----G----TC---.......---GC--------...---------GG-T--A-AC.........TT--------- 6598
50_A1D.GB.10.12792           ----------...A----TA--A------------AA---A----AA-------C-----T-----------A--TGC--GCC----G......A-C----G----GTT---.......--T----------...--------G-----TG-A---.........-T---G----- 6562
51_01B.SG.11.11SG_HM021      ----------...------A--A----C-------AA---C----AA-------C-------------G--C--T--T-A--------......-------C--------C---.......---TT----T--...--------G-C---T-CA---.........-T---G----- 6371
52_01B.MY.03.03MYKL018_1     ----------...------A--A----C-------AA---A----A--T---A-C-------------G--C--T--TMA--------......--G--C--------C---.......---TC--------...--------C---T--AG--.........-TG-------- 6512
53_01B.MY.11.11FIR164        ----------...------A--A----C-------AA---C----A--------C-------------G--A--T--T-A--------......-------C--C-----C---.......--AGG--------...--------C---TT-AAC-.........-T--------- 6526
54_01B.MY.09.09MYSB023       ----------...-----CA--A----C-------AA---C----AA-------C-G--A--------G--C--T--TGA--------......--------G-----C---.......---TC--------...------------T--A-A-.........-T--------- 6579
55_01B.CN.10.HNCS102056      ----------...------A--A----C-------AG---C----AA-------C-------------G--C--T--T-AC-------......AC---C-------C---.......---GC--------...--------C-G-T--AAC-.........-T--------- 6463
56_cpx.FR.10.URF5_patient_A  ----------...A-----A--A-G--C-------AA---A-A--AA-------C----T-------------AT----TGAG------......TC-------------C---.......--T-G--------...--------------C--A-A-.........TT---G----- 6507
57_BC.CN.09.09YNLX19sg       ---T-------...--T-GAA--A-----------CAAG--C-G-AAA-------C-----A-----------T-----T-A--------......------G-A----T-----------.......----G---T---...-G---------TG-AA-G.........-T--------- 6360
58_01B.MY.09.09MYPR37        ----------...-G----A--A----C-------AA---C----AA-------C-------------G--C--T--T-A--------......-------C-----G-C---.......---TC-G-------...--------C---TG-A---.........-T--------- 6549
59_01B.CN.09.09LNA423        T---------...---------A--G---------AGA---------AA-------------A-------------T----TGA--------......-T--------C--------.......--------------...-----G-----T--A-A-.........-TG-------- 6346
60_BC.IT.11.BAV499           ----------...-G----A--A-----------AAA-----A--A--------C-----A-----------T--T--TGA-G-----......R-C----C----GA---.........-----T--T---...-------G----T--AA-G.........-T--------- 7140
61_BC.CN.10.JL100010         ----------...-G---AA--A------------AA---C-GG-A------TC-----A-------------T--TCA--------......------GTA-----------.......--------T---...-------G----T--AA-G.........-T--------- 6573
62_BC.CN.10.YNFL13           ----------...A-----A--A------C-----AA---C-G--A--------C-----A--C--------T--T--T-A--------......-T----G-A-----------.......--------T---...------------T--AA-G.........-T--------- 6481
63_02A1.RU.10.10RU6637       --------G...-G-------ACG-----------AA---C----AA-------C--GGT-------------T--TGAG-------......-C-------------C---.......------T-----...------GT-G-TG-A-A-.........-T--------- 6691
64_BC.CN.09.YNFL31           ----------...------A--A-----------AAA---C-G--A--------C----TA-----------T--T--TGA--------......------GTA-----------.......---G---T---...------------T--AA-G.........-T--------- 6521
65_cpx.CN.10.YNFL01          ----------...-----CA--A------------AA---C-G--AA------TC-----A-----------T--T--TGA--------......------GTA-----------.......---C----T---...------------T--AA-G.........-T--------- 6565
67_01B.CN.11.ANHUI_HF115     ----------...------A--A----C-------AG---C--R-AA-------C----T---------G--C--T--T-A--------......------C--------C---.......---TT---M---...--------C---TG-AA-A.........-T--------- 6505
68_01B.CN.11.ANHUI_WH73      ----------...------A--A----C-------AG---C----AA-------C-------------G--C--T--T-A--------......------C--------C---.......---TC--------...--------C---TG-AA-A.........-T--------- 6487
70_BF1.BR.10.10BR_PE004      ----------...------A------------CAG----C-------T--A-----T---T--A---------TG---------......---------------------.......--------------...---------G-T--A-A-.........-TG-------- 6608
71_BF1.BR.10.10BR_PE008      ----------...-----------------------AA--------A----C-------AT------------T---A-A--------......A---------------G.........----G-----------...--------G---T--A-A-.........-T--------- 6828
72_BF1.BR.10.10BR_MG002      ----------...------A--A----C------CAA---A--T-A--T-----A-----T------------C--T--TGA--------......C-G------------------.......--------------...-------C--G-T--A-A-.........-T--------- 6681
O.BE.87.ANT70                --CT-TA--...-G-A-AA--AG---G-TGG-AAAAG--A-TTT---AGG--GA----AT--C-----GACC--A---T-TA-CC--......A-C--G-CC---GA---.......--AC-A-TAG-C......-T-C--GAG--GA-A.........-T---T----.. 7189
O.CM.98.98CMA104             --CT-CA--...-GGA-AA--AG-----TGGGAAAAG--A-T--A-CTGG-T--CGC-AT--CC----GACT--A--TT-TA--A--......A-T--A-CC---GAG---.......--ATT---TCTG...--GGT-G--CAG--GAAG.........-T---C----.. 6648
O.CM.98.98CMABB141           --TT-GA--...-G-A--A--AG-C-G-TGG-AAAA---A-TT-A--T---AGCC---AT--C-----GACC--A--T--TA--A--......AGC--G-CC--CCAG---.......--AGGA--T-CA.....--GT-C-GGAG--AAAA.........-T---T----.. 6707
O.CM.98.98CMABB212           --CT-TAG-...-G--G-A--AG-----TGGGAACA--CA-T--A----G---A--G-AT--C------ACC--A-----TC------......A-C--G-CC--C--G-------......GAAG-A-GA......-A--TTC--GAG--A-AC.........GT---C----.. 6694
O.CM.98.98CMU5337            --CT-T---...-G-A-AA--AG-----TGGGAAAA--CA-TT-A--TTC---GG---AT--CT---GACC--A--TTATC-------......A-C--A-C----CAG---.......--A...--TGT-.....-GGT-C--GAG--AAAA.........-T---T----.. 6652
O.CM.99.99CMU4122            --CT-T---...-G-A-AA--AG-----TGGGAAAA---A-T--A----G--GA----AT-------GACC--AG-TT-TG--A--......A-C--A-CC---GAG---.......--ATGG--TC-G...T--GT-C--GAGC-GAAA.........-T---T----.. 6629
O.FR.92.VAU                  --CT-TA-G...-GGA-TA--AC---C-TGGGGAAGG--A-TT-A----G--GGG-G-A---CC-GA--ACC-----T--TGA-A-G......AC---AGCA--GAG---.......--AGGA-GTC-G...----T-CG--AG--AATG.........GC---T----.. 6731
O.GA.11.11Gab6352            A--T-TA--...-G-A-AA--AG---A-TGGGAACA---A-TT---C-TC---GC---AT--CT-----ACT--A--T--TA--A--......A-G--G-CC---GA------.......--AGGGG----CGTG---GT-GG-CAGG-GG-C.........CC---T----.. 6388
O.SN.99.99SE_MP1299          --CT-T---...-G-A-CA--AG-----TGGGAAAA---A-TT-A------ATG----AT--C-----GACC--A--TT-TA--A--......A-C--G-CC---GTG---......-AAGGA--TC-G...T--GT-C--GAG--A-AA.........-T---T----.. 7230
O.US.10.LTNP                 A-CT-TA-G...-G-A-AA--AG---G-TGG-AAAA---A-T--A------A-G----AT--CC-----ACC--A--TTATA-CA--......-----G-TC---GAG---.......--ATGG-GAG-G......-T-GCTCCG--GA-C.........GT---G---.. 7086

N.CM.02.DJO0131              T---ATA--......--TA-T--T------AA-............A-T-G--G-GG--ATC-GT-G--G--ATG---TGAGATA--G......AC---G-----------G.......--AGGG--T---.....---G--GG-CAGG-A-AA.........-T-----T-C 6628
N.CM.02.SJGddd               T--AATAC-....--T--TA-TA-T------AA-............--T-G-AG-GGT-ATC--T-G--G--ATG---T-AGA-A--G......TC---A-----------G.......--AGGA--T---....------GG-CA-G-G-AA.........-T-----T-C 6609
N.CM.04.04CM_1015_04         ----ATAC-....--T--TA-T--TM-----YR-...........S-T...AAWG-T-AT--GT-G--G--ATG---TMAGA-A--G......TC---A-----------G.......--AGGG--T----...Y-G--GG-CAGG-G-AG.........-T-----T-C 6621
N.CM.06.U14296               T-------T.....--TA-T--TG----..............A-T---AGCGGC-AT--GT-G--G---TG--RTGAGA-A--G......TC---A-----------G.......--AGGA--T---.....---G-GGG-CAGG-G-AG.........-T-----T-C 6637
N.CM.06.U14842               ----ATAC-....--T--TA-T--T-----GAA-AGT-G-GAT-GCA-T-CGAG-GGTGATC-GT-G--G--ATG---T-AGA-A--G......TC---A-----------G.......--AGGA--T---.....---G--GG-CAGG-A-AA.........-T-----T-- 6662
N.FR.11.N1_FR_2011           T---GCAC-........--AA-T--TR---R-AA-............--T-G-CG-GGT-ATC-GTGG--G--ATG---TGAGA-A--G......TC---G-----------G.......--AGGA--T---.....---G-TGG-CAGG-G-AG.........-T-----T-C 6531

P.CM.06.U14788               ------A-G...-G----CCCT--G-A-TTA--CAG---G-GT-A--T-CA-GA---GT---C-------A-AT-A--TGAGAG---G......TC-C-C-CC---T----.......--AGGT--T--C.........------GG-CA-G-G-AA.........CT----GTC.. 6670
P.FR.09.RBF168               --------G...-------CCCT--G-A-TTA--CAA---G-GT-A----CAAGAT-TGTT--C-------A-AT-A--T-AGAA---G......TC---C-CC---GT-----.......--AGGT--T--C.........------GG-CA-G-G-AA.........-T-----C.. 7222

CPZ.CD.06.BF1167             G---A-------...-G---AAC-AAGGCATAC-A---AG-AGC--AAC-T-----A...C-AT--T--A--A-AT-C--T-A-G-AACGGTCTTTA-C--A-C----GT----.......---GG------C.........------GG-CAGG-A-AG.........-T---G----- 7284
CPZ.CM.05.SIVcpzMT145        AC------C...A-TC--ACT--GGCC---ATA...G-CCCTT-A--G---TTA.....GC----A-------C--AGATC-A--------......A-G---CA---GG---.......--AGGA--T--C.........------GGTCA---A-AA.........-T--------C 6680
CPZ.GA.88.GAB1               T----T------...-G-A-CA--ACTG-A---GTG-AG---AAGT-AA-A---A--G-TGT-TGG-----------AGTAGA-G-A--T......AGTC-A-----CAT---.......--AGGA-------.........------GGG-GAGG-G-AG.........-T---G----- 7220
CPZ.TZ.06.SIVcpzTAN13        TAT-A-A-G...--GA-AACGAAGG-GTATCA-ACA---G-A--AA-T---C-A...C-AC--T----TA--T-C--TGAGA-AA--......C-C--A-CC---GA----......A-AGGG-------......------GGTCAGG-A-AG.........-T--------- 6830
CPZ.US.85.US_Marilyn         T-----ACT...A--A-T--TACTG-A---AG-AAA--C--TG-A---.........-TT--TC-------AT-CTCAGAGGGA--C......A-T--G-C-----TT---.......--AGGA-------...---GTGGG--ATG-A-AA.........CT--------- 7194

GOR.CM.07.SIVgor2139_287     T---------...-GGA-T-CCAC---A-TCAGCAAA---AGT-GT--T-G--GA---GAT--T-----CA-A--AGC--G-C-A---.............CC---GA---G......A-AGGT--T---...------GG-CAG----AA.........-T---G--C.. 6733
GOR.CM.07.SIVgorCP2135con    T-----A--...--GC-T-CCAC---A-TTAGCAAA---AGT--T--T-G--GA---GAT--T-----CA-A--AGC--G---A---.............CC---GA---G......A-A-GT-------...------GG-CAG----AA.........GT---G--C.. 6714
```

```
                                                                                V3 loop end
B.FR.83.HXB2                GAGAGCATTTGTTACAATAGGAAAA..............ATAGGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATAACACTTTAAAACAGATAGCTAGCAAATTAAGA......GAACAATTTGGA.......................AAT. 7289
Env                         __R__A__F__V__T__I__G__K__.__.__.__.__I__G__N__M__R__Q__A__H__C__N__I__S__R__A__K__W__N__N__T__L__K__Q__I__A__S__K__L__R__.__.__E__Q__F__G__.__.__.__.__.__.__.__.__N__
A1.AU.03.PS1044_Day0        ACA-A----CTA-G---C---T............GAAATA-----GG----T---A-----------GTGG-C-A--A-A--G----------A-A---G---A--G---T-GTAC--------......AC-T-C---AAG........................ 6524
A1.CH.03.HIV_CH_BID_V3538   ACAG--T--CTA-G----G-AC............GACATAC----GG-C--A---A-------G-----TG-C-AC---A--GC----G-A---------C--A---------CAC--------......ATTATG-ACAAC........................ 6489
A1.CY.08.CY236              ACA------CTA-G---C---T............GATATT-----G-----A---A----C------G-GG-C-A-G--A--G----------AG--------GG-G---T-GAAC--------......AC-T-C---A-T........................ 6521
A1.ES.05.X1608_8            A---A--C-CTA-G--GC--AT............AAAATA-----G-----A-----------C-----TG-C--C---A--C----------A---A-GC-GG--G---GCGAAC-----GAG......AC-T-C---AAC........................ 6779
A1.KE.11.DEMA11KE001        ACA------CTA-G---G-AAT............GCCATA-----GG----A---A-------------TG-C--C-A-T-----------G--------CG----G-----GAAC--------......A--T-C---AA-........................ 6668
A1.RU.11.11RU6950           ACA-A-C--CTA-G---C---T............GATGTA-Y---GG-C--A---AG----T-------TG-C--C---A--G---------G---C---C--A-----AG--CAC--------......AG-T-C---CAT........................ 6828
A1.RW.11.DEMA111RW002       ACA-A----CTA-G---C---T............GAGATA-----GG----A-----------------TG-C-------A-G-C---TT-G-AG---------A-T---T-C-AAC-----G--......A----C-----G........................ 6652
A1.SN.01.DDI579             ACA------CTA-G---C---T............GACATA-----GG----A--G--------------TG-C-------A---T----C-G-G-----C-----G------AAC------A-......AG---C---AAC........................ 6499
A1.UG.11.DEMA110UG001       ACA-A----CTA-GT--C---T............GACATA-----GG----A-----------C-----TG-C------AAG----------G-------G-G-G--G---TC---G-------G......A--T-C-----G........................ 6636
A1.ZA.04.04ZASK162B1        ACA-A----CT--G---C---C............GACATC--T--GG----A-----------------TG-C--C-A-A.GG------TG--AGT-C-GC-----G---T---A---------......A-GG------A-........................ 6738
A2.CM.01.01CM_1445MV        ACA---C--CTA-----ATAAT............AAAATA-----GG----A-----------C-----T--C--C-A----G-C------GTT------C--A--G---T--AAC-----CAC......AGC--C--CCCT........................ 6476
A2.CY.94.94CY017_41         ACA---C--CTA-...-C-AAT............GAAATA-----GG-C--A------------------T--C-AC-A-A--TT-------G---------C--A--G-----GAAC--------.......--GA---CCCT........................ 6664

B.BR.10.10BR_MG029          ----------TA-GG--C-...............GAAATA------G---A---G---------C--------------TG-C------R-S-------R-R--G---T-GMT---------......A--M---A--R-........................ 6780
B.CA.07.502_1191_03         -------A--TA-G---C---G............GATRTA------G---A-----------C--C---------G-----C------GG---------G-A--G-----TA---------......---------AAT........................ 6745
B.CH.08.M2_0803101_NFLG8    A---A----T--G---CT---.............GATATA-----GG---A-----------------C-----A-G-----C-------GC-------C----------T-GCA---------......---------A-G.......................GGA. 6808
B.CN.10.DEMB10CN002         -G---T----TA-----C----............GATATA-----------A-----------------TG---A----A---GC---C--G-------------------------CA--------G......---A---GGCCT........................ 6671
B.ES.10.DEMB10ES002         -----T----TA-G---C----............GACATA------G---A---------T-C-------------CA----G-T------G-A-------C-TG------C-TA---------......---A-----AAC........................ 6683
B.FR.11.DEMB11FR001         ----------TA-----C----............GAAATA------G---A---------------------C---------A--G-T-----------A----G-AGT-------TA-------A-......---------AAT........................ 6495
B.GB.05.MM45d213_GN1        ACA------GTA-----C-...............AACATA------G---A----------------------C-----GAAGG--------CG-----------T-----T--AA-----------......-C-----A---G.......................-G-. 6695
B.HK.06.HK003               ----------TA-G---C----............GACATA------G---A-------------------------A-----------------C-------AG-----TG------------......---A--------........................ 6655
B.HT.05.05HT_129389         ---------CTA-G---C----............GCTATA------G---A---A-------------------G--T--G----G-A--G-TG--GGG-A--G-----CA-------A-......--G------AAT........................ 6519
B.JP.12.DEMB12JP001         -CA-A----ATA-G---C----............GACATA-----GG-C--A-----------------------CA----GC------------------G----G-G-T-GCA-------A-......---------AAT........................ 6705
B.KR.07.HP_18_07JHS10_3909  ---TA--A--TA-G---C----............GACATA------G---A-------------------------A---------C----------G--------T--TA-----------......------------G........................ 6943
B.PE.07.502_0525_wg5        --A-------TA-G---C----............GATATA-----CG---A-------------------------A--GGT------G----------TA-----AGGGAA-------A-......AGC---------........................ 6729
B.RU.11.11RU21n             --A-------TA-G---C----............GACATA------G---A------------------TG------A-AT----------------A----T----G---TGG-------------......---------AAG........................ 6823
B.TH.08.MERLBDTRC10         --A-------TA-----C----............GACATA------G---A---A-----------C-G-----A----A--G-T---G-A-------------------G-------A--------TG......------------G........................ 6545
B.US.11.ES38                ----------TA-G---CG---............AGCATA-----CG---A--G--G-----------------------AGG----------A----------T-TG-----GA-------G--......---------AAC.......................-CA. 7229

C.AR.01.ARG4006             ACA-A----CTA-G---C----............GACATA------G---A---G------------T-------C-A--GC----------A------C-TG-AT--AG--AA-------A-......A-G-TC--CCCT........................ 6480
C.BR.07.DEMC07BR003         ACA------CTA-G---C----............GACATAG-----G---A-----G-------------------------A--GC-------C--A--C---C--AGAG--AG--AA-G----TC-......--GT-C--CCCT........................ 6708
C.BW.00.00BW5031_1          ACA-A----CT--G---C----............GGCATA-----GG-A--A----G----------G----C-A-G--A-G-------CCG-A------C--A----AG-GAA-----------......-G-T-C---...................... 6642
C.CN.10.YNFL19              ACA-A----CTA-G---C----............GATATA------G-C--A----------T-------------GC-AGT-------------A------C--GG-G--AG--AA------GC-......----TC--CCCT........................ 6668
C.CY.09.CY260               ACA-A----CTA--A---G---............GAAGTA-----------A-------------------------GA-CA---C-----C------------C--AG-G--AG-GAA-----------......---T-C--CCCC........................ 6539
C.ES.08.X2363_2             ACA------CTA-G---C----............GAAATA------G---A---A-----------T------GGA--ATCT---------A------C--AT-G--AG--AA-------A-......-----C--CCCT........................ 6761
C.ET.02.02ET_288            ACA-A----CTA-G---C----............GAAATA------G---A-----------------------G---AC-G----------G---C--G--A-TG--AAGGAA-G----CA-......A-G-----CCCT........................ 6530
C.IN.09.T125_2139           ACA-A----CTA-G---C----............GAGATA------G-C--A--GA-----T-----G-G----A-GA-CAG------------A--A---CG-G--G------AA-GT---GC-......A----C--CCCT........................ 7306
C.KE.00.KER2010             ACA-A-----TA-G---C---T............GACATA------G-C--A---A-----T----------------A--AG-------------T------C------G--AGG-AA------GA-......-----C--CCCT........................ 6493
C.MW.09.703010256_CH256.w96 ACA-A----CT--G---C----............GACATA------G---A---------T-C----C-----A-G-GA-TC--------G-----------TAG-G--AGAGAA-C----CA-......AGG--C--CCCT........................ 7283
C.TZ.08.707010457_CH457.w8  ACA-A----CTA-G---C----............GAAATA------G---A---A-----------------------AGT--G------G-A------G-GAG----AG--A-------GA-......-----C--CCCT........................ 7270
C.YE.02.02YE511             ACA-A----CTA-G---C----............AGCATA------G---A-----------------------GA-AAG---------G----C---C--G------GA-AA-------A-......-----C--CCCT........................ 6481
C.ZA.08.705010534_CH534.w12 ACA-A-----TA-G---C----............GATATA------G-C--A--GG-----------------------GA--A--C----------A---C--C-CAT-G--AG-GAA-------A-......-----C--CCCT........................ 7264
C.ZA.10.DEMC10ZA001         ACA-A----CTA-G---C---T............GACATA------G-C--A---------T-------T-------A--AGG----------A---C--G--AG-G--AGG-AA-G----GA-......-----C--CCCT........................ 6639
C.ZM.11.DEMC11ZM006         ACA-A----CTA-G---C----............GACATA------G-C--A-----------------TG-----GA-AAG------G--G-A--C---C-TAG-G-------A-----------......---T-C--CCCT........................ 6656

D.CM.10.DEMD10CM009         AC-GT--C-CTA-----C-...............GACATA------G-C--A---A-C---T-------TG-----...A----T------------C--C-T--AG----AGAA------G--......A-G-TTC--AAC........................ 6636
D.CY.06.CY163               -CAG---C-CTA-G---C-AC-............AACATA---------C--A---G------------T--------T-G-G-C------------C--C----AG----A-AA------G--......A---TTC--AAC........................ 6502
D.KE.11.DEMD11KE003         -C-------CT------C-...............AACATAG---GTCC--A-----G-----------TC--------AT-G---------GA------G-TAG-G--AT-GAA-------A-......A-C-TTAC-AAC........................ 6649
D.KR.04.04KBH8              ACA----C-CT--G---C-AC-............AATGTA------G---A---AG-----------T--C-----CA-----------G-A------C------T----A-AA------G--......--C-TC---AAC........................ 7214
D.SN.90.SE365               -CAG-T-C-CCA-----C-A-G...............GTA-A----G---A-------------------------A--A--G----------A------C------G---T--AA-------A-......--C-TC---AAC........................ 7317
D.TZ.01.A280                -CA-A----CT-----GC-...............GAGGTA-C----G---A---A-----T---------------G-----G-----G----A------C----AG------CA---C--G--......--C-TTC--AAC........................ 6486
D.UG.10.DEMD10UG004         -GC----A-CTA-----C-...............AAGACA------G---A-----G-----------T---------T-G-C---G---C---C---C-----G------AA-----------......--C-TTC--AAC........................ 6593
D.UG.11.DEMD11UG003         -C-------CTA-...-C-AAC............GACATA------G---A-----------------TG--------A--GCT---G----A------C------G-----GCA------G--......A-C-TT-A-AAC........................ 6634
D.YE.02.02YE516             -CAG---C-CTA-----C-...............AATGTA------G-C--A-----------------T------G--A--C---------------C--C--A---------AA------G--......A---TT-A-AAC........................ 6466
D.ZA.90.R1                  -CA-A----CTA-G---GCA-G............ATGATA------G-C--A-------------------------GA--A---------G-GA------C------G------TA------G--......A-C-TTC--AAC........................ 6649

F1.AO.06.AO_06_ANG32        ACA------CTA-G---C---T............GACATA------G---A--GA-G-----------T------G-G-ACC---------------CC--C--A--G--AGGGCA--G----A-......TCT-TT...CT-........................ 6498
F1.AR.02.ARE933             ACA-------TACG---C----............GACATA---------C--C---A-G-----------TG-----G--ATGC----------CA--GC--G---G-G--AAGGCA-------AG......CCT-TT--CCCT........................ 6580
F1.BR.10.10BR_RJ015         -CA-------TA-G---C----............GACATA------G-C--T---A-G-------------------G--TT-C----------A--A---G------G--AAGGCA--GC----G......CCTTTT--CCCT........................ 6707
F1.CY.08.CY222              ACA------CTA-G---CG--T............GCGATA------G---A---A-G--------C----------G-GA--C----------A---A-GG---G-G-GCGGGAA--G---C-G......TCTT-T---AAT........................ 6482
F1.ES.02.ES_X845_4          ACA------CTA-G---C---T............GACATA------G---A---A-G----C-------------G--A--C-----------------C---CA-TG---AGGCA--G----AG......TCT--T--CCCT........................ 6735
F1.RO.96.BCI_R07            ACAC-----CCA-----C--AC............AGGATA------G---A---A-----------T-G---A-G--A--C---------C----G--C---G-G--CAAGCA--G----AG......-CTT-T--CCAG........................ 7361
F1.RU.08.D88_845            -GC------ATA-G---C---T............GCGATA------G-C--A---A-----------T------G--T---------G-AG-A---------A-YG--AG--AA--G-----G......--C--TG---TT.......................G 6773
F2.CM.02.02CM_0016BBY       ACA--T----TA-G---C---T............GAGATA------G-C--A---A-G----------C------A-G--A--CTG------GCA--------TAG-G-----GCAG-GG-C-A-......A-C-TC...ACT........................ 6472
F2.CM.10.DEMF210CM001       A--------CT--G-C-C---T............GACATA------G---A---A-G---T-------C-----A-G--A--CTG------G----------C-A-----AT-GCAG-G-AC-A-.......--CA--AGCAAT........................ 6570
F2.CM.10.DEMF210CM007       ACA--T---CTA-G---C---T............GAAGTA------G---A---A-----TT------T-----A-G--A--C-G------G----------C-A-----AAGGAA--G--CCA-AACCTAA-GAT-CAAAAT........................ 6637

G.BE.96.DRCBL               ACA------CTA-----C----............GAAGTA------G-C--A---A-G-----------TG-----T-GA------------G-G--GC---GGG-TG-CCAGGCA---C--CAG.........---T-C---ATT........................ 7249
G.CM.10.DEMG10CM008         ACA--TG--CTA-G---C---T............GCTATA------G---A---G-G---T------GTG---AATT-CA-C-G------G-G--GA--C-GA-TG-GA-AGAA---C---A-.........---GTCC--AA-........................ 6673
G.CN.08.GX_2084_08          ACA-A-G--CTA-G-G-C---G............GGCATA-----GG---A---G------------TG-A---GAGAAG------T--G-G-TA---C-GA--G--AAG-TAC--C--G--......A-GGTC---AAT........................ 6540
G.CU.99.Cu74                ACA---G--CTA-G---C---T............GACATA------G---A---------------G-GG---A-G--A-----------G-G--AC--C-GA--G--AAAGAA---C---AG.........---AGC-ACAAC........................ 6917
G.ES.09.X2634_2             A-A------CTA-G---C---T............GGCATA------G---A--G--------------T------G--A--G-C-----A--T-TGA--GG-A-TG--A--GAAGG-C---A-.........---ATC-----C........................ 6796
G.GH.03.03GH175G            ACA------CTA-G---C---T............GACATA---------C--A--------------TTG---A-G---A----------C-GGTGA--C--A--G-CAAA-CAC-TC--GAG.........---ATC-A-AAT........................ 7365
G.KE.09.DEMG09KE001         ACA-A-G--CTA-G---C---G............GCAATA------G---A-----G-----------TG-----G--A--GC-----GA--T-TGA-GC-GGGAG--AAAGCA---C--GA-......-G-ATC---AAT..................ACA-G-. 6689
G.NG.09.09NG_SC62           ACA-A-G--CTA-G--CA---T............GACATC-----------A-----------------TG-A---GC-----------A--T-TGA--C-GA-CG--A-GGAACT-C---A-......A-GACC-A-AAC........................ 6492
G.ZA.01.TV546               ACA------CTA-G---C---T............GACATA-----GG-C--A------------------G----C-A-A--G-----G-GGG--TG-----GA--G--AAAGAA---C----C......A-CATC---CAC........................ 6488

H.BE.93.VI991               ACA------CTA-G---C---T............GACATC------G---A----G----T-------T------G--AA-C--------G-G--C---C-CA--G--ATC-C-------G--......AGCT-C----AC........................ 6685
H.CF.90.056                 AC-------CTA-G---C---T............GACATC------G---A-----------------T--------A--G-C-------G------C-C---G---T--CAC-----G--......AT---C--GAAC........................ 6614
H.GB.00.00GBAC4001          ACA------CTA-GT--C---T............GAAGTC---------C--A-----------------T------G--A--C----------G--C---T-CA-TG---TC-A-C----------......A-G--C---AAT........................ 6815

J.CM.04.04CMU11421          ACA--TGC-CT--G---C--AG............AGAGTAG---GG---A---AG---T-C----TG-------C...--C-----C-GAGT----GGTACAG-----GCA---C----------......-G-T-T---AAT.......................G 6819
J.SE.93.SE9280_7887         ACA--TGC-CTACG---C----............GAAATA------------A---G--A-------------------GA-AG-G-T----G----------CGTAGAG-----CA---C-----......-----C---AAT........................ 6596
J.SE.94.SE9173_7022         ACA--TGC-CTACG---C---G............GAAATA------G---A--GA-----T-------------------AA-G-T-----------------CGTAGAG-----AA---C-----......-----C---AAT........................ 6603

K.CD.97.97ZR_EQTB11         A--------CTA-G---C---T............GACATA------G---A--G--G--G----------------CG---G-C-G---------A---G----C--AG--AAG-AAG-----G--......A----C---AAC........................ 6505
K.CM.96.96CM_MP535          A-A------CTA-----C---T............GATATA------G---A-----G--------C----------G---A-----------C-TG-------GCAGAG--AAGGAA--GC---A-......-----T---AAG........................ 6491

U.CA.01.TV749               ACA-A----CT--G---C-...............GACATA------G---A---------T--------T-------TAA---C----CAG--------C--A--G-----GA----C-----......-----C---AAG........................ 7305
U.CA.99.TV721               ACA-A----CTA-----C---C............GAAGTA------------A-------------------T-------AGAGC--T----CA--------------G--G-----TA---C---A-......---A----AAAT........................ 7334
U.CD.83.83CD003_Z3          A-A--T---CTA-G---A---T............GGAATA-C-.........G-------------------CAGAT-GGG------GG--T------C-----G-----T-GC---------......AG-------AAT........................ 6762
U.CD.90.90CD121E12          AC---T---CTA-G---C-A-T............AACATA------G---A-------------------C---CA-A-TGGC--------G-GAG-----C------G-----GATGC---------......AG---C---AAT........................ 6451
U.CY.05.CY090               ACAGA----CT--G---C---T............GAAATA------G---A-----G------------T-----------A-GC-----------A------C--------A--GAA--G---GA-......-TC--C--CAAT........................ 6493
U.CY.08.CY223               A-A-A-----TA-G---C---T............GACGTA-----GG---A-----------C-----TG-C------CAA-C--------GG-GG-----C--AG-G---T-GTAC-----GA-......A----G-----G.....................---. 6573
U.ES.10.DEURF10DZ001        -CA-A-----TA-G---C--AG............ACAATA-----GG---A-----------C-----T-------A-AA-G--------G----A---C--G--G-----GAAC-------G......A---GC---A-G........................ 6542
U.GR.99.99GR303             ACA---G--CT--G---C-...............GACATA------G---A-------------------TG--TCAG-------C-------C-------C-----AG-G---T-GCA---C-----......------G-----G........................ 6573
U.NL.95.U_NL_95_H10986_D1   AC----G--CTA-G---C---T............GACATA------G---A---A-----T-----G-GG---A----A------------------TG----G-A--G-----GTA---C-----......------G---AATCTTAAT.................. 7186

01_AE.AF.07.569M            ACA--TG---TA-GG--C----............GACATA------G--CCA---A-----T-----C-G----A-G--A----T----------AG----G-------G--A--G-A-------A-......--G--C---CAG........................ 6501
01_AE.CF.90.90CF11697       ACA--T---CTA--A--C----............AGCATA-C----G-C--A---A-----T-----G-G----A-G--A------------G-A--------------G--A-A-AA-----------......--G--C---AAG........................ 7253
01_AE.CN.10.YNFL03          ACA--T---CTA--G--C----............GAGATA------G---A--G------T------TG----A-G--A------------G-A--C----G----G------AA-------A-......-GG--C---AAT........................ 6679
01_AE.HK.04.HK001           ACA--T---CTA--G--C----............GAAATA-C----G---A--GA-----T-----C-G----A-GC-A----------G-A--AGT-----------G------AA-----------......--G-----AAT........................ 6659
01_AE.IR.10.10IR.THR48F     A-A--T---CTA--G--C----............GAAATA-C----G---A--GA--------------G-G----A-G--A--G---------R-A-------------G-----GAA-------A-......--G-GC---AAT........................ 6488
01_AE.JP.x.JRC77AE          ACA--T---CTA--G--C----............GACATAG-----G---A--GA-----TT----C-G----A-G--A--G------------A--C---G------G------AG-------A-......--G--T---AAT........................ 7289
01_AE.TH.09.AA090a_WG11     -CAC----GGTA--G--C----............GACATAG-----G---A---A-----T-----G-G----A-GA-----------------A------G------G------AA-----------......--G--C---AAG........................ 6720
01_AE.TH.90.CM240           AC---T---CTA--G--C----............GATATA-----------A---A-----T-----G-G----A-G--A----------------AGT----------G--A--GAA-------A-......--G--C---AAT........................ 6863
01_AE.US.05.306163_FL       AC---T---CTA--G--C----............GATATA-C---GG---A-A-A-----T-----G-G----A-GA-A----------G-G-AGT----------G-----G-A------GA-......--GAGC--CAAG........................ 6438
01_AE.VN.98.98VNND15        ACAT-T---CTA--G--C----............GACATA-----------A--GA-----T-----G-G----A-G--A---C---------AGT----------G--A--G-A---C---A-......--G--C---AAT........................ 6556
```

```
                                                              V3 loop end
B.FR.83.HXB2                   GAGAGCATTTGTTACAATAGGAAAA..............ATAGGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATAACACTTTAAAACAGATAGCTAGCAAATTAAGA......GAACAATTTGGA.......................AAT. 7289
Env                            __R__A__F__V__T__I__G__K__.__.__.__.__I__G__N__M__R__Q__A__H__C__N__I__S__R__A__K__W__N__N__T__L__K__Q__I__A__S__K__L__R__.__.__E__Q__F__G__.__.__.__.__.__.__.__N__
02_AG.CM.10.DE00210CM013       -----T---CTA-G---C---T...........GATATA-C---GG----A-------------------------A--G---------A------C--GG-G-G----CAC----G--G......-GG--C---CCG........................ 6719
02_AG.CY.09.CY256              -CA-A----CTA-G---C-..............GACATA-----G-----A-----G-----------TG-C--C-A--TTC---------GA------T-T--AG-G----CAC-------G......A-GT-C---AAC........................ 6472
02_AG.ES.06.P1423              ACA------CTA-G---G---T...........GACATA-----GG----A---------------G-GG-C-A--A-A--G-----GTG--A------C-CA--G--T--GCAC-----GAG......A-GT-C---AAG........................ 6705
02_AG.GW.05.CC_0048            ACA-A----CTA-G---C---T...........GACATA------G----A-----------------TG-C----A-AG--CC------G-G------C--A-TG-G----CAC-------G......A--T-C--CAAT........................ 6672
02_AG.KR.12.12MHR9             ACA-A----CTA-G---C---G...........GAGATA-----G-----A--G-------------CTG-C---GA-A--G------C----------CG-A--G------TAC-------G......-----C--CAAG........................ 6957
02_AG.LR.x.POC44951            -CA-A-----TA-G---C---T...........GACATA------G----A---------T-------TG-C------A--G---------CA------C--A--G------CACG------G......--GTCC--G--G........................ 7266
02_AG.NG.09.09NG_SC61          -CA-A----CTA-G---C---T...........GACATA-----------A-----------------TG-C---G--GGG-C-------CTG-----C----AG-G-T-GAAC-----G-G......A-GT-C--CAAC........................ 6484
02_AG.NG.x.IBNG                -CA------CTA-G---C---T...........GACATA----GG----A-----------------TG-C----A-A--G----------A-----C-T---G---T--CAC------AG......ACGT-C---AAG........................ 6807
02_AG.SE.94.SE7812             -CA-A----CTA-G---C---T...........GACATA----GG----A-----------------TG-C--CT-GCA-C-------C--A-----C-CG-TG-G----CA--------G......--GT-T---AAT........................ 6649
02_AG.SN.98.98SE_MP1211        -CA-A----CTA-----C---T...........GCAGTA----GG-C--A-----------------TG-C-------A-----------G-------C--A--G-G-T--CA------G-G......A-G--T---AAT........................ 6476
03_AB.RU.97.KAL153_2           ----------TA-G---C----...........GACATA-C----G---A-------------------------T-A-----------------A-------------T--T-------------......A----------G........................ 6481
04_cpx.CY.94.94CY032_3         ACT-A---GGTA-G---C---T...........GAAATA------G---A------------------------G--AATG-T------G----C------GT----AG-GAAG----G-A-......AG--TC--CCCT........................ 6671
05_DF.BE.x.VI1310              ACA------CTA-G---C---T...........GACATA------G-C--A---A-G-----------G-------G-A-C----------A--C----TC---G----G-AAG-G---CAG......TCT--C--CCCT........................ 6686
06_cpx.AU.96.BFP90             ACA---G--CA--G---C---T...........GACATA------G---A--------T--------TG-------------T----CAG-T-TAC--GG-G-AG-GAAGGTA---C--GA-..........---GTT---AAT........................ 7355
07_BC.CN.98.98CN009            ACA-A----CTA-G---C----...........GAAATA------G-C--A--------------------------GA--AT------------A---C--C--GG-G--AG-GAA------GC-..........-----C--CCAG........................ 6629
08_BC.CN.97.97CNGX_6F          ACA-A----CTA-G---C----...........GACATC------G-C--A--------------------------A--ATGC----T--G-A------C-GAG-G--AG--AA------GC-..........-----C--CCCT........................ 6471
09_cpx.GH.96.96GH2911          ACA-A-G--CTA-G---C---T...........GACATA------G---A--GA---T------C-G----A-G--A-----------G-A-TC---GG-A--G-----GAA-------A-......-G---C---AAT........................ 6482
10_CD.TZ.96.96TZ_BF061         ACA-A----CTA-G---C-..............GACATA--------C--A---------T---A----------G--A--G----------A------C----AG------AA------G--.......--C-TT---AAC........................ 6653
11_cpx.CM.95.95CM_1816         ACA---T--CTA-G---C---T...........GCTATA----GG----A---------------GTA--C-A-GC-A---T-------------------AG-G---T-GAAC-------GCAAATA---ATC---AAT........................ 6694
12_BF.AR.99.ARMA159            AC-G------TA-G---C----...........GACATA----GG-C--T---A-G-----------TG-----G-----C------G---A--GG--G-----G--AAGGCA--G-----------......TCT-----C........................... 7293
13_cpx.CM.96.96CM_1849         ACA------CTA-G---C---G...........GATATA------G---A-----G--------------C------CA---------------C---C-----G-----GCAC-G-----------......A--T-T--GAA-.......................G--. 6708
14_BG.ES.05.X1870              A----T---GTA-----C----...........SAAATA------G--R-A---A------------------C-A--A----------------C--GG--------T-GAA-----------------..............-------CAA-........................ 6774
15_01B.TH.99.99TH_MU2079       ----T---GGTA-----C----...........GAAATA------G---A-----------------C-------CA--------GG------------C------A--GAA-----------------..............-----------G........................ 6655
16_A2D.KR.97.97KR004           AC----C--CTA-----GGCA--C-TATACAAGGCAAGCA-A---GG----A-----------A-----T--C--C--CAG-C--------G---------C--A--G-----GAGC-G-----------......A--T-C--CTCT........................ 6667
17_BF.AR.99.ARMA038            AC-T------TA-G---C----...........GAAATA------G-C--C---A-G-----------TG------A-A--C----------A--A---G-----G--AGGGCA--G---GAG......TCT--T--CCCT........................ 6502
18_cpx.CU.99.CU76              ACA------CTA-G---C---T...........GACATA------G---A-----------------T-----CGA-AA-G-T------G---------C-CG--G------CAC-----G--.......-G---CC-GAAC........................ 6591
19_cpx.CU.99.CU7               ACA--T----TA-G---C----.....................----GG-C--A-----G-----------TG-C----A-AA-G-G-----AG-AG-C---C--G--G------AAC-------C-.......-TC-TC---AAC........................ 6421
20_BG.CU.99.Cu103              ACA---G--CTA-G---C---T...........GACATA-C----G---A---------TT----GTGG---A--A-A-------------G-G-TGC--C-GA--G--AAAGCA---C--GCT..........---AGC-ACAAT........................ 6753
21_A2D.KE.99.KER2003           -CA----C--TA-----CC..............AAGATA------G-C--A--G---------------T--C---G-----G-----------A------C--AG-G------TA-----------------......AG--TT---AAC........................ 6481
22_01A1.CM.01.01CM_0001BBY     ACA--TC--CTA-G---GG--T...........GATATA-----GG---A---A-----T-----G-GG-C-A--A-A--G--------G-A------C-CGG-G---TA-AAC-G-----------......A----T-GGAAC........................ 6474
23_BG.CU.03.CB118              ACA------CTA-G---C---T...........GACATA-----------A---G-----TT----GTGG---A----A--GCC------G-G--GA-GC-GA--G--AAAGCA---C--GCT..........---AGC-ACAAC........................ 6753
24_BG.ES.08.X2456_2            AC----G--CTA-----C---T...........GATATA------G---A-------------C----TAG---A-G--A--------GGGG-G-TGA----TA--G--AAAGCAG--C--TCT......A--A---ACAAC........................ 6773
25_cpx.CM.02.1918LE            ACA------CTA-G---C----...........GACATA------G---A-------------------T---------T-CC-------C--G------C--GGAG---T-GCAC--------------......A----C---AAG........................ 6500
26_AU.CD.02.02CD_MBTB047       ACA-T----CTA-G---C---T...........GATATA------G-C--A--------TGC---C-----C------AA-G-T---G----AG----G-GCA--G------AAC-G-----T.......--T--C--CCCT........................ 7286
27_cpx.FR.04.04CD_FR_KZS       ATCC-T--GGTA--G-CC----...........GAAGTA------G---A---------T-------T-----A----AA-G-G------GGA------T-CA--G---T-GCAC-----G--.......AG---CC-GAAC........................ 7276
28_BF.BR.99.BREPM12609         ----AT-A-GTA-GT--C--A-...........CGAATACA----GG---A--------------C--------------A--------G-A-------------G----G---T--CA------GA-.................---------AC-AAATTAGGAGAACAATTTAAAGATTT-A 6685
29_BF.BR.01.BREPM16704         ----------TA-G--GC----...........GAGGTA--------G--A----------------G-C-----GA-T----------A-GG------G-----------CA------G--.................-----G---AA-..........................-TC. 6707
31_BC.BR.04.04BR142            ACA-A----CTA-G---C----...........GACATA------G---A--------------------------A-A--GC------C-GA------C--G--G---G--AA------GC-.........--G--C--CCCT........................ 6751
32_06A1.EE.01.EE0369           -CA---G---T--G---C---T...........GACATA------G---A--------T---------------G------------------G-G-TG---C-TA--G-GATAGCA-------AG.........---TCC---CCT........................ 6954
33_01B.ID.07.JKT189_C          ACA--T---CTA--G--G----...........GACATA------G---A--GA----T-----G-GG---A-GA-A----T---T----AGT-----------G-GA--GAA-------A-..........-----C---AAT........................ 6602
34_01B.TH.99.OUR2478P          ACA------CTA--A-------...........GACATA------G---A---------T-----G-G----A-G--A------------G-A------------G-----G-A-----------------......A-G--C---AAT........................ 6476
35_AD.AF.07.169H               ACA------CTA-G---C-..............GATATA----GG-C--A-----------------TG-C------T--GCC-----G--A------G-----G-GAAG-CAC--------------......A----C-GGAAC........................ 6495
36_cpx.CM.00.00CMNYU830        ACA---T--CTA-G---G-AAT...........GATGTA----GG-C--A---A-----------G-GG-C-A-GC-A--G--------G-------GC-G---G---TACAA--------------......A--TTT-GGAAC........................ 6495
37_cpx.CM.00.00CMNYU926        ACA-A----CT--G---C---T...........GACATA----GG-C--A---A-------A---GCTG-C-A--A-A--G-G--------AG-----C--A--G---T-GCAC--------------......A-C--T-----G........................ 6477
38_BF1.UY.03.UY03_3389         ACA-------TA-----C----...........GATATA------G-C--C---A-G---T-Y------G-----G---A-C-T-----------A---G--AGAG--AAGGAA--------G......TCTT-T--CCCT........................ 6703
39_BF.BR.04.04BRRJ179          ACAT------CA-----C----...........AAGATA------G-C--C---A-G-------A---T------G--A--C----------A--A----G----G--AAGGAAG----G-AG......TCT--C--CCCT........................ 6775
40_BF.BR.05.05BRRJ055          AC--------TA-G---C----...........GACATAG-----G-C--C---A-G-----------TG--------A--C-----G----A--A---------G---GGGCAG------AG......CGT-TT--CCAT........................ 6813
42_BF.LU.06.luBF_18_06         AC--------TA-G---C----...........GACATA------G-C--C---A-G---T-------TG-------G-A-C----------A--A---G---G-G--AGGGAA--G---G-G......TCT-TT--CAAT........................ 6869
43_02G.SA.03.J11223            ACA-A-G--CTACG---C---T...........GATATA------G---A---G-------------TG-A------A---G-------G-G--G-----GA-TG--AAGGCAC--C--CAC......A-GATC---AAC........................ 6811
44_BF.CL.00.CH80               ACA-------TA-G---C----...........GACATA------G-C--C---A-G-----C-----TG------A-A--G----------A--A---G--AG-G--AAAGCAG------AG......TCT--C--CCCT........................ 6737
45_cpx.FR.04.04FR_AUK          ACA--TG--CTA-G---C---T...........GATATA-C---------A--GA-------------TG--GA---------C----CA--A---C----GG-AG----GCAGC-C----G.......-----G----AC........................ 7248
46_BF.BR.07.07BR_FPS625        AC--------TA-G---C----...........GATATA------G-C--C---A-G-----------G-----G-----C----------A--A---G--AG-G--AAGGCA-------AT......TCTT-T--CATT........................ 6707
47_BF.ES.08.P1942              C-AG-T----TA-G---C----...........GATATA-----------A-------------------------A--------------------G---------C--A-----G--.................-----------AAC........................ 6709
48_01B.MY.07.07MYKT021         ACAT-T---CTA--GC-C----...........GAAATA-C----G-AG-A-AGA-----T-----G-GG---A-G--A---C-----G---AGT-C---C---AG-----GAA---C---A-.........--G--C----AT........................ 6533
49_cpx.GM.03.N26677            -CA--TGC-CTA-G---C---T...........GCAGTTG-----C-G--A-------------------C---GA-T--G-C------------C---GTA--G------TA---C-----..............---T-C---AAT........................ 6722
50_A1D.GB.10.12792             ACA-A----CTA-G---ATAAT...........GACATA-----G-----A-----------C----TG-C---G--T-----------A------C-----G-------G-----G-G......A-C-TTC--AAC........................ 6686
51_01B.SG.11.11SG_HM021        ACA--T---CT--CA--C----...........GACATA------G---A---A----TT-----G-G----A-G--A-----------G-G------G-----G--A-A--A--------------......---------C--AAT........................ 6495
52_01B.MY.03.03MYKL018_1       AC---T---CT---G--C----...........GAAATA-C----G---A---A-----TT----G-G----A-G--A-------------AGT----------G--AAGGRA-------A-.........--G--CY--AAT........................ 6636
53_01B.MY.11.11FIR164          ACA--T---CTA--G--C----...........GACATA------G---A---AG----T-----G-G--G-A-G--A--G-------GG-AGT----G-----G--A---A--------A-.........--G--C---AAT........................ 6650
54_01B.MY.09.09MYSB023         AC---TG--CTA--G-TC----...........GATATA------G---A---------T-----G-G----A-G--A-------------A--------T---G--T---AA-------A-.........--G--C---AAT........................ 6703
55_01B.CN.10.HNCS102056        ACA--T---CTAC-G--C----...........GATATA------G---A---A----TT-----G-A----A----A--G-------GG-AGT---------AG--A--TCA-------A-.........-GG--C---AAT........................ 6587
56_cpx.FR.10.URF5_patient_A    ACA-A----CTA-G---C---T...........GACATA----GG----A-----------------TG-C------AACC------TG-CA------C--AG-G-----C-AC-------G......A-G--C---AA-........................ 6631
57_BC.CN.09.09YNLX19sg         ACA----A-CTA-G---C----...........GACATA----GG-C--A---------T-------T-----GGA-AAG--C----------A---------AT-G---G--AA------GC-..........-----C--CCCT........................ 6484
58_01B.MY.09.09MYPR37          ACA--T----TA--G--C----...........GACATA-CT---G---A---------T-----G-G----A-G--A--GT-------G-T-----C-----G--T--GAA-------A-.........--G--C---AAT........................ 6673
59_01B.CN.09.09LNA423          ----A-----T--G--CAC---...........GAGATA------G---A---A-----T----------C------GA--A-C---------------C-----AG------C-TA------CA-..........---------AG........................ 6470
60_BC.IT.11.BAV499             ACA-A----CTA-G---C----...........GATATA------G---A---------------C-----------------GC------C--A------C--GCAG---G--AA------GC-.........--G--C--CCCT........................ 7264
61_BC.CN.10.JL100010           ACA-A----CTA-G---C----...........GGCATA-C------C--A---------T-----------------A--AGGG------GG-AG-----G-CAG-G---AG--A-------C-..........-----C--CCCT........................ 6697
62_BC.CN.10.YNFL13             ACA-A----CTA-G---C----...........GACATA------G-C--A------------------------GC-AATG-T------G-AG------G-A-TG--AG--AA------GC-..........----TC---CCT........................ 6605
63_02A1.RU.10.10RU6637         ACA------CTA-G---C---T...........GCCATAG----G-----A---------T-------TG-C---G-----GG---------AG-----C--A--G-G-T--CAC-----C-G......A----T---AAG........................ 6815
64_BC.CN.09.YNFL31             ACA-A----CTA-G---C--A-...........AAAATA------G-C--A--------------------------A-AG---C----C-G-A------C--A--G---G--AA------GC-......A---TC..TAT........................ 6642
65_cpx.CN.10.YNFL01            ACA-A----CTA-G---C----...........GACATA-----------A---------T------------------A---T---------G-A------C--AG-G---G--AA------GC-......-G-T-C--CTCT........................ 6689
67_01B.CN.11.ANHUI_HF115       CCA--TG--CTA--G--C-AA-...........GRCATA--R---G---A---A-----T-----S-GR---A-G--A--G--------G-A--C---------G-TAAAGCA------G--.........--G-GC---AAT........................ 6629
68_01B.CN.11.ANHUI_WH73        CCA--TG--CTA-GG--C----...........GACATA------R---A---A-----T-----G-G----A-G--A--G--------G-A--C---------G--A--GCA-------A-.........--GT-C---AAT........................ 6611
70_BF1.BR.10.10BR_PE004        C----T----TA-----C----...........GRCATA------G-C--A--M-------------------R--RW-MRG------R---------C----A----YY-CA------G--..............-----W---AA-........................ 6732
71_BF1.BR.10.10BR_PE008        T-A-------TGGG---C----...........GGAATA------G---A-----------------TC---------A--G------------------------------T--TA---C-----..........-C-------AAT........................ 6952
72_BF1.BR.10.10BR_MG002        ACA-------TA-G----G---...........GACATA------G-C--C--GA-------------TG-----G--ATGC-----------A--A---G--GG-G--AAAG-A-----------......TCT-TT--CCCT........................ 6805
O.BE.87.ANT70                  .-TG--C-GGTAC-GC--G----T-GGGGGA...ACAGCAGG-AAC-GCTCA--GGC---TT----C--GTA--A-GCCA-TG-T---GGA--A-TA--------A-C----GAA-GG-ATTT-..........----T-G-AAAC........................ 7321
O.CM.98.98CMA104               .-TG--C-GGTAC-GC----AT-T-GGGAGA...GGAAAC-C-A-CG-CCCA---AG---TTT---C--AT--GAAGA-A-C---------AG-AG-C----R---A-C----GAA-GG-ATTT-..........----TTG-GAAC........................ 6780
O.CM.98.98CMABB141             .-TGT-C-GGTAC-GC--GA-C-C-.........GAAGGC-ACCCTG-GTCA---AGG--TT-----G-GTA--A--CCA-T--T---GTAG-GG-C----------C----GAA-GG-ATTT-..........---TT-G-AAAC........................ 6833
O.CM.98.98CMABB212             .-TGAG--GGCGC-GT--G-C--GTGTGCAG...GAGAAC-G-AA-TCATCA--GGC---TA------CATA--A---CA-TG------GA-----C---G--A-A-C-----AA-GG-ATCT-..........----TTG-AA-C........................ 6826
O.CM.98.98CMU5337              .-TG--C-GGTAC-GC--G-CC-T-.........GAGAAA-ACTC--CCTCA--GAT---TTC---C--GTA--A-G-CA-TG-G---G-A--AG-C--------A-C----GAA-GG-ATTT-..........----TTA-GAAC........................ 6778
O.CM.99.99CMU4122              .-TGT-C-GGTAC-GC--G--C-T-AGCTCA...CCACTCTC-AATGCCTCA--GTT---TT----CCGGTA--A-GCCA-TG-G-----A--AG-C--------A------GAA-GG-ACTT-..........----TTG-AAAC........................ 6761
O.FR.92.VAU                    .-TG--T-GGTAC-GC--G-CCCTTAGT......AATACA-AG--GG---CA--GGC---TT-------TTA----GCCA-TG-C-----C--AG-C-------A-C---A--GAA-G--ATTT-..........----TTG-A-A-........................ 6860
O.GA.11.11Gab6352              .-TG--TG-GA-C-GC--GAACCTGGCAGCAGAGGGCAGAGG-AAC--CTCA--GAG---TT----C--GTA--AC-CCT-T--T---G-AG-AG-A----G-G-A-CCA--GAA-GG-ATTT-..........----TTG-AAAC........................ 6523
O.SN.99.99SE_MP1299            .-TG--C-GGTAC-GC--GA-CCTGGCCCAAGAGGGAAAACC-AAT--CTCA--GAT---CT----C--GTA--AC-TCT-TG-T---G-A--AG-C----------C----GAA-GG-ATTT-..........--C-TTAGGAAC........................ 7365
O.US.10.LTNP                   .-TG--CC-GTAC-GCTACC-TCT-GCGCAA...TCAGGG-ATAAC-GCTCA--GG----T-----CCGGTA--A-CCCA-TG-T---G-A-GAG-AC-------A-C----GAA--G-ATTT-..........----TTG-CAAC........................ 7218

N.CM.02.DJO0131                T-TGA-----TA--AC----A-...........AAAATAG-----G-C--T-----------C-----TG-CTC--AC-A-...----GATCA-TG-GG--TA-A-C-AAAGAG---A---AG......AGC-TCC-G---........................ 6749
N.CM.02.SJGddd                 T-TGA-----TA--AC----AG...........AGAATAG-----G-C-CT--G-----G--C----TTG-CTC--A-AA-G-T----CA--A-TG-GG--TAGA-C-AAAGAG---A--G-G............................................ 6721
N.CM.04.04CM_1015_04           T-TGA-----TA--AC----AG...........AAAATA-----GG-C--T-----------C-----TG-CTC-GA--A-...-----ATCA-TG-GGG-TAGA-C-AAAGAG---A---AG......-GC-TCC-G---........................ 6742
N.CM.06.U14296                 T-TGA-----TA--AC----A-...........AAAATA------G-C--T---------------TTG-CTC---AA-G-T---G-ATCA-TG-GG--TA-A-C-AAAGCA---A---A-.......--GACCC-G---........................ 6761
N.CM.06.U14842                 T-TGA-----TA--AC----AG...........AGAATA------G-C--T-----------C---GTG-CTC---AA-G-C---GGAGCA-TG-GG--TAGC-C-AAAGAG---A----G......A-C-TCC-G--G........................ 6786
N.FR.11.N1_FR_2011             T-TGA-----TA--AT----AG...........AAAATA-C----G-C--T-------G--C-----TG-CTC--A--A-G-G-----AG-A-TG-GG--TA-A-C-AAAGAGG--A----G......A--AT-GGGAAC........................ 6655

P.CM.06.U14788                 .-TGA---GGTA--AT--GAA-...........CACTACG----CG----C--GGC---T--------TG--TC--G-A---T-----A-GA--C---G--TG-G-CAG-GAGGCCA-CT-G......A-GGCT-A-CCTCAACAT..................TCC. 6802
P.FR.09.RBF168                 .-TGA-C-GGTA--AT--GAA-...........TTCTAC-C---TG----C--GA----TT-----------TCC-TGCA---T----CAG-A--C------TCTG-CAG-GAGGCCA-CT-G......A--GCT-A-CCTCAACCC..................CCC. 7354

CPZ.CD.06.BF1167               C-TGA-C--CTA--AT---A--...........AATGTT--T---G-C-CT---A-G---TT---CT-TG-A-AC-A-A--CT----G-A--TG--ACTG-GGCAGC-AAAGTAGC-A---AC..........---ACTC--AA-..............AACACCACGTTCA 7421
CPZ.CM.05.SIVcpzMT145          A-TGA-T---TA--AC----A-...........AATGTAG---GG-C-CT---A-----T-----G-G--A-A-G-GA--C-G---GCA--AG-CC-C--TG-A-C-AAAGAAGTTC----G......A-CAT-C-GA-G........................ 6804
CPZ.GA.88.GAB1                 A-TGA-C---TA--AT----A-...........AATGTAG----G---CC---TCT--CT-C-----G--A-A-G-GA-C-C--------CGA--GG-GG--G-AG--AAA-AGGCT---GC-......ACCTC--CAAAT........................ 7344
CPZ.TZ.06.SIVcpzTAN13          --TGA-C--CTA--AC-G--A-...........AACATCT-T---G---CT--GA----CTT---CT-TG---A-GC-A-TG-G---CGG-GA---C-TG-GAT-GCCAAA-CAGCCC-C---.........---GCCACAAA-........................ 6954
CPZ.US.85.US_Marilyn           A-TGA-T---TA--AC---CC-...........AAGATAG-----G--G-A---G----------------CTCA-A-CTG-C----G-G--ACAGAG----T-C-CTTTAGAG-T-A---A-......A-GG--GCAAAC........................ 7318

GOR.CM.07.SIVgor2139_287       .-TGA--A--TA--ACTC--A-...........AACATAG-------C-CA---A-G--TTT---C--GTA--A-GAGA----C---C-GGGTG-G--G---G-C-C--TGCAGGCTC-C-AC.......--GAGC............................ 6850
GOR.CM.07.SIVgorCP2135con      .-TGA--A--TA--ACTC--A-...........AACATAG-------C-CA---A-G--TTTC--C--GTA--A--AGA----C---C-G--TG-A--G---G-T-C--TAC-GGCTC-C-A-.......-CGA-T............................ 6831
```

V4 loop start
B.FR.83.HXB2
Env
A1.AU.03.PS1044_Day0
A1.CH.03.HIV_CH_BID_V3538
A1.CY.08.CY236
A1.ES.05.X1608_8
A1.KE.11.DEMA11KE001
A1.RU.11.11RU6950
A1.RW.11.DEMA111RW002
A1.SN.01.DDI579
A1.UG.11.DEMA110UG001
A1.ZA.04.04ZASK162B1
A2.CM.01.01CM_1445MV
A2.CY.94.94CY017_41
B.BR.10.10BR_MG029
B.CA.07.502_1191_03
B.CH.08.M2_0803101_NFLG8
B.CN.10.DEMB10CN002
B.ES.10.DEMB10ES002
B.FR.11.DEMB11FR001
B.GB.05.MM45d213_GN1
B.HK.06.HK003
B.HT.05.05HT_129389
B.JP.12.DEMB12JP001
B.KR.07.HP_18_07JHS10_3909
B.PE.07.502_0525_wg5
B.RU.11.11RU2ln
B.TH.08.MERLBDTRC10
B.US.11.ES38
C.AR.01.ARG4006
C.BR.07.DEMC07BR003
C.BW.00.00BW5031_1
C.CN.10.YNFL19
C.CY.09.CY260
C.ES.08.X2363_2
C.ET.02.02ET_288
C.IN.09.T125_2139
C.KE.00.KER2010
C.MW.09.703010256_CH256.w96
C.TZ.08.707010457_CH457.w8
C.YE.02.02YE511
C.ZA.08.705010534_CH534.w12
C.ZA.10.DEMC10ZA001
C.ZM.11.DEMC11ZM006
D.CM.10.DEMD10CM009
D.CY.06.CY163
D.KE.11.DEMD11KE003
D.KR.04.04KBH8
D.SN.90.SE365
D.TZ.01.A280
D.UG.10.DEMD10UG004
D.UG.11.DEMD11UG003
D.YE.02.02YE516
D.ZA.90.R1
F1.AO.06.AO_06_ANG32
F1.AR.02.ARE933
F1.BR.10.10BR_RJ015
F1.CY.08.CY222
F1.ES.02.ES_X845_4
F1.RO.96.BCI_R07
F1.RU.08.D88_845
F2.CM.02.02CM_0016BBY
F2.CM.10.DEMF210CM001
F2.CM.10.DEMF210CM007
G.BE.96.DRCBL
G.CM.10.DEMG10CM008
G.CN.08.GX_2084_08
G.CU.99.Cu74
G.ES.09.X2634_2
G.GH.03.03GH175G
G.KE.09.DEMG09KE001
G.NG.09.09NG_SC62
G.ZA.01.TV546
H.BE.93.VI991
H.CF.90.056
H.GB.00.00GBAC4001
J.CM.04.04CMU11421
J.SE.93.SE9280_7887
J.SE.94.SE9173_7022
K.CD.97.97ZR_EQTB11
K.CM.96.96CM_MP535
U.CA.01.TV749
U.CA.99.TV721
U.CD.83.83CD003_Z3
U.CD.90.90CD121E12
U.CY.05.CY090
U.CY.08.CY223
U.ES.10.DEURF10DZ001
U.GR.99.99GR303
U.NL.95.U_NL_95_H10986_D1
01_AE.AF.07.569M
01_AE.CF.90.90CF11697
01_AE.CN.10.YNFL03
01_AE.HK.04.HK001
01_AE.IR.10.10IR.THR48F
01_AE.JP.x.JRC77AE
01_AE.TH.09.AA090a_WG11
01_AE.TH.90.CM240
01_AE.US.05.306163_FL
01_AE.VN.98.98VNND15

```
                                                                                                                                                       V4 loop start
B.FR.83.HXB2                ..AAT...AAA....................ACA...ATAATCTTTAAGCAA......TCCTCA......GGAGGGGACCCAGAAATTGTAACGCACAGTTTTAATTGTGGA...GGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGTACTTGGTTTAAT 7415
Env                         .__N__.__K__.__.__.__.__.__.__.__T__.__I__I__F__K__Q__.__.__S__S__.__.__G__G__D__P__E__I__V__T__H__S__F__N__C__G__.__G__E__F__F__Y__C__N__S__T__Q__L__F__N__S__T__W__F__N_
02_AG.CM.10.DE00210CM013    ..---GCC-C-....................G--...-----T---GCTA-C......CA---.........-------TT------CAC--A--T--------------...--A----------T--C-G-A--T--GC-----------C------C-GTT- 6848
02_AG.CY.09.CY256           ..-----..G-C...................---CGA----------TAG-..............--------TGT------AACT--A--T------------C-...--A-------T--T--CG--A--T--AG-----------C------GA--C- 6601
02_AG.ES.06.P1423           ..--C...---....................---......-----A-----CTA--.............K-------TAT------ACM--A--T------------...--A-----C-----T--CW--A--T--AG---A--------------GA---- 6831
02_AG.GW.05.CC_0048         .........-CC...................---......---G-----GCTA-C......---------.......--------TTT------AC--A--T--C-------C---...--A----------T--C---A--T--GC-----------C------AAGC-C 6795
02_AG.KR.12.12MHR9          ..----...--C...................---......--C-----GCTA-C.....C---T-.....--------TTT------AC--A--T--------------...--A----------T--C---A--T--A--T---------C------GAC--C 7083
02_AG.LR.x.POC44951         ..---AATC-C....................---......--------CTA-C.....CA----.....--------TTT------A---A--T--------------...--A----C-----T--C---A--T--AT-----------C------GAC--- 7395
02_AG.NG.09.09NG_SC61       ..--A...TCC....................---......---G-------A-C.....CA----.....--------TTT------AC--A--T--------------...--A----------T--C---A-----C--------G--CGT---AA---A 6610
02_AG.NG.x.IBNG             ..--C...-CC....................---......---------GCTA-C.....C---T-.....--------TGT------AC--A--T--------------...--A----------T--C---A--T--A-----------C------GA---- 6933
02_AG.SE.94.SE7812          ..----...-CC...................---......---------G-TG--.....C------.....--------TTT------AC--A--T--------------...--A----------T--C---A--T--A-T----------A------AA-C-- 6775
02_AG.SN.98.98SE_MP1211     ..-G-...TCC....................-A-.......----------CTA-C.....-------.....--------TTT------AC--A--T--C-----------...--A----------T--C---A--G--G------------C------GC--GC 6602
03_AB.RU.97.KAL153_2        ..----...---...................---......---G------T---.............-----------------------T-----------------...-----------------------A-----A------------C------AA-GG- 6607
04_cpx.CY.94.94CY032_3      ..----...---...................---......-----AA--GCT-C-.....C--GT-....----------T------AC--A--T--C--------AA-...--A----------T--C---A-----C---------------ACACA-GC-A 6797
05_DF.BE.x.VI1310           ..----...---...................---......-----AA----CTC-.....-------.....----------T------AC--T--T----------A--...--A-------------CG--A--T--A--T-------GCC--AGT------- 6812
06_cpx.AU.96.BFP90          ..--C...-C-....................CAC......----C-------TC-.....--TG--.....----------T------AC--A--T--------------...--A----C-----T-----A--T--A-C----------CAT-AGACC-GTT- 7481
07_BC.CN.98.98CN009         ..----...---...................---......-----AA--GCATC-.......----------T------AC--A--T--C--------A--...--A----------T-----A--T--GGC---------G----A-ACA-G... 6752
08_BC.CN.97.97CNGX_6F       ..----...---...................---......-----AA-CGCATC-.......----------T------AC--A--T--C--------A--...--A----------T-----A--T--GGC---------G----A-ACAA-GG- 6597
09_cpx.GH.96.96GH2911       ..----...--G...................---......---------CGA-C-......--------TGT------AC--A--T--------------...--A----------T--CG--A--T--GGC--A-------AC------AA-GG- 6608
10_CD.TZ.96.96TZ_BF061      ..CGA...CC-....................---......------T-----A-C-.......-----------------AC--A---CC------------...----G----------C---A--T--G--------------A---ACATCA 6779
11_cpx.CM.95.95CM_1816      ..--A...-C-....................---......----------TAGC.......--------TTT-----C-AC--A--T--C-----------...--A----------T--C---A--T--A------------C------AA--GC 6820
12_BF.AR.99.ARMA159         ..----...---...................---......-----AA-----CTC-......--------------AC--T--T-----C-----A--...--A-------------C---A--T--GG---------GA---A......... 7410
13_cpx.CM.96.96CM_1849      ..----...-C-...................---......---------GCTA-T.....C------......--------TTT---G--AC--A--T--------------...--A----------T--C---A--T--GCC----------C------GAA--- 6834
14_BG.ES.05.X1870           ..----...---...................---......--TG-------T-GT......------A-----G-------T--------C---------...--A--G-----T---------A--------T-A-----C---------AG--G- 6900
15_01B.TH.99.99TH_MU2079    ..----...---...................---......---G-------C---.....C------...----G--A--GT--------C------...----------A--T------C-G 6781
16_A2D.KR.97.97KR004        ..----...---...................---......---------CCA-C.........G.....AT------AC--A--T---G-C-----...--A----------T--CG--A--T--GGC----------C--G---CC-GCC 6793
17_BF.AR.99.ARMA038         ..----...---...................---......-----AA----CTC-...-----T------AC--T--T-----------A--...--A-------------C---A--T--A----------GACT-CC--GA---- 6628
18_cpx.CU.99.CU76           ..--G...-GC....................---......---TA-----A-C-.......--T---.....----A--AT------AC--A--T-C---...--A----------T-----A-----GCCT-----------C---AA---A 6717
19_cpx.CU.99.CU7            ..........---..................---......---G---GCTA-C.....----A--TTT-----AAC--A--T----...--A----------T-----A-----G-C-----------.......... 6532
20_BG.CU.99.Cu103           ..--G...--C...............ATAACA-AC......-----AA---GC-C-.....C--A--...-----------T------AC--A--T------------A--...--A----------T-----A--T--GG---A---------G-AAT...... 6879
21_A2D.KE.99.KER2003        ..--C...---....................---......-----CT-----T--C......--------T---------------A------...--------C-----------A---------------C--A---AAC-G- 6607
22_01A1.CM.01.01CM_0001BBY  ........---....................---......--------C-CTA-C.......--------TTT------AC--A--T----------A--...--A----------T--C---A--T--GGC----------C------CA-G-C 6597
23_BG.CU.03.CB118           ..-CG...-GC..............ACAGAA-AC......----CA---GGA-C-.....C-------------T------AC--A--T------T---A--...--A----------T-----A--T--GG-----------C-G-AAT...... 6879
24_BG.ES.08.X2456_2         ..-CA...-CC..............ACAGCG-AC......----CA----GC-C-.....CA------...-------T------AC--A--T-----------A--...--A----------T-----A--T--A-C------------G-AAC---... 6902
25_cpx.CM.02.1918LE         ..--C...-C-ATA.................---......---TCA----CAGT........--A--TTT---G--AC--A---T-------------...--A----C-----T--C---A--T--AGCT------------G----ACA--- 6629
26_AU.CD.02.02CD_MBTB047    ..----...---...................-GT......---GC-----T---.......----------T------AC--T--T----C----CA--...--A--G--------T-----CA--T--GGC-----------A------AAAG-- 7412
27_cpx.FR.04.04CD_FR_KZS    ..--C...-TGAGC.................---......---AT---C---C-...CCCC-------...----------AT------AC--A--T-C---...--A----------T--CG--A---A-G-----------------C---AA---G 7408
28_BF.BR.99.BREPMI2609      AA----...-G-...................---......---GC-----T---......-----------G-T----C-------...-------------------CA-----TG-------C---------AAGTGG 6813
29_BF.BR.01.BREPM16704      ..----..G--....................---......---GC----G-T----.CCTC--C--......--------T----------C-----...--------------------A--T--GG-----------GT----AGGCTG 6836
31_BC.BR.04.04BR142         ..----...---...................---......-----A----CAA-G......CA---......-------TT------AC--A--T--C--------A--...--A----------T--C---A--T--GGC--------------A-AC-A--G- 6877
32_06A1.EE.01.EE0369        ..GG-...--G....................-AC......----C------TTC-......--GG--.....----------T------AC--A--T--------------...--A----------T-----A--T--A-------------CT-A-CAAGA-G- 7080
33_01B.ID.07.JKT189_C       .........--T...................---......----G----G-A-C-......------A--T-T------AC--T--TCA---------A--..-------------T--C---A-----AG----------AC-----CC-A... 6722
34_01B.TH.99.OUR2478P       .........--G...................---......----G------A-C-......C-------A--T-T------AC--T--TCA---------A--...----G-------T-----A-----A--T-----------------CCAG-A 6599
35_AD.AF.07.169H            ..----...---...................---......---C------TG-C......--------TGT------AC--A--T------------A--...--A--G-------T-----A--T--GGC-------C-A--GCCT------C 6618
36_cpx.CM.00.00CMNYU830     ........-C-....................---......---------C-CTG--......C------..A--TTT---G--A---A--T-----------...--A----------T--C---A--T--GGC-----------C------GGGT-- 6618
37_cpx.CM.00.00CMNYU926     ..--C...-C-....................---......-------G---GCTA-C...A--TGT------AC--A--T------------C-...--A----C--------C---A--T--GGC-----C--------A---AAC-C- 6603
38_BF1.UY.03.UY03_3389      ..----..G--....................---......-----AA----CTC-...------T------AC--T--T-----------A--...--A-------------C---A--T--A-----------GAC--AGCT-CC--- 6829
39_BF.BR.04.04BRRJ179       ..----..GG-....................---......-----G------A-C-.....G----------T---G-CAC--T--T-----------A--...--A-------------C---A--T--A-T--------GAC--G......... 6892
40_BF.BR.05.05BRRJ055       ..----...---...................---......-----AA----TTC-...------T------AC--T--T-----------A--...--A-------------C---A--T--GG-T-----G---A--ACACAGGG-C- 6939
42_BF.LU.06.luBF_18_06      .........---...................---......-----AA----TTC-...-----A--------G--AC--T--T-A---------A--...--A-----------------A-----AT----------GAA--AG-A------ 6992
43_02G.SA.03.J11223         .........--G...................--T......----C-------CTC-.....---AT-...--------TR-----AC--A--T---------YA--...--A----------T-----CA--T--AG-------GGA............... 6919
44_BF.CL.00.CH80            ..----..GC-....................---......-----AA--G-ATC-...--------T---G--AC--T--T-----------A--...--A-------------C---A--T--GG----------GAC-G-ATACCC-TA 6863
45_cpx.FR.04.04FR_AUK       ..--C...---....................---......---GC------TTC-.....--TG--...-------TTT------AC--A--T--------------...--A----------T-----A--T-----T----------C------AGCC-C 7374
46_BF.BR.07.07BR_FPS625     ..----..G--....................---......-----AA----CTC-......--G--A----T------AC--T--T-----------A--...--A-------------C---A--T--GG----------GAC--AG-A-CC--C 6833
47_BF.ES.08.P1942           ..----...---...................---......---G------T---......------------T--T-C------C-----...--------------------A----TC---------------A---GAG--- 6835
48_01B.MY.07.07MYKT021      .........---...................-AC......-----C----C-A---.....C--G--...----A--TAT------AC--T--TCA---------A--...--------------T--C---A----A---A------GA------CA-AG-A 6656
49_cpx.GM.03.N26677         ..G-G...-C-ACA.................---......-----AA--G--TC-.....G-AA-T.....CC---A---AT---G---ACG--A--T-C-------------...--A----------T-----A--------T-----------GC---AA---G 6851
50_A1D.GB.10.12792          ..C-G...-C-....................---......-------T-----A-C-.........--------T-------CAC--T--T--------------...--A----------T--C---A--T--GGC-----------C-T----AA--GC 6812
51_01B.SG.11.11SG_HM021     .........--G...................---......---------CGA-C-.....C---T-.....----A--T-T-----------T--TCA---------A--...--------------T--C---A-----AG--------------A---ACAGGA 6618
52_01B.MY.03.03MYKL018_1    .........--G...................----.R--GA----C-A-C-.....C-------...----A--T-T------AC---A--TCA---------A--...--------------T--C---A------T-----------A--A---CACAG-A 6762
53_01B.MY.11.11FIR164       ..----..C-C....................---......----TC----C-A-C-.....C--------------T-T------AC--T--TCA---------A--...--------------T-----A-----A-GT-------G.........ACAG-A 6767
54_01B.MY.09.09MYSB023      .........--G...................---......---------C-C-C-.....C-T---......----A--T-T------AC--T--TCA---------A--...--------------T--C---A-----A--------G------T-----A-GGA 6826
55_01B.CN.10.HNCS102056     ..-C-...G-G....................-A-..------C--C-A-C-.....C-------...----A--TGT------AC--T--TCA---------A--...--A----------T--C---A-----A-GT-----------------A-A-TG 6713
56_cpx.FR.10.URF5_patient_A ....---GAC-C-..................GA-......----C----GCTAGT.........--------TGT------AC--A--T--C---------C-...--A----------T--C---A--T--GG-------------------AA-G-C 6760
57_BC.CN.09.09YNLX19sg      ..---------....................-T-......-----AA----TAT-.......--------T------AC--A--T--C--------A--...--A----C-----T---G--A--T--GGC---------G----A-ATGG--CA 6610
58_01B.MY.09.09MYPR37       .........--G...................---......----GAA---C-A-C-.....CA------...----A--T-T------AC--T--TCA------C--A--...-----G-------T--C---A-GT--AC---A--------G--A......... 6787
59_01B.CN.09.09LNA423       ..----...--G...................---......-----GA----T-G-.......--TG--...-----------T----G-----T--------C----...--------------------------C--------------T-----GGCC- 6596
60_BC.IT.11.BAV499          ..----...-G-...................---......------A----GAA-C.....CA------...--------TT----G--AC---A---.-C--------A--...--A----------T--C---A--T--GGC------G-------A-AC-ACTGG 7389
61_BC.CN.10.JL100010        ..----...---...................---......----GAA---GCA-C-...--------TT------AC--A--T--C--------A--...--A----------T-----A--T----T----------T--A-GATAC--GG 6823
62_BC.CN.10.YNFL13          ..----...---...................---......-----A----GCA-C-.....C-AGT-...GAA-----A----T----G--A---A--T--C--------A--...--A----------T-----A--T--GGC-----------A--A-ACA----- 6734
63_02A1.RU.10.10RU6637      ..----...-CC...................---......---------G-TA--.....CA---.....--------TTT------AC-T-A--T-C------------...--A----------T--C---A--T--GC---A---------C-G----AA--GC 6941
64_BC.CN.09.YNFL31          ..----...---...................---......-----AA---G-A-C-...----------T---G--AC---A--T--C--------A--...--A----C-----T-----A--T--G-C---------G----G-TTAA-GG- 6768
65_cpx.CN.10.YNFL01         ..GGA...---....................C--......------A---GCA-C-.......--A----------TT------AC---A--T--C--------A--...--A----------T-----A--T--GGC---------G-C---A-ACAA-GG- 6815
67_01B.CN.11.ANHUI_HF115    .........--G...................---......---------C-A-C-.....C---------A--TST------AC--T--TCA---------AA-...-----------Y--T--C---A-----A-C---------GAA--CAT-....... 6746
68_01B.CN.11.ANHUI_WH73     .........-CG...................---......---------C-A-C-.....C---------A--T-T------AC-----TCA---------A--...--------------T--C---A-----A--T--------------AGAA...... 6728
70_BF1.BR.10.10BR_PE004     .........--G...................---......---R-Y-----WKY-..........AC--T------...-------------CA-----A---------------------A----- 6858
71_BF1.BR.10.10BR_PE008     ..----...---...................---......-----T-----T----....--T-------GA-A-----G---------T------...-------------------------C---------------------AA--C- 7078
72_BF1.BR.10.10BR_MG002     ..----...--G...................---......-----AT-----CTC-.......-------A----T------AC--T---T-----------A--...--A----------T--C---A--T--A-CT----C----AC-GC......... 6922
O.BE.87.ANT70               ..----...-C-...........GGTAGTATT-AC...--G-CA--C--T--C......AG-AGC.......--T--A--T-T---GG-AACCCATTTACAC-----C---CAT...--A-----C--T--T-----CA--G-TA-GA--------TA---C-TT-CATG- 7456
O.CM.98.98CMA104            ..----...-C-ATA.........GATAATGTT-GC...G-G-----C--TAG-......AG-AAT.......-AT--A--T-----G--A-CCTTTTTGCA------C--CCAT...--A--G--C--T--T-----CA--T-GGGGA--------TA---C-TTACATG- 6918
O.CM.98.98CMABB141          ..----...-C-...........GGTAATATA-GT...--G--A--C---A-C......AGTA-T.......----A--T-T---GG-AACCCAATTGCA---------CAT...--A--G--C--T-------CA--T-T--GA--------TA---C-TTACATG- 6968
O.CM.98.98CMABB212          ..--G...-C-GGAAATGGTGCAGGAGATGTT--C...-----A-----CA--......A-TAGT.......--T--A--TGA-----AAGC-ATTTGCA------C---CAT...--------C--T--T-----CA--T-T---A-------CTA------TAACTG- 6973
O.CM.98.98CMU5337           ..----...--TGAG.........ACAAATATA---...--G-CA--C--T-GC......AG-CTT.......-AT--A---G----G-CAACCCATTTGCAC----C---CAT...--A--G--C--T--T-----CA---GTGG-A--------TA---C-TT-CATG- 6916
O.CM.99.99CMU4122           ..----...-C-AACAATACAGRTGACAATGTT-GC...--G--A--C-GTA--......AG-A-T.......--T--A--TT----G--A-CCCACTTGCA-----GC---CAT...--A--G--------T-----CA----C--GA--------TA---C-TC-CATG- 6908
O.FR.92.VAU                 ..T--...--TCAA.........ACTGATGTT--C...-----AA--CGGTA-T......CA-AGT.......--T--A--TG----G-AAC--ATTT-TT------C---CAT...--A----C--T--T-----CA---AT--G---------CA---C-TT-CCTGC 6998
O.GA.11.11Gab6352           ..----...-GT...........CAGTTAGTT----...G-GCAG--C--TAG-......AG-A-T.......-AT--T--TG----G--AACC--TTTACA------C---CAT...--A--G--C--T--T-----CA-CT-TA-GA-----TG-TA------TCAACTG- 6658
O.SN.99.99SE_MP1299         ..----...-C-...........AATACAGTT-AC...-----CA--CG---GC......AG-ATT.......--T--A--TT----GG-AACCCATTTGCA---C--C---CAT...--A--G--C-----T-----CA--T-TA-GA--------TA---C-TT-CATG- 7500
O.US.10.LTNP                ..----...-C-...........AGAAATATT--C...G-G--A--C---A-C......AG-AGTAATAAC-----A--T-T---G-CAACCCG-TTACA---------CCAT...--A-----C--T--T---G-CA--GGGA--A--------TA---CCT-GAATG- 7359
N.CM.02.DJO0131             ..--C..........................-AC...-----CA--C---GCTCAGGAAAAGAAT...........-----A---------G-AAC-CACTTA-TG--C-----------...--A--G---A----T------A-TT-CAG-T-A--C---GAG-GCAT-AA--CA 6878
N.CM.02.SJGddd              ..-GC..........................-AC...-AG-C---CC--GC-CGTAACAAGA-T.......-A---A---------G-AACCCACTTA-TG--C------A--...--A--G---------T-----CA-TT-CA--T-A---G---AA-A--TAC----C 6850
N.CM.04.04CM_1015_04        ..--C..........................-AC...-C--C--------TCTCGGGTGAA-ATT.......-----A---------G-AAG-CACTT--TG--C-----------...--A--G------CTT-----CA-TT-CAG-T-A---G--GAG-ACG--AC-GTA 6871
N.CM.06.U14296              ..G-A..........................-AC...-----CA--C-G-GCTCGGAAGAAAAAT.......-A---A---------G-AAY--ATTTA-TG--C------CAT...--A--G---------T------A-TT-CA--T-A---G--GAGGRCCA-AA---G 6890
N.CM.06.U14842              ..-GC...---....................-AC...-----C---C-G-GCTCGGGGGAAGAAT.......-AT--A---------G-AAC-CACTTA-TGC-C-------------...---------------T-----CA-TT-CA--T-A-----CGAG-ACGA-AC-... 6915
N.FR.11.N1_FR_2011          ..TGG..........................-AC...-----C---C---GTTCGGGAGAA-AAT.......-----A---------G-AACCCATTTA-TGC-C-------------...--A--G---------T-----CA-TT-CA-CA-A---G-CGAGGGCAA-AA-G-A 6784
P.CM.06.U14788              ..--CCAT-C-TCC......AAACACAACCAT----...T-TG-T-----AA-C......AGTA--........--G--A--------GG--TCTT---TGCAC--C-G----CAT...---------C--T--T-----CA-C-GCAGC--------CTT--G--ATAC-TGG 6946
P.FR.09.RBF168              ..--C..................CAAAACCAT----...T-TG-----CGAA-C......AGTA--........--G--A--------GG--TCTTTT-TGCAC-KC------CAT...---------C--T--T-----CA-C-GTACC----------TA--GC-ATAC-TG- 7486
CPZ.CD.06.BF1167            AA--CACC-CC.........TATACATTTAGG---GATGGGGATC-A---TGG......AT----.......-----A----A---G-C-CAG--T--TT-G--C--C---CA-...--A-----------T---G-CCTC-A-A-TTG-ACA--C............... 7552
CPZ.CM.05.SIVcpzMT145       .........--G...................-AC......---G-----T-GTC......C-T--T.......--G--A--T------G-GACT-AC--TCAC-----C--C--G...--A-----C-------------A---GTG-GA-CA-C----TC---AA-A-A--C 6927
CPZ.GA.88.GAB1              ..CGA...-C-...............GCAGCC-AC......----CAC-C--CAG-......G----T.......-----A--T----GG-CAC-CAT--T-TG-------------...--A----------T-----A-T-GC---A-A----C-GAC-AC......... 7467
CPZ.TZ.06.SIVcpzTAN13       ..--A..........................C-T...-CCC-GAA--TCAC-......CGGC--.......AAG----------GG-A-AG-ACTTTT-G---------CAG...--A--------------C---CT----A--TG-A----CGAG--C---C----C 7077
CPZ.US.85.US_Marilyn        ..CTG...-C-....................-AG...G--GAG--A-TT-C-......AATG--.......-----A---------G-G----ATATG-TGC--------------...--A--------T--T------A---TT-CC--A-----C-TG--C-ACAAC--- 7444
GOR.CM.07.SIVgor2139_287    ...............................CTG...--CCA---CCG-A-T.......--T---........-----A--T------G--ACCTTT-TGCA----------CAT...--A-----------------A-C--TA--A------GCTAC-A---CACA-CA 6970
GOR.CM.07.SIVgorCP2135con   ...............................CAG...--CCG---CC--A-T.......--T---........-----A--T------G--ACCTTT-TGCA----------CAT...--A---------------G--A-C--TA--A-------CTAC-A---TACA--G 6951
```

V4 loop end

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

V4 loop end

HIV-1 Genomes

```
                                                                                                                       V5 start
B.FR.83.HXB2                 GAATAAAACAAATTATAAACATGTGGCAGAAAGTAGGAAAAGCAATGTATGCCCCTCCCATCAGTGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAAT...AGCAACAAT.............................GAG 7616
Env                          R__I__K__Q__I__I__N__M__W__Q__K__V__G__K__A__M__Y__A__P__P__I__S__G__Q__I__R__C__S__S__N__I__T__G__L__L__L__T__R__D__G__G__N__.__S__N__N__.__.__.__.__.__.__.__.__.__.__E_
A1.AU.03.PS1044_Day0         -------G---G-------T---------C-CC----C---------------------AC-A---GC---A--G---GA------C--------A--AA------------------G----...GAT----CA.........................AGCACA 6854
A1.CH.03.HIV_CH_BID_V3538    -------G---T--G----T---------G-------C---------------------AA---A----ACAG--CAT---------------T--A-------------------G-----G...G-T--A-GC.........................AATATT 6807
A1.CY.08.CY236               A------G-----------T---------G--C----C-------A---C-----------C-G---AT---A--G---GA---C--C--------A--AT-------------------AG--...-CTG-A--C............................ACC 6833
A1.ES.05.X1608_8             A------G---G--G---GA---------G-------C-----G-A-----------------CA---GC---A--T---AAC----CC--------AT-AA---------------A--GG--...-AT-G--G-....................GGAAGG--A 7100
A1.KE.11.DEMA11KE001         -------G-----------T---------G-------C---------------------CAA---GT---ACAG---GT------C--------A--A--G---------------G--G---...GTG--T-GC....................AGTAAT-GC 7013
A1.RU.11.11RU6950            -------G-----------T---------G-------C------------------T-----A---AGT--A-AG---GT------C--------A--A---C---T------------AGG-...G-A--T---.........................AGCA-C 7143
A1.RW.11.DEMA111RW002        A------G-----------T---------G-AC----CG--------------------ACCA---AC---A-TG---AAC----GC--C-----A--------------------------G---...--TGGG--C.......................AACAGT 7000
A1.SN.01.DDI579              ------G-------G--C-T---------G-------C---------------------TCAA--GGT---A-AC---A-------C----------A--G--------G-------GG--...-AT-GT-G-............................ACA 6820
A1.UG.11.DEMA110UG001        -------G------G----T---------G--C----C----C--A-----------------AA---AT---A-AT---AG------C--------A--AA-------------------AC--...-AT--T..............................ACA 6942
A1.ZA.04.04ZASK162B1         -------G------T----T---------G-------C-G--G---------------CCA---G----A--G---GA------C--------A--AA-C-----------C--G--GCT-...GAGCTT--C....................AATACGA-T 7071
A2.CM.01.01CM_1445MV         A---------G------------------G-------CG--------------------TGCA---GT---A-AG---AT---------C-----AA--A-C--G--------------G--G...C----TG--.........................TCCAGT 6821
A2.CY.94.94CY017_41          -----------------------------G--------G--------------------TGCA---AT---A-AG---A-------C--------AA-AA----G-------------------C...-ATGGG..............................ACT 6988
B.BR.10.10BR_MG029           A-----------------------------C------------------------------A---A----------------------------------------------------GG-...-ATGGGTCC.......................AGTACT 7113
B.CA.07.502_1191_03          A----------C----------------GGG-------GG--------------------G---A-C-----C----------------A--A-------------------A--C...-ATGGG-CA.......................AAT-GT 7063
B.CH.08.M2_0803101_NFLG8     A------------------G------AG-----------------------------AA----T-----AC---GT----------------A-A-------------------C--...TCT--G---ACAAAT...........AGTACAGATAAT--T 7159
B.CN.10.DEMB10CN002          ------G-------------GA------G---------------------T-----------A---GT---C-AC---A----------------A--A--G---GT-----------G-T-...-A-----GC..........................ACT 6992
B.ES.10.DEMB10ES002          --------------------G-------GG------------------------------A---TC--------------G-------------------------------G--C...-CG--TG-G..........................ACC 7004
B.FR.11.DEMB11FR001          ----------------------------G----T--------------------------A--------------C----------------A------------------A-CT-...-A---T-CC..........................C-- 6813
B.GB.05.MM45d213_GN1         -------G------G-----C--------GG-------------A----------------G--G-T------C--------------------T--A--------------------...G-TR-G--CATA...................AGMAACA-T 7034
B.HK.06.HK003                ------G---------------------AG-------------------------G------AA---------A-------------------T------------------G--C-C-...-AT-C--CA.......................AAC-GA 6976
B.HT.05.05HT_129389          ---------------C--T---------GG-------------------------------C----C---A-----GA------------------------------------------G-C...CAG---GGGACG..................GACGCTACC 6840
B.JP.12.DEMB12JP001          --------------G-----------------------------------AA---------GCA----T----------------------AT----GTG-----------G--C...-CA-CG--........................ACAACC 7026
B.KR.07.HP_18_07JHS10_3909   --------------------G-------G-------------------------------A---------A--------------------------------------G--G-C-...GAG-----CAAG..............ACTGGGAGCAACAG- 7279
B.PE.07.502_0525_wg5         -G--------------------------G-----------------------------CAA-----------T-------------------A--C-----------------------......----GT--C..................ACTAAT-GA 7047
B.RU.11.11RU21n              ----------------------------GC----------------------------TCA----------A-----------------A--------------------------C.........................................AAT--C 7144
B.TH.08.MERLBDTRC10          ----------------------------G--------------------------------------C------------------AA-A--------------------GG-...-AT--T-CG...........................-GA 6866
B.US.11.ES38                 -------------A----------------A------------------------T--TCCA----C----AC-------------------A-------------------AGG-...GAT-GG-G-..................AATGCCACC 7565
C.AR.01.ARG4006              ---------------------------AG-G------CG--------------------T--A---ATG--A-C----AG---------C-----AA-A---C-G---C-T-----A--C-CA...-ATGG--CA......................AATA-C 6804
C.BR.07.DEMC07BR003          ----------------------------G-G------CG--------------------TGCA---G----A-C----AG---------C-----AT-A---C-------G-----A--AG-C.....................................ACGAGT 7038
C.BW.00.00BW5031_1           ----------------------------GGG------CG---------------C-----TGCA---A-C--A-T----AA---------C-----AG-A-----G---TAT-----A--AG-A...G-GGTA.........................AAG--A 6951
C.CN.10.YNFL19               -------G--------------------G-G------CG--------------AA-----TGCA---A-C--A-C----AA-------------A--A-----GGT-C-T-----A--AG-A...G-T-CA---........................GCTACA 6986
C.CY.09.CY260                T---T-T----C-GG---TA-G-------G-G-------G------A-CA-----C---CTGCA---G---A--TC-CAGG----T----C---C-AG-AT---GGC---CC-GA--G--A--A...-ATTGAGGG.................CAAAAT--C 6851
C.ES.08.X2363_2              ------------------------------G------C--------------------T--TGCA---A-C--A-C----AAC---------C-----A--A-----G---C-T-----A--A-GC...--T--T..........................GACACT 7088
C.ET.02.02ET_288             -------G--------------------G-G------CG--------------------T-AA---AT---A-C----AG---------C-----AT-AT--C-G---C-T-----G--G-TG...-A-----CA........................AACA-C 6824
C.IN.09.T125_2139            -------G--------------------G-G------CG--------------------TGCA---A-C--A-C----AAC----GC--C-----T--------G---C-G-----A--C----...GA---T--A.....................ACGTCAAC- 7648
C.KE.00.KER2010              ------------------------------G------CG--------------------TGCA---A-C--A-C----A-------A--C-----A--AT----GCA-C-T-----A--G--C...--T-CA..........................AATA-A 6805
C.MW.09.703010256_CH256.w96  A-----------C---------------GGG-------G--------------------TGCA---A-C--A-C----A----------C-----A--A-----G---C-T-----A--C----...GTT-CA..........................AAT--C 7577
C.TZ.08.707010457_CH457.w8   A-----------------------------G-----------------------------TGCA-GA-C--A-C----A-C--G----C-----AA-A---------C-T-----A--A----...-AT-GTG-C.......................AGT--T 7567
C.YE.02.02YE511              ----------------------------GGG------CG-----------------T--TGCA---A-C--A-C----A------C--C-----A--A-----G---C-T-----A--AGGA...-AAG--..........................AAC--- 6808
C.ZA.08.705010534_CH534.w12  ----------------------------GGG------CG--------------------TGCA---G-C--A-C----AA-----C--A-----A--AA-----G---C-T-----A--GG-A...GAA-G---A....................ACTAATA-T 7585
C.ZA.10.DEMC10ZA001          -------------C-----------------------CG--------------------TGCA---A-C--A-C----AGT---------C-----A--A-----G---C-T-----A--A-----...-CA-CAG--AAT..............ACAACAAATAATA-C 6948
C.ZM.11.DEMC11ZM006          -----------------------------GG------C---------------------TGCA---A-C--A-C----AA---------C-----A----------G---C-T-----A--AGCA...-AA--G---GAG..................AGTGACCCC 6986
D.CM.10.DEMD10CM009          ----------------------------GG-------------------------T--T-AA---------A-----------C-G-G---------AA--C-G---GT-----AA-T-----...-ATTCG-G-.........................CCT 6954
D.CY.06.CY163                C----------------T------C---GG---G---GT----------A--------CCTGCA---AT----CAG---AA--------------ATTACAGGGC-G-------------------...-AT-GT---.........................TGT 6826
D.KE.11.DEMD11KE003          ----------------------------GG-------GG--------------------T-GGGAA---TTG--C-C---------------C-----A--A--G--G---------CAA-AGGGC...TCA--T--C.........................ATT 6946
D.KR.04.04KBH8               ---------------------------AG---------G--------------------TGAA---GT---C-A------------C------------AA----G--------------CA---...-AT--T.............................AC- 7529
D.SN.90.SE365                ----------G---------C---------G--------------------------------TGAG---GTG--C--C------------------------A-----GCT------------G-TA...CAT--T-C-...................AGCAAT--- 7614
D.TZ.01.A280                 ----------------------------GG------------------------------GCA---T---C------A-----------------A--AT--C-------------------------...GTA--T---......................AGTAGA 6804
D.UG.10.DEMD10UG004          -------------------T----------G-A-------------------------------TGAA---T---C-A------C------------------A--AT-G--GGA----------A--GT-...GAA--T---.........................AGT 6926
D.UG.11.DEMD11UG003          ----------------------------GG------------------------------TGAA---G----C-A---AT-----------------A--AT-G--G-----G---------------...-AT--TGC-.........................ACC 6970
D.YE.02.02YE516              -------------------T--------GG-------G---------------------GCA---AT-----A------------------------A-----G----------------G-...-A---T-C-......................AACAGT 6796
D.ZA.90.R1                   ------G---------------------GG------------------------------AGAA----T-----------------------------AA----G-----G---------G--...-AT-GT-G-......................TCTC-C 6970
F1.AO.06.AO_06_ANG32         --------------G-------------GG-------C---------------G----TG-TGCA---A------CC---AAC-----C--------C--A-----G-------------------------..................AATAGTC-T 6783
F1.AR.02.ARE933              --------------G-------------G--------CG----ACA--CA-------CATGCA---A-C----CC---GGT---------------T--A-----G-------------C-...G-G--TTGCACTGAT.........GGTGGTACTGGGAATTGC 6886
F1.BR.10.10BR_RJ015          --------------G-------------AG---G---CGC-------------AA----G-TGCA---A------CC---AAC----C--------T--A-----G----------------...GAT--G-GCACT.................GRTAATA-C 7001
F1.CY.08.CY222               A---------------------------GG------GC----------------------TGCA---A------CC---AAC----G--------C--A-----G---------------C-A...--T--T--GAGTAAA..............GATAAT-G- 6782
F1.ES.02.ES_X845_4           A-------------G-------------GG-------C--------------G---T--TGCA---A-C----CC---AAC-------------C--A-----G-------------------...-AT--T-C-.......................AGT--- 7029
F1.RO.96.BCI_R07             ----------G-----------------GG------CC--------------G-------TGCA---AGC----CC---AAC-------------CT-A-----G---------------C-G..................................AATAAT--- 7670
F1.RU.08.D88_845             --------------G-------------GG-------CG--------------G-------TGCA---A-----GTC---AAC--M-GY-------T--A-----G-----G---------GR-...-ATG-T-C-.....................AGGACT 7102
F2.CM.02.02CM_0016BBY        --------------G-----G-------A-G------GCG--------------------TGCA---A-----CAG--CAAC-------C-----T--A-----G---------------A-G-...-AGGCA---..........................A-C 6775
F2.CM.10.DEMF210CM001        -------------------T--------A-G-------------------------------TGCA---A-----CAG---AAC---G---C-----T--AA---------T-------AA--GA...-CA--T---..........................A-- 6876
F2.CM.10.DEMF210CM007        ------G-------G-------------A--G------CG--------------------TGC----GT----C-G---AAC---G---C-----T---T-------------------GG-A...GAT--T-G-..........................AGC 6937
G.BE.96.DRCBL                AG------------G-G-GA---------G---G---C---------------------TGCA---A-C----C----AG------C--------A--AA-------------------GG--...-A---T-GC..........................ACT 7555
G.CM.10.DEMG10CM008          -G------------G---GA--------A-G-------C------------T--------TGCA---A-G----C---CAT------C--------A--A-----------G--------G--C...-A---T.............................AGT 6979
G.CN.08.GX_2084_08           -G------------G-G-GA-----------G---G---C------------T--C--T--TGCA---A-T----C----AG---T--C-----------AT-------------------CA---...-AT-CT............................-GA 6846
G.CU.99.Cu74                 AG--------G---G-G-GG------------G-------C--------------------GCA---A-C----C----GT------C--------A--AT---------------------A-C-...-AT--T-C-......................AATAGA 7235
G.ES.09.X2634_2              --------------G-G-GAT----------G---G-AGC--------------------TGAA---A-T--C-C----AGG-----C--------A--A------GT-------------------...-ATG-T-CA.....................AAG--- 7096
G.GH.03.03GH175G             A-------------G---GA-----------G---G---C--------------------GAA---G---------GAC----CA--------A--AA--------------------G--C...--T-----....................ACTAATA-T 7662
G.KE.09.DEMG09KE001          AG------------G-G-GA-----------G-------C--------------------GAG---A------C----AAC-----C--C-----AT-A--G--------------G--G-A...-AT--T-GC..........................ACT 6986
G.NG.09.09NG_SC62            -G------------G-G-GA---------------------C-----------------T---GAA---AGT----C----AG------C----------AT--C---TG-----------A---...-AT--T-CA........................AATA-- 6792
G.ZA.01.TV546                AT-----G------G-G-GA---------------G---C--------------------GAA---A-C----C----RT------C--------A--A--------------------A--G--G...-AT---.............................AGT 6803
H.BE.93.VI991                A-------------G-G-------------G-------C-----------------------A---A-C--C-C----AT--------------A--AA---G--TTT--CC-GAA----.........................................ACAA-T 7015
H.CF.90.056                  --------------G---------------G-------CG--------------------CAA---A-C----TG---GT--------------A--AA---G----TT--C-AG-----C...GCGTCTGCA..........................--A 6935
H.GB.00.00GBAC4001           A-------------------T---------G-------C---------------------CAA---A-T----C----GT--------------AT-AA---G---CTT--CA-GTA-GTA...GAT--TGG-..........................ACC 7124
J.CM.04.04CMU11421           A-----G------AG-G-GA-----------------C-------A--------------TGAA---A-C----C---C----------------AT-AA---G---C-T-----A--AG-C...-AG----GCACT................AGCAACAACAGT 7142
J.SE.93.SE9280_7887          A-------------G-G-GA--------A-GGAC----C-------C--------------TGCA---A-C----CG--CA---------------AT-A----G---C-T-----A--A--C...--GGGG---GGC.................AGTGAGAAT-GC 6923
J.SE.94.SE9173_7022          A-------------G-G-GA--------A-G-AC----C-------A----------------GCA---A-C----C---CAA---------------AT-AT---G---C-T-----A--A--C...--G--TGGC...................AGTGAGAAT-GC 6930
K.CD.97.97ZR_EQTB11          -G--------G--A--------------------G--GC-------A------------C-GCA---A-C----CC---AG---------------AA--A---G--------------------......GA---T---...................ACTAGGACT 6799
K.CM.96.96CM_MP535           -------G--G---------------------G-------------A--------------TGCA---AGT----AC---AGC---------------AA--A---G------------------...-AT-CT.............................C-T 6776
U.CA.01.TV749                --------------G-G-GA----------G-------C--------------------TGCA---A-T----C----GT------C-------CT-AT--C-G---C-T-----GAAAT--...--GCCA---.........................AGTACT 7656
U.CA.99.TV721                A-------------G-G-GA----------G-------C-----G--------------TGCA---A-T----C----A-------C-------CT-AT--C-G---C-T-----G-----C...-AT--T-G-..........................--T 7664
U.CD.83.83CD003_Z3           -G----------G-GG---GA-CA------GG-------C-------------------T---GAA-GAC-----G--C-----------------T--A-----G--------------------......G----TGCA........................ACTC-A 7092
U.CD.90.90CD121E12           --------------G---GA-CA------GG-------C-------------------T----AA-GAC-----TG--CAAC----G---------A--A------------------------G--...-ATG-G---.......................GATACA 6787
U.CY.05.CY090                C-T-T-GG-T----C---GT-----------GC-G---CT---CC-A--C---------A--TGCA--TA---G--C----A-C------TC------A-C-A----G------CA--A-A-TGG..................................ACTAG- 6790
U.CY.08.CY223                A-----G------AG---GA--------A-G-------C---------------------CAA--GGT---ACAT---CA------C--------A--A---------G--------A-AC--G.....................................AAT--A 6912
U.ES.10.DEURF10DZ001         -------G-----------A----------G-------C------A----------------C-A---A-C--A--T---CAC-----C-----------A--------------C--A--G--C...-AT-GT..............................ACA 6866
U.GR.99.99GR303              ----.--G-----------T----------G-------C----------------------CAA---G----A-AG---AAT---CC-----T--A----------------------------C...--T-GT---...........................--C 6897
U.NL.95.U_NL_95_H10986_D1    AC------------G-G-GA--------A-G-----A-C--------------------GCA---AG----CA----AT---------------AT-A--G------C-------G-----G...-AT--TG-A.........................ACTAAT-GA 7522
01_AE.AF.07.569M             -G-----G-------------------G---C----C-------------T---------------AT---A-AT---GT---------------AA-A-----G----------------G----...--T-GT..............................ACT 6807
01_AE.CF.90.90CF11697        -G-----G---G--G-------------G---------------------T-------------A-GC-G---AT---GT---------------AA-AA-------------------------------.......................GCGACT 7568
01_AE.CN.10.YNFL03           -G------------G---GA---------GG-------C-------------------------T-----A-C----TG---GT---------------AA-A-----G-----------------G--...-AT--T..............................AGC 6985
01_AE.HK.04.HK001            -G---------------------------G--AC---------------T------------G---AGC---CAG---GT---------------AA-A-----G-------------------...-AT--TGGG......................TCTAAC--- 6974
01_AE.IR.10.10IR.THR48F      AG-----G--------------------G--AC----C-------------T-----G----C---AT----AT----T---------------AA-A-----G-----T---AC---AGGA...-AA--T---.................GGGACAGGAGAAAAT-G- 6821
01_AE.JP.x.JRC77AE           AG-----G-------------------AGG--C----C-------------T----------G---A-G----AT---GT---------------AA-A-----G----------------------......--T-GTGCA........................A-T 7595
01_AE.TH.09.AA090a_WG11      -G-----G--------------------GG--C----C---------------------------G---A------AG---GT---------------AA-AT----G------------------...-TT--T---........................GCAACT 7044
01_AE.TH.90.CM240            AG-----G--------------------GG--C----C-------------T---------------AG-----AT---GT---------------AA-A-----G----------------GT-...-AT--T-CG...........................--T 7169
01_AE.US.05.306163_FL        AG-----G-----------T--------GG-------C-------------T---------------AG-----AT---GT---------------AA-A-----G----------------CT-...-AT--T..............................ACT 6729
01_AE.VN.98.98VNNDI5         -G-------------------------A-G-------C-------------T---------------AG-----AT---GT---------------AA-A-----G-----------------C-...-AT-G--GGACTAAC.........GAGACTAACGAGACTA-C 6883
```

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

```
                         V5 end                    Rev Responsive Element (RRE) region                          Env gp120 end   Env gp41 start
B.FR.83.HXB2             TCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTG..................CAGAGAGAAAAAAGA......GCA.. 7760
Env                      _S__E__I__F__R__P__G__G__G__D__M__R__D__N__W__R__S__E__L__Y__K__Y__K__V__V__K__I__E__P__L__G__V__A__P__T__K__A__K__R__R__V__V__.__.__.__.__.__.__Q__R__E__K__R__.__.__A__.
A1.AU.03.PS1044_Day0     AAT--A-C------------------------------------------------------G-------------GG---------C-----A------------C-------------------..................---------------......---.. 6998
A1.CH.03.HIV_CH_BID_V3538 A-A----------------A-------A------A------C---------------------------------------------C--------------------------------------..................G--------------......---.. 6951
A1.CY.08.CY236           A-A--A-----T--------------------------------------------------G------------------------C-----------------G------G-------------..................---------------......---.. 6977
A1.ES.05.X1608_8         AGT--A-C---------------------C---------------------------------------------------------C------------------G-------------------..................G--------------......---.. 7244
A1.KE.11.DEMA111KE001    AAT--A--------------------------------------------------------G------------------------C------------------G-------------------..................GG-------------......---AT 7159
A1.RU.11.11RU6950        AAT----C----------AT-------A--------------------------C-------G-----G------C-----------A-----A-------T----CA-----A------------..................G--------------......---.. 7287
A1.RW.11.DEMA111RW002    AAT--A-C------G-------------------AA--------------------------G-----G------------------C------------------G-------------------..................G--------------......---.. 7144
A1.SN.01.DDI579          AAT----CA------------------A------A----------------------------------------------------C----------------------------G---------..................G--------------......---.. 6964
A1.UG.11.DEMA110UG001    AAT--A-C------G-----------------------------------------------G---------------G--------C------------------C-----G-------------..................---------------......---GT 7088
A1.ZA.04.04ZASK162B1     AAT----C------G-----------------------------------------G--------C------A--------------C------------------G------A------------..................G--------------......---.. 7215
A2.CM.01.01CM_1445MV     AGT----C------G-----------------------------------------------G------------------------C-----A------------------G-------------..................G--------------......---.. 6965
A2.CY.94.94CY017_41      AAT----C----------------------------------------------------------------------C--------C-------------T----G------A------------..................G--------------......---.. 7132

B.BR.10.10BR_MG029       AAT----C---T----------------------A---T----------A-------------------------------------------A-------------------A------------..................--A------------......---.. 7257
B.CA.07.502_1191_03      A-T----C-----------A-------A------A---T----------A-------------------------------------------A------------------G-------------..................----------G----......---.. 7207
B.CH.08.M2_0803101_NFLG8 A-----G----------------------------------------------------------------G-----------G------------------------------------------..................---------------......---.. 7303
B.CN.10.DEMB10CN002      AAT----C---------------------------A-----------------------C-----C----------G-C----G---------------------G------------T-------..................---------------......---.. 7136
B.ES.10.DEMB10ES002      A-T--------------G--------------------T------------------------------------C-----------A-----A--------------------------------..................--------------G......---.. 7148
B.FR.11.DEMB11FR001      AA-----C----------------------------------------------------------------------------------------------------------------------..................--------G------......---.. 6957
B.GB.05.MM45d213_GN1     A-T--A-C---------A----------------------------------------------------------------------------------------T---------------------..................---------------......---.. 7178
B.HK.06.HK003            A-A-------------------------------------------------------------------------G----------------A------------------GA---GA-------..................--------G------......---.. 7120
B.HT.05.05HT_129389      AAT----C---T------------------------------------------C---------------------G----------------------------G--------------------..................---------------......---.. 6984
B.JP.12.DEMB12JP001      A----A-C---------------------------------------------------------C--------------------------------------T---------------------..................---------------......---.. 7170
B.KR.07.HP_18_07JHS10_3909 A-T--------------A--G-------------------------------------------------------G------G--------GA------------------G---G---------..................---------------......---.. 7423
B.PE.07.502_0525_wg5     A-T--------------------------------A--------------------------G--C---------------------------A-------------------------------A..................--A------------......---.. 7191
B.RU.11.11RU21n          A------C--------------------------------------------------------------------G----------------A--------------------------------..................--------G------......---.. 7288
B.TH.08.MERLBDTRC10      AAT----C------------G------A------AA-----------------------------------------------G---------A--------------------------------..................---------------......---.. 7010
B.US.11.ES38             AA-----C------------------------------------------------------------------------------------GT-----------------C--------------..................---------------......--T.. 7709

C.AR.01.ARG4006          A-A-----A-----G--------------------------------------------------Y-----G---G----CA-C---C-----A----------TR-T-----A--G---------..................G--------------......---AT 6950
C.BR.07.DEMC07BR003      GA------A-----G--------------------------------------------------------G---G--G--A-----------------------G-------A--G---------..................G--------------......---.. 7182
C.BW.00.00BW5031_1       AAT----CA----------C-------A------A------------------------------------G---G-----A-----------A----------T--------A--G---------..................G--------------......---.. 7095
C.CN.10.YNFL19           AAT----CA--------------------------------------------------------------G---G-----A-G-----G--GT----------TG-A-----A--G---------..................G--------------......---.. 7130
C.CY.09.CY260            A-A----CA-----------------------------T--------------------------------G---G-----A-----------A----------TG-------A--G---------..................G--------------......---.. 6995
C.ES.08.X2363_2          AAA-----A---------AC-------A-C----AA--------------------------------G--G---G-----A-G---------A----------T--------A--G---------..................----------G-GA-AAAAGA---GT 7240
C.ET.02.02ET_288         A-A----CA-----------------------------------------------------G--------G---G-----A-G--------------------TGG------A--GGC-------..................GG-------------......---.. 6968
C.IN.09.T125_2139        AAT----CA---------CT-------A---------------------A------------G--------G---G-----A-G---C-----A----------T-CA-----A--G---------..................G--------------......---.. 7792
C.KE.00.KER2010          A-T----CA-----G------------A------A------C-----------------C--------------GG-----A-G-----G---A----------TG-------A--G---------..................G--------------......---.. 6949
C.MW.09.703010256_CH256.w96 A-A-----A------------------A-------------------------------C-----------G---G-----A-G--------GA-------------------A--GC--------..................G--------------......---.. 7721
C.TZ.08.707010457_CH457.w8 C-A----CA---------AC-------------------------------------------------------G--G-A------------A----------TGGA-----A--G---------..................G-----C-G------......---.. 7711
C.YE.02.02YE511          --A----CA-----------------------------T--------------------------------G---G-----A-----------A----------TGCA--------G---------..................G--------------......---.. 6952
C.ZA.08.705010534_CH534.w12 A-A-----A--T-------C---------C-----------------G-----------------------G---G-----A-G--------------------TG-------A--G---------..................G-------G------......---.. 7729
C.ZA.10.DEMC10ZA001      A-A---G-A------------------A----A-AA-----------------------------------G---G-----A-----------A----T--A--T--A-----A--G---------..................G--------------......---.. 7092
C.ZM.11.DEMC11ZM006      C-A----CA--------------------------------C-----------------------------G---G--G--A-------G---T-----------------C-A------------..................GGA------------......---.. 7130

D.CM.10.DEMD10CM009      GAG-C--C------G--------------------------------------------------------G--G-G----------------A-------A------------------------..................G-A--------G---......---.. 7098
D.CY.06.CY163            AAT----C------G-----------------------------------------G-------------------G----------------A------------GA------------------..................G-A------------......---.. 6970
D.KE.11.DEMD11KE003      AAT----C-------------------A------AA-----------------------C----------------G----------C----TC------------G------A------------..................G-A------------......---.. 7090
D.KR.04.04KBH8           AAT----CT------------------A------AA------------------C--------------------C--G--------A-----C-------T------------------------..................G-A--G---------......---.. 7673
D.SN.90.SE365            A------C--------------------------------------------------------------------G-C-C------C------------------CA------------------..................G-A------------......---.. 7758
D.TZ.01.A280             GAG--------------------------------A-----------------------C----------------G---A------A------------------G-------------------..................G-A------------......---.. 6948
D.UG.10.DEMD10UG004      CAG----CTG---------T---------G-----AA----------------------C----------------G----A-----A----TC------------G-------------------..................G-A-A-C--------......---.. 7070
D.UG.11.DEMD11UG003      AAT----C---------------------------A-----------------------C----------------C----------C-G--TA-----------------------G--------..................G-A------------......---.. 7114
D.YE.02.02YE516          A-T----C------------------------------T-----------------G-------------------G----------C-------------------------A------------..................G-A------------......---.. 6940
D.ZA.90.R1               AAT----C---------------------------------------------------------C----------G----------A-----T------------G-------------------..................GCA--G---------......---.. 7114

F1.AO.06.AO_06_ANG32     A------C------------G------A------AA------------------C--------------------G--C---------------------------------GA------------GCGAGA............GG-------------......---.. 6933
F1.AR.02.ARE933          A-T----C------------G------A------AA--T--------------------------C---------G-----------C-------------------A----GA---CA------A..................A--------------......---.. 7030
F1.BR.10.10BR_RJ015      A-T--------T-----AT-G------A-------A-----------------------------C-----G---------------C-----A-------------------A---CC-------..................---------------......---.. 7145
F1.CY.08.CY222           AGT-----------------GA-----A----A-AA------------------C----------C---------G-----------C-----A-------------------A---CA------A..................A--------------......---.. 6926
F1.ES.02.ES_X845_4       AA---A-C------------G------A------AA------------------C--------------------G-----------C-G---A-------------------A---C--------GGGAGA............G--------------......---.. 7179
F1.RO.96.BCI_R07         A-T----C------------G------A------AA------------------C--------------------G-----------C-----A------------G----C-A---CA-------..................A-----------C--......---.. 7814
F1.RU.08.D88_845         AAT----C-----------C-------A-------A-----------------------C--------G--G----G----------C-------------------------A---CAT------..................G----GC--------......---.. 7246
F2.CM.02.02CM_0016BBY    A-A--C---C--------AT---G-----G-----A--------------------------------G------C-----C-----C-----A--------T---G------A---CA-------..................A---------G----......---.. 6919
F2.CM.10.DEMF210CM001    AGT----C-C-------------G----------AA------------------C-G-----------G-----G-------------------------------G------A---C--------..................AGC-----G------......---.. 7020
F2.CM.10.DEMF210CM007    A-T--C-C-C-------------G----------AA--T---------------C-G-----------G------------------C----T--G----------G------A---CA-------..................A-C------------......---.. 7081

G.BE.96.DRCBL            AGT-------------------------------AAA-------------------------G------AC---------CA--T--C-----A------------G-----G-------------..................G--------------......---.. 7699
G.CM.10.DEMG10CM008      ATT--A-C-----------C--------------A------------------------C--G------------------A-----C-----A------------G----CGA------------..................GGA------------......---.. 7123
G.CN.08.GX_2084_08       AAT----C------------------------------------------------------G------A-----------A-----C----GA------------------G-------------..................--------G------......---.. 6990
G.CU.99.Cu74             A-T---G------------C---------C-------------------------------GG------------------A-----C------------------R-----G-------------..................G--------------......---.. 7379
G.ES.09.X2634_2          AAT----CT----------C---------C--------------------------------G------------------A-----C-----A------------G-----G-------------..................GG-------------......---.. 7240
G.GH.03.03GH175G         A-A----C------------------------------------------------------G--C---AC---------CA--T--C-----A------------G-----G----C--------..................G--------------......---.. 7806
G.KE.09.DEMG09KE001      AA-----C-----------C---T-----------A--------------------------G---------A---G---CA-----C------------------------G-------------..................GGA-----------G......---GT 7132
G.NG.09.09NG_SC62        A------C----------TC------------------------------------------G------A-----------A---------A-A--------GG--------G-------------..................--------G------......---GC 6938
G.ZA.01.TV546            AGT-----------------------------------------------------------G------------------A-G---C------------------------G-------------..................GGA-----G------......---.. 6947

H.BE.93.VI991            AAT-TT-C-------------------------------------------------------------------------------C-G-----------A---G------G-------------..................G--------------......---.. 7159
H.CF.90.056              AATT-T-C------------------------------------------------G------------------------------C-G---A--------------A---G-------------..................G--------------......---.. 7079
H.GB.00.00GBAC4001       AAT-TT-C-C-----------------------------------------------------------------C-----------C-G---A------------------G-------------..................G--------------......---.. 7268

J.CM.04.04CMU11421       A-T--------T-----CAC--------------AA---------------------------------------G-G-----G---C------------------G------A------------..................G--------------......---.. 7286
J.SE.93.SE9280_7887      A-T----C----------AC-------A------A----------------------------------------G-G-----G---C-------------------------A------------..................G--------------......---.. 7067
J.SE.94.SE9173_7022      A-T----C----------AC-------A------AA--T------------------------------------G-GC----G---C-------------------------A------------..................G--------------......---.. 7074

K.CD.97.97ZR_EQTB11      GAG----C---------------------C-----A------------------C--------------------C-G---------C-----A------------G-----G-------------..................-----------G---......---.. 6943
K.CM.96.96CM_MP535       AAT----C---------------------C------------------------C-G------------------C-G---------C-----A------------G---C-G-------------..................---------------......---.. 6920

U.CA.01.TV749            AAT----C----------AC---------A----------T--------A------------T-----G-----GG-G---------C------------------------G-------------..................--------G--G---......---.. 7800
U.CA.99.TV721            AAT----C----------AC--------------------T---------------------T-----GA----GG-G---------C------------------------G-------------..................-----------G---......---.. 7808
U.CD.83.83CD003_Z3       AAT----C---T-----------------C----AA--T-----------------G-----G------------------------C------------------G------A------------..................G--------------......---.. 7236
U.CD.90.90CD121E12       AAA----C---T-----------------C-----A--T-----------------------G--C-----G---------------C------------------G------A------------..................G--------------......---.. 6931
U.CY.05.CY090            A-A----C---------------------------------------------GC---------------------GG-----G---C-----C----------T-G-----G----C--------..................G--------------......---.. 6934
U.CY.08.CY223            A-T--A-C----------AC------------------------------------------G------------------------C------------------G-----G-------------..................G--------------......---.. 7056
U.ES.10.DEURF10DZ001     AAT----CT-----G--------------C-----A--T-----------------------G------A-----------------C-T--T------------C-T-----A------------..................GG-------------......---.. 7010
U.GR.99.99GR303          -------C-------------------------------------S-----------------------------------------C-----------------GG------A------------..................G--------------......---.. 7041
U.NL.95.U_NL_95_H10986_D1 AAT----C-----------C-------A-C----A------C------------C----C-----------------G---------C-----A----------T-------G-------------..................---------------......---.. 7666

01_AE.AF.07.569M         AA-----C-----------A-------A----A-AA---------------------------------------C-----------C-----A------------G-------------------..................G--C--C--------......---.. 6951
01_AE.CF.90.90CF11697    AAT----C-------------------A------A----------------------------------------C-----------C-----A------------G----CG-------------..................G--------------......---.. 7712
01_AE.CN.10.YNFL03       A-T------------------------A------A----------------------------------------C-----------C-----A--------------------------------..................G--------------......---.. 7129
01_AE.HK.04.HK001        A------C---T--------------------A--A---------------------------------------C-----------C------------------G-------------------..................G--------------......---GT 7120
01_AE.IR.10.10IR.THR48F  A-T----C-----------C-------A------A----------------------------------------G-----------C-----A--------------------------------..................G-----C--------......---.. 6965
01_AE.JP.x.JRC77AE       AAT----C-------------------A----A-A------------------------C---------------C-----------C-----C------------G-------------------..................G-----C--------......---.. 7739
01_AE.TH.09.AA090a_WG11  AAT----C-------------------C-G----AA---------------AG------------------T---C-----------C------------------G-----G-------------..................G--------------......---.. 7188
01_AE.TH.90.CM240        AA-----C-------------------A-C--A-A----------------------------------------C-----------C-----A------------G-------------------..................G-------------G......---.. 7313
01_AE.US.05.306163_FL    AA-----C-------------------A----A-AA--------------K------------------------C-----------C-----C--------------------------------..................G-----C--------......---.. 6873
01_AE.VN.98.98VNND15     GAG-----------------------------A-AA---------------------------------------C-----------C-----A------------G-----G-------------..................G--------------......---GT 7029
```

V5 end
Rev Responsive Element (RRE) region
Env gp120 end
Env gp41 start
B.FR.83.HXB2
Env

```
B.FR.83.HXB2                .GTGGGA...ATAGGAGCTTTGTTCCTTGGGTTCTTGGGA...GCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGC 7923
Env                         __V__G__.__I__G__A__L__F__L__G__F__L__G__.__A__A__G__S__T__M__G__A__A__S__M__T__L__T__V__Q__A__R__Q__L__L__S__G__I__V__Q__Q__Q__N__N__L__L__R__A__I__E__A__Q__Q__H__L__L__
A1.AU.03.PS1044_Day0        .--T---...---------G-C---A----------A---...-----------------------G--G-----A--------------------G--------------C--------A-----A-G------------------A-----T-----A---A-----A 7161
A1.CH.03.HIV_CH_BID_V3538   .--T---...C-G------G----------------A---...-----C-----T-----------G--G-----A----------C---------G--------------------A--A-------G----C-------------A-----T---------------A 7114
A1.CY.08.CY236              .--T---...---------G-C--------------A---...-----------------------G--G-----A--------------------G--------------C--------A-----A-G------------------A-----T--------GA-----A 7140
A1.ES.05.X1608_8            .--T---ATG---------G-C--------------R---...--------------------------GG--C-A--------------------------------------------A-----A-G------------------A-----T---------A-----A 7410
A1.KE.11.DEMA1IKE001        T-GT--G...C--------G-T--------------A---...ATG--------------------G--G---------------------------A-C-----------C--------A-----A-G------------------A-----T-----A---------A 7323
A1.RU.11.11RU6950           .A-T---...C-G------GCC--------------A---...-----------T-----------G--G-----A-----------------------------------------A--A-----A-G------------------A-----T--G---------C--A 7450
A1.RW.11.DEMA111RW002       .--T---...--G------C-T---A--------------...-----------------------G--G---G----------------------------------------------A-----A-G-----------A------A-----T-----A---------A 7307
A1.SN.01.DDI579             .--T---...C--------G-C--------------A---...-----------T-----------G--G-----------------------------------------C-----A--A-------G----C-------------A-----T--------CA-----A 7127
A1.UG.11.DEMA110UG001       AT-T---...C-G------G-C---A----------A---...-----------------------G--G-----A-----------------------------------C--------A-----A-G------------------A-----T---------------A 7252
A1.ZA.04.04ZASK162B1        .--T---...C-G------G-T--------------A--G...--------------------------G-----A-----------------T--G--------------C--------A-----A-G------------------A-----T-----A---T-----A 7378
A2.CM.01.01CM_1445MV        .--T---...C-G------GCT------------------...--T--------------------G--G-----A-----------------------------------C--------A-----A-G-----------A------A-----T--------G------A 7128
A2.CY.94.94CY017_41         .--T---...C-G------G-C------------------...-----------------------G--G---C-A--------------------------------------------A-----A-G----------CA------A--A--T---------------A 7295

B.BR.10.10BR_MG029          .------...------------------------------...--------------------------G--C-----------------------------------------------A--------------------A--------------------C------- 7420
B.CA.07.502_1191_03         .-CA---...C-G------------A----------TC--...--------------------------A---C-A--------------A-----------------------------A--------------------A---------------------A------ 7370
B.CH.08.M2_0803101_NFLG8    .---ACG...C-------CA--------------------...--------------------------G--GC----------------------------------------------A-----------CC-----------------------------A------ 7466
B.CN.10.DEMB10CN002         .--A---...C--------A--------------------...--------------G-----------G-----A-------T------A---------------A-------------A-----A-G----------------------------------A------ 7299
B.ES.10.DEMB10ES002         .-C-......T-C-----A---------------------...--------------------------G---C-A--------------------------------------------A-------G----C-----------------------------A------ 7308
B.FR.11.DEMB11FR001         .------...---------C----------------A---...--------------------------G------------------------------C-------------------A------------C----A--------------------A---A------ 7120
B.GB.05.MM45d213_GN1        .A-A---...C--------G--------------------...--------------------------A---G----------------------R-----------------------A------------C------A---------------R------------- 7341
B.HK.06.HK003               .------...---------C----------A---------...--------------------------G---C----------------------------------------------A--------T---C-------A--------------G------A--C--- 7283
B.HT.05.05HT_129389         .------...C--------A----------------A---...--------------------------G-----AG---------C---------G-----G------------------------------C------------------------------------ 7147
B.JP.12.DEMB12JP001         .-----G...---------G--------------------...--------------------------G---C-A-T------------------------------------------A------------C----------C------------------A------ 7333
B.KR.07.HP_18_07JHS10_3909  .------ACG---------A--------------------...--------------------------G----------------------------T---------------------A--------T-----------A---------------------------- 7589
B.PE.07.502_0525_wg5        .------ACGC-G------G--------------------...--------------------------G-----A----------C----------------C----------------A------------A------A----------------------------- 7357
B.RU.11.11RU21n             .------...---------G--------------------...-----------------------G--G-----A--------------------G---------G-------------A------------C-------A-----------------A---------- 7451
B.TH.08.MERLBDTRC10         .------GTG---------A--------------------...--------------------------G---C----------------------------G-----------------A--------------------A---------------------------- 7176
B.US.11.ES38                .------...T-T-----GC--------------------...--------------------------G---G----------------------------------------------A--------------------A--------------------C------- 7872

C.AR.01.ARG4006             A------...C--------G--------------------...--------------------------A-----A--------------------------------------------------A-G------------------A---------------A------ 7114
C.BR.07.DEMC07BR003         .------...---------G-----------A--------...--------------------------A-----------------------------------------G--------A-----A-GT-----------------A---------------A------ 7345
C.BW.00.00BW5031_1          .-CA......C--------G--------------------...--------------------------A-----A--------------------------GC----------------A-----A-G------------A-----C---------------A----A- 7255
C.CN.10.YNFL19              .------...---------G--------------------...-----C-----------------G--G-----A-----------------------G--------------------A-----A-G------------------A---------------------- 7293
C.CY.09.CY260               .-----G...---------G--------------------...-----------------------G--G---C-A--------------------G-----GC----------------A-----A-G-----------A------A---------------------- 7158
C.ES.08.X2363_2             A------...C--------G--------------------...--------------------------A-----A--------------------------------------------A-----A--------------------A---------------------- 7404
C.ET.02.02ET_288            .-CA......C-----------------------------...--G-----------------------G-----A--------------------------------G-----------------A-G------------A-----A---------------A------ 7128
C.IN.09.T125_2139           .------ATG--G------A--------------------...--T--------------------G--G-----A--------------------G-----G-----------------A-----A-G------------------A---------------------- 7958
C.KE.00.KER2010             .------...---------G-AC-----------------...-----------------------G--G-----A------------------------C-G-----------------A-----A-G------------------A--A------------A------ 7112
C.MW.09.703010256_CH256.w96 .------...---------G-A------------------...--------------------------G-----AG-------------------------G-----------------A-----A--------------A-----A---------------A------ 7884
C.TZ.08.707010457_CH457.w8  .------...--G------A--------------------...--------------------------G-----A--------------------------------------------A-----A-G------------------A---------------A-A---- 7874
C.YE.02.02YE511             .-CA---...--G-----------------------TC--...--------------------------A-----A--------------------G-----G-----------------A-----A-G------------------A---------------------- 7115
C.ZA.08.705010534_CH534.w12 .------...------------------------------...A----------------------G--G-----A--------------------------------------------A-----A-------------A------A---------------A------ 7892
C.ZA.10.DEMC10ZA001         .------...---------A----------A---------...-----------------------G--G-----A------------------------C-GC----------------A-----A-GT-----------A-----A---------------A------ 7255
C.ZM.11.DEMC11ZM006         .------...---------G--------------------...-----------------------G--G-----C--------------------------GC-A-----C--------A-----A-G------------------A---------------------- 7293

D.CM.10.DEMD10CM009         .A-A---...C-G------A--------------------...--------------G-----------G-----A--------------------------------------------A-----A--T-----------------A---------------------- 7261
D.CY.06.CY163               .A-A---...T--------A--------------------...--------------G-----------G-----AG---------C-------------G-G-----C-----------A-----A--T--------------G--A-----------A---------- 7133
D.KE.11.DEMD11KE003         .A-A--C...C-G------A--------------------...--------------A-----------G---C-------------------------G---C----------------A-----A--------------A-----A---------------------- 7253
D.KR.04.04KBH8              .A-A--G...C--------A--------------------...--------------A-----------G---G-------------------------G--------------------A-----A--------------------A---------------------- 7836
D.SN.90.SE365               .A-C---...C------T-A--------------------...--------------G-----------G------G----------------------GC-------------------A-----A--------------A-----A---------------------- 7921
D.TZ.01.A280                .A-A---...C--------A--------------------...--------------G-----------G---T-------------------------G--------------------A-----A--------------------A---------------T------ 7111
D.UG.10.DEMD10UG004         .A-A---...C--------A--------------------...-----G--------G-----------G------G-----------------------G--C----------------A-----A--------------------A---------------------- 7233
D.UG.11.DEMD11UG003         .A-A---...C--------A--------------------...--------------G----------TG---T------------------T------GC----------C--------A-----A--------------------A---------------A------ 7277
D.YE.02.02YE516             .A-A---...C--------A--------------------...--------------G-----------G---G----------------------------------------------A-----A--T-----------------A-----------A---------- 7103
D.ZA.90.R1                  .A-A---...C---CGAGC-G-------------------...--------------A-----------G--TG-------------------------G--------------------A-----A--------------A-----A---------------------- 7277

F1.AO.06.AO_06_ANG32        .------...---------------T--------------...--------------------------G-----A-----------------------------------A--------A------------C-------A--------A-----G------A------ 7096
F1.AR.02.ARE933             .------...--G------------A------------C-...---------G-------------G--G-----A-----------------------G---C-------A--------A-----A--------------A--------A------------A------ 7193
F1.BR.10.10BR_RJ015         .A-A---GGAC-G------G--------------------...-----M-----------------G--G---C-------------------------------------A--------A-------G----C----------------A------------------- 7311
F1.CY.08.CY222              .----A-...---------G--------AC----------...--------------------------G---C-A-----------------------------------G--------------A-------------A---------A------------------- 7089
F1.ES.02.ES_X845_4          .------...------------------------------...--------------------------A-----A-----------------------------------A--------A------------C----------------A------------------- 7342
F1.RO.96.BCI_R07            .-C----...T-----CT----C-----------------...-----------------------G--G-----A-----------------------------------A--------A-----A-------------A---------A------------------- 7977
F1.RU.08.D88_845            .T-----...--G-----C---------------------...--------------------------G-----A-----------------------------------A--------------A--T--C-----------------A------------A------ 7409
F2.CM.02.02CM_0016BBY       .--A---...---------G--------------------...--------------------------G-----A--------------------------------------------A-----A-------------A------A--A------------------- 7082
F2.CM.10.DEMF210CM001       .------...C--------A--------------------...-----------------------G--G-----A--------------------------G-----------------A-----A-G-----------A---------A-----G-----CT------ 7183
F2.CM.10.DEMF210CM007       .------...---------G----------A---------...A----------------------G--G-----A----------------T---------------------------A-----A-G------------------A--A------------A------ 7244

G.BE.96.DRCBL               .--T---...G--------A-C--------------A---...A-------------------------G-----A----------------T------------------C--------A-----A-G------------A-----A--------G------------- 7862
G.CM.10.DEMG10CM008         .--T---...C-G------G-T--------------A---...-----------------------G--G-----A----------------T-C----------------C--------A-----A--------------A-----A---------------------- 7286
G.CN.08.GX_2084_08          .--T---ACG--G------C-C--------------A---...A----------------------G--G---C-A----------------T------------------C--------A-----A-G------------A-----A-----------A---------- 7156
G.CU.99.Cu74                .--A---...--GA-----G-C--------------A---...-----------------------G--G-----A----------------T---G--------------C--------A-------G------------------A--------G------------- 7542
G.ES.09.X2634_2             .-CA......C-G------A-C--------------R---...-----------------------G--G-----A----------------T------------------C--------A-----A-G------------------A--------G------A------ 7400
G.GH.03.03GH175G            .--T---...C-G------G-C--------------A---...-----------------------G--G-----A----------------T---G--------------C--------A-----A-G------------------A---------------A------ 7969
G.KE.09.DEMG09KE001         A--T---...G-G------G-C--------A-----ATC-...-----------------------G--G-----A----------------T------------------C--------A-----A-G------------A-----A--A-----G------------- 7296
G.NG.09.09NG_SC62           T--T---...--G------C-C--------------ATC-...-----------------------G--G---T-A-----------------------------------C--------A-----A--T-----------------A--------G--A-----AC--- 7102
G.ZA.01.TV546               .--T---...C-G------G-C--------------A---...-----C--------------------G-----A----------------T------------------C--------A-----A-G------------------A--------G--A---Y------ 7110

H.BE.93.VI991               .--A---...--G--------T------------------...--------------------------G--T--A---T----------------------------------------A-----A-G------------A-----AC--------------A------ 7322
H.CF.90.056                 .------...--G-------CT------------------...-----------------------G--G-----A--------------------G-----------------------------A-G------------A-----AC------G-------A------ 7242
H.GB.00.00GBAC4001          .------...--G--------T--------------R---...--------------------------G-----A------------------------C-------------------A-----A-G------------A-----AC----------A---A------ 7431

J.CM.04.04CMU11421          .C-----...---------G--------------------...A-------------------------G-----A----------------T---------G--------C--------A-----A-G-----------A------A-----T---------------A 7449
J.SE.93.SE9280_7887         .------...---------G--------------------...A----------T--------------G-----A----------------T---------G--------C--------A-----A-G-----------A------A---------------------A 7230
J.SE.94.SE9173_7022         .------...---------G--------------------...A-------------------------G-----A----------------T---------G--------C--------A-----A-G-----------A------AN-A------------------A 7237

K.CD.97.97ZR_EQTB11         .--A---...--------G---------------------...-----------------------G--G-----A--------------------------------------------A-----A--------------A-----A--A--------A--AA----A- 7106
K.CM.96.96CM_MP535          .--A---...T--------G-----T----A---------...-----------------------G--G-----A--------------------------------------------A-----A-G------------A-----A--A--------A--------A- 7083

U.CA.01.TV749               .------...--------------------------A---...--------------------------G-----A--------------------------------------------A-----A-GT-----------------A-----T-----A---------A 7963
U.CA.99.TV721               .------...--------------------------A---...A-------------------------G-----A-----------------------------------C--------A-----A-GT-----------------A-----T-----A---------A 7971
U.CD.83.83CD003_Z3          .-----G...--G-----------T--C------------...-----C-----------------G--G------G---------C---------------------------------------A--------------------A---------------------- 7399
U.CD.90.90CD121E12          .------...--G-----------T--C------------...-----C--------------------G-----AG---------C---------------------------------A-----A--------------------A---------------------- 7094
U.CY.05.CY090               .------...C-G---------------------------...--------------------------G---C-------------------------G--------------------A-----A--------------A-----A---------------A------ 7097
U.CY.08.CY223               .--T---...--G------A-C--------T-----A---...-----------------------G--G-----A-----------------------G------G----C--------A-------GT---C------AA-----A-----C--------A------A 7219
U.ES.10.DEURF10DZ001        .A-T---...--G--------C--------------A---...--------------------------G-----A----------C---------------------C-----------A-------G-----------A------A-----T-----AA--------A 7173
U.GR.99.99GR303             .--T---...--G-CG---A-CA-----------------GCA-----------------------G--G---G-------------------------------------C-------GA-------G----C-----CA---C--A-----T---------------A 7207
U.NL.95.U_NL_95_H10986_D1   .--A---...C-G------------T--------------...-----------------------G--G-----A--------------------------------------------A-----A--------------------A--A------------------- 7829

01_AE.AF.07.569M            .------...---------A--A--T----------A---...--------------------------G-----A-----------------------------A--------------A-----A-G------------------A--------G------A------ 7114
01_AE.CF.90.90CF11697       .------...---------A--A--T----------A---...-----------------------G--G---C-A--------------------------------------------A-----A-G------------------A--------G------------- 7875
01_AE.CN.10.YNFL03          .------...---------A--A--T----------A---...-----------------------G--G-----A--------------------------------------------A-----A-G----------C-------A---------------A------ 7292
01_AE.HK.04.HK001           A------...---------A--A--T----------A---...-----------------------G--G---T-A---------------------------C----------------A-----A-G------------------A---------------A------ 7284
01_AE.IR.10.10IR.THR48F     .------...---------A--A--T----------A---...-----------------------G--G-----A--------------------------------------------A-----A-G------------------A--------G------------- 7128
01_AE.JP.x.JRC77AE          .------...---------A--A--T----------A---...--------------------------G-----A--------------------------------------------A-----A-G------------------A--------G------------- 7902
01_AE.TH.09.AA090a_WG11     .------...---------A--A--T----------A---...-----------------------G--G-----A--------------------------------------------A-----A-G------------A-----A--------G------A------ 7351
01_AE.TH.90.CM240           .------...---------A--A--T----------A---...-----------------------G--G-----A--------------------------------------------A-----A-G------------------A--------G------------- 7476
01_AE.US.05.306163_FL       .--A---...------A-CA--A--T----------A---...-----------------------G--G-----A--------------------------------------------A-----A-G------------------A--------G----G-A------ 7036
01_AE.VN.98.98VNNDI5        A------...---------A--A--T----------A---...-----------------------G--G-----A--------------------------------------------A-----A-G------------------A--------G------------- 7193
```

HIV-1 Genomes

RRE end
B.FR.83.HXB2 AACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGT...AATAAA 8075
Env Q L T V W G I K Q L Q A R I L A V E R Y L K D Q Q L L G I W G C S G K L I C T T A V P W N A S W S N K
A1.AU.03.PS1044_Day0 7313
A1.CH.03.HIV_CH_BID_V3538 7266
A1.CY.08.CY236 7292
A1.ES.05.X1608_8 7562
A1.KE.11.DEMA11KE001 7475
A1.RU.11.11RU6950 7602
A1.RW.11.DEMA111RW002 7459
A1.SN.01.DDI579 7279
A1.UG.11.DEMA110UG001 7404
A1.ZA.04.04ZASK162B1 7530
A2.CM.01.01CM_1445MV 7280
A2.CY.94.94CY017_41 7447
B.BR.10.10BR_MG029 7572
B.CA.07.502_1191_03 7522
B.CH.08.M2_0803101_NFLG8 7618
B.CN.10.DEMB10CN002 7451
B.ES.10.DEMB10ES002 7460
B.FR.11.DEMB11FR001 7272
B.GB.05.MM45d213_GN1 7493
B.HK.06.HK003 7435
B.HT.05.05HT_129389 7299
B.JP.12.DEMB12JP001 7488
B.KR.07.HP_18_07JHS10_3909 7741
B.PE.07.502_0525_wg5 7509
B.RU.11.11RU2ln 7603
B.TH.08.MERLBDTRC10 7328
B.US.11.ES38 8027
C.AR.01.ARG4006 7266
C.BR.07.DEMC07BR003 7497
C.BW.00.00BW5031_1 7407
C.CN.10.YNFL19 7445
C.CY.09.CY260 7310
C.ES.08.X2363_2 7556
C.ET.02.02ET_288 7280
C.IN.09.T125_2139 8110
C.KE.00.KER2010 7264
C.MW.09.703010256_CH256.w96 8036
C.TZ.08.707010457_CH457.w8 8026
C.YE.02.02YE511 7267
C.ZA.08.705010534_CH534.w12 8044
C.ZA.10.DEMC10ZA001 7407
C.ZM.11.DEMC11ZM006 7445
D.CM.10.DEMD10CM009 7413
D.CY.06.CY163 7288
D.KE.11.DEMD11KE003 7405
D.KR.04.04KBH8 7988
D.SN.90.SE365 8073
D.TZ.01.A280 7263
D.UG.10.DEMD10UG004 7385
D.UG.11.DEMD11UG003 7429
D.YE.02.02YE516 7255
D.ZA.90.R1 7429
F1.AO.06.AO_06_ANG32 7248
F1.AR.02.ARE933 7345
F1.BR.10.10BR_RJ015 7463
F1.CY.08.CY222 7241
F1.ES.02.ES_X845_4 7494
F1.RO.96.BCI_R07 8129
F1.RU.08.D88_845 7561
F2.CM.02.02CM_0016BBY 7234
F2.CM.10.DEMF210CM001 7335
F2.CM.10.DEMF210CM007 7396
G.BE.96.DRCBL 8014
G.CM.10.DEMG10CM008 7438
G.CN.08.GX_2084_08 7308
G.CU.99.Cu74 7694
G.ES.09.X2634_2 7552
G.GH.03.03GH175G 8121
G.KE.09.DEMG09KE001 7448
G.NG.09.09NG_SC62 7254
G.ZA.01.TV546 7262
H.BE.93.VI991 7474
H.CF.90.056 7394
H.GB.00.00GBAC4001 7583
J.CM.04.04CMU11421 7601
J.SE.93.SE9280_7887 7382
J.SE.94.SE9173_7022 7389
K.CD.97.97ZR_EQTB11 7258
K.CM.96.96CM_MP535 7235
U.CA.01.TV749 8115
U.CA.99.TV721 8123
U.CD.83.83CD003_Z3 7551
U.CD.90.90CD121E12 7246
U.CY.05.CY090 7249
U.CY.08.CY223 7371
U.ES.10.DEURF10DZ001 7325
U.GR.99.99GR303 7359
U.NL.95.U_NL_95_H10986_D1 7981
01_AE.AF.07.569M 7266
01_AE.CF.90.90CF11697 8027
01_AE.CN.10.YNFL03 7444
01_AE.HK.04.HK001 7436
01_AE.IR.10.10IR.THR48F 7280
01_AE.JP.x.JRC77AE 8054
01_AE.TH.09.AA090a_WG11 7503
01_AE.TH.90.CM240 7628
01_AE.US.05.306163_FL 7188
01_AE.VN.98.98VNND15 7345

RRE end

```
B.FR.83.HXB2                 AACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGT...AATAAA............... 8075
Env                          Q__L__T__V__W__G__I__K__Q__L__Q__A__R__I__L__A__V__E__R__Y__L__K__D__Q__Q__L__L__G__I__W__G__C__S__G__K__L__I__C__T__T__A__V__P__W__N__A__S__W__S__.__N__K__.__.__.__.__.__._
02_AG.CM.10.DE00210CM013     G------G-----------T--A---------------G--------C-------------------------------A--AC-------C----------G------C---------A-C--A--C------T----C------...------............... 7514
02_AG.CY.09.CY256            G------G-----------T--A---------------G--------C---------------G---------------A--A--G-----C-----------------C---------AG---A--C------T-----------...-----G............... 7252
02_AG.ES.06.P1423            -------R-----------T--A---------------G--------C---------------G---------------A--A--------C-----------------C---------AA------C-----CT-----------...----G-............... 7515
02_AG.GW.05.CC_0048          -------------------T--A---------------G--------C---------------G---------------A--A--------C-----------------C--------------A--C-----CT-----------...----G-............... 7449
02_AG.KR.12.12MHR9           -------G-----------T--A---------------G--------C-----G---------G---------------A--A--------C-----------------C---------A----A--------CT-----------...------............... 7761
02_AG.LR.x.POC44951          G------------------T--A---------------G--------C---------------G---------------A--A--------C-----------------C---------AA---A--C------A-----------...------............... 8046
02_AG.NG.09.09NG_SC61        -------G-----------T--A---------------G--------C---------------G---C-----------A--A--------A----------G------C--------CAA---A--------CTA----------...------............... 7279
02_AG.NG.x.IBNG              -------G-----------T--A---------------G--------C---------------GA--------------A--A--------C-----------------C---------A----A--C-----CT----C------...------............... 7581
02_AG.SE.94.SE7812           G------G-----------T--A---------------G--------C-A---GC------------------------A--A--------C-----------------C---------A----A--C-----CT-----------...------............... 7465
02_AG.SN.98.98SE_MP1211      -------G-----------T--A--------A------G--------C-------------------------------A--A--------C-----------------C---------A----A--------CT-----------AAT------............... 7280
03_AB.RU.97.KAL153_2         ----------------------------------G---G-------------G---------------------------------------------------------------------------------A-----------...------............... 7252
04_cpx.CY.94.94CY032_3       G------G-----------T--A---------------G--------C-------C-----------------------A--A--------C-----------------C---------AA---A--------CT-----------...------............... 7445
05_DF.BE.x.VI1310            -G--------------T--T--A---------------G----------------------------------------A---C-------CA------------A---C---T-----AG------C-----CT-----------...------............... 7481
06_cpx.AU.96.BFP90           -------------------T--A--------------------------------------------------------A--A--------C-----------------C--TC-----AA-----TC-----CA-----------...------............... 8123
07_BC.CN.98.98CN009          -------G-----------T------------A-----G--------A-A-----------------------------A-----------C-----------------C-----T--------A--------CT-C---------...--C---............... 7397
08_BC.CN.97.97CNGX_6F        -------G-----------T--------A---A-----G--------A-A-----------------------------A-----------C--------------T--C-----T--------A--------CT-C---------...--C---............... 7254
09_cpx.GH.96.96GH2911        -------G-----------T--A---------------G----------------G-----------------------A--A--------C-----------G-----C---------AA------C-----CT-----------...------............... 7247
10_CD.TZ.96.96TZ_BF061       -------------------T--A---------------G-------------G--C-----------------------A--A--------A----------G-A--------------AA------C-----CT----C------...------............... 7427
11_cpx.CM.95.95CM_1816       -------T-----------T--A---------------G--------A-A-----------------------------A-----------C-----------------C---------AA------C-----C-T----------...------............... 7474
12_BF.AR.99.ARMA159          -G-----------------T--A---------------G----------------------------------------A---C-------C-----------------C---G-----AA------C-----CT-----------...------............... 8040
13_cpx.CM.96.96CM_1849       -------G-----------T--A--------A---------------A-A-----------------------------A--A--------A-----------------C---------AA------------CT----C------...------............... 7500
14_BG.ES.05.X1870            -------------------A------------------G--------A--------------C--------G----------A------------------------------------AA------------C------------...------............... 7557
15_01B.TH.99.99TH_MU2079     --------------------------------------G----------------------------------------------------------------------C------------------------------------...------............... 7447
16_A2D.KR.97.97KR004         G------G-----------T--A---------------G-----------------------C-------------------A--------C-----------------C--------CTT------C-----CT-----------...----GG............... 7441
17_BF.AR.99.ARMA038          -G-----------------T--A---------------G----------------------------C--G------A----A---------------------------------------------------A-----------...------............... 7264
18_cpx.CU.99.CU76            -------G-----------T--A---------------G----A-----------------------------------A--A--------C-----------------C---C-------G-----------CT-----------...------............... 7383
19_cpx.CU.99.CU7             G------C-----------T--A---------------G--------C----------T--------------------A--A--------C-----------------C---------A-------C-----CT----------C...----C-............... 7141
20_BG.CU.99.Cu103            -------------------T--A---------------G----------A-----------------C-----------A---C-------C-----------------C---------AA------C-----CA-----------...------............... 7509
21_A2D.KE.99.KER2003         -------------------T--A---------------G-TT-------------------------------------A--A--G-------------------A------T-------A------C-----CT-----------...------............... 7267
22_01A1.CM.01.01CM_0001BBY   -------------------T--A---------------G--------------------------------------T-A--A--------C-----------------C--------CA-------C-----CT-----------...--C---............... 7242
23_BG.CU.03.CB118            -------------------T--A---------------G--------C-A-----------------C-----------A---C-------C-----------------C---------AA------------C------------...------............... 7503
24_BG.ES.08.X2456_2          -------------------T--A---------------G----------A-----------------C-----------A---C-------C-----------------C---------AAC-----C-----CT-----------...------............... 7526
25_cpx.CM.02.1918LE          -------------------T--A---------------G----------------------------------------A--A--------C-----------------C---------AA-------------A-----------...------............... 7274
26_AU.CD.02.02CD_MBTB047     -G-----------------T--A---------------G-------------------------------G--------A-----------C--------G---A----C---------AA------------CA-----------...------............... 8063
27_cpx.FR.04.04CD_FR_KZS     -------------------T--A---------------G-------------------------A---------T----A--A--------C------------A----C-----------------C-----CT-----------...------............... 8074
28_BF.BR.99.BREPM12609       ----------------------------------G--------------A-----------------------------A-----------------------------C---------AA-------------T-----------...------............... 7476
29_BF.BR.01.BREPM16704       --------------------------------------G----------------C-------G---------------------------A------------------------------------------------------...------............... 7520
31_BC.BR.04.04BR142          -------G-----------T------------A-----G--------AC------------------A-----------A-----------C-----------------C--------CAA-------------T-----------...------............... 7525
32_06A1.EE.01.EE0369         -------------------T--A---------------G-------GT----R-----G--------------------A--A--------A---A-------------C---C-----AA------C-------A----------...------............... 7701
33_01B.ID.07.JKT189_C        -------------------A--A-----------G---G-------------------------A------A--T----A--AC-------C------------AC---C-----------------------CT-C-C-------...----G-............... 7367
34_01B.TH.99.OUR2478P        -------------------T--A---------------------------------------------------T----A--AC-------C------------A----C-----------------C-----CT-C-C------C...----G-............... 7241
35_AD.AF.07.169H             G------G-----------T--A---------------G--------C---------------G---------------A--A--------C-----------------C---------A-------------CT----------C...------............... 7257
36_cpx.CM.00.00CMNYU830      -------G-----------T--A---------------G--------C---------------G---------------A--A--------C-----------------C---------AA------C-----CT-----------...------............... 7272
37_cpx.CM.00.00CMNYU926      -------G-----------T--A---------------G----A---C-------------------------------A--A--------C-----------------C---CA----A-C-----A-----CT----------C...------............... 7257
38_BF1.UY.03.UY03_3389       -G-----------------T--A-----------R---G----------------------------------------------------------------------C---------A--------------A-----------...------............... 7471
39_BF.BR.04.04BRRJ179        -G-----------------T--A---------------C------------------T------A--------------A---C-------C-----------------C---------AA------C-----CT-----------...------............... 7525
40_BF.BR.05.05BRRJ055        -G-----------------T--A---------------G----A-----------------------------------A---C-------C-----------------C---------AA-------------AT----------...------............... 7572
42_BF.LU.06.luBF_18_06       --------------------------------------G------------------------G---------------A---C-------C-----------------C---------AA------------CT-----------...------............... 7580
43_02G.SA.03.J11223          -------------------T--A---------------G----------A-----------------------------A-----------C-----------------C---------AA------------CA-----------...------............... 7558
44_BF.CL.00.CH80             -G-----------------T--A---------------G----------------------------------------A-----------C-----------------C---------AA------C-----CT-----------...-----G............... 7490
45_cpx.FR.04.04FR_AUK        -G-----G-----------T--A---------------G----------------------------------------A-----------C-----------------C---------A-------C-----CT-----------...------............... 8034
46_BF.BR.07.07BR_FPS625      -G-----------------T--A--A-----------------------------------------------------A---C-------C-----------------C---------AA------------CT-----------...-----G............... 7451
47_BF.ES.08.P1942            --------------------------------------G-----------------------------------A--A--------------------------TA---C------------------------A----C------...------............... 7492
48_01B.MY.07.07MYKT021       ----------T-----------A--------A------G--------------------------------AC-T----A--AC-------C------------A----C---------AA------C-----CT-C-C----G--...------............... 7295
49_cpx.GM.03.N26677          -G-----T-----------T------------------G----T------------------C----------------A--A--------C------------T----C---------T-A-----C-----CT-----------...------............... 7502
50_A1D.GB.10.12792           -----------------G-T--A---------------G----------------T-----------------------A---C-------C--------G---A----C---------AA------C-----CA----C------...------............... 7481
51_01B.SG.11.11SG_HM021      -------------------T------------------G-------------------------A--------------A---------------------------------------AA-------------------------...------............... 7266
52_01B.MY.03.03MYKL018_1     -------------------T--A---------------G-------------------------A------A--T----A--AC-------C------------A----C-----------------------CT-C-C-------...------............... 7404
53_01B.MY.11.11FIR164        -------G-----------T--A---------------G--------------------------------A--T----AA-CC-------C--------G---A----C-----------------C-----CT-C-C-------...------............... 7391
54_01B.MY.09.09MYSB023       -------------------T--A---------------G-------------------------A------A--T----A--AC-------C------------A----C-----------------C-----CT-C-C-------...------............... 7483
55_01B.CN.10.HNCS102056      -------------------T--A---------------C----A--------------------A------A--T----A--AC-------C------------A----C---------AA------C-----CT-C-C-------...------............... 7397
56_cpx.FR.10.URF5_patient_A  G---------T--------T--A--------A------G--------C-------------------------------A--A--------C-----------------C---C----------C--------CT-----------...------............... 7402
57_BC.CN.09.09YNLX19sg       -------G-----------T------------------G--------A-A------------G----------------A-----------C-----------------C-----T--------A--------CT-C-------A-...--C-G-............... 7282
58_01B.MY.09.09MYPR37        --------------------------------------G-------------------------------G---------------------------------A--------------A--------------------------...------............... 7426
59_01B.CN.09.09LNA423        --------------------------------------G-T------C---------------G---------------A------------------------A----------------------------CAT----------...------............... 7253
60_BC.IT.11.BAV499           -------------------T--------T---A-----G-T------C-A-------------G------------T--A--R--------C-----------------C-----------------------CT-----------...------............... 8035
61_BC.CN.10.JL100010         -------------------T--------T---A-----G--------A-A-------------G---------------A-----------C-----------------C-----T--------A--------CA-C---------...--C---............... 7453
62_BC.CN.10.YNFL13           -------------------T------------A-G---G--------A-C-----------------------------A-----------C---------------------C-T-----G--A--------CATC---------...--C---............... 7385
63_02A1.RU.10.10RU6637       -------G-----------T--A---------------G----------------------------------------A--A--------------------------C---------AA------C-----CT-----------...------............... 7601
64_BC.CN.09.YNFL31           -------G-----------T------------------G--------A-A-----------------------------A-----------C-----------------C-----T--------A--------CT-C---------...--C---............... 7419
65_cpx.CN.10.YNFL01          ----------------------A---------------G--T-------------------------------A-----A---C-------C---------------------------A--------------TC----------...------............... 7460
67_01B.CN.11.ANHUI_HF115     -------------------T--A---------------G-------------------------A-----GA--T----A--AC-------C------------G----C---------AA------C-----CT-C-C-------...------............... 7373
68_01B.CN.11.ANHUI_WH73      -------------------T--A---------------G--------------------------------A--T----A--AC-------C------------A----C---------AA------C-----CT-C-C-------...------............... 7367
70_BF1.BR.10.10BR_PE004      -------------------T---------------------------------------------------------------C-------A---------------------------AA-------------------------...------............... 7512
71_BF1.BR.10.10BR_PE008      -------------------------------------GG-------------G----------G---------------------------A------------------------------------------A-C---------...------............... 7738
72_BF1.BR.10.10BR_MG002      -------------------T--A---------------G--------T-------------------------------A---C-------A-----------------C---------AA------C-----CT-----------...---CT-............... 7555
O.BE.87.ANT70                GG--AT-TR-A-----T----GA--A----GA--TC-CC-G--A--CT-A----CC-TA---C--A----G--A-----AA-CC-A-----C--TAAA-----G--AG-C---TA---AT-A--AAAA------AGA-CA----TA...GGA--C............... 8128
O.CM.98.98CMA104             -G--AT-T--A-----T----GA--A----GA--TC-CC-G--A--CT-A----CC-TTA--C--A----G--A-----AAACC-A-----C--TAAA-----T--A--C---TA---AT-A--AAAA-----CAA--CA---G-C...GGAG-TAGG............ 7626
O.CM.98.98CMABB141           GG--AT-T--A-----T----GA--A----GA--TC-CC-G--A--CT-A----CC-TAA--C--A----G--A-----AAACTCA-----C--TAAA-----G--AG-C--TTA---AT-A--AAAG-----CAAG-CA----CA...GGA--T............... 7646
O.CM.98.98CMABB212           GGT-AT-T--A-----T--T-GA--A----GA--TC-CC-G-AA--CT-A----CCCTTA--C--AT---G--A-G---A-ACC-A-----C--TAAA----G-A-A--C--TTA---AT-A--AGAG-----CAAA-CA---G-A...G-CT-TAAT............ 7657
O.CM.98.98CMU5337            GGT-AT-T--A-----T----GA--A----GA--TC-CC-G--A--CT-A----CC-TTA--C--A----G--A-----AAACC-A-----C--TAAG----G---A--C--TTA---AT-A--AG-A-----CAAC-CA----CC...--GT-TAAT............ 7579
O.CM.99.99CMU4122            GG--AT-T--A-----T----GA--A----GA--TC-CC-G-AA--CT-A----CC-TAA--C--A----G--A-----AAACC-A--------TAAG----GG--A--C--TTA---AT-A--AAAA------AAC-CA---GCA...--CGGCACT............ 7598
O.FR.92.VAU                  GG--AT-T--A-----T--T-GA--A----GA--TC-CC-G--A--CT-A----CC-TTA--C--A----G--A----ATAACC-G-----C---AAGAAT-G---A--C---TA---AT-A--AAAG------AAA-CA---G-A...GGAG-TAAT............ 7694
O.GA.11.11Gab6352            GG--AT-T--A-----T----GA--A----GA--TC-CC-G--A--CT------CC-TAA--C--A----G--A-----AAACC-A-----C--TAAG----GG--A--C---TA---CT-A--T-AG-----CAAG-CA---GTA...--C---ACA............ 7354
O.SN.99.99SE_MP1299          GG--AT-T--A-----T----GA--A----GA--TC-CC-G--A--CT-A----CC-TAA--C--A-C--G--A-----AAACC-A-----C--TAAG-----G--A--C---TA---AT-A--AAAA-----CA-G-CA----CA...---TGCACA............ 8181
O.US.10.LTNP                 GG--GT-T--A-----T----GA--A----GA--TC-CC-G-AA--CT-A---TCC-TAA--C--A----G--A-----AAACC-G-----C--TAAA-----G-AAG-C--TTA---AT-A--AAAA-------ACTCA----CA...GGA--T............... 8052

N.CM.02.DJO0131              -----T--A----------T--A-------GT----A-G----T---A-A-----------T-G------G--AA----AA-TC-A-----C-----A------ACT--A---TAT--CA---------------AG-C------C...-----T............... 7499
N.CM.02.SJGddd               -----T-------------T--A-------------A-G----T---A-A-------T---T-G------G--AA----A--TC-A-----C-----A------AST--A---TAT--CA---------------AG-C------C...-G---T............... 7471
N.CM.04.04CM_1015_04         -----T--A----------T--A--------A----A-G----T---A-A-----------T-G------G--AA----AA-TC-A-----C-----A------ACT--A---TAT--CA---------------AG-C------C...---C-T............... 7477
N.CM.06.U14296               -----T-------------T------------------G----T---A-A-----------T--R-----G--AA----A--TC-A-----C-----A-------CT--A---TAT--CA---------------AG-C------C...-G---T............... 7523
N.CM.06.U14842               -----T--A----------T--A-------GA----A-G----T---A-A-----------T--------G--AA----AA-TC-A-----C-----A------ACT--A---TAT--CA---------------AG-C-------...-G---T............... 7545
N.FR.11.N1_FR_2011           -----T-------------T--A-------------A-G----T---A-A-----------C-GA-----G--AA----A--TC-A-----C-----A------ACT--A---TAT--CA-------------C-AG-C------C...-----T............... 7405

P.CM.06.U14788               G----T-T--G--------A-GA-------GA--TC-CC-G--T--CA-T----CC--TT---G------G--A-----A--AC-G-----A-----A---C--A-AR-A--TTAT---AA------A------AG--CC----CC...--C---AAT............ 7627
P.FR.09.RBF168               G----T-T--G--------A-GA-------GA--TC-CC-G--T--CA-T----CC---T---G------G--A-----A--AC-G-----A-----A---C--ACAG-A--TTAT---AA------A------AGG--C----CC...--C---AGT............ 8167

CPZ.CD.06.BF1167             -G---T-G--G--------A--A--------A--C---C-G--A--A-----G-A--------GA-----------------TC-A-----A--TG--AAC---T-G--C--TCAT-G-T-AT-AGAA------AGCTC----TTA...--C---ACCCTGTGCGGTAAT 8215
CPZ.CM.05.SIVcpzMT145        -G--TT-T--------A-----A-----------G---G-T--T-----------G-------G------G--A-----A--TC-A-----C--TA-A------ACT-----TC----------A-GC-----CAAG-CA---G-A...----TT............... 7584
CPZ.GA.88.GAB1               ----AT--A-------AG-A--A--A--A--A--C---T-G--T-----A-----G-----GC-------G---A-T-----CC-G-----C-----------GGCTG-C--TTAT--CA-G-----------CAACTC----CC-GGG-GC--T............... 8115
CPZ.TZ.06.SIVcpzTAN13        -G--T--T--A------G----A--------A--C---C-G--T--AA----G-AG---A-C-GA--C---------T-AA-TG-------A--TG-CAAC---T-GG-G--TCAT-G-A-----GTA-------AC-C----GC-...GC--GGAAACCCTGT...... 7704
CPZ.US.85.US_Marilyn         ----AT----T--------T--A-----T--A--C---G-A--T-----------------T--------G---A----A--TC-G-----C-----A------ACA-----TTAT--CA---------------A--CC------...--C--C............... 8080

GOR.CM.07.SIVgor2139_287     G----T-T--T--------A-GA-------GA--TC-CC-G--A--CA-A----CC-------GA--C--G--------A--AC-G-----A---A-A--G-G-T-A-----TTA----AA------A-----CAAG-CC----CC...GGG---AAT............ 7636
GOR.CM.07.SIVgorCP2135con    G----T-T--T--------A-GA-------GA--TC-CC-G--A--CA-A----CT-------GA--C--G--------A--AC-G-----A---A-A--G-G-T-A-----TTA----AA------A-----CAAG-CC----CC...GGG---AAT............ 7620
```

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

B.FR.83.HXB2
Env
02_AG.CM.10.DE00210CM013
02_AG.CY.09.CY256
02_AG.ES.06.P1423
02_AG.GW.05.CC_0048
02_AG.KR.12.12MHR9
02_AG.LR.x.POC44951
02_AG.NG.09.09NG_SC61
02_AG.NG.x.IBNG
02_AG.SE.94.SE7812
02_AG.SN.98.98SE_MP1211
03_AB.RU.97.KAL153_2
04_cpx.CY.94.94CY032_3
05_DF.BE.x.VI1310
06_cpx.AU.96.BFP90
07_BC.CN.98.98CN009
08_BC.CN.97.97CNGX_6F
09_cpx.GH.96.96GH2911
10_CD.TZ.96.96TZ_BF061
11_cpx.CM.95_95CM_1816
12_BF.AR.99.ARMA159
13_cpx.CM.96.96CM_1849
14_BG.ES.05.X1870
15_01B.TH.99.99TH_MU2079
16_A2D.KR.97.97KR004
17_BF.AR.99.ARMA038
18_cpx.CU.99.CU76
19_cpx.CU.99.CU7
20_BG.CU.99.Cu103
21_A2D.KE.99.KER2003
22_01A1.CM.01.01CM_0001BBY
23_BG.CU.03.CB118
24_BG.ES.08.X2456_2
25_cpx.CM.02.1918LE
26_AU.CD.02.02CD_MBTB047
27_cpx.FR.04.04CD_FR_KZS
28_BF.BR.99.BREPM12609
29_BF.BR.01.BREPM16704
31_BC.BR.04.04BR142
32_06A1.EE.01.EE0369
33_01B.ID.07.JKT189_C
34_01B.TH.99.OUR2478P
35_AD.AF.07.169H
36_cpx.CM.00.00CMNYU830
37_cpx.CM.00.00CMNYU926
38_BF1.UY.03.UY03_3389
39_BF.BR.04.04BRRJ179
40_BF.BR.05.05BRRJ055
42_BF.LU.06.luBF_18_06
43_02G.SA.03.JI1223
44_BF.CL.00.CH80
45_cpx.FR.04.04FR_AUK
46_BF.BR.07.07BR_FPS625
47_BF.ES.08.P1942
48_01B.MY.07.07MYKT021
49_cpx.GM.03.N26677
50_A1D.GB.10.12792
51_01B.SG.11.11SG_HM021
52_01B.MY.03.03MYKL018_1
53_01B.MY.11.11FIR164
54_01B.MY.09.09MYSB023
55_01B.CN.10.HNCS102056
56_cpx.FR.10.URF5_patient_A
57_BC.CN.09.09YNLX19sg
58_01B.MY.09.09MYPR37
59_01B.CN.09.09LNA423
60_BC.IT.11.BAV499
61_BC.CN.10.JL100010
62_BC.CN.10.YNFL13
63_02A1.RU.10.10RU6637
64_BC.CN.09.YNFL31
65_cpx.CN.10.YNFL01
67_01B.CN.11.ANHUI_HF115
68_01B.CN.11.ANHUI_WH73
70_BF1.BR.10.10BR_PE004
71_BF1.BR.10.10BR_PE008
72_BF1.BR.10.10BR_MG002
O.BE.87.ANT70
O.CM.98.98CMA104
O.CM.98.98CMABB141
O.CM.98.98CMABB212
O.CM.98.98CMU5337
O.CM.99.99CMU4122
O.FR.92.VAU
O.GA.11.11Gab6352
O.SN.99.99SE_MP1299
O.US.10.LTNP
N.CM.02.DJO0131
N.CM.02.SJGddd
N.CM.04.04CM_1015_04
N.CM.06.U14296
N.CM.06.U14842
N.FR.11.N1_FR_2011
P.CM.06.U14788
P.FR.09.RBF168
CPZ.CD.06.BF1167
CPZ.CM.05.SIVcpzMT145
CPZ.GA.88.GAB1
CPZ.TZ.06.SIVcpzTAN13
CPZ.US.85.US_Marilyn
GOR.CM.07.SIVgor2139_287
GOR.CM.07.SIVgorCP2135con

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

HIV-1 Genomes

Tat premature stop in HXB2
Tat end
B.FR.83.HXB2
Rev exon 2
Tat exon 2
Env
A1.AU.03.PS1044_Day0
A1.CH.03.HIV_CH_BID_V3538
A1.CY.08.CY236
A1.ES.05.X1608_8
A1.KE.11.DEMA11KE001
A1.RU.11.11RU6950
A1.RW.11.DEMA111RW002
A1.SN.01.DDI579
A1.UG.11.DEMA110UG001
A1.ZA.04.04ZASK162B1
A2.CM.01.01CM_1445MV
A2.CY.94.94CY017_41
B.BR.10.10BR_MG029
B.CA.07.502_1191_03
B.CH.08.M2_0803101_NFLG8
B.CN.10.DEMB10CN002
B.ES.10.DEMB10ES002
B.FR.11.DEMB11FR001
B.GB.05.MM45d213_GN1
B.HK.06.HK003
B.HT.05.05HT_129389
B.JP.12.DEMB12JP001
B.KR.07.HP_18_07JHS10_3909
B.PE.07.502_0525_wg5
B.RU.11.11RU21n
B.TH.08.MERLBDTRC10
B.US.11.ES38
C.AR.01.ARG4006
C.BR.07.DEMC07BR003
C.BW.00.00BW5031_1
C.CN.10.YNFL19
C.CY.09.CY260
C.ES.08.X2363_2
C.ET.02.02ET_288
C.IN.09.T125_2139
C.KE.00.KER2010
C.MW.09.703010256_CH256.w96
C.TZ.08.707010457_CH457.w8
C.YE.02.02YE511
C.ZA.08.705010534_CH534.w12
C.ZA.10.DEMC10ZA001
C.ZM.11.DEMC11ZM006
D.CM.10.DEMD10CM009
D.CY.06.CY163
D.KE.11.DEMD11KE003
D.KR.04.04KBH8
D.SN.90.SE365
D.TZ.01.A280
D.UG.10.DEMD10UG004
D.UG.11.DEMD11UG003
D.YE.02.02YE516
D.ZA.90.R1
F1.AO.06.AO_06_ANG32
F1.AR.02.ARE933
F1.BR.10.10BR_RJ015
F1.CY.08.CY222
F1.ES.02.ES_X845_4
F1.RO.96.BCI_R07
F1.RU.08.D88_845
F2.CM.02.02CM_0016BBY
F2.CM.10.DEMF210CM001
F2.CM.10.DEMF210CM007
G.BE.96.DRCBL
G.CM.10.DEMG10CM008
G.CN.08.GX_2084_08
G.CU.99.Cu74
G.ES.09.X2634_2
G.GH.03.03GH175G
G.KE.09.DEMG09KE001
G.NG.09.09NG_SC62
G.ZA.01.TV546
H.BE.93.VI991
H.CF.90.056
H.GB.00.00GBAC4001
J.CM.04.04CMU11421
J.SE.93.SE9280_7887
J.SE.94.SE9173_7022
K.CD.97.97ZR_EQTB11
K.CM.96.96CM_MP535
U.CA.01.TV749
U.CA.99.TV721
U.CD.83.83CD003_Z3
U.CD.90.90CD121E12
U.CY.05.CY090
U.GH.08.CY223
U.ES.10.DEURF10DZ001
U.GR.99.99GR303
U.NL.95.U_NL_95_H10986_D1
01_AE.AF.07.569M
01_AE.CF.90.90CF11697
01_AE.CN.10.YNFL03
01_AE.HK.04.HK001
01_AE.IR.10.10IR.THR48F
01_AE.JP.x.JRC77AE
01_AE.TH.09.AA090a_WG11
01_AE.TH.90.CM240
01_AE.US.05.306163_FL
01_AE.VN.98.98VNND15

```
                                                                        Tat premature stop in HXB2                                   Tat end
B.FR.83.HXB2                CCCCG......AGGGGA...CCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCCTTGGCACTTATCTGGGACGATCTGCGGAGCCTGTGCCTCTTCAGCTACCACCGCTTGAGAGACTTACTCTTG 8552
Rev exon 2                   P  .  .  E  G  .  T  R  Q  A  R  R  N  R  R  R  R  W  R  E  R  Q  R  Q  I  H  S  I  S  E  R  I  L  G  T  Y  L  G  R  S  A  E  P  V  P  L  Q  L  P  P  L  E  R  L  T  L
Tat exon 2                  P  R  .  .  G  D  .  P  T  G  P  K  E  *  K  K  K  V  E  R  E  T  E  T  D  P  F  D  *
Env                         T  P  .  .  R  G  .  P  D  R  P  E  G  I  E  E  E  G  G  E  R  D  R  D  R  S  I  R  L  V  N  G  S  L  A  L  I  W  D  D  L  R  S  L  C  L  F  S  Y  H  R  L  R  D  L  L  L
02_AG.CM.10.DE00210CM013    A--A-......G---A-...-------------G-A---CC-------G----------CA-------A--------G-A----------G-----T---A--G---T----------C-------A-------A-T--------------A-----------TG----- 7991
02_AG.CY.09.CY256           A--A-......C---A-...-----------A---A----C-------G----------CA-------AG-------G-G--C-------G-----T---AA-----G--------------A-------------T--------------A-----------TGC---- 7729
02_AG.ES.06.P1423           A--A-......----A-...---------------A----C-------G------C---CA---------------AG-GA---------G-----T---A--T---GC------A--C--A----A-------------T----------A------------G----- 7992
02_AG.GW.05.CC_0048         AT-A-......----AT...----G----------A----C-------G------C---CA---------------GG-G----------G-----T---C--G---GC---------------------------------T--------------------T-----A 7926
02_AG.KR.12.12MHR9          A--A-......----A-...---------------A----C-------G------C---CA--G--------A---AG-G----------G-----T---A-----------------C--------------------------------A------------G----- 8238
02_AG.LR.x.POC44951         G--A-......----A-...---------------A----C-------G------C---CA------------A---TG-G------T---G---G-T---A-----GC-------------A----------------------------A-----------TGC---- 8523
02_AG.NG.09.09NG_SC61       A--A-......----A-...--------------------C-------G----------CA-------A--------G-G----------G-----T---A------GC------------------------------------------A-----------T------ 7756
02_AG.NG.x.IBNG             A--A-......----A-...--------A------A----C-------G----------CA-----A----------G-G----------G-----T---A------GC------------------------------------------A-------------A---- 8058
02_AG.SE.94.SE7812          A--A-......----A-...----G----------A----C-------G------C-G-CA----------------G-G----------G-----T---A------GC------------------------------------------A-----------CG---C- 7942
02_AG.SN.98.98SE_MP1211     A--A-......------...G-A-----------------C-------G------C---CA------------G---G-G----------G-----T---A------GC---------C---------------------CT---------A-----------TG----- 7757
03_AB.RU.97.KAL153_2        ---A-......------...-------------------------------A--------------------C-------------------A---T---A-----------------C---A-----------T------T--------A------------------- 7729
04_cpx.CY.94.94CY032_3      -AA-CCAA...C-----...-T-----------G-----C---------------C---CA-------AG------------C-------------T----C----------------C-------A--------------------G--A-------A----------A 7925
05_DF.BE.x.VI1310           -----......------...----------------A--C--------G----------CA--------G------A--------G--------C-T--CA---------------------A---A------------------------A------------AC---A 7958
06_cpx.AU.96.BFP90          A---A......-C----...G-A-----------G-A----C-------G----------CA--G----AC------G-----------------G-T---A------GC---------------------------------T--------A-----------CGG---- 8600
07_BC.CN.98.98CN009         A---A......G-----...----------T--G-A----C------------------CA-----A-----------------------------T---A--G---GC-----------------A------------------------A-----------CA-A--A 7874
08_BC.CN.97.97CNGX_6F       A---A......G-----...--A-G-----T--G-A----C------------------CA-----A-AC---C----------------------T---A--G---GC---------C--A----A---------------T--------------------CA-A--A 7731
09_cpx.GH.96.96GH2911       A----......----A-...--------------------C-------G----------CA-----A--T------A-------------G-----T---C--G---GC---------C---A---A------------------------------------CA----A 7724
10_CD.TZ.96.96TZ_BF061      -----......------...-------------------------------------------G-----G---------------G----G-----T----A-G--------------C---A---A-------------------------------------A----A 7904
11_cpx.CM.95.95CM_1816      G--AC......-A--A-...G-------T----G------C-------G----------CA-------AC------G-----------C-G-----T---A--T---GC---------C-------A---------T-A-----------R-------------A----- 7951
12_BF.AR.99.ARMA159         G----......----A-...--------------------C-------G----------CA--G--A--T-------G-G----------------TA---A----------------C---A-------------T--------------------------------- 8517
13_cpx.CM.96.96CM_1849      --AA-......----A-...--A------------A----C-------G----------CA------------------------G--C-G-----T---A------GC------------------------------------------------------CA----- 7977
14_BG.ES.05.X1870           A--A-......----A-...-------------G------C------------------CA-----A-----------------------G-----T---A--G---GCG------------A---------------------------W------------CA----- 8034
15_01B.TH.99.99TH_MU2079    AG-A-......----A-...G-------A------A----C-------G------C---CA--G---A--G------G-GA---------G-----T---A--C---GC-----------------------------------------------A-------G---C- 7924
16_A2D.KR.97.97KR004        ----A......GA--AT...----------A--G-A----C-----C-G----------CA--G-------------G-G----------G-----T---A--G---G---------A--C--A--------------------------------------T-GCA---C- 7918
17_BF.AR.99.ARMA038         G----......----A-...--------------------C-------G----------CA--G--A---------A-AA-----G----------T----C----------------C---------------------T---------------------------G- 7741
18_cpx.CU.99.CU76           A--A-......CT----...-------------------C-----G---------C---CA--G-----G------AG-G----------G-----T---AA-----GC------G--C-------------------------------------------------G- 7860
19_cpx.CU.99.CU7            A-.........--A--G...----------A----A----C------------------CA-----A-----------------------------T---A------GC-------------A-------------T--------------------------CA---CA 7615
20_BG.CU.99.Cu103           A--A-......----A-...--------------------C----G--G----------CA-----AC---------G-G--C-------G-----T---AC-----GC---------C--C------------------------------------------------ 7986
21_A2D.KE.99.KER2003        -----......------...---------------------------------------C---G---------------------GC---------TG-CA--TA-C-----------------A---A------------------------A------------A----A 7744
22_01A1.CM.01.01CM_0001BBY  A--AA......GA----...-T-----------G-A----C-------G----------CA-------A---A---G-----C-------G-----T---A--G---GC------------------------------------------------------------CA----- 7719
23_BG.CU.03.CB118           A--A-......----A-...--------------------C-------G----------CA--------G-----------------CC-G-----T---C-GA---G--------------C--C------------------------------------------------ 7980
24_BG.ES.08.X2456_2         A--A-......----A-...-------------G-A----C----G--G------C---CA-----A--G------------------C-G-----T---AA-T---GC---------C--------------------------------------------------- 8003
25_cpx.CM.02.1918LE         A--A-......----A-...--A----------G------C--------------C---CA------------------A--C-------------T---C--T---G-------A--CT-A----A------------------------------------CA----- 7751
26_AU.CD.02.02CD_MBTB047    G---A......GA----...GT--C--------G------C------------------CA--------ACA----G-------------------T---A------GC------G--C--C----A-----------G-----------A------------T------ 8540
27_cpx.FR.04.04CD_FR_KZS    A---A......GA----...-T--------T--G-A----C------------------CA--G--AGAG------G-----------A-G---G-T---A------GC---------C-------A-------------CT--------------A----GTGC----- 8551
28_BF.BR.99.BREPM12609      T--A-......----A-...-------------G------C-------------------------AC--------------C----CAG-T----T---A---A-C--------A--C---A-------------T--------------A------------------ 7953
29_BF.BR.01.BREPM16704      TT---......------...--------------------C--------C-------------G------------------------AG-T----T---A---A-C-------------C--A-----------------------------A------------------ 7997
31_BC.BR.04.04BR142         T----......------...--A-------T--G--A---C------------------CA----------------G-G----------G-----T---A------GC---------C-----------------T---T----------A------------ACA--A 8002
32_06A1.EE.01.EE0369        A---A......-C----...G------------G-A------------G----------CA--G--ACAC--A---G----------CC-C-----T---A------G--------------C--------------------------------A-----------CGG---- 8178
33_01B.ID.07.JKT189_C       AG-A-......----A-...--------A----G-A----C--------------C---CA---------------GG-G----------G-----T---A--G---GC---------C--------T------------------------------------------ 7844
34_01B.TH.99.OUR2478P       AT-A-......----A-...--A------------A----C-------G------C---CA--G----A--------G-GA---------C-----T---AT-T---GC-----------A------------------------------------------------ 7718
35_AD.AF.07.169H            G--AA......G---AT...-T--G------A-G-A----C------------------CA-------AC------G-----------C-G-----T---A------GCT-----------C--------------------------------------------CAG---- 7734
36_cpx.CM.00.00CMNYU830     AT-A-......----A-GAA---------------A----CA------G------C---CA-----------A----G-G----------------T---A------GC----------T--A----------------------------------------CA----- 7752
37_cpx.CM.00.00CMNYU926     A--A-......----A-...---------------A----C-------G------C---CA--G-------------G-G--C-------G-----T---A------GC------------------------------------------------------------CA----A 7734
38_BF1.UY.03.UY03_3389      G----......----A-...--------------------C-------G----------CA-----A----------G-G--C-------G-----T---A--T---GC---------C---A---A--------------------G--A------------CACA--A 7948
39_BF.BR.04.04BRRJ179       G---A......CC--A-...--------------------C-------G----------CA--G--A----------G-GA---------G---T-T---A--G---G----------CG--A---A--------------------G--A------------CA-A--A 8002
40_BF.BR.05.05BRRJ055       -AG--......------...-------------G------C------------------------------------GGAA----------T----TA--A---A-C-----------C--A--------------T--------------------------------- 8049
42_BF.LU.06.luBF_18_06      -A-A-......------...-------------------C--------------------------------------GA-CT------G-T----T---AA-------------T---C--A----------T---------T-------A------------------- 8057
43_02G.SA.03.J11223         A--A-......C-A-A-...--------------------C-------G----------CA-----------------------------------T---A--G---GC------A--CG------A------------------------------------CA----- 8035
44_BF.CL.00.CH80            A----......------...--------------------C-------G----------CA--G-------------G-G--C-------------T---A--T---G----------CT--A---A--------------------G--T------------CA-A--A 7967
45_cpx.FR.04.04FR_AUK       A----......------...--------------------C-------G----------CA-----A-----------------------G-----T---A------G----------C-------A------------------------------------CA----- 8511
46_BF.BR.07.07BR_FPS625     G---A......GA--A-...-----------AG-------C-------G----------CA--G---------G---G-GG---------------T---C--T---G----------C---A---A--------------------G--A------------CA-A--A 7928
47_BF.ES.08.P1942           -T---......------...-------------G------C------------------------------------GG-----CG-C--G-----T---AA-----------AT---C--A----------------------------A------------------- 7969
48_01B.MY.07.07MYKT021      AG-A-......----A-...--------A------A----C--------------C---CA--G---------G---G-G----------G-----T---A------GC---------C--------------------------------------------------- 7763
49_cpx.GM.03.N26677         A---A......-C--A-...G------------G------C-------G----------CA--G--GAC-------G-----------C-------T---AT-GA--G----------C-----------------T-G--------GT-A------------TG----- 7979
50_A1D.GB.10.12792          -----......------...-----A---------------------------------CA--G-----G------A--------G----------T----A-------------A--C---A--------GT-------------------------------A----A 7958
51_01B.SG.11.11SG_HM021     -----......------...G-------------------C----G--------------------------C--T-GAGTCG-C--C--CT----T---A-------------T------A----------T--------------G---------------------- 7743
52_01B.MY.03.03MYKL018_1    AT-A-......----A-...---------------A----C-------G------C---CA--G--C----------G-G----------G-----T---T------GCG--------C--A------------------CT---------------------------- 7881
53_01B.MY.11.11FIR164       AR-A-......----A-CAA-T------A------A----C-------G------C---CA--G---------R---G-G----------G-----T---A--R---GCG-----M-----A-----------------------------A------------A----- 7871
54_01B.MY.09.09MYSB023      A--A-......----A-...--------------------C-------G------C---CA--G---------G---G-GA---------------T---A-AGA----------G-----A-----------------------------A-----------CA----- 7960
55_01B.CN.10.HNCS102056     AT-A-......----A-...--------A------A----C-------G------C---CA--G-------------G-G----------G-----T----C-G---GCG-----------A------------------CT---------------------CA-A--- 7874
56_cpx.FR.10.URF5_patient_A AT-AA......C---A-...-----------A-G-A----C--------------C---CA----------------G-GA---------G-----T---A------GC-------------A------------A-----CT---------A-----------TG----- 7879
57_BC.CN.09.09YNLX19sg      A---A......G-----...-A--------T--G-A----C------------------CA-----A-AC------G------------C-A----T---A------GC---------C-------A---------T--------------------G-----CA-A--A 7759
58_01B.MY.09.09MYPR37       A--A-......----A-...--A----AA-----------C-------G------C---CC----------------G-G--C-------------T---A------GC---------C--------------------------------------------CA----A 7903
59_01B.CN.09.09LNA423       AT-A-......----A-...--------------------C-------G------C---CA--G---------------A--C-------G---G-T---A------GCG--------C--A-----------------------------------------CA---C- 7730
60_BC.IT.11.BAV499          G----......------...--A-----A-TTAG--A---C---A--------AA----CA-----A-----------------------G-----T---A------GC---------C--------------------------------A------------A-A--A 8512
61_BC.CN.10.JL100010        A---A......G-----...-A--------T-AG-A----C------------------CA-------AC------G-------------------T---A--T---GC-----------------A------------------------A------G----CA-A-CA 7930
62_BC.CN.10.YNFL13          AT--A......C-----...----------T--G------C------------------CA-----A---------------C-------------T---A--G---GC---------C--------------------------------A-----------CA-A--A 7862
63_02A1.RU.10.10RU6637      A--A-......----A-...-------------G------C-------G------C---CA---------------AG-G----------------T---A--T---T---------------------------------------------AA------------GC---- 8078
64_BC.CN.09.YNFL31          G---A......G-----...----------T--G-A----C------------------CA-----AGAG------A-------------------T---C------GC---------C---A---A------T-----------------------------CA-A--A 7896
65_cpx.CN.10.YNFL01         AT-A-......----A-CAA--------A--A--------C-------G------C---CA--G--------------CA----------C-----T---C--G---GC---------C---A---A------------------------A-----------TA-A--A 7940
67_01B.CN.11.ANHUI_HF115    AG-A-......----A-...-------------G------C--------------C---CA--G-------------G-G----------G-----T---R--C---T--------------A----------------------------------------CA----- 7850
68_01B.CN.11.ANHUI_WH73     AT-A-......----A-...-------------G--A---C--------------C---CA---------------A-GGA---------G-----T----C----Y--------------A-------------------------G--------T-------G----- 7844
70_BF1.BR.10.10BR_PE004     -----......------...-------------------CC-G---------------------------------AGGA-A--------CT----T---A---A-C-------T---C--A--------------T--------------A------------------ 7989
71_BF1.BR.10.10BR_PE008     -T-A-......------...------G-------------C-G--G-----------G----------A---C--TGGG-A--C-----G-T----T---A---A-C-------T----------------A---------T-------A-------------------- 8215
72_BF1.BR.10.10BR_MG002     -T---......------...-------------G------C-------------------------C---------AGAAAC-----C-GGT----T---AA-----------AT---C---A------------------------T--G--A---------------- 8032
O.BE.87.ANT70               A--AA......GA--A-...G-A-GA-C---A-G-A---C--G--G--G---------AGA--G--GCC----G-GG--A-CC-CGCC-C-A--G-T----C-----GT-G-ACACG--C--CA-A-CAA-AATAT-G-GG-CT-------T------CA----GCA-CA 8599
O.CM.98.98CMA104            A--AA......TC--A-...GTA-GAGT---A-G-----C--G-----G---------AG------GCC---G-TG--A-C---GCA-C-A-----T----CA----GT-G-ACACG--C--CA-A-CAA-AAT-T-G-GG-CT-------T------CA----GCA-CA 8103
O.CM.98.98CMABB141          A--RA......TC--A-...G-A-GA-C---A-G-----C--G-----G---------AGA--G--GCC------TG--AGC--CGCC-C-A--G-T----C-----GT-G-ACACG--C---A---CAA-AAT-T-G-GG--T-------T------CA----GCA-CA 8117
O.CM.98.98CMABB212          A-AG-......CA--A-...G-A--AGC---A-G-A---C--G----------------CG-------CC---G-GG-CG-C---GCCTCCA----T-----CA--AGT-G-ACACG----CA---CAA-AAT-T-G-GG-TT--------T------CA----GCA-CA 8128
O.CM.98.98CMU5337           A--AA......CC--A-...G-A-GA-C------G----C--G-----G---------AGA--G--GCC--AG-TA--A-CC-CGCCTC-A-----T----CM----GT-G-ACACG--C--CA---CAA-AAT-T-G-GG----------T------CA----GCA-CA 8062
O.CM.99.99CMU4122           A--AA......TC--A-...G-A-GAGC---A-G-----C--G----------------AG-------GCC----TG--A-CC--GCC-CCA--T-T----CAT---GT-G-ACACG--C--CA---CAA-AAT-T-G-GG--T-------T------CA----GCA-CA 8075
O.FR.92.VAU                 AA-AA......GC--A-...GTA-GA-C---A-G-----C--G-----G----------CGA------GCG--G-GG-C--C---GCC-C-A--G-T----CAT---GT-G-ACACG--C--CA---CAA-AAT-T-G-GG-TT-------T------CA----GC--CA 8165
O.GA.11.11Gab6352           A--GA......TC--A-...G-A--A-C---A-G-A---C--G-----G---------AGA--G--GTC----TG-CAGCC-CGCCAC-A-----T----C------GT-G-ACACG--C--CA---CAA-AGTTG-G-GG--T-----T-T------CAT---GCA-C- 7837
O.SN.99.99SE_MP1299         A--AA......CC--A-...G-A--AGC---A-G-----C--G-----G---------AC----G-TGCC-CG-TG--A-CC-GGCC-C-A-----T----C-----GT-G-ACACG--C--CA---CAA-AAT-T-G-GG----------T------CA----GCA-CA 8667
O.US.10.LTNP                A--AA......GT--A-...G-A-GA-C---A-G-----C--G-----G---------AGA--G--GCC---G-GG-GA-C---GCA-C-A--G-T-----------GT-G-ACACG--C--CA-A-CAA-AATAT-G-GG--T-------T------CA----GCA-CA 8523
N.CM.02.DJO0131             -AA-AACAACC------...--A--------A----A---C-------GC-T-----GACA--G-------------G-GA-------C-G-----T----A-----G---------G--CT-CA---A-----TGA-----CT--------------C-----C------- 7985
N.CM.02.SJGddd              -AA-A......------...S-A--------A--------C-------GC-T-------CA--------G-------G-G----------G-----T---A--T---G---------G--C--C----A-----CGA-----CT--------------C-----C------- 7951
N.CM.04.04CM_1015_04        -AA-C......------...--A--------A--------C-------GC-T-------CA--------G------TG-G----------------T---A------G---------G--CTGCA---A---GTGAC-C--CT----------------------------- 7957
N.CM.06.U14296              -AA-A......------...--A--------A-G--A--C--G------C-T-------CA--------G-------G-G----------G-----T--CT------------------G--C--CA---A---GTGA------CT-----G--------C-----C------- 8003
N.CM.06.U14842              -AA-A......C-----...--A--------A----A--C--G-----GC---------CA----------------G-G----------------T---A------G---------G--C--CA---A----TGAC-----CT--------------C-----C------- 8025
N.FR.11.N1_FR_2011          -AG-A......C---AT...--A--------A--------C-------GC-T-------CA--------G-------G-G----------------T-CA-------G---------G--C--CA---A-----TGA-----CT--------------C-----C------- 7885
P.CM.06.U14788              A-GG-......G-CCC-...G-A-GA-TAG-ACC-----C---C-------------AGC---G--AC-G------G--CA----GC--G-A----TA---C-T--C-----------C---A---A-T--GTGG-GCAG-T------GATA--AGTCAG--GCA-A--- 8110
P.FR.09.RBF168              A-GG-......GATCC-...G-A-GA-TAG-ACC-----C-A-C-------------AGC---G--AC-G------G--CA----GC--G------T----C-T--CG----------C---AA--A----GTGG-GCAG-T------GAT---AGTC-G--GCA-A--- 8650
CPZ.CD.06.BF1167            -AG--......------...--A-GTCA---A-------C--------G------C-ATC---G---------G-GG-GGG--ATCCAGTA-----T---CA-------------A--CTGC--ACAG--CCT---G-GG-T------GAC--CAGT-A-------T-G- 8713
CPZ.CM.05.SIVcpzMT145       --AAT......CA----...--A---G---T--G--A--C---A--G-----------AC-T-------G--A------CA----G----G---G-T---C-T----G---------A-----T----A---CCT-AG----CT--------AA----------GTGCA-CT 8064
CPZ.GA.88.GAB1              T--A-......----A-...-AA-GA----T--G--A--C--C---G--------C---CA-------AG------AG-GA--------G-A--G-G----C----------------C--C--A-A----G-GA-T-GG--------GA------C-AG----GC--GC 8592
CPZ.TZ.06.SIVcpzTAN13       -GGA-......CA--AT...--A--GCA---A-G-----C-C----G------------C---G--GAT---G-GG-CG-CC-CGCCC-GA-----T--CAGCA-CG--------G--C--C----A---CCT-----GG-T------GAC--GTCTCA----CG---G- 8202
CPZ.US.85.US_Marilyn        -AA-T......C---A-...--A--------A-------C--------GC-C----A-A-C-----AC-T------A-CGA-------C-G-----T---A------G---------A-----T----A---CCTGA-----CT--------A--GAG----------A--A 8560
GOR.CM.07.SIVgor2139_287    A--A-......GTCCC-...G-A-GA-TAGTACC--A-----G------------AG-GC------AC--------G--CA----GC--GCA----T--AC--T--C------------A-CA-A-C----GTTG-G-GG-T-----GGACA--AGTCAT------A-C- 8119
GOR.CM.07.SIVgorCP2135con   A--A-......GTCCC-...G-A-GA-TAGTACC--A-----G------------AG-GC-------C--------G--CA----GC--GCA----T----C-G--C--------------CA-A-C----ATTG-G-GG-T-----GGACA--AGTCAT------A-C- 8103
```

Rev end (TAA) in some lineages
Rev end
B.FR.83.HXB2
Rev exon 2
Env
A1.AU.03.PS1044_Day0
A1.CH.03.HIV_CH_BID_V3538
A1.CY.08.CY236
A1.ES.05.X1608_8
A1.KE.11.DEMA11KE001
A1.RU.11.11RU6950
A1.RW.11.DEMA111RW002
A1.SN.01.DDI579
A1.UG.11.DEMA110UG001
A1.ZA.04.04ZASK162B1
A2.CM.01.01CM_1445MV
A2.CY.94.94CY017_41
B.BR.10.10BR_MG029
B.CA.07.502_1191_03
B.CH.08.M2_0803101_NFLG8
B.CN.10.DEMB10CN002
B.ES.10.DEMB10ES002
B.FR.11.DEMB11FR001
B.GB.05.MM45d213_GN1
B.HK.06.HK003
B.HT.05.05HT_129389
B.JP.12.DEMB12JP001
B.KR.07.HP_18_07JHS10_3909
B.PE.07.502_0525_wg5
B.RU.11.11RU21n
B.TH.08.MERLBDTRC10
B.US.11.ES38
C.AR.01.ARG4006
C.BR.07.DEMC07BR003
C.BW.00.00BW5031_1
C.CN.10.YNFL19
C.CY.09.CY260
C.ES.08.X2363_2
C.ET.02.02ET_288
C.IN.09.T125_2139
C.KE.00.KER2010
C.MW.09.703010256_CH256.w96
C.TZ.08.707010457_CH457.w8
C.YE.02.02YE511
C.ZA.08.705010534_CH534.w12
C.ZA.10.DEMC10ZA001
C.ZM.11.DEMC11ZM006
D.CM.10.DEMD10CM009
D.CY.06.CY163
D.KE.11.DEMD11KE003
D.KR.04.04KBH8
D.SN.90.SE365
D.TZ.01.A280
D.UG.10.DEMD10UG004
D.UG.11.DEMD11UG003
D.YE.02.02YE516
D.ZA.90.R1
F1.AO.06.AO_06_ANG32
F1.AR.02.ARE933
F1.BR.10.10BR_RJ015
F1.CY.08.CY222
F1.ES.02.ES_X845_4
F1.RO.96.BCI_R07
F1.RU.08.D88_845
F2.CM.02.02CM_0016BBY
F2.CM.10.DEMF210CM001
F2.CM.10.DEMF210CM007
G.BE.96.DRCBL
G.CM.10.DEMG10CM008
G.CN.08.GX_2084_08
G.CU.99.Cu74
G.ES.09.X2634_2
G.GH.03.03GH175G
G.KE.09.DEMG09KE001
G.NG.09.09NG_SC62
G.ZA.01.TV546
H.BE.93.VI991
H.CF.90.056
H.GB.00.00GBAC4001
J.CM.04.04CMU11421
J.SE.93.SE9280_7887
J.SE.94.SE9173_7022
K.CD.97.97ZR_EQTB11
K.CM.96.96CM_MP535
U.CA.01.TV749
U.CA.99.TV721
U.CD.83.83CD003_Z3
U.CD.90.90CD121E12
U.CY.05.CY090
U.CY.08.CY223
U.ES.10.DEURF10DZ001
U.GR.99.99GR303
U.NL.95.U_NL_95_H10986_D1
01_AE.AF.07.569M
01_AE.CF.90.90CF11697
01_AE.CN.10.YNFL03
01_AE.HK.04.HK001
01_AE.IR.10.10IR.THR48F
01_AE.JP.x.JRC77AE
01_AE.TH.09.AA090a_WG11
01_AE.TH.90.CM240
01_AE.US.05.306163_FL
01_AE.VN.98.98VNND15

Rev end (TAA) in some lineages
Rev end

```
                                                                                                                         Env gp41, gp160 end   Nef start
B.FR.83.HXB2                  CAGCCATAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAGTAGTACAAGGAGCTTGTAGAGCTATTCGCCACATACCTAGAGAATAAGACAGGGCTTG.GAAAGGATTTTGCTA.......................................TAAG.ATGGGT 8802
Nef                                                                                                                                                                                   _M__G_
Env                           T__A__I__A__V__A__E__G__T__D__R__V__I__E__V__V__Q__G__A__C__R__A__I__R__H__I__P__R__R__I__R__Q__G__L_#_E__R__I__L__L__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__*_
A1.AU.03.PS1044_Day0          T---A-------------GCT-----------------------A--G------AT-------------TTAA----------------C---------C-C.-----AGC--------.......................................---C.-----G 8061
A1.CH.03.HIV_CH_BID_V3538     ----A---A---------G-T-----------------------A--G---A--TT---C--------A-T---T--------------C-----------A.-----GC---AA--.......................................---C.------ 8014
A1.CY.08.CY236                T---A-------------GCT----------A-------A---------A---------AT-G----------AT-A-T--------------C---------C-T.-----GC--------.......................................---C.-----G 8040
A1.ES.05.X1608_8              T---A-----------A-GCT---------------W---A--G---A--CT-------------T-A-----------------T-----------A.-----AGC--------.......................................---C.-----G 8310
A1.KE.11.DEMA1IKE001          T---A-------------GCT-----------------------A--AG---A--AT-G----------C-T------------------C-----------T.-----AGC--------.......................................---C.-----C 8223
A1.RU.11.11RU6950             T---A----T--------G-T-----------------------A--G---------------------T-TA---------G----C---------GCA.-----GC-----A-.......................................---C.-----G 8350
A1.RW.11.DEMA111RW002         T---A---A---------GCT----------A-------A---------A---G---A--AT-------------TTA------A-C-----C-----------C.-----GC--------.......................................---C.------ 8207
A1.SN.01.DDI579               ----A----T----------T-----------------------A--G---A-----G--------C---T---T-----------G-C-----------A.-----GC--------.......................................---C.------ 7994
A1.UG.11.DEMA110UG001         T---A-------------GCT-----------A-------A--------A--G---A-CAT-G---G-------TT-----------------C--------TGCT.-----GC--------.......................................---C.-----G 8152
A1.ZA.04.04ZASK162B1          T---A---A---------GCT------G------------------A--G---A--AT-GT----------T-----------------T-----------A.-----AGC-----A-.......................................---C.-----A 8257
A2.CM.01.01CM_1445MV          T---AG--A-----------------------------------C--G---A--AT---C----------T-A---------C-----C-----------T.-----GC--------.......................................---A.-----G 8028
A2.CY.94.94CY017_41           T---AG-------------T------------------------A--GG--A-----TC----------T-A-----------------C-----------A.-----GC---A---.......................................---A.-----G 8195
B.BR.10.10BR_MG029            ----------------------A---------------------------A--AT--.---------C-T-A-T------C-C---GG----------G-----GC--------.......................................---.-----G 8299
B.CA.07.502_1191_03           ----T---------------------------------------T----A----G-------G---T---G---C----------G--------C.-----A-G--A---.......................................---.------ 8249
B.CH.08.M2_0803101_NFLG8      T---T-------------------------A--------T---C---A---G--T-------T---T--------------------------C-T.-----GC-----A-.......................................---AG----C 8346
B.CN.10.DEMB10CN002           ----T------------A-----------A---------C-------A--AT-G----------AT---------------------------C-A.-----A-----G--.......................................---.------ 8178
B.ES.10.DEMB10ES002           ----T----T-------------------A------------C---A------G-----T---T----------CT--------------GCA.--G---G--------.......................................---.------ 8187
B.FR.11.DEMB11FR001           ----------------------A-------------------A-------A--AT-------------AT---------GT-----C-----------------.-----GC--------.......................................---.-----A 7999
B.GB.05.MM45d213_GN1          -C----------------------C---C-------C------CA--A-----A---Y--T-----------T--------------------------------.------G---AG--.......................................---.-----G 8220
B.HK.06.HK003                 T---T-------------------------A---------C-------A---G-GT--------C-T--------------------------------.-----GC---G--.......................................---A.------ 8162
B.HT.05.05HT_129389           ----AG---------G---A---------A---------A-----AGGA-G-TGGT---------C-T---G-----C--------------------.-----GC------G.......................................---A.------ 8026
B.JP.12.DEMB12JP001           ----T-------------------C--AA-----------A----A--CT-------------TT---------GT----------------A.-----GC--------.......................................---.-----G 8215
B.KR.07.HP_18_07JHS10_3909    ----T-G----A--------------------C-------A--T----A-----A-----------T----------------------------T--T.-----GC--------.......................................---.------ 8468
B.PE.07.502_0525_wg5          T---T-------------------------AA-C-----------G--T-AACAT--T------AT--AT----------GCT-----C--------------.-----C--C------.......................................---.------ 8236
B.RU.11.11RU21n               ----T----------G------A-------------------A--A---T-------G-A-T-----------------------------C.-----GC--------.......................................---.------ 8330
B.TH.08.MERLBDTRC10           ----------------------------------------------A----C-A----------AT-----------------------------A.-------C---AG--TAAAATAAAA..ATAAGACAGGGCTTAGAAAGGACTTTAGTA---A.------ 8095
B.US.11.ES38                  T------------------------------A----------G-T-G---A------G-----T--AT--------------------------------.------------.......................................---.------ 8754
C.AR.01.ARG4006               T---A---------------A--A---------A-----------A-----C--AT---GGA-AT---CT-TA---------------------------T.---GCAGC-----A-.......................................---A.-----G 7993
C.BR.07.DEMC07BR003           T---A---------------A--A---------AA--------A--A-------AT---G--------CT--A---------------------------T.---GCAGC-----A-.......................................---A.-----G 8245
C.BW.00.00BW5031_1            T---A---------------A--A---------AA-------------------A--CT--A---------C-T-A---------------------------T.---GCAGCA--A-A-.......................................---A.-----G 8155
C.CN.10.YNFL19                ----T----------------------------A-C----C-A--A-R---A-------G----------..............-C-------G------A-CCT.---GTAGC-----A-.......................................---A.-----G 8181
C.CY.09.CY260                 --A-A-C-----C---G-G--CAG-------AA-AG-----T-GC---G-A--AT-A--G--AT---CT--AG---T-G-----------C-G-G-T---A.-C-GCAGC---A-A-.......................................---A.-----C 8061
C.ES.08.X2363_2               T---A---------------------A------A---------T-CA----A--AT------------CA--A--C-------------------T--T.---GCAGC----T--.......................................---A.-----G 8283
C.ET.02.02ET_288              ----A----T-----------A--A--------AA---------T--------A--TT---G--------CT--A--GC-------------------------T.---GCAGC--------.......................................---T.-----G 8028
C.IN.09.T125_2139             T---A---------------A--A----------A--------T-CC-G---A--CT------G-----CTA-A---------------------------A.---GCAGC-----A-.......................................---A.-----G 8858
C.KE.00.KER2010               T---A---------------A--A----------A--------T--A----A--AT-------------C-TTA---------C---------------C-T.---GCAGC--------.......................................---C.-----A 8012
C.MW.09.703010256_CH256.w96   ----A---------------A--A----------A--------T-CC---GA--AT-------------CTA---T---------------------------T.---GCAGC-----A-.......................................---A.-----G 8784
C.TZ.08.707010457_CH457.w8    T---A------AC-------A--A----------A--------C-------A--AT------G------CT--A--G------C---------------C-T.---GCAGC-----A-.......................................---C.-----G 8774
C.YE.02.02YE511               T---A---------------A--A---------AA--------A-------A--AG----------C-CT--A---------------------------T.---GCAGCA--A-A-.......................................---A.-----G 8015
C.ZA.08.705010534_CH534.w12   ----A-------------ACA--A----------A-----------A----A---GG------------CT--A--------------------A-----T.---CCGCA--------.......................................---C.-----G 8792
C.ZA.10.DEMC10ZA001           T---A---------------A--A---------AA--T-T----C-----T-A--TT------------C---A---C----C----------A-----T.---GCAGC--------.......................................---A.-----G 8155
C.ZM.11.DEMC11ZM006           T---A---AG-AC-------A--A----------A--------T------T-AATAT-----------C-T-AG---------------------------T.---GCAGC-----A-.......................................---A.-----G 8193
D.CM.10.DEMD10CM009           ----A-----------------A-----------------------C-----T-A-----T-----------TT-----C--C-----------------T--A.-----AGC--------.......................................---A.------ 8140
D.CY.06.CY163                 ----A------------------------------------AC----------A--A-----------G---TT-G-----C-----------------T--C.-----GC--------.......................................---A.------ 8015
D.KE.11.DEMD11KE003           T---A----T-A-------A----------AC--------A---GG-T-A--AT-G-----------ATTA-------C-----------------------.-----GC---A---.......................................---A.------ 8132
D.KR.04.04KBH8                ----A----------------------------------------------G-A------T-------G---TT-----------------------------C.-----GC----A--.......................................---C.------ 8715
D.SN.90.SE365                 T---A------------------------A-------C-----------A--------------------TT--------C-C-C-------------G---.---G--GC---A---.......................................---A.------ 8800
D.TZ.01.A280                  ----A------------------------A--------A---A-----A------T-------G---TT--------C-----G-------------------.-----GC---A---.......................................---A.------ 7990
D.UG.10.DEMD10UG004           ----A----T-----------A-------A-----------T-----A--AT-G-------------TTA---------C-C-C-------------GC-.-----GC---T---.......................................---A.------ 8112
D.UG.11.DEMD11UG003           ----A----T----------------A--------------T---G---A-----G-------G---TTA-T------C--G------------------T.--G---GC--------.......................................---A.------ 8156
D.YE.02.02YE516               ----A---------------A--------AA--------A---------A-G---G-----------TTA-------C----------------A--T--A.-----GC--------.......................................---A.------ 7982
D.ZA.90.R1                    T---A------AC-------------------------------------A-------C-------G---TTA--G----CGT-------C-----------A.-----C--------.......................................---A.------ 8156
F1.AO.06.AO_06_ANG32          ----A----T------------------------------------TT-G---A-----G--------G---T---------C------------A-----A.---------A-------.......................................---A.------ 7975
F1.AR.02.ARE933               ----A----T----------A--A----------A---------CT--G---A-----GT-------G--T-TA---------------------------.-----AGC---C---.......................................---A.------ 8072
F1.BR.10.10BR_RJ015           ----A----T-----G-----A----------AA---------CTT-G---A-GTT-G-----------C-T-A---------------------------T.-----GC----A--.......................................---A.------ 8190
F1.CY.08.CY222                T---A----T-------------------------A---------CTT-G---A-----G--------G---T-A--G-G-------A------------------------------.......................................---A.-----G 7968
F1.ES.02.ES_X845_4            ----A----T----------A-----------AA---------CTT-G---A--A--G-------------T----------------G-C----A-----A.-----AGC---A---.......................................---A.------ 8221
F1.RO.96.BCI_R07              T---A----T-------------------------A---------CTT---------TGCT--C--------T-A-------------------C-----------A.-----GC--------.......................................---A.------ 8856
F1.RU.08.D88_845              ----A----T---------T---------A---T------CTT-G---A-----G--------G-A-TT-------------------------------A.-----C--------TAAAA.................................---A.------ 8293
F2.CM.02.02CM_0016BBY         ----A----TG---------A-----------A--C------A-TC-G---A-----G------------T-------------------------T--C.-----GC--------.......................................---A.------ 7961
F2.CM.10.DEMF210CM001         ----A----T----------A-----------AA-C---------TT-G---A-----G------------T---------------G---------TGCA.-----GC--------.......................................---A.------ 8062
F2.CM.10.DEMF210CM007         ----A----T----------------------AA-C-----C--TT-G---A--AT-G-------G---TT---------------------------T.-----AGC--------.......................................---A.------ 8123
G.BE.96.DRCBL                 T---A-----------A-CT------------------------------C---A----G-------G---T-A------------------------------A.-----AGC--------.......................................---A.-----A 8762
G.CM.10.DEMG10CM008           ----A-------------TT------------------------------C---A---T-G-------T---TTA------------------A-----A.-----A---------.......................................---A.-----A 8165
G.CN.08.GX_2084_08            T---A-----------A-CT------------------------------C---CA----C----------CT--A-------------G--------A---C-A.-----AGC-C--G--.......................................---A.----G 8055
G.CU.99.Cu74                  T---AG----------A-CT------------------A-------A------A--A------G---T--T-A--------------G--------A-----T.------GCGC---AG.......................................---A.-----A 8442
G.ES.09.X2634_2               T---A-----------A-CT----------A-------A---------A---G---A---T-G---G----C-T-A-T------------------------.-----AGC--------.......................................---C.-----G 8300
G.GH.03.03GH175G              T---A-----------A-CT----------------------------A----A----G-C-----------TTA-T------C---G--------A-----T.-----AGC--------.......................................---A.--A--A 8869
G.KE.09.DEMG09KE001           T---A-----------A-C-----T-----------------A-----G---A---------G--G--T-------------------G--------A-----T.-----CGC----T--.......................................---.-----A 8175
G.NG.09.09NG_SC62             T---A----T-----G-A-CT-------------------------C---A-----G--------T---TT----------------G--------A---C-T.-----AGC--------.......................................---A.----AG 8002
G.ZA.01.TV546                 T---AG----------A-TT----------------------A--C---A--A--G-----------T--A-T--------------G--------------A.-----AGC--------.......................................--G........ 8003
H.BE.93.VI991                 ----A-----T----------A----------A-----------C---G---A-------G---------C-T------------------------------T.-----GC--------.......................................---CA----A 8202
H.CF.90.056                   ----A----------------A------G--A------T-A----G---A-------G---------T---------------------------------T.-----A-GC-------.......................................---TA----A 8122
H.GB.00.00GBAC4001            ----A----T-----------A----------A----M---C----G---A-----G-G---------AT-----------------------------T.-----GC--------.......................................---CM----A 8311
J.CM.04.04CMU11421            ----A----T--W------A--A---------A-C------A--W--AG-A------T-------G-A-TT--T--------------------T--T.-----GC--------.......................................---A.------ 8328
J.SE.93.SE9280_7887           ----A---------------A--A---------AA-C------A---C----A------T-----------TT------------------------T--A.-----AGC--------.......................................---A.-----G 8109
J.SE.94.SE9173_7022           ----A---------------A--A---------AA-C------A---C----A------T-----------TT------------------------T--A.-----AGC--------.......................................---A.-----G 8116
K.CD.97.97ZR_EQTB11           ----A---------------A--A---------AA-C-----GA-----T-CA------T-------T-A-TT-----------------------------T.-----AC----A---.......................................---A.------ 7985
K.CM.96.96CM_MP535            ----A--------------GA--A---------AA--------A---G----A------T-------T-A-TT------------------------TC-C.-----AGC---A---.......................................---A.------ 7962
U.CA.01.TV749                 T---A----------------------------A-------------G---A--AT--------------A-T--C---------------------------C.-----GC--------.......................................---A.------ 8863
U.CA.99.TV721                 T---A----T-----------------------AA------------G---A--AT-----------T--A-T--C---------------------------C.-----GC--------.......................................---A.------ 8871
U.CD.83.83CD003_Z3            ----A----T-----------A----------AA----------G-G---A--AT------G-------T-A-T------------------------T--C.-----GC---A--G.......................................---A.------ 8266
U.CD.90.90CD121E12            ----A----T----------------------A---------T-------A---T-G---G----C-T---T-------G--------------T--C.-----GC---A--G.......................................---A.------ 7961
U.CY.05.CY090                 ----A----T----------A--A---------AA--------A---G---A--AT--T--------T--ATTA--------------------------A.-----AGC--------.......................................---A.------ 7976
U.CY.08.CY223                 T---A---------------A--A---------AA--------A---G------------AT-----------TGA----------------T----------A.-----GC--------.......................................---CA------ 8099
U.ES.10.DEURF10DZ001          T---A------AC-----A-CT---T---C----C---------G-A-----A--AT-G----------T-A-TTA----------G---------A-----A.---AGCA--------.......................................---A.-----G 8073
U.GR.99.99GR303               ----A----T---------G-T--------------C------AC--G---GA--AT-G---------------T-----------------C-----A-----C.-----GC--------.......................................---T.------ 8092
U.NL.95.U_NL_95_H10986_D1     ----A---------------A--A---------AA--------A---G---A------T--------G---TG---------------------------T--T.-----GC--------.......................................---A.------ 8708
01_AE.AF.07.569M              T---A-----------G-G-T-------------------------A---------------G-----T--AT-----------------T-----A-----A.------------------.......................................---C.-----G 8014
01_AE.CF.90.90CF11697         T---AG----------G-G-T-------------------------------GA-------G-----C--AT-----------------C-----------A.-----GC--------.......................................---C.------ 8775
01_AE.CN.10.YNFL03            ----A---------T--G-G-----------------------------------A-------G----CC---TT-----------------C--G------C.-----GC--------.......................................---C.-----G 8195
01_AE.HK.04.HK001             ----A-----------A---T-----------------------------C---A-------G--G----C-TT-----------------CC---------A.------C----G--.......................................---C.-----A 8184
01_AE.IR.10.10IR.THR48F       ----A-----------G-G-T-----------------------------C---A-------G-----CC---T-----------------C-----------A.-----AGCA--------.......................................---C.-----G 8028
01_AE.JP.x.JRC77AE            ----A----------------T-----------------G--T----A--------------C---T------------------C---------C-A.------C--------.......................................---C.-----G 8802
01_AE.TH.09.AA090a_WG11       T---A-----------G-G-T-----------A------------A-----------C---A-------G-----C---AT---A-------G-----C-----------A.-----GC---AG--.......................................---T.-----A 8230
01_AE.TH.90.CM240             ----A-------C---G-G-T-------C-----------------------C------------G-----C---T-----------------C-----------A.------C--------.......................................---C.-----A 8376
01_AE.US.05.306163_FL         ----A------AC---G---T-----------------------------C---A--A--G-G-----C---T-A-----------------C-----------A.-----GC---A---.......................................---C.-----A 7936
01_AE.VN.98.98VNND15          ----A-----------G-G-T---------------C-------C---A-------G-----CT---T-----------------C-----------A.-----GC-----A-.......................................---C.-----A 8093
```

Alignments

Env gp41, gp160 end — Nef start

```
B.FR.83.HXB2                CAGCCATAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAGTAGTACAAGGAGCTTGTAGAGCTATTCGCCACATACCTAGAAGAATAAGACAGGGCTTG.GAAAGGATTTTGCTA.............................................TAAG.ATGGGT 8802
Nef                                                                                                                                                                                                _M__G_
Env                         T__A__I__A__V__A__E__G__T__D__R__V__I__E__V__V__Q__G__A__C__R__A__I__R__H__I__P__R__R__I__R__Q__G__L_#_E__R__I__L__L__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__*_
02_AG.CM.10.DE00210CM013    T---A------------A-CT-------C--A--------A--------A-----------------TA-A------------------T--------C-T.-----AGC-------.............................................---C.------ 8262
02_AG.CY.09.CY256           T---AG---------G-A-CT-------CG-A-------------G---A-----------------T-TA-----------C------C-------TC-A.-----AGC-----A-.............................................---C.------ 8000
02_AG.ES.06.P1423           T---A------------A-CT---------------------A--C---A----G-------T--TT-----------------C----A----C.------GC-------.............................................---C.------ 8263
02_AG.GW.05.CC_0048         T---A----T-------A-CT-------C--A--------GA--G---A-C---------------T-TA---------C------C--------C-C.------C--------.............................................---C.------ 8197
02_AG.KR.12.12MHR9          T---A------------A-CT----------A--C------A--G---A----G------------TTA-----------------T--------T--C.------GC-------.............................................---C.------ 8509
02_AG.LR.x.POC44951         T---A------------A-CT-------C--A---------A--G---A----G-------------TA---T---C------C--------TC-T.----A--C-------.............................................--GC.------ 8794
02_AG.NG.09.09NG_SC61       T---A------------A-CT-------C--AA--------A--G---A----G----........................................................................................................---C.------ 7973
02_AG.NG.x.IBNG             T---A----T-------A-CT----------A-C-------A--G---A--TGG-------------A-----------------C----------C.-----GC--------.............................................---C.------ 8329
02_AG.SE.94.SE7812          T---A----T-------A-TT-------C--A---------C--------A----G-----------T-A---------G----C----------T.-----GC--------.............................................---C.------ 8192
02_AG.SN.98.98SE_MP1211     ----AG-----------A-CT----------A---------A--------A--A--G------A--T--A-----------------C----------A.-----GC-----A--.............................................---C.------ 8028
03_AB.RU.97.KAL153_2        T---A-------------G-T---------------------A--G---A--TT------------G--TA----------G----C---------GCA.----A-GC-----A-.............................................---C.-----G 7979
04_cpx.CY.94.94CY032_3      ----A-----------------A----------A---------C------GA-----------------T--A----------------C---------C-T.-----AGC------T.............................................---A.-----A 8175
05_DF.BE.x.VI1310           ----AG---T-------------------A----------CTT-G--A----G-------------T-A-----------------------------A.------GC-------.............................................---A.------ 8208
06_cpx.AU.96.BFP90          ----A------------A-CT---------------GC-------------A--AT--T-------T--TTA--G-------------------A--A--T.-----AGC-------.............................................---A.-----A 8850
07_BC.CN.98.98CN009         T---A--------------A--A---------A-------T-----------CT-----------CTA-A--------------------------T.--GCAGC-----A--.............................................---A.-----G 8145
08_BC.CN.97.97CNGX_6F       T---A--------------A--A---------A-----A-CAG---------AT-----------C-A-A--G-----------------------T.--GCAGC-----A--.............................................---A.-----G 8002
09_cpx.GH.96.96GH2911       ----A----T------C---------------------------TT-G--A-G--G----------C-TT--T---------------C----------T.-----GC--------.............................................---C.------ 7974
10_CD.TZ.96.96TZ_BF061      T---A----A--------GCT---------------------A-------A----GTC------------TTA----------C-C-----------------.--G--GC--A-----.............................................---A.------ 8154
11_cpx.CM.95.95CM_1816      ----A----T---------A--A---------A-------G--A---TA--Y-TA------G---T---T-----------------C----------A.-----GC--------.............................................---C.-----A 8201
12_BF.AR.99.ARMA159         ----A----T--------G-------------AA--T------CTT-G---A--AT-G---------G---T-A--GC-----------G----------A---.-----GC--------.............................................---A.------ 8767
13_cpx.CM.96.96CM_1849      ----A--T------------T----------A-------GAC--G----A--AT-G-G---GC----T----------------G--C------------C.------GC--A-A--.............................................---C.------ 8248
14_BG.ES.05.X1870           T---A------------A-CT---------A-C----CW------G-------T-G----G--------T-A--G----GT-------------A--------.-----AT--C-TT--.............................................---C.-----G 8305
15_01B.TH.99.99TH_MU2079    ----A-----------G-G-T-------------------------C----A-------G-----C----T--G------------------C--------T--C.-----GC--------.............................................---C.-----A 8195
16_A2D.KR.97.97KR004        ----A---------------T---------A----------A--------A-------C---------AT-A-----------------C----------A.--G--AGC--A----.............................................---A.-----G 8189
17_BF.AR.99.ARMA038         ----A----T--------------------A----------CTT-G--A-----G---------G--T-A-----------------------------C.-----GC--------.............................................---A.------ 7991
18_cpx.CU.99.CU76           T---A------AC----A-CT------------------------A--------A--A--G------------TTA-----------------C--------GCC.----AGC-----A--.............................................---T.-----G 8131
19_cpx.CU.99.CU7            ----A----T------G-G-T---------------C--------A--A-----A--------G-------G--T--------------------C----------A.------GC-------.............................................---A.-----G 7864
20_BG.CU.99.Cu103           T---A------------A-CT-----------AA----------A--------A---GA-T--------T---TG-------------G-------A-----T.-----GCGC------.............................................---A.-----G 8257
21_A2D.KE.99.KER2003        ----A---------------------------------------A------G-A-------C-----------T-A-------------------C----------C.-----GC--A-----.............................................---A.-----G 7994
22_01A1.CM.01.01CM_0001BBY  ----A---A-------G-GAT-----------------------A--AC---A--------G------T---TG--T--------C-------T-------------.-----GC--------.............................................---C.------ 7990
23_BG.CU.03.CB118           T---A------------A-C-----------A---------A---------A--GA-T--------T---TG----------------------A-----T.-----GCGC------.............................................---A.-----G 8251
24_BG.ES.08.X2456_2         T---A------------A-C-----------A---------A---------A-G-GA-T--------T---TG-------------G--------A-----T.-----GCGC------.............................................---A.-----A 8241
25_cpx.CM.02.1918LE         T---A------AC------CT---------------------A--C---A----G-------------TTA--G-----GT----------------------A.----AGC--------.............................................---A.-----G 8022
26_AU.CD.02.02CD_MBTB047    T---AG--------------T-C-------------------A--C---A--G-G-------------T-------------------------------A.--G--GC-----A--.............................................---A.-----G 8811
27_cpx.FR.04.04CD_FR_KZS    T---A------AC----A-CT----------------------C---A-----G----------ATTA--------------G--------A---GCT.----AGC-----A--.............................................---A.-----A 8822
28_BF.BR.99.BREPM12609      ----T----------------------------A---------A--T----A--AT----------------A--------C------------------------A.------GC-------.............................................----.------ 8203
29_BF.BR.01.BREPM16704      ----T------------------------------------------------A--AT---G-----------T--T----C-C-------------------A.------GC-------.............................................----.------ 8247
31_BC.BR.04.04BR142         T---A--------------A--A---------A----------A---------AT-G-G--------CTA-A------------------------T.--GCAGC-----A--.............................................---A.-----G 8273
32_06A1.EE.01.EE0369        T---A--------------CT----------A-----------G---A---T--T--------T--TTA-T-------------------A--A--A.---TAGC--------.............................................---A.-----A 8428
33_01B.ID.07.JKT189_C       ----A-----------G-G-T-------------G-----------A----AT--T--G--------C-T----------------C----------T.-----TC--------.............................................---C.-----A 8115
34_01B.TH.99.OUR2478P       ----T---------------------------A------------------A------A-----------T----------------C----------A.-----GC--AA----.............................................---C.-----A 7968
35_AD.AF.07.169H            T---A-------------GCT---------------------------CA--G---A--TT-G-------------T----------------T----------A.--G-AGC------G..............................................---T.------ 8005
36_cpx.CM.00.00CMNYU830     ----AG--A-------A-G-T----------A--------------C---A--A--G------------TTA--G------C------C--------AC-A.-----GC--------.............................................---C.------ 8023
37_cpx.CM.00.00CMNYU926     ----A----T------A-G-T-------------------------A--C---A--------------------T------------------C--------T--TG-----GC-----A--.............................................---C.------ 8006
38_BF1.UY.03.UY03_3389      ----A----T----------A----------AA---------CTT-G--A----G----------G--T-A-------------------------C-T.----AGC----A---.............................................---A.------ 8198
39_BF.BR.04.04BRRJ179       ----A----T---------------------AA--C------CTT-G--AC-AT-G------G--C-T---------------------------A.-----TC--A-----.............................................----.------ 8252
40_BF.BR.05.05BRRJ055       T----------------------A--------A----------C-------A--AT-G-C--G-G--C-T------------------A-----A.-----A-G-------.............................................----.------ 8299
42_BF.LU.06.luBF_18_06      ----T---------------------------A-----------T----A--AT-G---G--G-C-T--T------C-----------------A.-----GC--------.............................................----.------ 8307
43_02G.SA.03.J11223         T---A-------------GCT-------------------------------C---A-------------------AT-A-------------G--------A-----T.----AGC--------.............................................---A.-----A 8306
44_BF.CL.00.CH80            ----A----T--------G-T----------A----------TT-G--A----G--------G---T----------------------------T.-----GC--A-----.............................................---A.------ 8217
45_cpx.FR.04.04FR_AUK       ----A-------------G-T----------------------AC-------A-TCT---------T---T-A------------------A---C-C.------GC-------.............................................---C.------ 8761
46_BF.BR.07.07BR_FPS625     ----A----T----------A----------AA---------CTT-G--A---GG--------G---T-A-------G------------------------.-----GC--------.............................................---A.------ 8178
47_BF.ES.08.P1942           ----A----T---------------------AA---------CTT-G--A-T--G--------G---T-----------------------------A.-----GC--A-----.............................................---A.-----A 8219
48_01B.MY.07.07MYKT021      ----A---A------G--A-------------A-------G---C--G-A-------G-----C--AT-------------------C----------A.-----GC--------.............................................---C.-----A 8034
49_cpx.GM.03.N26677         T---A--------------A--A--------A-C----------C--G-A------T--G--GC--TTA-------------C--------GC-T.-----GC--------.............................................---A.-----A 8229
50_A1D.GB.10.12792          T---A----T-------AGCT-------------C----------G---A--AT-G-------------T-------------------------TGCA.-----AGC-------.............................................---C.-----G 8208
51_01B.SG.11.11SG_HM021     ----AG------------------------------------------A--------A---T---G------T---TT----------C----------A-----A.-----A-G--A----.............................................---A.------ 7993
52_01B.MY.03.03MYKL018_1    ----T---------------------------A------------------A------A-------------T-----------C-------------------A.-----GC--------.............................................----.------ 8131
53_01B.MY.11.11FIR164       ----A-----------A-G-T-----------------------------A----C--G-----C----T-------------------T---------T--A.------GC-----AG.............................................---C.-----A 8142
54_01B.MY.09.09MYSB023      ----AG------------G-T----------A-------------------A--------G-----A----T-------------------C----------A.------GC-----AG.............................................---C.-----A 8231
55_01B.CN.10.HNCS102056     T---A----T------G-G-T-----------------------C--C------------G-G------------T--------------------C----------A.-----TC--------.............................................---C.-----A 8133
56_cpx.FR.10.URF5_patient_A TT--A------A-----A-CT----------A---------C---GG----C----G----------G-C-T---G-----------------------T.-----GC--------.............................................---C.------ 8150
57_BC.CN.09.09YNLX19sg      T------------------A--A---------A--------T-TA----A--AT-------AG----T-A------------A----------------T.A--GCAGC-----A-.............................................---A.-----G 8030
58_01B.MY.09.09MYPR37       ----A-----------G-G-T----------A---------------------A-CATC--G-----C----T--------------------C----------A.-----GC--------.............................................---C.-----A 8174
59_01B.CN.09.09LNA423       T---A-----------G-G-T-------------------------T---C----A-------G--G--CC---T----G---------------C--------T--A.------GC-----A-.............................................---C.-----A 8001
60_BC.IT.11.BAV499          ----T-------G-----A-----------A-A----------A--T-----A--TT-----AG---------T----------GC--A----------A-A----.-----A---------.............................................----.------ 8783
61_BC.CN.10.JL100010        T---A--------------A--A---------A--------T---------A--CT--A----------C---A--G------C--------------------T.-----GC--------.............................................----.------ 8201
62_BC.CN.10.YNFL13          T---A--------------A--A---------A--------AC-A-----A------A----------CT--A-----------G------------------T.---GCAGC-----A-.............................................---A.-----G 8133
63_02A1.RU.10.10RU6637      T---A------------A-TT-------------------------G---A---T-G----------C--TT--------A--------C--------C-C.-----GC--------.............................................---C.------ 8349
64_BC.CN.09.YNFL31          T---------------------------------------------A----A-----------------C---T--------------------------T.---GCAGC-----AG.............................................---A.-----G 8167
65_cpx.CN.10.YNFL01         ----AG--A-------A-G-T-------------------------C-------------G---------T---G----------------C----------A.-----GC--------.............................................---C.-----A 8211
67_01B.CN.11.ANHUI_HF115    ----A---R--A----G-G-T---------------------C--C---A-------G-------T--T-----------------C----A-----C.------GC---G---.............................................---C.-----A 8121
68_01B.CN.11.ANHUI_WH73     ----AG-----AC---G-G-T---------------------C--C---A-------G----------T-----------------C----A----T.------GC-------.............................................---C.-----A 8115
70_BF1.BR.10.10BR_PE004     ----A----T-----------A---------A---G-G---CTT-G--A-----G---------G--TTA--G---------G----------------C.------GC-------.............................................---A.------ 8239
71_BF1.BR.10.10BR_PE008     ----A----T---------A-----------A--C----G--CTT-G--A-CAT-G-----------T---TT--------------G------------------.------GC------G.............................................---A.------ 8465
72_BF1.BR.10.10BR_MG002     TT--A----T---------A-----------A---T------CTT-G--ACGCT-G----G-----CAT-----------------------------------.------GC---A-A-.............................................---A.------ 8282
O.BE.87.ANT70               TT--AG-G--------CA-TT----T--CG-CA-A--C-C--GGA-----A--ATAG-A-C--GA-----TA---C--A--G-----T-------------A.-----A-G---AT-G.............................................---C.-----A 8870
O.CM.98.98CMA104            TT--AG-G-------GCA--T----T--CGACA-A--CTT--GGA-----A-GATAG-A---GA--A-TTA---C--A--------T-------M--A.-----A-GC--AT-G.............................................---C.-----A 8374
O.CM.98.98CMABB141          TT--AG--------C--A-TT----T----ACA-A--CTT--GGA----GA-GTTAG-A---GA---T-GA---C--C-CG-----T-------G--T.---C-A-GC--T---.............................................---C.-----A 8388
O.CM.98.98CMABB212          TT--AG-G----CT--CA-TT----T--C-ACA-C--YTT--GTA-----A--ATAG-AC---GRY-C-T-A---C--A--------T----A--KGCA.-----AM-C--AK-G.............................................---C.-----G 8399
O.CM.98.98CMU5337           TT--AG-----ACT--YA-TT----T--C--CA-A--CTT--GGA-----A--ATAG-A---GA--YTATA---C--A--------T-------------A.--------G--AT-G.............................................---C.-----A 8333
O.CM.99.99CMU4122           TT--AG-------T--CA-TA-A--T--C--CA-AT-CTT--GGA-----A-GATAG-A---GA----TTA---C--A--------T-------------A.---CA-GC--AG-G-.............................................---C.-----A 8346
O.FR.92.VAU                 TT--AG-G-----T---A-TT----T--C--CA-C--CCT--GTA-----A-CATAG-G---GA----TTA-------A------G-T-----------C-T.--GC-AC-C---T--.............................................---C.-----A 8436
O.GA.11.11Gab6352           TT--AG-G-----G--CA-TT-------CG-CA-A--CTT--GGA-----A--ATAG-A---GA----T-A---C--A-C------T----------------.--G--A-G---A--G.............................................---C.-----A 8108
O.SN.99.99SE_MP1299         TT--AG-G-----T--CA-TT----T-TC-CAA-A--TTT--GGA-----A-GATAG-A---GA----TTA---C--A--------T-------------A.----A-G---AT---.............................................---C.-----A 8938
O.US.10.LTNP                TT--AG-G--------CA-TT----T--CG-CA-A--CTC--G-A-----A--ATAG-A---GA-----TA---C--A--G-----T-------------A.----A-G---AT-G-.............................................---CA-----A 8795
N.CM.02.DJO0131             T---T--T-T------A-----A-----C---C--------T---C----A--ATAG-A--G-GA--ATTA-----------------C---------C-A.-----AGCAC-T---.............................................----.-----A 8256
N.CM.02.SJGddd              ----T--T-T------A-----A---------C--------T---C----A--ATAG-A--G-GA--ATTA--T--------------C-----A---GCA.-----AGCA--AG--.............................................----.-----G 8222
N.CM.04.04CM_1015_04        ----T--T--------A-G---R---------C--------T---C----A--ATAG-A--G-GA--ATTA-----------------C---------C-A.-----A-CAC-TT--.............................................----.-----G 8228
N.CM.06.U14296              ----T--T-T------A-----A---------C--------T---C-------ATAG-A--G-GG--ATTA-----------------C---------C-A.----AGCAC--T---.............................................----.-----G 8274
N.CM.06.U14842              ----T--T-T------A-----A---------C--------T---C----A--ATAG-A--G-GA--CTTA-----------------C---------C-A.----A--AC-TG---.............................................----.-----A 8296
N.FR.11.N1_FR_2011          ----T--T-T------A-----A--------AC-------GT---C----A---TAG-A--G-GA--ATTA-----------------C-----A---C-A.-----AGCAC-TA--.............................................----.-----A 8156
P.CM.06.U14788              TT--T---T----T--AA-TT----T---CAA--A----C-----GG---CA-ATAG-A---GCT-CTTGA-------A-----GT-------A--GC-A.-----A-GC--A--G.............................................---A.-----G 8375
P.FR.09.RBF168              TT--T---T----T--A-G-T----T---CAA--A---AT-----GG---CA-ATAG-A---GCT--TTGA-------A-----G--------A--GA-A.-----A-GC------.............................................----.-----G 8919
CPZ.CD.06.BF1167            T---AGAGTTCACT-GCTG-T-------CG-----------C-C--AG--T---AGTAGACATA--AA-A--T-----A-CT-----C--G-----A--A.----TTGC-C-AAAC.............................................---CA-----A 9006
CPZ.CM.05.SIVcpzMT145       -C--T--T------GA--A--A--T------AA-CT------C-A--AC-C-CTTGG-C--G-GC---TTA--T-------------------T-----A-----A.----AGC-------T.............................................----.-----A 8335
CPZ.GA.88.GAB1              ----T--T--------A--A--T---------AA-A-------CTT-T----TTA-ACT----AT-----A-AA----C------GC-C--------A--GC--.-----AGC---A--TTAAGCACTGCAT..................................----.-----A 8875
CPZ.TZ.06.SIVcpzTAN13       TT--A-ATTTCAC---CTG-T----T--C-TAC-A-------G--C-GTTAACAT--T-----GC----A-AA----C--A-CC-----C-------A--T--A.----TAGC---AAAT.............................................----.-----A 8494
CPZ.US.85.US_Marilyn        ----A--------------A--A--T-----AA--------T--AC-AG-A-GCTC-T-CT--G----ATA-------A--------T--G--A---C-A.--G--A-GC--AT--.............................................----.------ 8829
GOR.CM.07.SIVgor2139_287    TT--TG--TGGACC---A-TT----T---CAA--A----C-A---CT----A--ATAG---G-GC--CTTAA--------C-----G--T-----A--A--A.-----A-GCC-AT--.............................................---A.-----A 8390
GOR.CM.07.SIVgorCP2135con   TT--TG--TGGACT-----TT----T---CAA--A----C-A---CT----A--ATAG---G-GC--CTTAA--------C-----G--T-----A--A--A.-----A-GCC-AT--.............................................---A.-----A 8374
```


Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546

```
B.FR.83.HXB2                 GGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCCTACTGTAAGGGAAAGAATGAGACGA....................GCTGAG......................................................CCAGCAGCAGATAGGGTGGGAGCAGCATC 8897
Nef                          _G__K__W__S__K__S__S__V__I__G__W__P__T__V__R__E__R__M__R__R__.__.__.__.__.__.__.__A__E__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__P__A__A__D__R__V__G__A__A__S
A1.AU.03.PS1044_Day0         -----------------A--CAGC-AG------T--GAAA-T-------A---A----A-....................A--CCT......................................................A---------AG-A--A-------T--- 8156
A1.CH.03.HIV_CH_BID_V3538    -----------------C--CA-AG-G--------GGAGA-T-----G--G---------....................A--CCA......................................................G---------AG-A--A--G----T--- 8109
A1.CY.08.CY236               -------------------CA-AG-G---------GAAA-C-----------A----A-.....................A--CCT......................................................G--------AG-G-A--A-------T--- 8135
A1.ES.05.X1608_8             -------------------CA-TG-G---------GAAA-T-----C-----A------.....................K--CCT...CCAGCA.............................................---------AG-G-A--A-------T--- 8411
A1.KE.11.DEMA1IKE001         -------------------CA-AG-G---------GAG--T-----------------T.....................---CCC......................................................T--------CCAG-A--A-------T--- 8318
A1.RU.11.11RU6950            -----A-------------CA-AG-G---------CAG--T-----------A------.....................---CCT...GCTCCT..........................................GCT---------AGCG---A---C---T--- 8454
A1.RW.11.DEMA111RW002        A---------------C---A-AG-G---------GAG--T--A--G---------A-......................---CGA...CGAGCTACT...............................CCTCCT----------CAG-A--A-------T--- 8317
A1.SN.01.DDI579              -----------C--------ACAAGAG-------A--AG--T-----T---------A-.....................A--CCC...CCAACA.............................................-----------AG-A--A-------T--- 8095
A1.UG.11.DEMA110UG001        -----A-------------CA-AG-G--T------GAG--T-----------A---AAC.....................A--CCT...................................................GCTCCT---------A-AG-A--A-------TC-- 8253
A1.ZA.04.04ZASK162B1         -------------------CA-AG-G---------CAGA-T-----------A----AG.....................A--CCT...ACAGCA.............................................G---------AG-A--A-------TT-- 8358
A2.CM.01.01CM_1445MV         A----A-T-------............---------G---C-----G---------A--.....................A--CCT......................................................-----------AG-A--A-------TG-- 8111
A2.CY.94.94CY017_41          -----------------G--CA-ACCA---------G--A-T-----G---------A--....................A--CCT...CCAACAGCACAAAGAACAGAA.......................GCAGTGTCT---------CCAG-A--A-------TG-- 8320
B.BR.10.10BR_MG029           --------------GC------G--AG--------------------------------.....................------......................................................-----------AG-----------T--- 8394
B.CA.07.502_1191_03          --------------GG--G--A--GG---G---AA-G-G--C-----T-----------.....................A--.........................................................-----------AG---------------- 8341
B.CH.08.M2_0803101_NFLG8     A-T----T-------T-C---CC-G--------T-----A-------------------.....................A-----...CCACGAAATGAA...........................CAAAATGAG-AG-------T-G--------------- 8462
B.CN.10.DEMB10CN002          ---------------CC---GC--GG-------T--G--A-------------------.....................A-----...CCA.....................................GTTGAA-----------AG---------------- 8282
B.ES.10.DEMB10ES002          -----------------C------G-G---------G--A-------C------------....................------...CCAGCA..................................GCTGAG----------CAG----------G----- 8294
B.FR.11.DEMB11FR001          ----TA-TT------C-C----G-GG----------G---------------------G.....................A---G-...ACTGAGCCAAAAGCAGAAAGGCTGGGGAAGGAAAGAGTGAGAAGAACTGAG-----------AG----A----------- 8145
B.GB.05.MM45d213_GN1         -----------------------G-----------R-----------------------.....................--ACGA..............................................GCTGAG-----------G-----------T--- 8321
B.HK.06.HK003                --------------C--G--AG-GA-------GA-G--A-T-----------A------.....................------...CCACGA..................................GCTGAG----------CCG-----------T--- 8269
B.HT.05.05HT_129389          -C-------------------AGA-----------G--A-------G------------.....................------...CCAGCT..................................GCTGAA-----------G---------------- 8133
B.JP.12.DEMB12JP001          -------------------CA-AG-----------G----------------A------.....................------...CCACGA..................................GCTGAG-----------AG-----------T--- 8322
B.KR.07.HP_18_07JHS10_3909   ---------------C--------TCC---G---GA-G--A----------------A--....................A-----......................................................-----------AG-----------T--- 8563
B.PE.07.502_0525_wg5         -----A--------GT----CCC-G---------------A-------C-----------....................------......................................................-----------AG---------------- 8331
B.RU.11.11RU21n              -------------------CT---------------GAA-------------CA-A--......................A-----......................................................-----------AG-----------T--- 8425
B.TH.08.MERLBDTRC10          --------------------C--G-----------G--A---------------A---......................------......................................................------------G-----------T--- 8190
B.US.11.ES38                 --------------G-TC----C-GA-------T---GA--------------------.....................----GA...ACA.....................................GCTGAG-----------AG---------------- 8858
C.AR.01.ARG4006              AA-----------.........T-G---------G------------------------.....................---CCT...CCTGGTGTAGGAGACAGMAGG...................AGACAAGATGCT-----T-----GG-A--A---C----T-- 8112
C.BR.07.DEMC07BR003          AA---A--------GT-----CCAG-G------T--G--A-----------------A--....................A-----......................................................-----------GG-A--A---C------- 8340
C.BW.00.00BW5031_1           ---------------TCC---C-AG----G------GAG-----A--T-----A---A--....................A----T......................................................-----------GG-A--A----------- 8250
C.CN.10.YNFL19               -C------------------CA-AG-----------G--A-------G--G---------....................A-----......................................................-----T------G-A--A--G-------- 8276
C.CY.09.CY260                AA-----------...--C--CT-A-GG---------G--A---A--G--G---------....................A---C-...CCAGCAGCAGAGAGGATGAGA.....................CAAACTGAG-----------GG-A--A-------G-- 8183
C.ES.08.X2363_2              -----------------C--AAGAG--------A--GAA-----A-----------A--.....................A--A--...CCAGCAGCAGAGCCA...........................GCAGTAGAG----------TGG-A--A--GC---G-- 8402
C.ET.02.02ET_288             A--GC----------T----CCCA-GG------T----CA----A-------------A-....................A--CCA......................................................-----------GG-A--A-------G-- 8123
C.IN.09.T125_2139            A--------------T-C--CA-AG-----------G-------A--C--G---------....................A-----......................................................-----------AG-A--A-------G-- 8953
C.KE.00.KER2010              -----------------C--CA-AG-----------G----------------------.....................------......................................................--------G--GG-A--A----------- 8107
C.MW.09.703010256_CH256.w96  -----------------C---A-AG------------A------------A--G-----.....................A----T......................................................-----------GG-A--A----------- 8879
C.TZ.08.707010457_CH457.w8   ---------------T-----A--G-G---------GAG-----A---------A--GA--...................A---CA......................................................-----------GG-A--A----------- 8869
C.YE.02.02YE511              ---------------TCC---A-AG----G------GTG--------C-----A--GA--....................A--A-T...CCAGCAGCAGCAGGAGTAGGA.........GCA......GCGTCTCAAGACT---------CAG---A-------G-- 8146
C.ZA.08.705010534_CH534.w12  -----A-----------C---A-AG-----------GAG--C--A-----------A------.................A--AG-...CCAGCAGCAGAGGGAGAG...................................AAAA----------GG-A--A---A----G-- 8908
C.ZA.10.DEMC10ZA001          AAT--------------C--CA-AG-----------G--------A--------------....................A--AG-...CCAACAGCAAAAAGAGAAAGA.....................CGAGCTAAG----------TAG-A--A-------G-- 8280
C.ZM.11.DEMC11ZM006          AA---A--------G............---------G--A----A--G------A-----G..................A-----......................................................-----------GG-A--A-------G-- 8276
D.CM.10.DEMD10CM009          -----A---------------A-AG-A-----------G--A-----------------A--..................A--C--...CCAACT..............................................GAG------------G----A-------T--- 8244
D.CY.06.CY163                -----A-------------CA-AG-----------G--A-------------A-A-A--.....................A----T......................................................------------G----A-------T--- 8110
D.KE.11.DEMD11KE003          ---------------------A-AG-----------G--A-------------A---A--....................A----A......................................................-----------CAG-------------TG-- 8227
D.KR.04.04KBH8               -----A---------..............------G--A-----------------A--.....................A-----......................................................------------G----A----------- 8795
D.SN.90.SE365                -----A---------------A-AG-A-A-------G--A-------------A---A--....................A-----......................................................------------G-------------T--- 8895
D.TZ.01.A280                 -----A---------------A-AG-----------G--A-------------A---AA-....................-----T......................................................-----------CG-------------T--- 8085
D.UG.10.DEMD10UG004          -----A---------------A-AG-----------G--A-------------A---A--....................A----T......................................................-----------AG---------------- 8207
D.UG.11.DEMD11UG003          -------------------CA-AG----C------G--A-------G-----A---A--.....................A---C-......................................................-----------AG-------G--C----- 8251
D.YE.02.02YE516              -----A---------------A-AG--------------A---------A---A---AA-....................A----T......................................................------------G----A-------T--- 8077
D.ZA.90.R1                   -----A---------------A-AG-----------G--A-------------A---A--....................A----T......................................................------------G-------------T--- 8251
F1.AO.06.AO_06_ANG32         -----------------C---A-AG-----------G--A-C-----------------G....................A-CCCT......................................................G----------AG----------C-TG-- 8070
F1.AR.02.ARE933              ---------------------A-AG----------GAAA-------------------G.....................A-CCCT......................................................-----------AG-------------TG-- 8167
F1.BR.10.10BR_RJ015          --------------G-A--AGA-ACCA----------G--A--------------------...................--CCCT...CCAGCAGAAAGGATG..............................GGAG-----------GG---------------- 8303
F1.CY.08.CY222               ---------------T-C---A-AG-C----------G---------------------G....................A-CCCT......................................................-----------AG-------------TG-- 8063
F1.ES.02.ES_X845_4           --------------G------A-AG-A----------G---------C-----------G....................A-CCCT......................................................G----------AG-------------TG-- 8316
F1.RO.96.BCI_R07             ---------------------C-AGC---G------G-------------G--A-----G....................A-CCCT...CCAGCAGCAGAA...............................AGGGGG---------AG-------------TG-- 8966
F1.RU.08.D88_845             -C-------------------A-AG-G----------G--A-------T------------...................--CCCT...CCAGCAGCCCCTCCA.....................GCAGCAGAAGGGRG---------AG----A-------TG-- 8415
F2.CM.02.02CM_0016BBY        ---------------------A-AG------------AG-------------------A-....................A-CCCT....................................................GT---------AG--------AAG-TG-- 8056
F2.CM.10.DEMF210CM001        -------------------CA-AG----G-----------A-----------C-----A--...................A-CCCA...GCAGCAGCAGCAGAA..............................GGGGG----------AG----------T-TG-- 8175
F2.CM.10.DEMF210CM007        -------------------CA-AG-----------------A-------G-----A-A--A-..................A-CCCT....................................................GT---------AG-A-----G----TG-- 8218
G.BE.96.DRCBL                AA---------------G-AA--AGC---------CGAG------------C-A----A-...................CACCCT......................................................G----------AG-A--A-------T--- 8857
G.CM.10.DEMG10CM008          -------------------CA-AG-----------GAG-----------------A------..................--CCCT...CCAGGA..................................CAAACCCCT-A---------CAG-A--A-------T--- 8275
G.CN.08.GX_2084_08           --------------G---T-CA-AG----G-----CGAGA-------------A----A-....................A-CCCT......................................................-----------AG-A--A----------- 8150
G.CU.99.Cu74                 -----A---------------A-AG-----------G-G---------------A--G-A-...................A-CCCT......................................................----------CAG-A--A-------T--- 8537
G.ES.09.X2634_2              ---------------C-CTCCA-AG----G------GAGA-----------C-A------....................A-CCCT...TCA.................................................G----------AG-A--A-------T--- 8398
G.GH.03.03GH175G             -----------------A--CA-A-AG-A--------AG--G-A----CAG--A----A-....................--CCCT......................................................G----------AG-A--A----------- 8964
G.KE.09.DEMG09KE001          AA---A-----------C-----AG-----------CAG---------------A-----G...................A-CCCT...CCAGCAGCAGCAGGAGAAGGA.....................GCAGCATCTCCT----------CAG-A--A----------- 8303
G.NG.09.09NG_SC62            AA---------------..........-----CG-G--------G-----A--G---.......................A-CCCA......................................................G----------AG-A--A----------- 8085
G.ZA.01.TV546                ................................................................................................................................................................... 8003
H.BE.93.VI991                -----A---------G--T--A-A-G-----------G---------------A--G-A-....................A-----......................................................-----------AG----A-------TG-- 8297
H.CF.90.056                  -----A---------------GA--GG---G---T-----A---------------G---....................-----A......................................................-----T------AG----A-------TG-- 8217
H.GB.00.00GBAC4001           ----GA---------------A-AG-G----------GRA--------------A--G--C...................------......................................................-----------RG----A---C---T--- 8406
J.CM.04.04CMU11421           -------------------CA-AG-----------CAG--------G--------G---.....................--A-CT...........................................GCTCCTGTT------------G-A--A-------TG-- 8432
J.SE.93.SE9280_7887          AA---A---------...........------CAG-----------------G---........................---CCC...................................................GCT------------G-A-----------TG-- 8195
J.SE.94.SE9173_7022          AA---------------...........------CAG--------C---------G---.....................----CT...CCT...............................................GCT------------G-A-----------TG-- 8205
K.CD.97.97ZR_EQTB11          ---------------------A-AG--------T----------------------AA-.....................A-ACCT......................................................------------G-------------T--- 8080
K.CM.96.96CM_MP535           ---------------------A-AG-----------G--A--------------------....................--ACGA...CCAGCAGCAGACAGGGTG........................GGAACACAA-----------CG-------------T--- 8084
U.CA.01.TV749                -----------------A--CA-AG-G---------G--R-C--R-----------A-G.....................---YCT...CCARCTCCA...........................................G---------YAG-A------------TG-- 8967
U.CA.99.TV721                -C------------G............---------G--A-C----------------A-G...................A--CCT...CCAACTCCA.........................................ACT-----------AG-A-----------TG-- 8966
U.CD.83.83CD003_Z3           AA---A---------..........CA--------G--A-------G----------.......................---AG-...CCAGCAGCT.........................................GAG------------G-A--A----------- 8361
U.CD.90.90CD121E12           AA---A---------..........CA--------G--A------------------A-.....................---AG-...CCAGCAGCT.........................................GAG------------G-A--A-------T--- 8056
U.CY.05.CY090                A-------T------------A-AG-----------G--A-------------A------....................A-CCCT...GCAGCAAGGGAAAGAATAAGA.....CAAACCCCTGCAGCAGCAGAAGGGG----------AG----------G-TG-- 8116
U.CY.08.CY223                AA---------------C...........------CAG--C-------------------....................A-CC--...CCACAAACC.........................................CAG------------G-A--A-------TG-- 8194
U.ES.10.DEURF10DZ001         -----------------C--CA-AG-----------GAGA----------G--A------....................A--CCA...ACAGCAGCAGAA...................................GGAG----------AG-A-CA-------TG-- 8183
U.GR.99.99GR303              ---------------------A-AG-G---------GAAA-T-----------A------....................A--CCT...CCAAAGGATAGA..............................CAACCTCTC-----------AG-A--A-------T--- 8208
U.NL.95.U_NL_95_H10986_D1    ---------------------A-AG-----------G-------A-----G--A--G---....................A--CCA...GCAGCGGAGGGAAGG..................................CGAGCT------------G-A--A--G----TG-- 8824
01_AE.AF.07.569M             A-T-----------------CA-AG-G---------CAG--C-------A---A-AGAA-....................A--CCT...CCAGCAGCAGAAGGAGTAGGAGCAGTATCTCAAGATCTAGATAAACAT...GG---------AG-A--A-------T--- 8157
01_AE.CF.90.90CF11697        ---------------A---AA-AG-G---------CAG--C-----------A--G---.....................A--CCT......................................................G---------AG-A--A----------- 8870
01_AE.CN.10.YNFL03           AA-------------............---------CGA--C--A-----G--ACAG-A-....................A--CCC......................................................-----------AG-A--A-------T--- 8278
01_AE.HK.04.HK001            -----A-------------CA-AG-G---------CAG--C--A---------A--G-A-....................A--CCC......................................................------A----AG-A--A-------T--- 8279
01_AE.IR.10.10IR.THR48F      A--------------------A-AG-G---------CAG--T-----------AGAG-AG....................A--CCT......................................................-----------AG-A--A-------T--- 8123
01_AE.JP.x.JRC77AE           -----------------C--CA-AG-G---------CAGA-C-----------A--G---....................A--CCT......................................................-----------AG-A--A-------T--- 8897
01_AE.TH.09.AA090a_WG11      A-----------------..........-----CAG--C-----------A--GAA-.......................A-CCCT...CCAGCAGCA.............................GAAGAAGTA...------A-AG-A--A-------T--- 8328
01_AE.TH.90.CM240            A-T-----------------CA-AG-G---------CAG--C-------A---A-AG-A-....................A--CCT......................................................------A---AG-A--A-------T--- 8471
01_AE.US.05.306163_FL        -----------------C...........------CAG--C--------G--A--GA--.....................A--CCT......................................................------A---AG-A--A-------T--- 8019
01_AE.VN.98.98VNND15         -----A-------------CA-AG-G---------CAGA-C-----------A--G---.....................A--CCT......................................................------A---AG-A--A-------T--- 8188
```

```
B.FR.83.HXB2                 TCGAGACCTGGAAAAACATGGAGCAATCACAAGTAGCAATACAGCAGCTACCAATGCTGCTTGTGCCTGGCTAGAAGCA.........CAAGAGGAG......GAG......GAGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAG 9046
Nef                          __R__D__L__E__K__H__G__A__I__T__S__S__N__T__A__A__T__N__A__A__C__A__W__L__E__A__.__.__.__Q__E__E__.__.__E__.__.__E__V__G__F__P__V__T__P__Q__V__P__L__R__P__M__T__Y__K__A__
A1.AU.03.PS1044_Day0         -AA---T--A--T-----------G-------C------GT-......-ATC-CC--AG---CA-------G------..........--------A......-G-..........----A--C---------GG--C----G---C--------------T-----G-- 8299
A1.CH.03.HIV_CH_BID_V3538    -A----T--A---G-----------T-----C--T-------------T------AG-----------G-----G..........-------AGAG...-G-..........----A--C--------A-GG--A-------------G----------TT----G-- 8261
A1.CY.08.CY236               --AG--TT-A--T---T---------------........................................................................................................................................ 8168
A1.ES.05.X1608_8             --A---T--A--T----G-------G-------C-------T-......-ATC-CC-YAG---C--------GA-T--G..........-------M..............--T--A--C---------GG--A--------C-----------------------G-- 8554
A1.KE.11.DEMA11KE001         --A---TT-A-CT-----------G-T--------T--C-T-......-A-C--C--AG------------G------..........-------C..............-----A--C---------GG--A--A----CC-C-----------A---------- 8461
A1.RU.11.11RU6950            --A---TT----T-GG---------G-------C--T-----------C--T-----------C--------G--G---..........-------A..............-----A--C---------GG--A-----------G----------A--T----G-- 8603
A1.RW.11.DEMA111RW002        -------T-A-------------------C--C--T--CGT-......-ATC-CC--AG---C--------G------........-G----C..............-----A--C---------G---A--A-----C-----------------A----- 8460
A1.SN.01.DDI579              --A---TT-AAGT-----------G-A-------CT---CA-----......-G----AG---CA-------G-----G..........-------A...............--A--A--C---------G---A--A----------G--C--------T--A-G-- 8238
A1.UG.11.DEMA110UG001        --A---TT-A--T-------------------C--T--CCT-......-ATC-CC--AG-AACA-------G-----T..........-------A.......-A..........----A--C---------GG--A--A-----AC----------------------- 8396
A1.ZA.04.04ZASK162B1         ------T--A---G------------------C--T-------A----AT---AAAG-CT-------G-GAG---...................GAG..............-----A--C---------GG--A--A--G---C--------------T--A-G-- 8510
A2.CM.01.01CM_1445MV         --A---TT-A-GT-C--G-------G------TCC-T-------------T---T---AC--C--------G-----G..........-----A--AGT....-T..........----A--C--C------GG--A---------C-------------C-------G-- 8263
A2.CY.94.94CY017_41          --A---TT-A-CT-CT---------G-------C--T-------------T---C---A---C------G-G-----G..........-----A--GAG...AGT........-A--A--C--C------GG--A-------------G----------C-T-----G-- 8472
B.BR.10.10BR_MG029           ---G----------C--G------T---------------------AG------A-------------------.................A--.................-A---------------G------------------------------------ 8543
B.CA.07.502_1191_03          --------------G-------------------------------------A----C-A-A------------------..................AAAT..............-A--A----------------G--------------------------C--------- 8493
B.CH.08.M2_0803101_NFLG8     ---------A--G-G-------------------------------------T-------A-------------------..........---------.............-A--A--C---------G-----A-------------------C-------G-- 8611
B.CN.10.DEMB10CN002          -A-G---T-----G---G-------C------C------------T--C-AT------------------------------..........-------T..............--C----------------GG--------------------------------A-G-- 8431
B.ES.10.DEMB10ES002          -------T----C-G-----------C-------------------------A--------A--A------------------........-G----TGAT...-G-.........----A--C---------A-----------C-------------A-------G-- 8446
B.FR.11.DEMB11FR001          -A-G-----A----G------C----------C--------------A--A---------A-A-----------------..........-------T..............--A----------------G-----------------------------------G-- 8294
B.GB.05.MM45d213_GN1         C-A--------G--G----------G----------------C-------C-A---------C--------G--------..........---ACA--AGGG...-A........-A----G---------G---K-------G------------C---------- 8473
B.HK.06.HK003                -----------C-C---G-------------------CCAG......-AT-------A-C---T----------------..........---------AGAGGAG---........-----C---------G---C--------------------------------AT- 8418
B.HT.05.05HT_129389          ---G--T-----------------C-------C----------T---AA-AT-------A-------------------..........---A-----GATGAT--A.........-A--A--------------G---------------R----------------------- 8288
B.JP.12.DEMB12JP001          -------------------C---------------------------T-......-AT-------A---C-T-------------------..........-------A........-A..........-----A-----------------G----------------------G-T-----A-- 8465
B.KR.07.HP_18_07JHS10_3909   -------------C---G-------------C---T----------AG--A--------------------------..........-------A........................-----------------G-----------C------------------------G-- 8712
B.PE.07.502_0525_wg5         -------------C-------------C-----GC-------------------AT------------------------..........------------------.......-A..........-----A--G---------A-----------C---------------T---A-G-- 8480
B.RU.11.11RU21n              --------------------------------C--T----------------------------------T-------..........-------............................................-T-G-------------------------------------G-- 8574
B.TH.08.MERLBDTRC10          --------------G-----------------A--A-------C------AT-------A------------G-------..........-----A--GAG...---............................-A--C-----------G--------------------G-- 8342
B.US.11.ES38                 --------------G---------------------------------T-A---A------------------G---C---..........-----A--GATGAT---........-A-------C-------G-----------C--------------T---A---- 9013
C.AR.01.ARG4006              --A----T-A--T-G-T----G---T-G-----C--T--C---C-T--C-AT-------AC-----------G-----G..........---C----GAG...-A..........----A--C---------G--------------------------TT----G-- 8264
C.BR.07.DEMC07BR003          --A----T-A--T-G-T----G--GC-G-----C-----C---C---G--AT-------------------G-----G..........---C---AGAA...-A..........----A--C---------G-M--------------G----------T------- 8492
C.BW.00.00BW5031_1           --A----T-A--T--------G---C-T----AC-----C---------C--T---AAA-AC-----------G-----G..........-------GAA...-GA..........----A--C---------G-----------------------------------G-- 8402
C.CN.10.YNFL19               -A-------------G--------------------T-----T-......-AT---A-------A-------C----G..........-------GAA...-A........-A--A--C---------G---------G---G-G--------------T-----G-- 8422
C.CY.09.CY260                -------T-A---G-----G---C-G-----............................................................................................................................... 8216
C.ES.08.X2363_2              --A------A--T---T--------C-T-----------C---C-CAGC-AT-------A---------T---C---..........-------GAA...-CGAG...-A--A--C-----T---G---------G--C-----------------T--A-G-- 8557
C.ET.02.02ET_288             C------T-A---C-------G---C-A---CC-----C---C---AG-AT----------------------GC----G..........-------GAA...-A..........----A--C---------G-----------------------------T------- 8275
C.IN.09.T125_2139            --A----T-A--T---T-----------T-----C-----C----GCA-C--T-------A-GT---T-----G---A-G.........-G-----GCA...-GA........-A--A--C---------G---------G-------------------T-----G-- 9105
C.KE.00.KER2010              --A----T-A--T---T-------C-T--C--C-----C-----T-GC-A---------AC-----------GC-----..........---------GAA...-A........-A--A-----------G---------G-----------------CGT----G-- 8259
C.MW.09.703010256_CH256.w96  -------T-A--G-G------G---C-T--T-CC-----C----TCAGC-A---------A----------G------..........-----A--GAA...A-A........-A--A--C---------G-----A--G-------------------T--A-G-- 9031
C.TZ.08.707010457_CH457.w8   ---------A--T-GG-----------T-----C-----C----CA-C--T------------------G------..........---------GCA...-A........-A--A--C--------T-G---------G-------------------T-----G-- 9021
C.YE.02.02YE511              --A----T-A-CT--------G---C-T-----C-----C--C-CA-C-AT----A---------------G-----G..........----C--GAC...-A..........----A--C---------G-----------------------------------G-- 8298
C.ZA.08.705010534_CH534.w12  --A------A--T--------G---C-T----CC-----C----CAGA-A-----A-------------------G.........-G--A--AGAC...-G-........-T--A--C---------G---------G--------------C-TT------- 9060
C.ZA.10.DEMC10ZA001          --A----T-A--T---------------T-----------CA-.......-----------------GC------..........-G-----AGAT...-C..........----A--C--C------G---------G--C-----------------T-----G-- 8426
C.ZM.11.DEMC11ZM006          --A----T-A--T---TT-------C-T----CC-----------CAAA-AT-------A-GT---------GC----G..........--------GAAGGA--CGGA...-A--A--C---------G---A--A--G-----------------T-----G-- 8434
D.CM.10.DEMD10CM009          ------T-------------G----------------------CAG--T-------AC-------------------..........-----A--AGAT...-G-.....-------------------G------------------------------------ 8396
D.CY.06.CY163                ------T-------------G-----------.............................................................................................................................. 8143
D.KE.11.DEMD11KE003          ---G---------CGC----G----------------------T-CCAG--T-------AC-------------------..........-----A--A.......-T.......-A-------C------G----------G------------------------- 8376
D.KR.04.04KBH8               ---G----------G-----G------------A------------------T------A-------------------..........-----A--AGAG......-....-A-----------------GG--------G---AC---------------------- 8947
D.SN.90.SE365                A-------------------G--G----------------AA--A---T-----G-AC-------------------..........-----A--GAT...-GA.....A-A-----------------GG------------------------------------ 9047
D.TZ.01.A280                 --------------G-----G-----------T-------CAC--T-------AC----------------G..........-----A--GCA...-C.......-A-----------------G---------CG--A---------------T-------- 8237
D.UG.10.DEMD10UG004          -AA-----------------G----------------------CAG--T---C----C----------G-------..........-----A--A..............-GA-------------T-G---------G--A-----------------------A-GG- 8356
D.UG.11.DEMD11UG003          -A-G--T-------------G-----------T---------CAG-----T-A-----------------------..........-----A--AGAA...-G-.......-A--T---------A---G-----------GC------------------G-- 8403
D.YE.02.02YE516              ------T-------------G-----------------T......-AT---C---C-------------------..........-----A--C...............-A-----------------G------------------------------------- 8220
D.ZA.90.R1                   --------------------G---------------C-----CAG---T-------AC-------------------..........-----A--AGT............-A---------------T-A--------------------------T---A-A-- 8403
F1.AO.06.AO_06_ANG32         --A---TT-A--T-G--GG--G-----T----TC-A------TAG------T---T--AC-TG--------G------.........-G----C.................-A--A--C---------G------------C-----------------TT----G-- 8219
F1.AR.02.ARE933              --A----T-A--C-G--GG--------T-----C--------TAA-AA--AT---C--AC-TG--------G------..........-------A................-A--A--C---------G---------------------------TT----G-- 8316
F1.BR.10.10BR_RJ015          -------T-A-C--GCAGA--------T-----C--------CAG--A---T---C--ACGTG--------G------..........-------C................-A--A--C---------G-------------------T--------T------- 8452
F1.CY.08.CY222               --A----T-A--T-G--GG--G---G-T-----............................................................................................................................. 8096
F1.ES.02.ES_X845_4           --A----T-A--T-G--GG--------T-----C--------T-------T---T--AC-TG--------G------..........-----------.....-C.......-A--A--C---------G-------------------C--------T------- 8465
F1.RO.96.BCI_R07             -------T-A--T-G-AGG--G-----T-----C--------T-G------T---C--AC-TG------G-G------..........---------GAT...-G-.......-A--A--C---------G----------GCCG------------T-----G-- 9118
F1.RU.08.D88_845             -------T-A--TC--AGG--G-----T--------------TAG------T---C--AC-TG--------G--G---..........-----R--A...............-A--A--C--------T-G---------G--------------------T--A---- 8564
F2.CM.02.02CM_0016BBY        --A----T-A--T--------G-----T-----C--------T-GG-----T---AA--ACATA-------G-----G..........-----A--T...............-A--A--------------G---------------------------T-----G-- 8205
F2.CM.10.DEMF210CM001        ---------A--T---T----G-----T-----C--------TAGG-----T-----A-AC-TG-----------G-A-.........------------------.....-A.......-A--A--------------G---C--------C--------------TT----G-- 8324
F2.CM.10.DEMF210CM007        --A----T-A--T---T----G-----T-----C--------TAAG-AA--T---C---AC-TG------T---------..........-----T--A...............-A--A--------------G---C---------------------C-TT----G-- 8367
G.BE.96.DRCBL                --A---TT-A--T-GG--------C-----------A---------GC-----C---A------------G------..........------------------.......-CTCA.......-A--C---------G---A---------------T--------T-----T- 9009
G.CM.10.DEMG10CM008          --A---TT-A--T-GG-----T-----T-----C-------------C-A---C---A------------G------..........-----A--.......-TTCA...-T--A--C--------T-GG--A----------G-------------T--A-GT- 8427
G.CN.08.GX_2084_08           --A---TT-A--T-GG----------------C-----C---------A--T---C---A------------G------..........---------........-TCA.......-A--C---------G---A----G-----G----------C-TT----GT- 8302
G.CU.99.Cu74                 --A---TT-A-CT-GG-----------T-----C--T-------------C-AT--CC---A------------GC-----..........-------A.......-TTCA...-A--A--C---------G---A----------G----------C--T-----T- 8689
G.ES.09.X2634_2              --A---TT-A-CT-G------------T-----C-------------C-AT--CC---A-C----------G--G---..........-----A--A.......-CTCA.......-A--C---------G---A--------C-G---------------T----G-- 8550
G.GH.03.03GH175G             C-A---TT-A--T--G--------GC-------C-------------A-AT---C---AC-----T----------------..........-------A.......-TTCA.......-A--C---------A---A--A-------G-----T--------T-----T- 9116
G.KE.09.DEMG09KE001          --A---TT---CT--G-----G--------G--C-------------A-C-AT--CC---A--------------------.........-C----C.......-ACCA.......-A--C---------G---A---------------------C--T-----T- 8455
G.NG.09.09NG_SC62            --A---TT-AACT-G--------------C--C---------T---A--GG---C---A-GT-------------------..........-C------.......-TTCA...-A--A--C---------G---A---------------------------T------- 8237
G.ZA.01.TV546                ............................................................................................................................................................. 8003
H.BE.93.VI991                --A---TT-A--T-G--G-------G------TC-AT--C-T----T---AT-------AC-CC--------G------..........-------AGAA...---.......-A--A--C---------GG--G-----------------------C--T----G-- 8449
H.CF.90.056                  ------TT----T-G--GC--G---G------T--AT---------T----T---CG--A-GCC--------G------..........-------CGGG...---.......-A--A--C---------GG------------------------C--T----G-- 8369
H.GB.00.00GBAC4001           ------TT----T-R--R-------G------TC-AT----------A--GG---C---AC-CC----------------..........-------AGAA...---.......-A--A--C---------G-------------------------C--T----G-- 8558
J.CM.04.04CMU11421           --A---TT-A-CT-CG-----G-----------C------------A---A-------AC-------------------.........---ACA--.................--A--A-----------------A-------------------------G-- 8581
J.SE.93.SE9280_7887          --A---TT---CT--G-----G-----------C----------------A-------AC----------G--------..........---ACA--.................--A--A----------A------A---A-------------A-------G-- 8344
J.SE.94.SE9173_7022          --A---TT---CT--G-----G-----------C----------------A---A--AC----------G--------..........---ACA--.................--T--A----------G---------W--A-------------A-----A-G-- 8354
K.CD.97.97ZR_EQTB11          --A---------T--------G--G-------C--------T---TTC-AT---C---A-----------G--------..........-G--A--C..............--C------------------G-----------A----------------TT----G-- 8229
K.CM.96.96CM_MP535           --A--------CT-G------G--G-------C--------TT--CAC-AT---C---A-----------G--------..........-G-----A...............-A--A--------------G-----------A-----------------T------- 8233
U.CA.01.TV749                -AA---TT-A--G--G-----------A----C---T...------T-C-AT---C--ACCTG------G--------..........-------AGAG...-G-.........----A--C---------G---A----R----G--------------T------- 9116
U.CA.99.TV721                -AA-AGTT-AC-G--G---------G------C---T...G-------C-AT---C--A-CTG------G------G..........-------AGGA.................-A--C---------G---A----G---AC-----------------T------- 9115
U.CD.83.83CD003_Z3           -------T-A--G---T----G-----------C--------------AA--T---AG--C------CA---------.........GCGC----GAC...-G-.......-A--A--C---------G-----------------------------C--T------- 8513
U.CD.90.90CD121E12           -------T----T-G------G-----------C-----------------T---AG---C------CA---------..........--G--A--AGAA...-GA.......-T--A--C---------A-----------------------------C--T--A---- 8208
U.CY.05.CY090                --A---TT-A-CT--------G---G------T..................................................................................................................... 8149
U.CY.08.CY223                --AG--TT----T-GG-----------------......................................................................................................................... 8227
U.ES.10.DEURF10DZ001         --A---TT-A-CT-GG-----G---G-------C------------A---A---C--AC-TG-------G-----G..........-----A--...................----A--C---------G---------G------------------C--T--A--T- 8332
U.GR.99.99GR303              --A----T-A--T-G----------G----------------AGG---AT-----AA-CTC-------G-----G..........------------------...........-A--C-----------G---G--T----C----------------T----A-- 8357
U.NL.95.U_NL_95_H10986_D1    ------------T-GG-----G-----------C-----------------G-----A-AC-------------------.........---ACA--.........-A.......-A--A--G---------G-----------A-----------------T------- 8973
01_AE.AF.07.569M             ------T--A--T-----------G-A--------T----TG......-AT-------A-AA--T-----GAG---..........-----A--AGAAGGG--CGACGAG-G---A--C---------GG--G--------C-----------------TT----G-- 8312
01_AE.CF.90.90CF11697        --A---T--A--T-------------A--------T----TG......-AT-------A------------G-----G..........-------A.................-A--C---------GG--A----G---C-----------------T-----G-- 9013
01_AE.CN.10.YNFL03           --A---T--A--T-----------G-G--------T----TG......-AT-------A-A--A-------GAGG---..........-------AGAT...........-A--A--C---------AG--A--------C-----------------T-----G-- 8424
01_AE.HK.04.HK001            --A---T--A--T-----------G-A-----C-AT----TG......-AT---A--A-A--T-----GAG---..........-------AGATGAG--A.....-G---A--C---------GG--A--------C-----------------T----AC- 8428
01_AE.IR.10.10IR.THR48F      --A---T--A--T-----------G-A--------T--C-TG......-AT-------A-----T-----G-----..........------ARC.....-A..........----A--C---------GG--G--------C-----------------T--A-G-- 8266
01_AE.JP.x.JRC77AE           --A---T--A--T-----------G-A---------------TG......-AT-------AAA---TT----GAG---..........---------GACGAG--AGAG...-GA--A--C---------G---G--------C----------------TT----A-- 9049
01_AE.TH.09.AA090a_WG11      --A---T----CT-----------G-A--------T----TG......-AT---C--A-A--T-----GAG---.........--T-----AGATGAA---..........----A--C--C------GG--G--------C-----------------T----A-- 8477
01_AE.TH.90.CM240            --A---T--A--T-------------A--------T----T-......GAT-------A-----T-----GAG---..........-------C..................----A--C---------TG--G--------C-----------------T----G-- 8614
01_AE.US.05.306163_FL        --A---T--A--T-------------A--------T---.TG......-AT-------A---A-------GAG---..........-----A--A.......-T..........----A--C---------AG--A--------C-----------------TT----G-- 8161
01_AE.VN.98.98VNND15         --A---T--A--T-----------G-A--------T----TG......-AT-------A-A---T-----GAG---.........--T-----AGAGGAT---.......-G---A--C---------GG--A--------C-----------------TT----G-- 8337
```

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546

```
                                                                      3' LTR U3 start                                           Nef premature end in HXB2
B.FR.83.HXB2                CTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAA.AGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGG 9215
Nef                         A__V__D__L__S__H__F__L__K__E__K__G__G__L__E__G__L__I__H__S__Q_#R__R__Q__D__I__L__D__L__W__I__Y__H__T__Q__G__Y__F__P__D__*__Q__N__Y__T__P__G__P__G__V__R__Y__P__L__T__F__G
A1.AU.03.PS1044_Day0        --T-T-----C---TT---------------------------T---T-----T-----ARG.-AG--------A---------------G-A---A------------------------G------T------------------A----G--C-----A--A----- 8468
A1.CH.03.HIV_CH_BID_V3538   -A--------G-----------------------------------T---T-----T-----AGG.-A---------A---------------G-----A----------------------C-G------T------------------AC-----TC-----A--A----- 8430
A1.CY.08.CY236              .......................................................................................................................................................................... 8168
A1.ES.05.X1608_8            -AT-T-----C----------------------------------T-----T-----AGG.-AG--------G---------------G-T---A------------------------G------T-----------------------G--C-----A--A----- 8723
A1.KE.11.DEMA11KE001        --A-------C--------------------------------T---T-----T-----AGG.-A---------AC--------T-A---G------------------------------G---T-----------------------------C------TG------ 8630
A1.RU.11.11RU6950           --T-C-----C--------------------------------T---------T-----A-G.-----------C---------------G----T-------------------------G------T--T--C------------A-T-----C-----A--A----- 8772
A1.RW.11.DEMA111RW002       --T-T-----C---TT-------------------------------T-----T-----A-G.-----------A---------------G------------------T-----------G------T--------------------------C--GT-A--A----- 8629
A1.SN.01.DDI579             --T-T-----C--------------------------------T---T-----T-------G.CA---G-----C---------------G-------------G----------------G------T--------------G---ACT-----C-----A--A----- 8407
A1.UG.11.DEMA110UG001       ----G-----C------------------------------------T-----T-----AGG.-A---G-----A---------T-----G------------------T-----------G------T------------------A------TC-----A--A----- 8565
A1.ZA.04.04ZASK162B1        --T-T-----C--------------------------------T---T----CTGG--TAG-.-A---------A---------------G-----A----------T-------------G------T--------------------------C-------------- 8679
A2.CM.01.01CM_1445MV        --T-T-----C--------------------------------T---T-----T-------G.-A---------A---------T-----G----------G-------------------G------------------------------G-TC-----A--A----- 8432
A2.CY.94.94CY017_41         -GT-T-----C---TT---------------------------T---T-----T-------G.-A---------C---------A-----G------------------------------G------T------------------A-------C---T-A--A----- 8641
B.BR.10.10BR_MG029          ----G------------------------------------------------T--------.-----G-----C--T------------G------------------------------G------------------------AA-------C-------------- 8712
B.CA.07.502_1191_03         ----------G-----------------C------------------T---C-TGG--T---.-AG------------------------G-T---A-T----------------------G-------------------------AC--------------------- 8662
B.CH.08.M2_0803101_NFLG8    --T----------------------------------------------G-----------G.-A--------------------------------------------------------G------T--------------------------C------TG------ 8780
B.CN.10.DEMB10CN002         --T-G-----C---------------------------------------------------.-AG--------G---------------G----------------A-------------G---C---------------------ACT----TC------TG------ 8600
B.ES.10.DEMB10ES002         --C-C------------------------------------------T-----T-------G.CA-------------------------G-G--TA-T--------------------C-G---A----------A------------------C-----CTG------ 8615
B.FR.11.DEMB11FR001         --C---------------------G----------------------------T-------G.-A-------------------------G------------------------------G-------------------------A---------------------- 8463
B.GB.05.MM45d213_GN1        --T-T------------------------------------------------TGG------.CA-------------------------G-----A------------T-----------G------------------A------AC---G-TC-----A-------- 8642
B.HK.06.HK003               --T-C---------------------------------------------G--T-------G.-A----A----C---------------G----T-----G-------T-----------G---TG--------------------A-------C-----------C-- 8587
B.HT.05.05HT_129389         -----------------------------------------------T--G--T-------G.-A------------T------------G-----A--------------........................................................... 8398
B.JP.12.DEMB12JP001         --T-T------------------------------------------T-----T--------.-A----A-----------A-C------G-----------------AT---------C-G---AG-A------------------AC-----TC-------------- 8634
B.KR.07.HP_18_07JHS10_3909  ----C---------------------------------------------G--T--------.-A-------------------------G------------------------------G-------------------------AC-----T--------------- 8881
B.PE.07.502_0525_wg5        ----G---------------------------------------------G--TGG------A-A-------------------------G-T----------------------------G-------------------------AC-----TC------TG------ 8650
B.RU.11.11RU21n             -----------------------------------------------------T-------G.-A--------------------------------------------------------G------T------------------ACT-------------------- 8743
B.TH.08.MERLBDTRC10         --C-G-----G-------------G----------------------T--G--T-------G.-A---------C--T-----------------T-------------T---------C-G-------------------------AC-----T--------------- 8511
B.US.11.ES38                --T----------------------T------------------------G--TGG--T--G.-A-------------------------G-----A------------------------G---A---------------------CC---------C----------- 9182
C.AR.01.ARG4006             -AT-C-----C---TT-------------------------------T-----T----T--G.CA---GA----C---------T-----G----TA------------------------G---A--------------A-----------------C----------- 8433
C.BR.07.DEMC07BR003         -AT-C-----C---TT-------------------------------T-----T----TA-G.-A---G-----C---A-----T-----G----TA------------------------G---A---------------------A----------C----------- 8661
C.BW.00.00BW5031_1          -AT-T--C--CG--TGG--------------------------T---T-----T----TA-G.-A---------G---------T-----G----T-------------------------G---A-----------G--A----------------------------- 8571
C.CN.10.YNFL19              -AT-T-----C---TT-------------------------------T-----T----TA-G.-A---G-----G---------T-----G----TA------------T-----------G-AGC-----------G--A--------------C--------T----- 8591
C.CY.09.CY260               .......................................................................................................................................................................... 8216
C.ES.08.X2363_2             -AT-C-----CG--TT---------------------------T---T---------YTA-G.-A---------C---------T-----G----T-------------------------G---A--------------A------A---------------------- 8726
C.ET.02.02ET_288            -AT-C------G--TT-------------------------------T-----T----T--G.-A---GA----G--T------T-----G----TA------------T-----------G---A--------------A--------------C-------------- 8444
C.IN.09.T125_2139           -AT-T-----C---TT-------------------------------T-----T----TA-G.-A---G-----G---------T-----G----T-----G-------------------G---T-----------G--A-----A--------C--------T----- 9274
C.KE.00.KER2010             -AA-C-----C---TT---------------------------T---T-----T----TA-G.-A---G-----C--A------T-A---G----TA------------T-----------G---A-----------G--A----------------------------- 8428
C.MW.09.703010256_CH256.w96 -AT-C-----C---TT-------------------------------T-----T-----A-G.-A---G-----C---------T-----G-T--TA------------T---T-------G---A-----------G--A----------------------------- 9200
C.TZ.08.707010457_CH457.w8  -A--C-----C---TT-------------------------------T-----T-----A-G.-A---G-----C---------T-A---G----T-----------T-------------G---A--------------A--------------------------G-- 9190
C.YE.02.02YE511             -A--C-----C---TGG------------------------------T-----T-----A-G.-A---G-----G---------T-------G--TA----------T-------------G---A-----------G--A------AC--------------------- 8467
C.ZA.08.705010534_CH534.w12 -AT-T-----C---TT-------------------------------T-----T-----AGG.-A---G-----G---------T-----G--C-TA------------T-----------G---A--------------A----------------------------- 9229
C.ZA.10.DEMC10ZA001         -AT-C-----CG--TT---------------------------T---T-----TGG--T--G.-A---G-----C---------T-----G----T-------------------------G---A--------------A------------------T---------- 8595
C.ZM.11.DEMC11ZM006         -AT-C-----CG--TT---------------------------T---T-----T-----A-G.-A---G---A-A---------T-----G----T-------------T-----------G---A-----------G--A----------------------------- 8603
D.CM.10.DEMD10CM009         -------------------------------------------T------G--T--------.-AG--------A-----------T---G------------------------------G-------------------------AC------C-----A-------- 8565
D.CY.06.CY163               .......................................................................................................................................................................... 8143
D.KE.11.DEMD11KE003         ----------G------------------------------------T-GG--TGG---AG-.-A---------G-----------T---G------------------T---------C-G------------G--------G---A-T-----C-------------- 8545
D.KR.04.04KBH8              --C-------------T-----------T---------------------G--TGG-----G.-A------------------C--T---G-----A------------TT--------C-G---T---------------------A---------------------- 9116
D.SN.90.SE365               ----T---------------------------------------------G--TGG---A-G.-A---------G-----------T---G------------------T-----------G---T---------------------ACT--------C---TG------ 9216
D.TZ.01.A280                --T-G---T-G-----------------T------------------T--G--TGG--T--G.-----G-----C-----------T---G------------------T-----------G---C--T------------------ACT-------------------- 8406
D.UG.10.DEMD10UG004         --T-------G--------------------------------T---A-----TGG-----G.-A---------------------T---G-T--T-------------------------G---A---------------------ACT-----------------C-- 8525
D.UG.11.DEMD11UG003         --T-G---T-G--------------A---------------------T--G--TGG----CG.-A---------A---------------G----------------T-------------G---A---------------------A-T--G----------------- 8572
D.YE.02.02YE516             --------------------------------------------------G-----------.--G--------------------T---G------------------------------G-------------------------A-----C---------------- 8389
D.ZA.90.R1                  -----------------------------------------------------TGG---A--.-AG--------G-----------T---G-----A-----------AT------C----G-------------------------A---------------------- 8572
F1.AO.06.AO_06_ANG32        -GC-C-----C-----------------------...................................................................................................................................... 8257
F1.AR.02.ARE933             --C-.-----C--T---------................................................................................................................................................... 8338
F1.BR.10.10BR_RJ015         --C-------C--T----------------------------................................................................................................................................ 8494
F1.CY.08.CY222              .......................................................................................................................................................................... 8096
F1.ES.02.ES_X845_4          --C-------C--------------------------------T---------------AGG.-A---------C-----------A---G-T----------------T-----------G---A---------------------A-----------------C-G-- 8634
F1.RO.96.BCI_R07            --C-------C-------------G----------------------T-----T-------G.-A-----G---G---------------G-T---A------------------------G-------------------------ACT-----C--G----------A 9287
F1.RU.08.D88_845            --T-T---Y-C------------------------------------T-----T-----A-G.-A---G-----G---------T-----G-T--------------------------C-G------T-----------A------------CTC-------------- 8733
F2.CM.02.02CM_0016BBY       ----C-----C--T-----------------------------T---T----CT-----AGG.-A---------A-----------A---G----T-------------------------G-----------------------......................... 8349
F2.CM.10.DEMF210CM001       --T-C-----CG--TT-------------------------------T-----T-----A-G.-A---------A---------T-----G-T--------G-------------------G-------------------------A-------C--C----------- 8493
F2.CM.10.DEMF210CM007       --T-C-----C---TT-------------------------------T-----T-----A-G.-A---------A---------T-----G-------------------T----------G-------------------------AC-----T--------------- 8536
G.BE.96.DRCBL               -GT-T-----C---TT---------------------------T------G--T-----A-G.-A---------G--------C--A---G----T--T--------A-T------A----G-------------------------ACT---GTC-------------- 9178
G.CM.10.DEMG10CM008         --T-T-----C---TT---------------------------T---------T-----A-G.-AG--------A--------C------G----TA-T--------A-T------A-C--G---A--T------------------AC------C---T---------- 8596
G.CN.08.GX_2084_08          -CT-T-----C---TT---------------------------T---------TGG--TA-G.-A---------C--------C--A---G----TA-T--------A--------A-C--G------------T------------ACT----TC-------------- 8471
G.CU.99.Cu74                --T-T-----C---TT---------------------------T---------T-----A-G.-A------G--C--------C------G----TA-T--------A-T------A-C--G-------------------------ACA-------------------- 8858
G.ES.09.X2634_2             --T-T-----C---TT---------------------------T---------T-----A-G.CA------------------C------G----TA-T--------A--------A----G-------------------------ACT----T--------------- 8719
G.GH.03.03GH175G            --T-T-----C---TT---------------------------T---------TGG---A-G.-A---------G---------------G----TA----------A-T------A----G-------------------------ACT---CTC-------------- 9285
G.KE.09.DEMG09KE001         --T-T------G--TT---------------------------T---------T-------G.-A---G-----C---------------G----TA-T--------G-T------A-C--G-------------------------A-T-----C-------------- 8624
G.NG.09.09NG_SC62           -CT-T-----CG--TT---------------------------T-------C-T-----A-G.CA---G-----C-----------A---G----TA----------A--------A----G------------T------------A-T--G----------------- 8406
G.ZA.01.TV546               .......................................................................................................................................................................... 8003
H.BE.93.VI991               --T-T-----C-----T------------------------------T-----T-----A-G.-A---------G---------------G----TA----------T-----------C-G---T----------------------AG--------C----------- 8618
H.CF.90.056                 --T-T-----C-----T--------------------------T---T-----T-----A-G.CA------G--C---------T-A---G----TA----------------------C-G--------------------------AG----T---C----------- 8538
H.GB.00.00GBAC4001          --T-G-----C---------------------------------------G--T-----A-G.-A---------G---------------G----TA----------------------C-G---T--T--------------------G--------C---TG------ 8727
J.CM.04.04CMU11421          --T-T------G--TT---------------------------T---T-----T-----A-G.-A------R--A---------Y-----G----T-----------T-T-----------G---A---------------------RY---R--C---Y---------- 8750
J.SE.93.SE9280_7887         ----T---------TT---------------------------T---T-----T-----A-G.-A---------G---------------G-TC-TA----------T-----------C-G---A---------------------A-------C---T----A----- 8513
J.SE.94.SE9173_7022         --T-T---------TT---------------------------T---T-----T-----A-G.-A---------G---------------G-TC-TA----------T-----------C-G---A---------------------AC--B---C---T----A----- 8523
K.CD.97.97ZR_EQTB11         --T-T-----GG--TT---------------------------T---------T-----A-G.-----------G---------------G-T--T--T----------T-----------G---A---------------------A-------C-------------- 8398
K.CM.96.96CM_MP535          --T-T-----CG--TT---------------------------T---------T-----A-G.-A---------A---------------G-T--T-----------T-T-----------G---A---------------------A-------C-------------- 8402
U.CA.01.TV749               -C--------C--------------------------------W---T-----T-------G.MA---------C---------------G-----A----G-------T-----------G------------------------AA-------C-----A--A----- 9285
U.CA.99.TV721               -CT-C-----CG--TT---------------------------T---T-----T-------G.CA---G--G--C---------T-----G-----A----G-------T-W---------G------------------------AA-------C-----A--A----- 9284
U.CD.83.83CD003_Z3          --T-T-----CG--TT---------------------------T---------T-----A-G.-A---------G---------------G----TA----------T--------A----G-------------------------A---------------------- 8682
U.CD.90.90CD121E12          ----T-----C---TT-------------------------------------T-----A-G.-A---------G---------------G----TA-------------------C----G-------------------------A-------------------A-- 8377
U.CY.05.CY090               .......................................................................................................................................................................... 8149
U.CY.08.CY223               .......................................................................................................................................................................... 8227
U.ES.10.DEURF10DZ001        --T-T-----C---TT---------------------------T---T-----T-----A-G.-A------------A------------G----TA----------T-T-----------G---C--------------A------A-------C--------T----- 8501
U.GR.99.99GR303             -AT-------C---------------------------------T--T-------------G.CA---------C---------------T-----A-------------------A----G------T-----------------AA-------C-----A--A----- 8526
U.NL.95.U_NL_95_H10986_D1   --T-T-----A---TT---------------------------T-----------------G.-A------G------------------G----T-----------T-------------G---A---------------------A-------C-----A--A----- 9142
01_AE.AF.07.569M            --T-T---------TT---------------------------T---------T-----A-G.-A---------G---------T-A---G-------T----------T-----------G---A-----------G---------AC-----TC------TGT----- 8481
01_AE.CF.90.90CF11697       --T-T---------TT---------------------------T---------T-----A-G.-A---------G--------CT-A---G----T--T----------------------G---A---------------------A-------C------TGT----- 9182
01_AE.CN.10.YNFL03          --T-T---------TT--------G------------------T---T-----T-----A-G.-A---------G--------CT-A---G----T-----G-------T-----------G------------------------A--------C------TGT----- 8593
01_AE.HK.04.HK001           --T-T---------TT---------------------------T---------T-----A-G.-A---------G--------CT-A---G-T--TA------------T-----------G-------------------------A-T------------TGT----- 8597
01_AE.IR.10.10IR.THR48F     --T-T-----C---TT---------------------------T---------T-----A-G.-A---------C--------CT-A---G--C-TA-T--------A-T-----------G---A--------------R------A-------C------TGT----- 8435
01_AE.JP.x.JRC77AE          --T-T---------TT-------------------------------------T----TA-G.-A---------G---------T-A---G----TA-T----------T-----------G---A---------------------AC------C------TGT----- 9218
01_AE.TH.09.AA090a_WG11     --T-T---------TT---------------------------T---------T-----A-G.-A---------G--------CT-A---G----TA-T-----G----T---T-------G------------------------A--------C------TGT--C-- 8646
01_AE.TH.90.CM240           --T-T---------TT---------------------------T---------T-----A-G.-A---------G--------CT-A---G----TA-T----------T-----------G---A---------------------A------TC------TGT----- 8783
01_AE.US.05.306163_FL       --T-T---------TT---------------------------T---------T-----A-G.-A---------G---------T-A---G-----A-T----------T---T-------G---A---------------------AC-----TC------TGT----- 8330
01_AE.VN.98.98VNND15        --T-T---------TT---------------------------T---------T-----A-G.-A---------C--------CT-A---G----TA-T----------T-----------G---C---------------------A-------C------TGT----- 8506
```

```
                                                    3' LTR U3 start                                          Nef premature end in HXB2
B.FR.83.HXB2                 CTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAA.AGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGG 9215
Nef                          A__V__D__L__S__H__F__L__K__E__K__G__G__L__E__G__L__I__H__S__Q#R__R__Q__D__I__L__D__L__W__I__Y__H__T__Q__G__Y__F__P__D__*__Q__N__Y__T__P__G__P__G__V__R__Y__P__L__T__F__G
02_AG.CM.10.DE00210CM013     --CAT-----C---TT-------------------------------T-----T-----A-G.-A---------G---------------G----T-----------A-T------A----G-------------------------ACT----TC-------------- 8705
02_AG.CY.09.CY256            .......................................................................................................................................................................... 8128
02_AG.ES.06.P1423            --A-------C--------------------------------T---T--G--TGG-----G.-A------G--A---------------G-Y--TA----------A--------A----G------T-----------A------ACT----TC-----A-------- 8676
02_AG.GW.05.CC_0048          --T-C-----C---TT-------------------------------T-----T-----AGG.-A---------G---------------G----T-------------T------A----G---A---------------------A-T-----C-----A-----C-- 8610
02_AG.KR.12.12MHR9           --T-C---T-C-------------G----------------------T-----T-----A-G.-A---------G---------------G----T-----------A-T------A----G---A---------------------A-T-----C-------------- 8964
02_AG.LR.x.POC44951          --T-T-----C--------------------------------T---T-------------G.-A-------------------------G----T-----------A-T------A----G-------------------------ACT----TC-------------- 9216
02_AG.NG.09.09NG_SC61        -CT-C-----C---TT-------------------------------T----AT-----A-G.-A---------G-----------A---G----TA----------A-T------A----G------------C------------ACT----TC-------------- 8386
02_AG.NG.x.IBNG              ----C-----C--------------------------------T---T-----T-----A-G.-A---------A---------------G----TA----------A-T------A----G-------------------------ACT----TC-------------- 8742
02_AG.SE.94.SE7812           ----C-----C------------------------------------T-----T-----A-G.-A---------A---------------G-T--T-----------A-T------A----G-------------------------ACT----TC-------------- 8605
02_AG.SN.98.98SE_MP1211      -G--------C------------------C-----------------T-----T-----A-G.-A---------G---------------G----T-----------A-T------A----G------T------------------A-T-----C-------------- 8450
03_AB.RU.97.KAL153_2         --T-C-----C--------------------------------T---T-----T-----A-G.-A---------A---------------G----T-------------------------G---................------A-T----TC-----A--A----- 8380
04_cpx.CY.94.94CY032_3       --T-------C--------------------------------T---------T-----A-G.-A---------G---------------G----T-----------T-T-----------G-G-T--------------A-------AG----TC------TG------ 8627
05_DF.BE.x.VI1310            ----------C--T---------------------------------T-----T-----A-G.-A-------------------T---G-----A----------T----C--------G-------------------------A-T-----------------G-- 8624
06_cpx.AU.96.BFP90           --T-T-----C---TT---------------------------T---------T-----A-G.-A------G--G---------------G----TA----------T-T-----------G---A---------------------AC-----TC--------T----- 9260
07_BC.CN.98.98CN009          --C-----------TT-------------------------------T-----T----TAGG.-A---G-----G---------T-----G----T-------------------------G---C--------------A-------------TC--------T----- 8558
08_BC.CN.97.97CNGX_6F        -AT-C-----C---TT-------------------------------T-----T----TA-G.-A---G-----G---------T-----G----T----------------------C-G---C--------------A-------------TC--------T----- 8415
09_cpx.GH.96.96GH2911        ----C-----C--TTT---------------------------T---T-----T-----A-G.-A---------A---------------G-------------------------A----G---A--------------A-------------T----T---------- 8390
10_CD.TZ.96.96TZ_BF061       -AT-C-----CG--TT---------------------------T---T-----T----T--G.-A---GA----C---------T-----G----TA------------T-----------G---A--------------A------AC---------G----------- 8570
11_cpx.CM.95.95CM_1816       --T-C------G--TT---------------------------T---------T-----A-G.-A---------G---------T-A---G----T-------------------------G---A---------------------A-------C------TGT----- 8623
12_BF.AR.99.ARMA159          --T-C------------------------------------------T-----T-T---A-G.-A---------G-----------------------------T-------------G--------------------------------TG-------------- 9180
13_cpx.CM.96.96CM_1849       --T-T------G--TT---------------------------T---------------A-G.-A---------C--------C------G----T--T----------------------G---A-----------------G-----------C------TGT----A 8640
14_BG.ES.05.X1870            --T-C-----C---TT---------------------------T---------T-----A-G.CA-------------------------G----TA----------A-T------A----G---------T--T------------A-T-----------C--------- 8721
15_01B.TH.99.99TH_MU2079     --T-T---------TT---------------------------T---------T-----A-G.-A---------G---------T-A---G-T--TA-T-----------T-----------G---A---------------------A-------C------TG------ 8590
16_A2D.KR.97.97KR004         -GT-T-----C-------------G------------------T--------------A--G.-A---G-----C---------------G-----------------T-----------G----GGT-----------------AA-T-----C-----A--A----- 8632
17_BF.AR.99.ARMA038          --C--------------------------------------------------T-----A--.-A---------A---------------G-----A-------------------------G-------------------------A-------------------G-- 8398
18_cpx.CU.99.CU76            --T-T--C--C---TT---------------------------T---T-----TT--T---G.-A---G-C---G---------T-----G-T-T-A------------TTC----------G------T-----------------A--------TC--C--A-------- 8544
19_cpx.CU.99.CU7             --T-------C------------------------------------A--G--TGG-T---G.-A---------C-----------T---G----T-----------T-------------G---T--T------------------A----G----------------- 8280
20_BG.CU.99.Cu103            --T-T-----C---TT---------------------------T---------T-----A-G.CA---------C--------C------G----TA-T-----------------A----G-------------------------ACA----TC-------------- 8661
21_A2D.KE.99.KER2003         --T-T--C--C--------------------------------T---T-----T-------G.-A---------A---------------G-----------------T-----------G------T------------------AC-----TC-----A--A----- 8410
22_01A1.CM.01.01CM_0001BBY   --T-T---------TT---------------------------T---------T-----A-G.-A---------G---------T-A---G----TA------------------------G---A-------------------......................... 8372
23_BG.CU.03.CB118            --T-T-----C---TT---------------------------T---------TGG---A-G.CA---------C--------C------G----TA-T----------T------A----G-------------------------A-A-------------------- 8670
24_BG.ES.08.X2456_2          --T-T-----C---TT---------------------------T---------T-T---A-G.CA---------C--------C------G----TA-T-------A---------A----G-----------------A------ACA----TC-------------- 8705
25_cpx.CM.02.1918LE          --T-T-----C---TT-----------------------------------G--TGG-----G.-AG--------A-----------------------------G----------------G------T-----G-------------......................... 8425
26_AU.CD.02.02CD_MBTB047     -CT-T-----CG--TT---------------------------T---------T-------G.CA----A----C--------C------G----TA-T-----------T---------C-G------T------------------A-T-----C-----A--A----- 9227
27_cpx.FR.04.04CD_FR_KZS     ----G-----C--------------------------------T---------TGG---A-G.-A---------A---------T-----G----TA----------T-----------C-G---A---------------A-------AG----T---C-------A-- 9238
28_BF.BR.99.BREPM12609       --T-------C-------------G-----------------------------------C-.-A---------G---------------G-T---------------------------G-------------------------AC-----TC-------------- 8628
29_BF.BR.01.BREPM16704       --A-----A-A---------------------------------------------------.-A---------G---------T-----G-------------------------------G-----------------A----------------------------- 8675
31_BC.BR.04.04BR142          -AT-T-----C---TT-------------------------------T-----T----TA-G.CA---G-----C---------T-----G-T--TA------------------------G---A---------------------AC-----T---C----------- 8734
32_06A1.EE.01.EE0369         --T-T-----CG--TT---------------------------T---------T-----A-G.-A---------G---------------G----T-----------T-------------G---A--------C----------A-------C---T---------- 8841
33_01B.ID.07.JKT189_C        --T-T-----A---TT--------G------------------T---T-----T-----A-G.-A------------T-----CT-A---G-T--TA------------T-----------G---A--------------------A--------C------TGT----- 8522
34_01B.TH.99.OUR2478P        --T-T---------TT---------------------------G---------T-----A-G.-------------G---------T-A---G----TA-T-----------T-----------G---A---------------------AC-----TC------TG------ 8381
35_AD.AF.07.169H             -CC-C-----C------------------C-------------T---T-----T-----AGG.-A---G-----A---------T-----G-----A-------G----------------G------T--T---------------A-----------T-A--A----- 8415
36_cpx.CM.00.00CMNYU830      --T-T---------TT------------------A--------T--------C--------G.-A------------------CT-A---G----T--T-----------T-----------G---A--------------A---.......................... 8410
37_cpx.CM.00.00CMNYU926      --T-T-----C---TT---------------------------T---T-G--------T---.-A---------C---------------------A-------------------------G------T---------------.......................... 8396
38_BF1.UY.03.UY03_3389       -----------------------------------------------------T-----A-G.-----------G---------------G-------------------------------G------T------------------A-----------------A-G-- 8611
39_BF.BR.04.04BRRJ179        --T-------C----------------------------------------G--TGG-----G.-A---------C---------T-----G-------------------------------G------T-----------------------GTC-------------- 8692
40_BF.BR.05.05BRRJ055        -CA-------C-------------G-A--G-----------------T-----------A-G.-----------C--A-----C------G-T--T-------------T---------C-G-------------------------A--------C----------- 8730
42_BF.LU.06.luBF_18_06       -AT--------------------------------------------------T-------.-----------C--------------------T-------------------------G------------C------------A--------C---TG------ 8720
43_02G.SA.03.J11223          -CT-T------G--TT-----------------------------------G--T-----A-G.-A------G--A---------T-A---G----T-----------A-T------A----G---T-----T--T------------A-T-----C-------------- 8743
44_BF.CL.00.CH80             ----------C--T---------------------------------T-------------G.-A---------C---------------G-----------------T-------------G------------C------------A-------------------G-- 8630
45_cpx.FR.04.04FR_AUK        -GA-T-----C---TT-------------------------------T--G--T-------G.-A---------C-----------A---G-------------------------------G------T------------------A-T-------------------- 9177
46_BF.BR.07.07BR_FPS625      --T-G-----C--T---------------------------------T---C-T-----A-G.-A------T--G--------C--T---G-------------------------------G-------------------------A----------------------- 8591
47_BF.ES.08.P1942            --T-C---------TT-------------------------------T-----T-T---AGG.-A---------G-----------A---G-T---------------A-------------G------------------------------G------T---------- 8632
48_01B.MY.07.07MYKT021       --T-T---------TT--------------------------GT---------T-----A-G.-A------------T------T-A---G----TA-T-----------T-----------G---A---------------------ACT----TC------TGT----- 8444
49_cpx.GM.03.N26677          -CT-T-----CG-------------............................................................................................................................................. 8519
50_A1D.GB.10.12792           -CT-T-----CG--TT---------------------------T---T-----T-----A-G.-A---------G---------------G-------------------T----------G------T------------------A-------C---T-A--A----- 8615
51_01B.SG.11.11SG_HM021      -CT-T---------------------------------------------G--TGG--T---.-A-------------------T-----G----T-------------------------G--------------------------AA----T---C----------- 8418
52_01B.MY.03.03MYKL018_1     -CT-T-----C---TT--------G------------------T---------T-----A-G.-A---------G---------T-A---G----T--T-----------T-----------G---C---------------------A-------C-----ATGT----- 8538
53_01B.MY.11.11FIR164        --T-T-----A---TT---------------------------T---T-----T-----AGG.-A---------G--------CT-A---G-T--TA-T-------A-T-----------GAGC----------------------ACT-----C-----ATGT----- 8552
54_01B.MY.09.09MYSB023       --T-T---T-AG--TT---------------------------T---------TGG---A-G.-A---------C--------CT-A---G----TA-T-----------T-----------G---A---------------------AC-----TC-----ATG------ 8662
55_01B.CN.10.HNCS102056      --T-T---------TT---------------------------T---------T-----ACG.-A---------G--------CT-A---G-T--TA-T-------T-T-----------G---C------------------------------C-----ATGT--C-- 8543
56_cpx.FR.10.URF5_patient_A  .......................................................................................................................................................................... 8396
57_BC.CN.09.09YNLX19sg       -AT-T-----C---TT-------------------------------T-----T----TA-G.-A---G-----G---------T-----G----T-------------------------G---A--------------A-----A--------C--------T----- 8446
58_01B.MY.09.09MYPR37        ----T---------TT---C---------------------------------T-----A-G.CA------------------CT-A---G-T--TA-T-----------T-----------G---A---------------------A-------C-----ATGT----- 8584
59_01B.CN.09.09LNA423        --T-T---------TT---------------------------T---------T-T--TA-G.-A---------G--------CT-A---G-T--T--T-----------T-----------G---C----------.................................. 8380
60_BC.IT.11.BAV499           --C--------------------------------------------------T--------.CA-----G------A------------G-----A-----------T--T----------G---A-----------G------A-A------------T--TG------ 9193
61_BC.CN.10.JL100010         --C--------------------------------------------T-----T-----A-G.-A---G-----G---------T-----G----T-------------------------G---A--------------A------A-T----TC--------T----- 8614
62_BC.CN.10.YNFL13           -AT-T-----C---TT-------------------------------T-----T----TA-G.-A---G-----G---------T-----G----T-------------T-----------G-G-C-----------G--A-------------TC-------------- 8549
63_02A1.RU.10.10RU6637       --C-C-----C------------------------------------------T-----A-G.-A---------A---------------G----T-----------A--------A----G-------------------------ACT----TC-------------- 8765
64_BC.CN.09.YNFL31           -AT-T-----C---TT-------------------------------T-----T----TA-G.-----G-----G---------T-----G----T-------------------------G---A-----------G--A-----AA------TC--------T----- 8583
65_cpx.CN.10.YNFL01          -AT-T-----C---TT-------------------------------T-----T-----A-G.-A---G-----G---------T-----G----T-------------------------G---A--T--------G--A-----A--------TC-------------- 8639
67_01B.CN.11.ANHUI_HF115     --T-T---------TT---------------------------T------C--T-----A-G.-A---------G--------CT-A---G-T--TA----T-------T-----------G---A---------------------A-------TC------TGT----- 8537
68_01B.CN.11.ANHUI_WH73      --T-T---------TT---------------------------T------G--T-----A-G.-A---------G--------CT-A---G-T--TA----T-------T-----------G---A---------------------A-------TC------TGT----- 8531
70_BF1.BR.10.10BR_PE004      --T-----------------------------------.................................................................................................................................... 8524
71_BF1.BR.10.10BR_PE008      --T-T--C-----------------------------------------------TGG-----G.-A-------------------------G-------------------------------G------T------------------AC-----GG-------------- 8872
72_BF1.BR.10.10BR_MG002      --T---------------------G-----------------------------T----T---.-A---------C---------------G-------T---------------------C-G-------------------------A--C------------------- 8740
O.BE.87.ANT70                -AT-T--C--C---TT------------------A------------T-----T-------T.-A----GC---A-----G-----T---G-G--TA----T--G--A-T-----------G------------------A-----AAC---G-TC--------A----- 9301
O.CM.98.98CMA104             -AT-C--C--C---TT------------------A--------C---T-----T----AG-T.GC----GC---A-----G-----T---G-GC-TA----T--G--A-T-M---------G------------------A-----AAC-----TC--C-----A----- 8805
O.CM.98.98CMABB141           -AT-T--C--C---TT------------------A--------T---T-----T-T-----T.-A----GC---G-----G-----T---G-AC-TA----T--G--A-T-----------G------------------A-----AAC-----TC--------A-A--- 8816
O.CM.98.98CMABB212           -CT-T--C--C---TT------------------A------------T-----T-------T.-AG---GC---A-----G-----T---G--C-T-----T--G--A-T-----------G----TG---------G--A------A-T-----C--------A----- 8824
O.CM.98.98CMU5337            -AT-T--C--C---TT------------------A------------T-----T---------.GCT---GC---G-----G-----T---G-A--TA-T--T--G--A-T-----------G------------------A-----AACT----TC--------A----- 8752
O.CM.99.99CMU4122            -AT-T--C--C--TGT---------------T---------------T-----T----T--T.G-G---GC---G-----G-----T-----A--TA----T--G--A-T-----------G-------------------------AACT----TC--------A----- 8771
O.FR.92.VAU                  -AT-T--C--C---TT------------------A--------T---T-----T-------T.-A----GC---G---T-A--C--T---G-GC-T-----T--G--A-T-----------G---A--------------A-----AAC-----TC---T-A--A----- 8873
O.GA.11.11Gab6352            -GCAT--C--C---GT------------------A------------T-----T-------T.GCT---GC---G-----G-----T---G-G--T-----T--G--A-T-----------G----C-------------A-----AACT----TC--------A----- 8539
O.SN.99.99SE_MP1299          -AT-T--C--C---TT------------------A--------T---T-----T-------T.-A----GC---G-----G-----T---G-A--T-----T--G--A-T-----------G----GG------------A------CCA----TC--------A----- 9363
O.US.10.LTNP                 -G.-T--C--C---TT-----A---GA--GGA.......................................................................................................................................... 9103
N.CM.02.DJO0131              ----T------TC-TT-------------------------------------TGG---AG-.-A---G-----CG-TT-A-----C---G-T--T-------------T---------C-G------------------------AACT----TC--C---------- 8690
N.CM.02.SJGddd               -CT-T--C---TC-TT-------------------------------------TGG---AGG.-A---G-----G--T--A-----C---G-T--T-------------T---------C-G-............................................... 8615
N.CM.04.04CM_1015_04         --T-T------TC-TT----------------------------------G--TGG---AG-.-A---G-----G--T--A--C--C-----T--TA----------------------C-G------------------------A-----G-----C-----A----- 8659
N.CM.06.U14296               --T....................................................................................................................................................................... 8545
N.CM.06.U14842               --T-T------TC-TT------------T----------G-------T--G--TGG--TAG-.-A---G-----G--T--A--C--C---G-............................................................................. 8637
N.FR.11.N1_FR_2011           G-T-T------TC-TT----------------------------------G--TGG---AG-.-AG--G--------T--A--C--C---G----T------------AT---------Y-G------------------------AACT-----C--C---------- 8584
P.CM.06.U14788               -AT-T--C--C---TT------------------A--------T---T--G--T----A-CT.GAG---GC---G-----A-----C--------T-----T--G--A-T---------C-G------T-----T----------A-AA--------C----------- 8818
P.FR.09.RBF168               -AT-T--C--C---TT-------------------------------T-----T-T--A-CT.G-G---GC---A--A--A-----C---G----T-----T--G--G-T---------C-G----C-T-----T----------A-AA-----TC-------------- 9338
CPZ.CD.06.BF1167             -A-----C--CTCATGG---------------------------------TA-TGG--T-C-.-AG---GC--CA--A--A---ACC-----G-TTA----T-----AATA----------G------T------------------A--T--------T----TGTAGA-- 9407
CPZ.CM.05.SIVcpzMT145        -CT-T-----CTCT---------------------------------T----CTGG--AAGG.-A------G--G---------T-A---G-T--T-----------T---------C-G---A--------C--G--------A--GC-C--C-----A-----C-- 8748
CPZ.GA.88.GAB1               --T-T------TCA---------------------------------T--G--T-----AGG.-----------G--------C--C---G----T-------------T---------C-G-------------A-----A-----AACA----TC-----CTGT----- 9288
CPZ.TZ.06.SIVcpzTAN13        -AAC------GTCATGG--C-----------------------T---T-----T----TAC-.C-T---GC--C------G---ACC-----G--TA----------AGTA-----A--C-G------T-----C--T--------A--A-----C-----CTG-AGA-- 8892
CPZ.US.85.US_Marilyn         -AT-C---T-AG--TT-------------------------------T--G--T-T---AG-.-----------G-----A-----C---G-T-----T--------AAT---------C-G------T-----C--------T--A--A--------C--A--A-A--- 9254
GOR.CM.07.SIVgor2139_287     -AT-T--C--C---TT---------T---------------------T--G--TT---T--G.GCC---AGT------T-A--C--C---G-T--T--T--T--G--A-----------C-G------T------------------A-A--------C----------- 8812
GOR.CM.07.SIVgorCP2135con    -AT-T--C--C---TT-------------------------------T--G--TT---T--G.GCC---AGT--G---T-A-----C---G-T--T--T--T--G--A-----------C-G------T--T---------------A-A--------C----------- 8796
```

B.FR.83.HXB2
Nef
A1.AU.03.PS1044_Day0
A1.CH.03.HIV_CH_BID_V3538
A1.CY.08.CY236
A1.ES.05.X1608_8
A1.KE.11.DEMA11KE001
A1.RU.11.11RU6950
A1.RW.11.DEMA111RW002
A1.SN.01.DDI579
A1.UG.11.DEMA110UG001
A1.ZA.04.04ZASK162B1
A2.CM.01.01CM_1445MV
A2.CY.94.94CY017_41
B.BR.10.10BR_MG029
B.CA.07.502_1191_03
B.CH.08.M2_0803101_NFLG8
B.CN.10.DEMB10CN002
B.ES.10.DEMB10ES002
B.FR.11.DEMB11FR001
B.GB.05.MM45d213_GN1
B.HK.06.HK003
B.HT.05.05HT_129389
B.JP.12.DEMB12JP001
B.KR.07.HP_18_07JHS10_3909
B.PE.07.502_0525_wg5
B.RU.11.11RU2ln
B.TH.08.MERLBDTRC10
B.US.11.ES38
C.AR.01.ARG4006
C.BR.07.DEMC07BR003
C.BW.00.00BW5031_1
C.CN.10.YNFL19
C.CY.09.CY260
C.ES.08.X2363_2
C.ET.02.02ET_288
C.IN.09.T125_2139
C.KE.00.KER2010
C.MW.09.703010256_CH256.w96
C.TZ.08.707010457_CH457.w8
C.YE.02.02YE511
C.ZA.08.705010534_CH534.w12
C.ZA.10.DEMC10ZA001
C.ZM.11.DEMC11ZM006
D.CM.10.DEMD10CM009
D.CY.06.CY163
D.KE.11.DEMD11KE003
D.KR.04.04KBH8
D.SN.90.SE365
D.TZ.01.A280
D.UG.10.DEMD10UG004
D.UG.11.DEMD11UG003
D.YE.02.02YE516
D.ZA.90.R1
F1.AO.06.AO_06_ANG32
F1.AR.02.ARE933
F1.BR.10.10BR_RJ015
F1.CY.08.CY222
F1.ES.02.ES_X845_4
F1.RO.96.BCI_R07
F1.RU.08.D88_845
F2.CM.02.02CM_0016BBY
F2.CM.10.DEMF210CM001
F2.CM.10.DEMF210CM007
G.BE.96.DRCBL
G.CM.10.DEMG10CM008
G.CN.08.GX_2084_08
G.CU.99.Cu74
G.ES.09.X2634_2
G.GH.03.03GH175G
G.KE.09.DEMG09KE001
G.NG.09.09NG_SC62
G.ZA.01.TV546
H.BE.93.VI991
H.CF.90.056
H.GB.00.00GBAC4001
J.CM.04.04CMU11421
J.SE.93.SE9280_7887
J.SE.94.SE9173_7022
K.CD.97.97ZR_EQTB11
K.CM.96.96CM_MP535
U.CA.01.TV749
U.CA.99.TV721
U.CD.83.83CD003_Z3
U.CD.90.90CD121E12
U.CY.05.CY090
U.CY.08.CY223
U.ES.10.DEURF10DZ001
U.GR.99.99GR303
U.NL.95.U_NL_95_H10986_D1
01_AE.AF.07.569M
01_AE.CF.90.90CF11697
01_AE.CN.10.YNFL03
01_AE.HK.04.HK001
01_AE.IR.10.10IR.THR48F
01_AE.JP.x.JRC77AE
01_AE.TH.09.AA090a_WG11
01_AE.TH.90.CM240
01_AE.US.05.306163_FL
01_AE.VN.98.98VNND15

Draft of 30 September 2015, not for public release
hiv1dna_main.tex r8876–10546


```
B.FR.83.HXB2                 ATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGATA.GAAGAG...GCCAATAAAGGAGAGAACACCAGC.TTGTTACACCCTGTGAGCCTGCATGGGATGGATGAC......CCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCA 9374
Nef                          __W__C__Y__K__L__V__P__V__E__P__D__K__I_#E__E__.__A__N__K__G__E__N__T__S_#L__L__H__P__V__S__L__H__G__M__D__D__.__.__P__E__R__E__V__L__E__W__R__F__D__S__R__L__A__F__H__H
02_AG.CM.10.DE00210CM013     G-------T---A---------C-G--T----TAG-----.--GAGA..------G-----------A---T.--A--------CA-CT---A------A-----------.....GAA--C---------C-G-TC-----A---------AC---G------A-A-- 8864
02_AG.CY.09.CY256            ........................................................................................................................................................................ 8128
02_AG.ES.06.P1423            ---------------------A-G------TCAG-AG--.--GA-A..------G-----------A-T-T.--A-----T--CA-CT-T-A------A--------T.....GAA--C---------C---TC-----A--------T-T---G-----CA-A-- 8835
02_AG.GW.05.CC_0048          G-------T---A---------A-G--T---ACAG--G--.--G--A..CTT-------A--------A---T.--A--------C--C--T-A------A-----G--.....GAA--T---------C-G-TC-----A---------R----G-----CA.... 8765
02_AG.KR.12.12MHR9           --------T---A---------T-G--T----CAG-AG--.--G--A..----------A--------A---T.--A--G--------AT-T-A------A-----G--.....GAA--C---------C-G-TC-----A----G---T-----T------ACA-- 9123
02_AG.LR.x.POC44951          G-------T---A---------A-G--T----CAG--G--.--G--A..-------------------A---T.--A--------CA-CT-T-A------A--A--G--.....A-C--C---C-----C-G-TC-----A---------A----G-----CA-A-- 9375
02_AG.NG.09.09NG_SC61        G-------T---A---------A-G--T----CAG-T---.--G--A..-------------------A---T.--A--------CA-CT---A------A-----G--.....GAA--A---------C-G-TC-----A---------A--------CGGAGA-- 8545
02_AG.NG.x.IBNG              G-------T---A---------A-G--T----CAG--G--.--G--A..------G-----------A---T.--A--------CA-CT-T-AA-----A-----G--.....GAC--T---------C-GATC-----A--------TA-A--G-----CAGA-- 8901
02_AG.SE.94.SE7812           G-------T---A---------A-G--T----CAG-C---.--GA-A..-A--C-G-----------A---T.--A--------CA-CT-T-A------A-----G--.....GAA--C---------C-G-TC-----A--------TA-A--G-----CACA-- 8764
02_AG.SN.98.98SE_MP1211      G-----------A---------A-G--T----CAG--G--.--G--A..------C-----------A---T.--A-----T--CA-CT-T-A------A-----G--.....GAA--C-A-------C-G-TC-----A---------A-G--G-----CAGA-- 8609
03_AB.RU.97.KAL153_2         G----------------------C--C----C-G-AG--.-------..----CAG-G----------A----.C-A---------A-AT---AA--C--A--------T.....GA---A-A----------AT-----A---------T-----G--TC-AACA-- 8539
04_cpx.CY.94.94CY032_3       --------T---------------A--T---C-GG--G-G.-------..----C-G-------------T-T.---C-G------A-A----A------A-----G--T.....GAA-----------A---A-----A---------T-----G----ACA-G-- 8786
05_DF.BE.x.VI1310            T-------T----T-----------A-T-----GG--G--.--GA--..------G--------G---A-T--.---C-------CA--------------A-----G--T.....GAT--C--------AC-GC------A---------A----G-----AAGA-- 8783
06_cpx.AU.96.BFP90           -------------------------T-----GG-AG--.------A..CTT-C---------------A-T--.C----------CA-AT---A------AGCA--G--T.....GAA--A-------------A------A----------TC-------AGGAGA-- 9419
07_BC.CN.98.98CN009          G-------T-----------------C---AGGG-AG--.-------..-----CG-------A---A-T--.----------------T---A---C--A-----G--T.....GAAC-C---------A---A-A----A---------T-AA------CACAGA-- 8717
08_BC.CN.97.97CNGX_6F        G-------T-----------------C---AGGG-AG--.-------..-----CG------AG---A-T-T.---C----------CT---A------A-----G--T.....GAAC-C---------A---A------A---------T-AA------CACAGA-- 8574
09_cpx.GH.96.96GH2911        --------T----T----G-----A--T---A-GG-AG--.--G---.----C-GG-----------A----.-------------A--T---AA-----AG--..--T.....T-A-GAG---------C-GAT----------------AC-------C-GA-A-- 8546
10_CD.TZ.96.96TZ_BF061       ------T-T----T-----------C---AGAG-AG-G.-------.------------------A-T--.---C-------CA-----------A-----G--T.....AAAC-TG--------A----T-----A---------AC-------CACA-A-- 8729
11_cpx.CM.95.95CM_1816       --------T--------------G--T---AGGG-AG--.--G---..------GC-----------A-T--.C---------CT------A------A-C------T.....GAA--A------------AT-----A---------TTC---C---CGAAGA-- 8782
12_BF.AR.99.ARMA159          G-------T-----------------C----TGG--G--.---A--..------G-----------A-T--.---C-------CA-----AA------A-----A--T.....GAA--C--------AC-GATA----A-----------------G--GC-AAGA-- 9339
13_cpx.CM.96.96CM_1849       --------T----T--------A----T--CA-AG-AG--.--G---..------G-------A---A-T--.C--C--------AA-----A------A--A--G--T.....G-A--A---------A---AT-----A---------TTC-------CGAA-A-- 8799
14_BG.ES.05.X1870            G-------T---A----A------A--T----CAG--G--.--G--A..--T-C-GG-----------A---T.--A--------CA-CT---A------A--A--G--.....G-A--C-AG--G---C-GAGA----AA--------TG--------CGGAGA-- 8880
15_01B.TH.99.99TH_MU2079     --------T-----------------C---AG-G-TG--.--G---..AA---C-------A---A-T--.C----------CA-----A------A-----G--.....GAA--A---------C-GA-----CA---------TGC-------CGAA-A-- 8749
16_A2D.KR.97.97KR004         --------T---------------CAGT----C-G--G--.--G---..--T-C-G-G-------C--A----.-----------A-AT---AA------A--A--G--.....--T---------------A-----GT------T--T-A---T---C-GGTA-- 8791
17_BF.AR.99.ARMA038          G-------------------------A-----GG------.--GA-A..-----CG-------A---A-T--.C--C--------CA-----AA----AA--------T.....GAA--C---------C-GA------A----------A----G------AGA-- 8557
18_cpx.CU.99.CU76            --------T---A--------------TT--A--G-AG--.-------..--T---G-G----------A---T.C-----------A-AT---AA------A--------T.....GA-----------A---A------A---------TGC---G---CGGAGA-- 8703
19_cpx.CU.99.CU7             --------T-G--------------G--T---A-GG--GC-.--GA-A..-A--C-G-----------A-T--.---------------C-T-T-AA-----A-----------.......--A----A-----------T------A---A------A-A---------G-R-- 8439
20_BG.CU.99.Cu103            G-------T---A---------A-G--T----CAG--G--.--G--A..-------T-----------A----.--A-----T--CA-CT---A-------------G--T.....AAT--A---------C-GATA-----A--------TA--------CGGAGA-- 8820
21_A2D.KE.99.KER2003         --------T---A--------------T---TC-G--G--.-------..-AT-C---G-----------A----.--A---------A--T---AA--A--A--A-------.....--T---------------AG-----A-------CCT-A---G---C-GA-A-- 8569
22_01A1.CM.01.01CM_0001BBY   ........................................................................................................................................................................ 8372
23_BG.CU.03.CB118            G-------T---A-----------A--T----CAG-TG--.--G--A..----G---T-----------A---T.--A--------CA-CT---A------A-----G--.....GAT--C-AG------C-GAT------A--------TA--------CGGAGA-- 8829
24_BG.ES.08.X2456_2          G-------T---A-----------G--T---ACAG--G--.--G--A..--------GT-------C--A---T.--A-----T--CA-CT---A------A-----A--T.....AGT--C-A-------C-GATA----AA--------TA----G---CGGAGA-- 8864
25_cpx.CM.02.1918LE          ........................................................................................................................................................................ 8425
26_AU.CD.02.02CD_MBTB047     G-------T------------------T----C-G--G--.-----A..--T----G----A---A----.C-AC-G--TG--A-------A----A----G--T.....AGT--AG---------C-CAT-----A---------T-AA------CGGA-G-- 9386
27_cpx.FR.04.04CD_FR_KZS     --------T---------------A--C--CA-GG--G--.-------..------G-------A---A----.C--C----T--A-------------A-----G--.....GAA--A-------------AT-----A---------T---------C-AAGA-- 9397
28_BF.BR.99.BREPM12609       --------TT-----------------T----AC--G--.----C-..------G---------T-A----.C----------A------A------A----------------.....--A----A-------C--AT-----A------------------------ 8787
29_BF.BR.01.BREPM16704       ---------------------------T----CC--G--.-------..-AG-----------------A----.C----------A------A------A-----------------.....--A---------------AT-----A-----------------G------- 8834
31_BC.BR.04.04BR142          C-------T-----------------C---AGGG-AG-G.-------..------GC-----------A----.---C-T------A--------------A-----G--T.....A-AC-CG-------AC--AT-----A---------T-T-------CGCAGA-- 8893
32_06A1.EE.01.EE0369         ---------------G-----------T---AGGG--G--.-----A..--T-C-------------A-T--.C----------CA-AT----------AGCA--G--T.....GAA--AG-------------AT-----------------TC-T-G---CGGAGA-- 9000
33_01B.ID.07.JKT189_C        --------T-----------------C---ACAG-AG--.--G---..-AA--C-------A---A-T--.C----------CA-----A------A-----G--.....GAA--C---------C-GA-----CA---------TTC-------CGAAGA-- 8681
34_01B.TH.99.OUR2478P        --------T-----------------C----GAG-AG--.A-G---................A---G-T--.C----------CA-----A------A--A--G--.....GAA--A---------C-GAT-----A---------TGC-------CGAA-A-- 8528
35_AD.AF.07.169H             ---------T--A-------------T----C-G-AG--.--GA--...-AT---G-----------A----.--A---------A--T---AA-----A--------T.....GA---------GACA---AT-----A-----------------G------ACA-- 8574
36_cpx.CM.00.00CMNYU830      ........................................................................................................................................................................ 8410
37_cpx.CM.00.00CMNYU926      ........................................................................................................................................................................ 8396
38_BF1.UY.03.UY03_3389       G-------T---------------G--C------G--G--.--GA--..------G-G----------A-T--.---C-G-----CA-----AA-----AGCA--C--T.....GA---C-A------AC-G-TC-------------------G---C-GAGA-- 8770
39_BF.BR.04.04BRRJ179        G-------T----------------AGC-----GG--G--.---AG-..------G--------G---A-T--.--------------CA-----AA-----A--------T.....GAA--C----------C-G-T-----A----------AAG--------C----- 8851
40_BF.BR.05.05BRRJ055        G-------T-----G-----------C-----GG--G--.---A--..------G------------A-T--.---C-------CA------A-----A-----G--T.....GAA----A------AC--C--------------GC-ACG-------------- 8889
42_BF.LU.06.luBF_18_06       --------T-------------C------A----C--G--.---A--..----C-G---------------------.--------------A-A-----------A-----G--.....A-A---------------G-T-----A----------AAG--------C----- 8879
43_02G.SA.03.J11223          G-------T---A---------A-G--T----CAG-TG--.--G--A..-----------A--------A---T.--A---------A-CT---AA-----A-----A--T.....GAA--C---------C-G-TA-----A--------TA--------CGGAGA-- 8902
44_BF.CL.00.CH80             G-------T-----------------C-----GG--G--.---A--..------G------------TG--.---C-------CA-----AA----------------T.....G-A--C-A------AC-GAT-----A-----------A----G---C-GAGA-- 8789
45_cpx.FR.04.04FR_AUK        --------TT---T----------G--T---A-GG-AG--.-------..----C-G------------TT-T.--------------A-A---AA-----A--A--G--.....A-A--A---------C--ATA----A---------T-A-------CAAAGA-- 9336
46_BF.BR.07.07BR_FPS625      G-------T---------------A-C--C--GG--G--.---AG-..------G-----------A-T--.---C-------CA-A---AA-----A-----G--T.....GAA--A--------AC-GC------A-----------A-------------C-- 8750
47_BF.ES.08.P1942            G-------T-----------------C-----GG--G--.---A--..------G-------A---A-T--.---C-------CA-----AA-----A-----G--T.....GAA--C--------AC-GAT-----A---------A-A--------C-A-G-- 8791
48_01B.MY.07.07MYKT021       --------T-----------------C---C-AG-AG--.--G---..-A---C-------A---G-T--.C----------CA-----AA-----A-----G--.....GAA--A---------C-GAT-----A---------TA--------CGAAGA-- 8603
49_cpx.GM.03.N26677          ........................................................................................................................................................................ 8519
50_A1D.GB.10.12792           ------------A-----------------A---GG-AG--.--GA--..--T-CCG-G---G-----A-T--.C----------GA-AT---AA-----A--------------.....GA-----A------A---C------A-----------A----G---CGAA-G-- 8774
51_01B.SG.11.11SG_HM021      --------T-R---------------T-----C--G--.---A--..------G-GAA-------------.C----------GA-----A-------AT------------.....--A----A----------AT----A-------------------C---G--- 8577
52_01B.MY.03.03MYKL018_1     --------T-----------------C---A-AG-AG--.--G--A..-A---C-------A---G-T--.C----------CA-A---A-----ACAA--G--.....GAA--AG--------C-GAT-----A---------TGC-------CGAACA-- 8697
53_01B.MY.11.11FIR164        --------T-----------------C----CAG-AG--.--G---..-----C----A---A---A-T-T.C----------CA-----AA-----A-----G--.....GAC--A---------C-GATA----A---------TTC-------CGAAGA-- 8711
54_01B.MY.09.09MYSB023       C-------T-----------------C---AGAG-AG--.--G---..-----C-------A---A-T--.C----------CA-----A------AGC---G--.....GAA----A-------C-GAT-----A---------TGC-------CGAACA-- 8821
55_01B.CN.10.HNCS102056      G-------T-----------------C---AGAG-AG--.--G---..A---CC---A---A--T-A-T--.C----------CA-----A------A-----------.....GAA----A-------C-GAT-----AA--------TGC-------AGAA-A-- 8702
56_cpx.FR.10.URF5_patient_A  ........................................................................................................................................................................ 8396
57_BC.CN.09.09YNLX19sg       --------T-----------------C---AGGG-A---.-------..------G-------A---A-T-T.---C---T------T--A------A-----G--T.....GAAC-C---------A---AGA----A---------T-AA------CACAGA-- 8605
58_01B.MY.09.09MYPR37        --------T-----------------C---AGAG-AG--.--G---..-----C-------A---A-T--.C----G-----CA-----A------A--A--G--.....AAC--A---------C-GAT-----A---------TGC-------CGAA-G-- 8743
59_01B.CN.09.09LNA423       ........................................................................................................................................................................ 8380
60_BC.IT.11.BAV499           --------T---------G-----G---A--A-GG-C.................T-G--A-----G--G-T--.------------A--CAA-AA--A-AG----G--................-----------------AT-----A----------A-------G--C----- 9340
61_BC.CN.10.JL100010         G-------T-----------------A---A-GG-AG--.-------..-----CCG-------A---A-T-T.---C-----------T---A------A-----G--T.....GAAC-C----------C--AT-----A-----------A-T----G--CAG--- 8773
62_BC.CN.10.YNFL13           --------T-----------------C---AGGG-AG--.--G---..-----CG-------A---A-T-T.C--C-------AA--------------A-----G--T.....GAAC-CG-------A---AG-----A---------T-AA------CACA-G-- 8708
63_02A1.RU.10.10RU6637       G-------T---A---------A-G--T----CAG-----.-----A..----C-G-----------A-TCT.--A--------CA-CT---A------AC----G--.....GAA--C---------C-G-TC-----A---------TAAA--G------G---- 8924
64_BC.CN.09.YNFL31           --------T-----------------C---AGGG-AG--.-------..-----CG----G--A---A-T--.C--C-----------T---AA-----A-----G--T.....GAAC-C---------A---AT-----A---------T-AA------CACAGA-- 8742
65_cpx.CN.10.YNFL01          --------T-----------------T---AGAG-AG--.-------..AG---CG-----------A-T--.---C---T---A------A------A--------T.....GAAC-CG-----------G-T-----A---------TATG------CACAGA-- 8798
67_01B.CN.11.ANHUI_HF115     --------T-----------------T---AGAG--G--.--G--A..-A---C-----G--A---AAT--.C----------CA-----A------A-----G--.....GAA--A---------C-GAT-----A-----------GC-------CGAA-A-- 8696
68_01B.CN.11.ANHUI_WH73      --------T-----------------T---AGAG-TG--.--G---..-A---C---A-G--AG---G-T--.C----------CA-A---A------A--A-------.....GAA--A---------C-GAT-----A-----------GC-------CGAA-A-- 8690
70_BF1.BR.10.10BR_PE004      ........................................................................................................................................................................ 8524
71_BF1.BR.10.10BR_PE008      G-------T---A-------------C-----A---G--.-------..----------AAG-------G----.--A---------A------------------G--.....G------A------------TA----AA----------A----G-G--------- 9031
72_BF1.BR.10.10BR_MG002      --------T-------------C----------AG-AG--.---A--..----G-G----.-----T-ATT--.C------------CA-AT------------------CCA...GGA--AG----------G-TA-----A-----T--------G----------- 8901
O.BE.87.ANT70                -----TG-TT--A-----------GTCAGA---AG--GC-.---AGACTA-GA----C-T-T----GGG-T-AT.C-CC-G--T--A-CAT-TGCC-----AT-T--A--T.....A-AC-T-A----A-AC-GAT-----A------T--ATCT----GCAACACC-- 9463
O.CM.98.98CMA104             ----ATG-TT--------------GTCAGA--C-G--GC-.-----ATTG-G------GCAT--A-GAG-T-AG.C-CC-G--T--A-CCT-TGCT-----AT-T-GA--T.....G-AC-CG-G---A-A---A-----CA------T--ATCA----GCC-CACC-- 8967
O.CM.98.98CMABB141          G----TA--T---T----------GACAAG---GG--GC-.---AGACTA-GA---G-GT-T----GGG-T--T.C-CC-G--T--A-CTT-TAAC--C--CT-T--A--T.......--AC-CG-G--GA-AC-GA-A---CA---------ATCA--G-GCAACACC-- 8978
O.CM.98.98CMABB212          -----TA-----------------GACAG----AG--GC-.--GAGACTG-GT----CGA-G--AG-TG-T--T.C-TC----T--A-CTT-TAAT-----CGCT--G--T.....G-AC-CG-G---A-AC-GA-A---CA------T--AACAT---GCC--ACA-- 8986
O.CM.98.98CMU5337           ----ATG-TT-------G------GTCAGA--C-G--GC-.--T--WCTG-GA-----GCAT--A-GGG-T-AA.C-CC-G--T--A-CTT-TAAC-----CT-T--G--T.......--TC-TG-----A-AC-GA-A---CA------T--ATCA----GCGGCACC-- 8914
O.CM.99.99CMU4122           G----TA-TT--------------GTCAGA--C-G--GC-.-----ACTA-GA-----GT-T-T--CAG-T-AA.C-CC-G--T--A-CTT--AAC-----TTCT--A--T.......--AC-CG-----A-AC-GA------CA------T--ATCA----GCAACATA-- 8933
O.FR.92.VAU                  ----CTA-TT--------------GTCAGA--C-G--GC-.----CACTG-G------GT-T--A-GGG-T---.C-CC-G--T--A-CTT-TAAT-----TTAT--A--T......-AAC-C-AG---A-AC-GA-A---CA------T--ATCA-----CAACACC-- 9035
O.GA.11.11Gab6352            ----CTA-TT--------------GACAGA----G--GCT.A--AG-CTA-GA---G--T-T-T--GGG--CAT.C-CC-G--T--A-CTT-TGCC-----CT-T--A--T.......--AC-TG-----A-AC-GA------A------T---TCA----GCC-CACC-- 8701
O.SN.99.99SE_MP1299          G----TA-TT---T----------GTCAGA--C-G--GC-.--G--ACTA-GA--C--GT-T----GGG-T---.C-CC-G--T--A-CTT--AAC-----CT-T--A---.....AACC-CG-GC--A-AC-GA-A---CA------T--ATCA----GCAGCACC-- 9525
O.US.10.LTNP                 ........................................................................................................................................................................ 9103
N.CM.02.DJO0131              ------T-T---A---------T-GTCAG-T--AGCAG--.-------..--T---G-------C----ATGC-.C-C-----T--CC-AT---AA-----AG-A-----T.....GGAC-C-A-------C-GC-----CA-----------TC---G---AGAAGA-- 8849
N.CM.02.SJGddd               ........................................................................................................................................................................ 8615
N.CM.04.04CM_1015_04         --------T---A---------T-GTCAG-T--AGCAG--.-----A..--T---G-------C----ATGC-.C-C-----T--CA-AT-T-AA-----AG-A-----T.....GATC-C-A-------C-G-T----C------------TC---G---AGAAGA-- 8818
N.CM.06.U14296               ........................................................................................................................................................................ 8545
N.CM.06.U14842               ........................................................................................................................................................................ 8637
N.FR.11.N1_FR_2011           --------T---A---------C-GTCAG-T--AGCAG--.-----A..--T-C-G-G----------ATTC-.C-C--G----CA-AT---AA-----AG-A-----T.....GATC-C-A-------C-G-T----C-A---------TC-T-----AGAAGA-- 8743
P.CM.06.U14788               G---CTG-TT--A-----------CTCTGAG-TAG-AGCT.--G--AATG-GAG----GCAG-----AG-T-AG.C--C----T--A-CCT--ACTT-----T-TTCA--T.......--TC-T-AG--GA-CC---T-----A-----------TCA--T-G-AGAG-A-- 8980
P.FR.09.RBF168               G---CTG-TT--A-----------CTCTGAG-TAG-AG-T.--G--AATG-GAG--G-GCAG-----AG-T-AG.C--C----T--A-CCT--ACTT-----GATTCA--T.......-ATC-T--G--GA-C---AT-----A----------ATCA--G-G-GAAAGA-- 9500
CPZ.CD.06.BF1167             T----TA-TT---T---------CA--C---CCAG-A....................G-G-AT--A-GG-ATCTA.C-CA-G-------CTT--TCCT----A-GA-----------.......--TA-TG-----C----G-TA----A-------TCTGAA--G--TAGAAGA-- 9551
CPZ.CM.05.SIVcpzMT145        G-------TT--A--G--C--TT-GCCA--T--GG--G-G.--GA-A..------G--------A---A-T--.C--C----T--CA--T---A-------A-----G--T.....GA---A--------AC-G-TA----CT-----T---AAG--G--TC-CAGA-- 8907
CPZ.GA.88.GAB1               --------T---A-----G--CC-GACAGAG--GC--G--.----C-A..------G-------T----A-T--.C----G--T--CA-AT-T-A---------------A--T.....GA---C-A---G--C--G-TC---C-C-----------A-G--G--C---AAGA-- 9447
CPZ.TZ.06.SIVcpzTAN13        G----TA-TT--------------AACT---CC-G--....................G-C-AT--A-GG-A--T-.---C-G--T--A-CT-A-TCT-----A-GA--G---CCAGATGGA--CG-C--GACTC-CAT----CA---------C-A---------AGAAGA-- 9042
CPZ.US.85.US_Marilyn         G-------TT---T----C--TC-CACAGA---GG-AG--.--GC--..-------------------C--A--TA.C--C-G-----CA--T---A------A-----A--T.....GAAC-TG-C-----C---ATC---CA---------CTGAA--G--TCGGAGA-- 9413
GOR.CM.07.SIVgor2139_287     ----CTG-TT---T----G-----CTCAGAG-CAG-AGCG.--G---ATG-GA-CATC-AAT-----GG---TG.C--C-G--T--A-C-T--AGC-------GA--C---.......--AT-TG----GA-CC-T-TC-----A--C--T--AGCAT-G-G-ACAAC--- 8974
GOR.CM.07.SIVgorCP2135con    ----CTG-TT---T----G-----CTCAGAG-CAG-AGCT.--G---ATG-GA-CAGC-A-T----GAG-T-TA.C--C-G--T--A-C-T--AGC-------GA--C---.......--AC-TG----GA-CC-T-TC-----A--C--T--AGCAT-G-G-ACAAC--- 8958
```

```
                                                         TCF-1 alpha binding
                                                                                    Nef end
B.FR.83.HXB2                CGTGGCCCGAGAG.CTGCATCCGGAGTACTTCAAG....................................AACTGCTGACA.................................TCGAGCT........TGCTAC.A................................ 9433
Nef                         __V__A__R__E_#_L__H__P__E__Y__F__K__.__.__.__.__.__.__.__.__.__.__.__.__N__C__*_
A1.AU.03.PS1044_Day0        ---A---------.-------------T--A---A....................................G----.............................................................................................. 8667
A1.CH.03.HIV_CH_BID_V3538   -T-T--TAA----.-------------T--A---A....................................G----------.................................CA--AG-........----GA.C...............................A 8649
A1.CY.08.CY236              .......................................................................................................................................................................... 8168
A1.ES.05.X1608_8            -AGA---------.A-A-------------A---A....................................G----------.................................CA--AG-........----GA.C................................ 8941
A1.KE.11.DEMA11KE001        -AC----------.A---------------A-G--....................................-----------.................................CA--AG-........----GA.C...............................A 8849
A1.RU.11.11RU6950           TAGA---------.A---------------A---A....................................G----------....................CAGACTGCTGACA.T--AG-........--A-GA.C................................ 9002
A1.RW.11.DEMA111RW002       ---A---A----A.GA-----------T--A---A....................................G-T--T-----GGAA..............CAAAGACTGCTGACACA--AG-........----GA.C..............................CA 8868
A1.SN.01.DDI579             -ACA--T------.A------------T--A---A....................................G----------..................CAGAAGTTGCTGACACA--AG-........----GA.C................................ 8631
A1.UG.11.DEMA110UG001       TACA---------.A------------T--A---A....................................G--------T-............................CAGTACAA--AC........----GA.CAC...........................AGA 8793
A1.ZA.04.04ZASK162B1        -C-A---------.A------------T--A---A....................................G----------..................CAGAGACTGCTGACACA--AG-........----GA.CAGGGACTTTCT..................GCT 8926
A2.CM.01.01CM_1445MV        -AGA--------A.---T-------T-T--A---A....................................G----------..................CAGAAGCTGCTGACACA--AG-........----GA.C................................ 8665
A2.CY.94.94CY017_41         -AGA---------.----------------A---A....................................G----------.................................CA--AG-........----GA.C................................ 8859
B.BR.10.10BR_MG029          T------------.A-------A--A--T-A----ACCCAAGAACTGCTGACAGCAAGCAACCAAG.....-----------.................................G-A----........-T----.-................................ 8961
B.CA.07.502_1191_03         T--A---AA----.---A-G-------T-----GAGACCGCTGACAT...............CGAGAATTAT----------.................................-----A-........-T----.-................................ 8901
B.CH.08.M2_0803101_NFLG8    -A-----------.--A----------T--A---A....................................G----------..................TCGAGACTGCTGACAC------........-T----.-................................ 9013
B.CN.10.DEMB10CN002         -A-------C---.-------------------GA....................................G----------.................................CT-----........-T----.-................................ 8818
B.ES.10.DEMB10ES002         -------------.-----------------T--A....................................G----------.................................-----TC........-T----.-................................ 8833
B.FR.11.DEMB11FR001         -A-----------.AA-----------TT-A----....................................G--------T-.................................-------........-T----.-................................ 8681
B.GB.05.MM45d213_GN1        ---A--------AG----------------A----.......................................AA----AC................................................----GA.................................. 8850
B.HK.06.HK003               -A-----------.----------------A----....................................-----------.................................C------........-T----.-................................ 8804
B.HT.05.05HT_129389         .......................................................................................................................................................................... 8398
B.JP.12.DEMB12JP001         ---A---A-----.AAA-------------A----....................................G----------.................................-----A-........-T----.-................................ 8852
B.KR.07.HP_18_07JHS10_3909  ---A---------.AAA-------------A----....................................G----------.................................-------........-T----.-................................ 9099
B.PE.07.502_0525_wg5        -AAA---A-----.A------------T--A----GACTGACAC.........................TG-----------.................................CT-----........-T----.-................................ 8879
B.RU.11.11RU21n             -------------.AA-----------TT-A---A....................................G----------.................................-----A-........-T----.-................................ 8961
B.TH.08.MERLBDTRC10         -AG----------.-------------T--A---A....................................G----------.................................C------........-T----.-................................ 8729
B.US.11.ES38                ---A---------.----------------A----....................................-----------.................................CT-----........-T----.-................................ 9400
C.AR.01.ARG4006             -A-------C---.--A-----------T-A---A....................................G----------.................................CA--AGGG.....AC-TTCCG.C................................ 8654
C.BR.07.DEMC07BR003         -A-------C---.--A-------------A---A....................................G----------.................................CA--AGGG.....AC-T-CCG.C................................ 8882
C.BW.00.00BW5031_1          -A-C--------A.-GA----------TT-A---A....................................G----------.................................CA--AGGG.....AC-TTCCG.C................................ 8792
C.CN.10.YNFL19              -A-------C---.--A----------TT-A---A....................................G----------CAGAAGGACTGTTACTTGCTGACA.........CA--AGGG.....AC-TTCCG.C................................ 8836
C.CY.09.CY260               .......................................................................................................................................................................... 8216
C.ES.08.X2363_2             -AA------C---.--A-----------T-A---A....................................G----------.................................CA--AGGG.....AC-TTCCA.C................................ 8947
C.ET.02.02ET_288            -C-------C---.--A-----------T-A---A....................................G----A-----.............................CAGAAG-GA--T.......-T--G-.TGACACAGAAGGGACTTTCTGT........... 8688
C.IN.09.T125_2139           -C-------T---.-AA----------TT-A---A....................................G----------.................................CA--AGGG.....AC-TTCCG.C................................ 9495
C.KE.00.KER2010             -C-------C---.-AA-----------T-A---A....................................G-T--------.................................CA--AGGG.....AC-TT-TG.C................................ 8649
C.MW.09.703010256_CH256.w96 -A-------C---.--A----------TT-A---A....................................G----------.................................CAA-AGGG.....AC-TTCCG.C................................ 9421
C.TZ.08.707010457_CH457.w8  -A-------C---.AAA-----------T-A---A....................................G----------.................................CA--AGGG.....AC-TTCCG.C................................ 9411
C.YE.02.02YE511             ---------C---.--A----------TT-A---A....................................G----------.................................CA--AGGG.....AC-TTCCG.C................................ 8688
C.ZA.08.705010534_CH534.w12 -A-------C---.A-A-----------T-A---A....................................G----------.................................CA--AGGG.....AC-TTCCG.CG..............................C 9452
C.ZA.10.DEMC10ZA001         -C-------C---.--A-----------T-A---A....................................G----------.................................CA-GAGGG.....AC-TTCCGCC................................ 8817
C.ZM.11.DEMC11ZM006         -C-----------.--A----------GG-A---A....................................G----------.................................CA--AGGG.....AC-TTCCGCC................................ 8825
D.CM.10.DEMD10CM009         TAA----A--C--.AA------A----T---T--A....................................G----------...........CCTTCTGTGAAGACTGCTGACAC------........-T----.-................................ 8805
D.CY.06.CY163               .......................................................................................................................................................................... 8143
D.KE.11.DEMD11KE003         TAA------TC--.--A-------------A---A....................................-----------.....................AGACTGCTGACA.-TGAG-........-T----.-................................ 8774
D.KR.04.04KBH8              -AA-------AT-.-------------TT-A---A....................................G----------.................................CT---T-........-T----.-................................ 9334
D.SN.90.SE365               -AA----A--A--.-------------T--A---A....................................G----------.................................CA--AG-........----GA.C................................ 9434
D.TZ.01.A280                -AA-------AGC.T---------------A---A....................................G----------..................TCGAGACTGCTGACAC----T-T.......------.-................................ 8640
D.UG.10.DEMD10UG004         TAA----AAGATA.-AA-------------A---A....................................G----------.................................CT---T-........-T----.-................................ 8743
D.UG.11.DEMD11UG003         -AA----A---T-.T-A--------T-T--A---A....................................G----------.................................CT-----........-T----.-................................ 8790
D.YE.02.02YE516             -AAA---A--A--.AC-----------T---T--A....................................G----------...........CTGAG..TGAAGACTGCTGACAC-A--T-........-T----.-................................ 8627
D.ZA.90.R1                  -AA----------.--A-------------A---A....................................G----------.................................C----T-........-TG---.-................................ 8790
F1.AO.06.AO_06_ANG32        .......................................................................................................................................................................... 8257
F1.AR.02.ARE933             .......................................................................................................................................................................... 8338
F1.BR.10.10BR_RJ015         .......................................................................................................................................................................... 8494
F1.CY.08.CY222              .......................................................................................................................................................................... 8096
F1.ES.02.ES_X845_4          -A-A---A-----.AGATT-----------A---AGACTGAG.............................-----------................CGCAGAACTTGCTGACACA--AGA........-T---A.................................. 8875
F1.RO.96.BCI_R07            T-AA---A-----.AGA----------T--A-C-AGACTAAGACCTC...............TGACACAGAG-T--------T..........AGAAGATTAGAGATTGCTGACA-A--AGA........-T---A.................................. 9548
F1.RU.08.D88_845            -C-C---A-----.AA---C--A----T--A----GACTAAGTTGTT................GACACAGA--TC-------.................................CA---A-........----GA.CWCAGAAGAATYT...................A 8984
F2.CM.02.02CM_0016BBY       .......................................................................................................................................................................... 8349
F2.CM.10.DEMF210CM001       -A-A---A-----.AAA-----A----T--A---AGACTAAG.............................-----------.................................CA--AGC........-T---A.GGGGACTTTGCTGACACAGAAGAATCTAAG... 8747
F2.CM.10.DEMF210CM007       -A-A---AA----.AAA----------T--A---AGACTGAG.............................-----------..................CAGAAACTGCTGACA.---AGA........-T---A.-................................ 8775
G.BE.96.DRCBL               -T-A---------.----------------A---A....................................G----------.................................CA--AG-........----GA.CAAGGGGA........CTTTAAGTTGCTG.ACA 9419
G.CM.10.DEMG10CM008         -A-A---------.AAA----------T--A---A....................................G----------.................................CA--AG-........----GA.C..............................AA 8816
G.CN.08.GX_2084_08          -CGA--------A.-------------T--AT---....................................G----------.................................CA--AG-........----GA.CATTGCTGAT......ACAGATGTTGCTGACAA 8715
G.CU.99.Cu74                -T-A---------.-----------T-T--AT--A....................................G----------.................................CA--AG-........----GA.C...............................A 9077
G.ES.09.X2634_2             -T-A---------.-------------T--AT--A....................................G----------.....................CAGCTGCTGACGCA--AG-........----GA.C..............................AA 8951
G.GH.03.03GH175G            TC-T---------.----------------A----....................................-----------.................................CA--AGG........--T-GA.CTGCTGAC........ACAGAAGTTGCTG.ACA 9526
G.KE.09.DEMG09KE001         -A-A---------.---T------------A---A....................................---------T-................................TAA--AG-........----GA.C..............................AA 8845
G.NG.09.09NG_SC62           -A-A---------.-----------T----A----....................................-----A-----.................................CA--AG-........----GA.C...............ACAAATGTTGCTG.ACA 8640
G.ZA.01.TV546               .......................................................................................................................................................................... 8003
H.BE.93.VI991               -AGA---AA----.-------------T--A---A....................................G----------.................................CA--AG-........----GA.CACAGAAGATTCT................ATAC 8852
H.CF.90.056                 -T---------TA.AA----------...-A---A....................................G----------.................................CA--AGA.........T--GA.-................................ 8752
H.GB.00.00GBAC4001          -AGA---A-----.-AA----------T--A---A....................................G----------........................CTGCTGACACA--AG-........----GA.C................................ 8954
J.CM.04.04CMU11421          -R-A---------.-------------TT-A---A....................................G----------.................................AG--AGWACAAAGAC----GA.C..............................AA 8978
J.SE.93.SE9280_7887         -A-A---------.-------------TT-A---A....................................G----------.................................AA--AG-........-T---G.C................................ 8731
J.SE.94.SE9173_7022         -A-A---------.--A----------TT-A---A....................................G----------.................................AA--AG-........-T---G.C................................ 8741
K.CD.97.97ZR_EQTB11         T--A---------.A---------------A---A....................................G--------.......................................................................................... 8600
K.CM.96.96CM_MP535          ---A---------.-----------C----A---A....................................G--------.......................................................................................... 8604
U.CA.01.TV749               -A-A---AA----.-----------C-T--A---A....................................G----------.................................CA--AG-........----GA.C................................ 9503
U.CA.99.TV721               -A-A---A-----.---T-------C-T--A---A....................................G----------.................................CA--AG-........----GA.C................................ 9502
U.CD.83.83CD003_Z3          -C-A---------.-------------TT-A---A....................................G----------....................TAGACTGCTGACACA--AG.........-T---A.G................................ 8912
U.CD.90.90CD121E12          -AA----------.--------------T-----A....................................G----------....................CAGACTGCTGACACA--AG.........-T---A.................................. 8606
U.CY.05.CY090               .......................................................................................................................................................................... 8149
U.CY.08.CY223               .......................................................................................................................................................................... 8227
U.ES.10.DEURF10DZ001        -C-A---------.AA-----------TT-A---A....................................G----------GGAGT......CTT..A.TAAAGACTGCTGACAAA--AGA........-T---A.C................................ 8743
U.GR.99.99GR303             -AGA---A---GA.GAA-------------A---A....................................G----------.................................CT--A-.........CT---A.-................................ 8745
U.NL.95.U_NL_95_H10986_D1   -CAA-------T-.--A-------------A---A....................................G--------T-.................................AA--AG-........-T---A.C................................ 9360
01_AE.AF.07.569M            -A-A--------A.--------A-------AT--A....................................G----------.................................AA--AG-........-T---A.C................................ 8699
01_AE.CF.90.90CF11697       -A-A---------.----G---A-------A---A....................................G-T--------.................................GA--AG-........-T---A.C................................ 9400
01_AE.CN.10.YNFL03          -A-A--------A.-A---C--A----T--AT--A....................................-----------.................................GA--AG-........-T---A.C................................ 8808
01_AE.HK.04.HK001           -C-A--------A.--------A----T--AT--A....................................G----------.................................AA--AG-........-T---A.C................................ 8815
01_AE.IR.10.10IR.THR48F     ---A--------A.--------T-------AT--A....................................G----------.................................AA--AG-........-CT--A.C................................ 8653
01_AE.JP.x.JRC77AE          -A-A--------A.--------A-------AT--A....................................G----------.................................AA--AG-........-T---A.C................................ 9436
01_AE.TH.09.AA090a_WG11     -A-T--------A.----G---A-------AT--A....................................G----------.................................AA--AG-........-T---A.C................................ 8864
01_AE.TH.90.CM240           ---A--------A.-A------A-------AT--A....................................G----------.................................AG--AG-........-T----.................................. 9000
01_AE.US.05.306163_FL       T--A--------C.--------A-------AT--A....................................G---............................................................................................... 8527
01_AE.VN.98.98VNND15        -A-A--------A.--------A--T----AT--A....................................G----------.................................AA--AG-........-T---A.C................................ 8724
```

Draft of 30 September 2015, not for public release

hiv1dna_main.tex r8876–10546


```
                                                         TCF-1 alpha binding
                                                                                    Nef end
B.FR.83.HXB2                   CGTGGCCCGAGAG.CTGCATCCGGAGTACTTCAAG..................................AACTGCTGACA................................TCGAGCT........TGCTAC.A............................... 9433
Nef                            __V__A__R__E_#_L__H__P__E__Y__F__K__.__.__.__.__.__.__.__.__.__.__.__N__C__*_
02_AG.CM.10.DE00210CM013       -AAA---------.--------T----T--A---AGACTGAAGATGTTGCTGACAGAA............G----------CAGGAGATTGCTGACTGCTGACA..........CA--AG-...........----GA.C.............................. 8969
02_AG.CY.09.CY256              ............................................................................................................................................................................ 8128
02_AG.ES.06.P1423              -ACA--T------.----------------A---A..................................G----------................CAGAAGTTGCTGACACA--AG-...........----GA.C.............................. 8909
02_AG.GW.05.CC_0048            ............................................................................................................................................................................ 8765
02_AG.KR.12.12MHR9             -AA---TAA----.-------------TT-A---A..................................G----T---................................................................................... 9166
02_AG.LR.x.POC44951            -AGA--T-A----.--------------T--A---A..................................G----------..............................CA--AG-...........----GA.C.............................. 9434
02_AG.NG.09.09NG_SC61          -A-A--------A.-----C--A-------A----..................................G----------..............................CA-GAG-...........----GA.CACAGAAGTTGCTGACACAGTTGTTGCTGACAA 8636
02_AG.NG.x.IBNG                -ACA--TA-----.----------------A---A..................................G----------..............................CA--AG-...........----GA.C.............................. 8960
02_AG.SE.94.SE7812             -AAA--T------.A------------TT-A---A..................................G----------.........CTTCTA.CAAAGACTGCTGACACA--AG-...........----GA.CAGGGACTTTCT..................GAC 8858
02_AG.SN.98.98SE_MP1211        -ACA--T------.----------------A---A..................................G----------..............................CA--AG-...........----GA.C.............................A 8669
03_AB.RU.97.KAL153_2           -AGA---------.-------------T--A---A..................................G----------...................CAGACTGCTGACA.T--AG-...........----GA.C.............................. 8610
04_cpx.CY.94.94CY032_3         ---A---------.-------------TT-A---A..................................G----------..............................CA--AG-...........----GA.C............................AA 8847
05_DF.BE.x.VI1310              -A-A---A-----.AGA-----A----T--A-C-AGACTGAGACTGC..............TGACACAGA-----------................CAGGAAGTGCTGACACA--AGCT........-T----.-............................. 8879
06_cpx.AU.96.BFP90             -A-A---A-----.AAA----------TT-A---A..................................G----------...................AGGCTGCTGACAAA--AG-........-T---A.T............................. 9490
07_BC.CN.98.98CN009            -AG------C--A.--A----------TT-A---A..................................G----------..............................CA--AGGG.....AC-TTCCG.C.............................. 8779
08_BC.CN.97.97CNGX_6F          -AG------C--A.--A----------TT-A---A..................................G----------..............................CA--AGGG.....AC-TTCCG............................... 8635
09_cpx.GH.96.96GH2911          YA-A---A-----.AAA----------T--A---A..................................G----------..............................CA--AG-.........-CTAG-.-............................. 8605
10_CD.TZ.96.96TZ_BF061         -A-A-----C---.--A-----------T-A---A..................................G-T--------..............................CA--.................................................. 8778
11_cpx.CM.95.95CM_1816         -A-A---------.-----------C-T--A---AGACTAA............................G----------..............................CA--AG-...........----GA.C.............................. 8847
12_BF.AR.99.ARMA159            -T-A---A-----.AAA-----A----GG-AT--AGACTGA............................G----------..............................CA---AC...........----GA.CACAGAAGAATC..................TAA 9418
13_cpx.CM.96.96CM_1849         -A-A---------.-------------T--A---A..................................G----------..............................AGAGAT-.........-TAC-............................... 8856
14_BG.ES.05.X1870              -A-A---------.----G------T-T--A----..................................G----------...................CAGCTGCTGACACA--AG-...........----GA.C............................AA 8953
15_01B.TH.99.99TH_MU2079       -A-A--------A.--------T-------AT--A..................................G----------..............................AA--AG-.........-T---A.C.............................. 8808
16_A2D.KR.97.97KR004           -AAA---------.-------------T--A---A..................................G----------................CAGAAGCTGCTGACACA--AG-...........----GA.C.............................. 8865
17_BF.AR.99.ARMA038            ---A---A-----.AGA-----A----T--A-C-AAACTGA............................G----------..............................CA---A-...........----GA.CACAGAAGAATCT...................A 8635
18_cpx.CU.99.CU76             -A-A--AA----A.-AA----------T--A---A..................................G----------..............................CA--AG-...........----GA.C.............................. 8762
19_cpx.CU.99.CU7              -AA----A--ATA.AA---C--A----T--A---A..................................G----------..............................C-A--T-.........-T----.-............................. 8498
20_BG.CU.99.Cu103              -T-A---------.-------------T--A---A..................................G----------..............................CA--AG-...........----GA.C............................AA 8881
21_A2D.KE.99.KER2003           -AGA----A----.AA-----------T--A---A..................................G-T--------.........CAGAAGCCCAAAGCTGCTGACACT---T-........GT----.-............................. 8650
22_01A1.CM.01.01CM_0001BBY     ............................................................................................................................................................................ 8372
23_BG.CU.03.CB118              -A-A---------.-G---------C-T--A---A..................................G----------..............................CA--AG-...........----GA.C............................AA 8890
24_BG.ES.08.X2456_2            -A-A---------.-----------C-T--A---A..................................G----------..............................CA--AG-G.............CG..........................ACA 8922
25_cpx.CM.02.1918LE            ............................................................................................................................................................................ 8425
26_AU.CD.02.02CD_MBTB047       -ACA-----C---.----------------A---A..................................G----------..............................CA--AG-...........----GA.C.............................. 9445
27_cpx.FR.04.04CD_FR_KZS       T--A---------.AAA----------T--A---A..................................G----------..............................CA--AG-...........----GA.C............................AA 9458
28_BF.BR.99.BREPM12609         -A-----------.--------------T-A---AGACTAA...........................AG----------..............................-----A-........-T--G-.-............................. 8853
29_BF.BR.01.BREPM16704         -A-----------.-------------TT-----A..................................G----------..............................-----T-........-T----.-............................. 8893
31_BC.BR.04.04BR142            -A-------C---.--A----------TT-A---A..................................G----------..............................CA--AGGG.....AC-TTCCG.C.............................. 8955
32_06A1.EE.01.EE0369           -A-A---A-----.AA-----------TT-A---A..................................G----------..............................AA--AG-........-T---A................................ 9058
33_01B.ID.07.JKT189_C          -A-A--------A.-AA-----T-------AT--A..................................G----------..............................AA--AG-........-T---G.C.............................. 8740
34_01B.TH.99.OUR2478P          -A-A--------A.-------------T--AT--A..................................G----------..............................AA--AG-........-T---A.C.............................. 8587
35_AD.AF.07.169H               -AAA---------.-------------T--A---A..................................G----------..............................CA--AG-...........----GA.C.............................. 8633
36_cpx.CM.00.00CMNYU830        ............................................................................................................................................................................ 8410
37_cpx.CM.00.00CMNYU926        ............................................................................................................................................................................ 8396
38_BF1.UY.03.UY03_3389         -A-A---A-----.AGA----------T--A-C-AGACTGAGACTGC..............TGCCACGAG-----------..............................CA---A-...........----GA.CACAGAAGAATC..................TAA 8864
39_BF.BR.04.04BRRJ179          -------------.----------------A----..................................G----------..............................CT-----..........-T----.-............................. 8910
40_BF.BR.05.05BRRJ055          T--------C---.-----C-------T-CA----..................................-----------..............................-T-----..........-C----.-............................. 8948
42_BF.LU.06.luBF_18_06         --------AC---.-------------T--A---A..................................G----------..............................C-A--T-.........-TA---.-............................. 8938
43_02G.SA.03.J11223            T--A--------A.-----C--A-------AT---..................................G---------G.....................CTGCTGACACA--AG-...........----GA.C................ACAGATGTTGTTGACAA 8987
44_BF.CL.00.CH80               -A-A---A-----.AGATT--------T--A-C-AGATGACTGA.........................GG-C-------..............................CA---A-...........----GA.CA............................. 8858
45_cpx.FR.04.04FR_AUK          -T-----A-----.AAA----------T--A---A..................................G----------..............................CA--AG-...........----GA.C.............................A 9396
46_BF.BR.07.07BR_FPS625        -AGA---------.-----------C-T--A----AACTGTTGCATC.................GAGCTG-----------..............................-T-----..........-T----.-............................. 8827
47_BF.ES.08.P1942              ---A---------.--A-----------T-A----..................................G----------..............................--------..........-T----.-............................. 8850
48_01B.MY.07.07MYKT021         -A-A---------.----G---A-------AT--A..................................G---------G...............................AA-GAG-........-T---A.C.............................. 8662
49_cpx.GM.03.N26677            ............................................................................................................................................................................ 8519
50_A1D.GB.10.12792             -A-A---------.-------------TT-A----..................................G----------..............................CA--AG-...........----GA.C.............................. 8833
51_01B.SG.11.11SG_HM021        -A-----AA----.AA---------T----A----..................................G----------..............................-----A-........-T----.-............................. 8636
52_01B.MY.03.03MYKL018_1       -ACA--------A.-----C--A--A-T--AT--A..................................G----------..............................AA--AG-........-T---A.C.............................. 8756
53_01B.MY.11.11FIR164          -A-A--------A.-----C--A-------AT--A..................................G----------..............................AA--AG-........-T---A.-............................. 8770
54_01B.MY.09.09MYSB023         ---A--------A.A----C--A--A-----T--A..................................G----------..............................AA--AG-........-T---A.C.............................. 8880
55_01B.CN.10.HNCS102056        -AGA---------.---------A----T--AT--A..................................G----------..............................AA--AG-........-T---A.C.............................. 8761
56_cpx.FR.10.URF5_patient_A    ............................................................................................................................................................................ 8396
57_BC.CN.09.09YNLX19sg         -AGA-----C---.--A----------TT-A--GA..................................G----------...................................CAG-................................................ 8654
58_01B.MY.09.09MYPR37          -A-A--------A.AA---C--A-------AT--A..................................G----------..............................AA--AG-........-T---A.C.............................. 8802
59_01B.CN.09.09LNA423          ............................................................................................................................................................................ 8380
60_BC.IT.11.BAV499             -A-----------.AA------------------A..................................G---------G.................................---.AG-........-T--M-.-............................. 9398
61_BC.CN.10.JL100010           ---A---------.AAA----------GG-A---A..................................G----------..............................CA--AGGG.....AC-TTCCG............................... 8834
62_BC.CN.10.YNFL13             -AG------C---.-------------TT-A---A..................................G---------CAGAGGGACTTTTGCTGACA............CA--AGGG.....AC-TTCCG............................... 8789
63_02A1.RU.10.10RU6637         T--A--T------.-G-----------TT-A---A..................................G----------..............................CA--AG-...........----GA.C.............................A 8984
64_BC.CN.09.YNFL31             -AG------C---.--A----------TT-A---A..................................G----------..............................CA--AGGG.....AC-TTCCG.C.............................. 8804
65_cpx.CN.10.YNFL01            -A-------C---.-AA----------TT-A---A..................................G----------..............................CA--AGGG.....AC-TACCG.C.............................. 8860
67_01B.CN.11.ANHUI_HF115       -A-A--------A.--------A--K----AT--A..................................G----------..............................AA-GRG-........-T---A.C.............................. 8755
68_01B.CN.11.ANHUI_WH73        -A-A--------A.--------A--T----AT--A..................................G----------..............................AA-GAG-........-A---A.C.............................. 8749
70_BF1.BR.10.10BR_PE004        ............................................................................................................................................................................ 8524
71_BF1.BR.10.10BR_PE008        T------------.T---------------A----..................................-----------..............................-T-----..........-T----.-............................. 9090
72_BF1.BR.10.10BR_MG002        -A-----------.----------------A----AACTAA...........................AG----------..............................-----T-........-T----.-............................. 8967
O.BE.87.ANT70                  T--T--TATGATA.ACT--C--A---CT----C--AAGGACTAA........................A----------CT..............................GA--A-...........----GA.CA...........................CTG 9536
O.CM.98.98CMA104               T--T--TAAGATA.ACA--C--A---CT----CCCAAGGACTGA..........................-C-------TCT.AAAGGGACTTTCCAG..............CAA--AC...........----GA.CA...........................CTG 9055
O.CM.98.98CMABB141             T--T--TAAGATA.ACT--C--A---YT----CYCAAGGACTAA..........................----------CT..............................GA--A-...........----GA.CA...........................CTG 9050
O.CM.98.98CMABB212             T--A--T-TGC-A.AA---C--A---CT----CCCAAATAG.............................--------G-C..............................T--TA-...........----GA.CA...........................CTG 9050
O.CM.98.98CMU5337              T--T--TATGATA.ACT--C--TC-TCT------CAAGGACTAA..........................----------CT..............................GA--GW...........----GA.CA...........................CTG 8986
O.CM.99.99CMU4122              T--T--TAAG-TA.ACC--C--A---CT----CTCAAGGACTAA..........................----------CT..............................GA--A............----GA.CA...........................CTG 9004
O.FR.92.VAU                    T--A--T-T-ATA.ACT--C--A---CT----TCCAAAGACTAA..........................-----TCA---G..............................GA--AC...........----GA.CG...........................CTA 9107
O.GA.11.11Gab6352              T--TT-TAA--TA.A-C--C--A---CT-----CCAAGGACTAA..........................-----------T..............................GA--A-...........----GA.CG...........................CTG 8773
O.SN.99.99SE_MP1299            T--T--TATG-TA.ACCA-C--A---CT---T--CAAGGACTAA..........................----------CT..............................GA--A-...........----GA.CA...........................CTG 9597
O.US.10.LTNP                   ............................................................................................................................................................................ 9103
N.CM.02.DJO0131                -A-A--AA-----.-----------C-TT-A----..................................-----------A..............................G-GA--T.......-A--...............................AAT 8909
N.CM.02.SJGddd                 ............................................................................................................................................................................ 8615
N.CM.04.04CM_1015_04           -A-A--AA-----.-----------C-T--A----..................................-----------A..............................G-GA--TT....AC----GA.CAAGGGACTTTAC...............ACAT 8897
N.CM.06.U14296                 ............................................................................................................................................................................ 8545
N.CM.06.U14842                 ............................................................................................................................................................................ 8637
N.FR.11.N1_FR_2011             T--A--AAA----.-------------TT------..................................G---------A..............................GAGA--TT....AC----GA.CAAGGGACTTTAT...............GCTT 8822
P.CM.06.U14788                 T--T---TT-C-A.AA---C-----ACTG--T-TTAAAGACTAA.............................-T.................................................----GA.CG...........................CCA 9036
P.FR.09.RBF168                 T--T---TT-C-A.AA---C-----ACTG--T-CTAAAGACTAA.......................................................................AG...........----GA.CG...........................CTA 9556
CPZ.CD.06.BF1167               -A-A---A-----.AGA----------------ATAA...............................GG------TC-GG.......CAACCACA................CAAGAAC..........---AGA.CG.............................. 9624
CPZ.CM.05.SIVcpzMT145          TC-A--AA-----.AAA--------A--T-A--GA..................................G---ACC---G..CGGCGCATGCGCACTAAGAACTGCTGAT.CAACAT-...........----GA.CA.............................. 8996
CPZ.GA.88.GAB1                 TA-T---A----A.-AA-------------A---A..................................G---ACTTT-G.ACTGGCGCATGCGCA..............CAAGAAC...........----GA.CT...........................CTG 9526
CPZ.TZ.06.SIVcpzTAN13          -A-A---A-----.AGA----------------ATAAGGACTT.......................CCGGG---CCTAGT................................AC-AAC...........----GA.CA.............................. 9113
CPZ.US.85.US_Marilyn           -AGA--TAA----.--------A---------CG-..................................-----A--G-TGTAACCGCGCAGGCGCA................ATA-AAC...........----GA.CTGAGGACTTT.................CTA 9502
GOR.CM.07.SIVgor2139_287       T--T---TT-C--.AA---C--A--ACTG------AGAGACTAA.................................................................-A--............----GA.CA...........................GCA 9032
GOR.CM.07.SIVgorCP2135con      T--T--ATT-C--.AA---C--A--ACTG------AGAGAATAA........................AG---AA..................................AA-A--............----GA.CA...........................GCA 9025
```

NF-κ-B-II
NF-κ-B-I
Sp1-III
Sp1-II
Sp1-I
TATA box
TAR element start
3' LTR U3 end
3' LTR repeat start

NF-κ-B-II
NF-κ-B-I
Sp1-III
Sp1-II
Sp1-I
TATA box
TAR element start
3' LTR U3 end
3' LTR repeat start
B.FR.83.HXB2 AGGGACTTTCCGCT GGGGACTTTCCA GGGAGGCGTGGCC TGGGCGGGACTGGGGAGTGGCGAGCCCTCAGAT CCTGCATATAAGCAGCTGCTTTTTGCC TGTACTGGG TCTCTCT GGTTAGACCAGATCT 9563
02_AG.CM.10.DE00210CM013 9099
02_AG.CY.09.CY256 8128
02_AG.ES.06.P1423 8983
02_AG.GW.05.CC_0048 8765
02_AG.KR.12.12MHR9 9166
02_AG.LR.x.POC44951 9564
02_AG.NG.09.09NG_SC61 8766
02_AG.NG.x.IBNG 9090
02_AG.SE.94.SE7812 8988
02_AG.SN.98.98SE_MP1211 8789
03_AB.RU.97.KAL153_2 8739
04_cpx.CY.94.94CY032_3 8978
05_DF.BE.x.VI1310 9009
06_cpx.AU.96.BFP90 9620
07_BC.CN.98.98CN009 8906
08_BC.CN.97.97CNGX_6F 8762
09_cpx.GH.96.96GH2911 8735
10_CD.TZ.96.96TZ_BF061 8917
11_cpx.CM.95.95CM_1816 8978
12_BF.AR.99.ARMA159 9548
13_cpx.CM.96.96CM_1849 8916
14_BG.ES.05.X1870 9028
15_01B.TH.99.99TH_MU2079 8937
16_A2D.KR.97.97KR004 8991
17_BF.AR.99.ARMA038 8765
18_cpx.CU.99.CU76 8893
19_cpx.CU.99.CU7 8512
20_BG.CU.99.Cu103 8935
21_A2D.KE.99.KER2003 8777
22_01A1.CM.01.01CM_0001BBY 8372
23_BG.CU.03.CB118 8964
24_BG.ES.08.X2456_2 8925
25_cpx.CM.02.1918LE 8425
26_AU.CD.02.02CD_MBTB047 9575
27_cpx.FR.04.04CD_FR_KZS 9589
28_BF.BR.99.BREPM12609 8886
29_BF.BR.01.BREPM16704 8979
31_BC.BR.04.04BR142 9057
32_06A1.EE.01.EE0369 9189
33_01B.ID.07.JKT189_C 8871
34_01B.TH.99.OUR2478P 8650
35_AD.AF.07.169H 8698
36_cpx.CM.00.00CMNYU830 8410
37_cpx.CM.00.00CMNYU926 8396
38_BF1.UY.03.UY03_3389 8935
39_BF.BR.04.04BRRJ179 9041
40_BF.BR.05.05BRRJ055 9042
42_BF.LU.06.luBF_18_06 9056
43_02G.SA.03.JI1223 9085
44_BF.CL.00.CH80 8858
45_cpx.FR.04.04FR_AUK 9525
46_BF.BR.07.07BR_FPS625 8957
47_BF.ES.08.P1942 8926
48_01B.MY.07.07MYKT021 8792
49_cpx.GM.03.N26677 8519
50_A1D.GB.10.12792 8963
51_01B.SG.11.11SG_HM021 8767
52_01B.MY.03.03MYKL018_1 8885
53_01B.MY.11.11FIR164 8899
54_01B.MY.09.09MYSB023 9010
55_01B.CN.10.HNCS102056 8889
56_cpx.FR.10.URF5_patient_A 8396
57_BC.CN.09.09YNLX19sg 8781
58_01B.MY.09.09MYPR37 8931
59_01B.CN.09.09LNA423 8380
60_BC.IT.11.BAV499 9528
61_BC.CN.10.JL100010 8960
62_BC.CN.10.YNFL13 8917
63_02A1.RU.10.10RU6637 9114
64_BC.CN.09.YNFL31 8931
65_cpx.CN.10.YNFL01 8987
67_01B.CN.11.ANHUI_HF115 8884
68_01B.CN.11.ANHUI_WH73 8878
70_BF1.BR.10.10BR_PE004 8524
71_BF1.BR.10.10BR_PE008 9198
72_BF1.BR.10.10BR_MG002 9062
O.BE.87.ANT70 9683
O.CM.98.98CMA104 9198
O.CM.98.98CMABB141 9182
O.CM.98.98CMABB212 9177
O.CM.98.98CMU5337 9134
O.CM.99.99CMU4122 9152
O.FR.92.VAU 9255
O.GA.11.11Gab6352 8921
O.SN.99.99SE_MP1299 9745
O.US.10.LTNP 9103
N.CM.02.DJO0131 8938
N.CM.02.SJGddd 8615
N.CM.04.04CM_1015_04 8926
N.CM.06.U14296 8545
N.CM.06.U14842 8637
N.FR.11.N1_FR_2011 8892
P.CM.06.U14788 9193
P.FR.09.RBF168 9712
CPZ.CD.06.BF1167 9758
CPZ.CM.05.SIVcpzMT145 9135
CPZ.GA.88.GAB1 9656
CPZ.TZ.06.SIVcpzTAN13 9247
CPZ.US.85.US_Marilyn 9628
GOR.CM.07.SIVgor2139_287 9180
GOR.CM.07.SIVgorCP2135con 9174

```
                                                       TAR element end           Poly-A signal   3' LTR R repeat end  3' LTR U5 start                                3' LTR U5 end
B.FR.83.HXB2                GAGCCTGGGAGCTCTCTGGCTAACTAGG....GAACCCACTGCTTAA.GCCTCAATAAAGCTTGCC.TTGAGTGCTTCAA.GTAGTGTGTGCCCGTCTG.TTGTG.TGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCA 9719
A1.AU.03.PS1044_Day0        ..................................................................................................................................................................... 8667
A1.CH.03.HIV_CH_BID_V3538   ..................................................................................................................................................................... 8750
A1.CY.08.CY236              ..................................................................................................................................................................... 8168
A1.ES.05.X1608_8            ..................................................................................................................................................................... 9017
A1.KE.11.DEMA11KE001        ----------------------G-G---....--------............................................................................................................................. 9015
A1.RU.11.11RU6950           ----------------------G-----....---------------.------------------.----------T--.................................................................................... 9205
A1.RW.11.DEMA111RW002       ---T-C----A---------C------A....-------.............................................................................................................................. 9033
A1.SN.01.DDI579             ----------------------G-A---....------------......................................................................................................................... 8801
A1.UG.11.DEMA110UG001       ---T------A-----------G-G---....-----................................................................................................................................ 8956
A1.ZA.04.04ZASK162B1        ----------------------G-A---....---------------A---------............................................................................................................ 9109
A2.CM.01.01CM_1445MV        ------------------------A--G....---------............................................................................................................................ 8832
A2.CY.94.94CY017_41         -----C----------------G-AG--....---------------.------------------.----A------....................................................................................... 9060
B.BR.10.10BR_MG029          --------............................................................................................................................................................. 9100
B.CA.07.502_1191_03         ---------------------G------....------------......................................................................................................................... 9072
B.CH.08.M2_0803101_NFLG8    ----------------------G-----....---------------.------------------.------------..-CG-................................................................................ 9221
B.CN.10.DEMB10CN002         -----C----------------G-----....----................................................................................................................................. 8980
B.ES.10.DEMB10ES002         ----------------------------....---------------.-------.............................................................................................................. 9013
B.FR.11.DEMB11FR001         ----------------------G-----....-------.............................................................................................................................. 8846
B.GB.05.MM45d213_GN1        ..................................................................................................................................................................... 8850
B.HK.06.HK003               ---------------------------.......................................................................................................................................... 8962
B.HT.05.05HT_129389         ..................................................................................................................................................................... 8398
B.JP.12.DEMB12JP001         ---------------------G------....-----................................................................................................................................ 9016
B.KR.07.HP_18_07JHS10_3909  ..................................................................................................................................................................... 9123
B.PE.07.502_0525_wg5        -----C---------------------A....---------------.------------------.-----------....................................................................................... 9081
B.RU.11.11RU21n             ---------------------G------....---------------.------------------.---------CT-...................................................................................... 9165
B.TH.08.MERLBDTRC10         ..................................................................................................................................................................... 8793
B.US.11.ES38                ----------------------G-----....-----.------A-....................................................................................................................... 9572
C.AR.01.ARG4006             ----------------------T-----....------G-----......................................................................................................................... 8822
C.BR.07.DEMC07BR003         ----------------------T-----....----------........................................................................................................................... 9047
C.BW.00.00BW5031_1          -----C----------------T--G--....--------------G.------------------.---------CT--.------------------.-----.-----------------------------.............................. 9047
C.CN.10.YNFL19              ----------------------T-----....---------------.-------.............................................................................................................. 9013
C.CY.09.CY260               ..................................................................................................................................................................... 8216
C.ES.08.X2363_2             ..................................................................................................................................................................... 9023
C.ET.02.02ET_288            -----C--............................................................................................................................................................. 8822
C.IN.09.T125_2139           --------------------C-T-----....---------------.------------------.---------CTG-.-C----------------.-----.------------------------------T----.-GT-------------------- 9777
C.KE.00.KER2010             -----C----------------T-----....------------......................................................................................................................... 8817
C.MW.09.703010256_CH256.w96 ---------------------GC-----....---------------.------------------.---------CT--.------------------.-----.------------------------------A----.-GTT-C----------------- 9704
C.TZ.08.707010457_CH457.w8  ----------------------T--G--....---------------.------------------.---------CT--.------------------.-----.------------------------------AA--GT-GT-------------------. 9693
C.YE.02.02YE511             --------------------C-------....----------........................................................................................................................... 8854
C.ZA.08.705010534_CH534.w12 -----C----------------T-----....---------------.------------------.----------G--.------------------.-----.---------------------------------G-T-GT-------------------. 9733
C.ZA.10.DEMC10ZA001         ----------------------T----A....------............................................................................................................................... 8978
C.ZM.11.DEMC11ZM006         -----C----------------T--G--....----................................................................................................................................. 8984
D.CM.10.DEMD10CM009         -------A-----------T--G-----....---.................................................................................................................................. 8965
D.CY.06.CY163               ..................................................................................................................................................................... 8143
D.KE.11.DEMD11KE003         -------A--------------G--C--....---.................................................................................................................................. 8934
D.KR.04.04KBH8              T---------------GT--AT-GC-........................................................................................................................................... 9490
D.SN.90.SE365               -------A--------------G-----....---------------.------------------.-------........................................................................................... 9633
D.TZ.01.A280                ------............................................................................................................................................................... 8775
D.UG.10.DEMD10UG004         -------A--------------------....---.................................................................................................................................. 8904
D.UG.11.DEMD11UG003         ----------------------G-----....---.................................................................................................................................. 8950
D.YE.02.02YE516             -------A--------------G-----....------------......................................................................................................................... 8797
D.ZA.90.R1                  ..................................................................................................................................................................... 8916
F1.AO.06.AO_06_ANG32        ..................................................................................................................................................................... 8257
F1.AR.02.ARE933             ..................................................................................................................................................................... 8338
F1.BR.10.10BR_RJ015         ..................................................................................................................................................................... 8494
F1.CY.08.CY222              ..................................................................................................................................................................... 8096
F1.ES.02.ES_X845_4          ..................................................................................................................................................................... 8953
F1.RO.96.BCI_R07            ----------------------G-----....---------------.------------------.-------........................................................................................... 9747
F1.RU.08.D88_845            ..................................................................................................................................................................... 9049
F2.CM.02.02CM_0016BBY       ..................................................................................................................................................................... 8349
F2.CM.10.DEMF210CM001       ---------C-----------G------....---.................................................................................................................................. 8907
F2.CM.10.DEMF210CM007       ----------------------------....----................................................................................................................................. 8936
G.BE.96.DRCBL               ----------------------G-AG--....-----ACTGCT-AG-.------------------.-------------.------------------.-----.------------------------------AC-C---AT-----.A------------- 9705
G.CM.10.DEMG10CM008         ------------------A---G-AG--....-----................................................................................................................................ 8979
G.CN.08.GX_2084_08          -----C-------------T--G-A---....---------------.-------.............................................................................................................. 8895
G.CU.99.Cu74                ..................................................................................................................................................................... 9098
G.ES.09.X2634_2             ..................................................................................................................................................................... 9007
G.GH.03.03GH175G            ----------------------GTAGA-....---------------.------------------.-------........................................................................................... 9725
G.KE.09.DEMG09KE001         ----------------------G-AG........................................................................................................................................... 9001
G.NG.09.09NG_SC62           ----------------------G-AG-A....--------------....................................................................................................................... 8812
G.ZA.01.TV546               ..................................................................................................................................................................... 8003
H.BE.93.VI991               ----------------------G-----....---------------.------------------.-----G--GCAT.GC................................................................................... 9056
H.CF.90.056                 ----------------------G---A-....---------------.------------------.----------T....................................................................................... 8953
H.GB.00.00GBAC4001          ..................................................................................................................................................................... 9081
J.CM.04.04CMU11421          ..................................................................................................................................................................... 9077
J.SE.93.SE9280_7887         ----------------------------....---------------.------------------.-----G--GCAT..-C-AGCC-............................................................................ 8943
J.SE.94.SE9173_7022         ----------------------------....---------------.------------------.-----G--GCAT..-C-AGCC-............................................................................ 8953
K.CD.97.97ZR_EQTB11         ..................................................................................................................................................................... 8600
K.CM.96.96CM_MP535          ..................................................................................................................................................................... 8604
U.CA.01.TV749               ----------------------G-AG--....---------------.------------------.-------------.------------------.-----.------------------------------AC-C---A-------A------------- 9788
U.CA.99.TV721               ------------------A---G-AG--....---------------.------------------.-------------.------------------.-----.------------------------------AC-C---A--------------------- 9787
U.CD.83.83CD003_Z3          -----C------------................................................................................................................................................... 9060
U.CD.90.90CD121E12          -------------------T--G-----....---------------.----................................................................................................................. 8784
U.CY.05.CY090               ..................................................................................................................................................................... 8149
U.CY.08.CY223               ..................................................................................................................................................................... 8227
U.ES.10.DEURF10DZ001        ----------------------G-AG--....--................................................................................................................................... 8904
U.GR.99.99GR303             ----------------------G-A---....--------............................................................................................................................. 8909
U.NL.95.U_NL_95_H10986_D1   ----------------------G-----....---------------A------------------.----------A--.------------------.-----.------------------------------A--.------------------------- 9646
01_AE.AF.07.569M            ..................................................................................................................................................................... 8762
01_AE.CF.90.90CF11697       -----C----------------G-A---....---------------A------------------.----------A--.-G----------------.-GA-A.G.......................................................... 9628
01_AE.CN.10.YNFL03          -----C----------------G-A---....----------........................................................................................................................... 8975
01_AE.HK.04.HK001           -----C----------------G-A---....--------------GA------------------.----------G--.-G----------------.-GT-A.G------------.............................................. 9056
01_AE.IR.10.10IR.THR48F     ..................................................................................................................................................................... 8667
01_AE.JP.x.JRC77AE          -----C----------------G-A---....-------------G-A------------------.-------........................................................................................... 9634
01_AE.TH.09.AA090a_WG11     -----C----------------G-A---....-------------G....................................................................................................................... 9035
01_AE.TH.90.CM240           -----C----------------G-A---....---------------A------------------.----------A-...................................................................................... 9203
01_AE.US.05.306163_FL       ..................................................................................................................................................................... 8527
01_AE.VN.98.98VNND15        ----------------------............................................................................................................................................... 8875
```

```
                                                          TAR element end          Poly-A signal   3' LTR R repeat end  3' LTR U5 start                                     3' LTR U5 end
B.FR.83.HXB2                  GAGCCTGGGAGCTCTCTGGCTAACTAGG....GAACCCACTGCTTAA.GCCTCAATAAAGCTTGCC.TTGAGTGCTTCAA.GTAGTGTGTGCCCGTCTG.TTGTG.TGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCA   9719
02_AG.CM.10.DE00210CM013      -----------------A---G-AG---....-------------..........................................................................................................................   9140
02_AG.CY.09.CY256             .....................................................................................................................................................................   8128
02_AG.ES.06.P1423             .....................................................................................................................................................................   8983
02_AG.GW.05.CC_0048           .....................................................................................................................................................................   8765
02_AG.KR.12.12MHR9            .....................................................................................................................................................................   9166
02_AG.LR.x.POC44951           -----------------A---G-AGA--....---------------.------------------.----C---.........................................................................................   9633
02_AG.NG.09.09NG_SC61         ---------------------G-A----....-------------..........................................................................................................................   8807
02_AG.NG.x.IBNG               ---------------------G-GGA--....---------------.------------------.-------------.------------------A-----.---------------............................................   9201
02_AG.SE.94.SE7812            -----------------A---G-AGA--....---------------.------------------.-----G--GCAT..-C..................................................................................   9063
02_AG.SN.98.98SE_MP1211       ---------------------G-AGA--....---------------.------------------.-.................................................................................................   8851
03_AB.RU.97.KAL153_2          ---------------------G---A--....---------------.------------------.--------..........................................................................................   8808
04_cpx.CY.94.94CY032_3        ---------------------G------....---------------.------------------.----------........................................................................................   9050
05_DF.BE.x.VI1310             ---------------------G------....---------------.G-----------------.-----G--GCAT..-C..................................................................................   9083
06_cpx.AU.96.BFP90            -----------------A---G-AG---....---------------.------------------.----------AT-.------------------.-----.-----------------------------AC-C---A------.A--------------   9775
07_BC.CN.98.98CN009           ----------------------------....---------------.------------------.-----------.......................................................................................   8978
08_BC.CN.97.97CNGX_6F         ---------------------G------....------------...........................................................................................................................   8802
09_cpx.GH.96.96GH2911         ---------------------G-A----....---------------.------------------.--------C-A--.------------------.-----.-----------------------------AC-C---A-GAG--.A--------------   8890
10_CD.TZ.96.96TZ_BF061        -----C---------------T--G---....-----------C---.----------------.......................................................................................................   8976
11_cpx.CM.95.95CM_1816        ---------------------G------....-----------............................................................................................................................   9017
12_BF.AR.99.ARMA159           ---------------------G------....---------------.------------------.-------------.-------------A-T--.-----.-----------------------------A--C--------------------------   9704
13_cpx.CM.96.96CM_1849        .....................................................................................................................................................................   8916
14_BG.ES.05.X1870             .....................................................................................................................................................................   9028
15_01B.TH.99.99TH_MU2079      -----C---------------G-A----....---------------A-G-------------TG-...................................................................................................   8999
16_A2D.KR.97.97KR004          ---------------------G-AG---....---------------A------------------.---------G-.......................................................................................   9064
17_BF.AR.99.ARMA038           ---------------------G------....---------------........................................................................................................................   8808
18_cpx.CU.99.CU76             ---------------------G------....R------YK-Y-W--.R-T---TA-............................................................................................................   8945
19_cpx.CU.99.CU7              .....................................................................................................................................................................   8512
20_BG.CU.99.Cu103             .....................................................................................................................................................................   8935
21_A2D.KE.99.KER2003          -------A-------------G------....------------...........................................................................................................................   8816
22_01A1.CM.01.01CM_0001BBY    .....................................................................................................................................................................   8372
23_BG.CU.03.CB118             .....................................................................................................................................................................   8964
24_BG.ES.08.X2456_2           .....................................................................................................................................................................   8925
25_cpx.CM.02.1918LE           .....................................................................................................................................................................   8425
26_AU.CD.02.02CD_MBTB047      -----------------A---G-A----....---------------.------------------.-------------.--G---------------.-----.-----------------------------AC-C---GT-----.A--------------   9730
27_cpx.FR.04.04CD_FR_KZS      ---------------------G------....---------------.------------------.-------------.------------------.-----.-----------------------------A-----------------------------   9745
28_BF.BR.99.BREPM12609        .....................................................................................................................................................................   8886
29_BF.BR.01.BREPM16704        .....................................................................................................................................................................   8979
31_BC.BR.04.04BR142           .....................................................................................................................................................................   9057
32_06A1.EE.01.EE0369          -----------------A---G-AG-A-....------------...........................................................................................................................   9229
33_01B.ID.07.JKT189_C         -----C---------------G-AG---....---------------A------------------.---.....................................................................................................   8936
34_01B.TH.99.OUR2478P         .....................................................................................................................................................................   8650
35_AD.AF.07.169H              .....................................................................................................................................................................   8698
36_cpx.CM.00.00CMNYU830       .....................................................................................................................................................................   8410
37_cpx.CM.00.00CMNYU926       .....................................................................................................................................................................   8396
38_BF1.UY.03.UY03_3389        .....................................................................................................................................................................   8935
39_BF.BR.04.04BRRJ179         ---------------......................................................................................................................................................   9057
40_BF.BR.05.05BRRJ055         .....................................................................................................................................................................   9042
42_BF.LU.06.luBF_18_06        ---------------------G------....---------------.------------------.--R--.............................................................................................   9122
43_02G.SA.03.J11223           .....................................................................................................................................................................   9085
44_BF.CL.00.CH80              .....................................................................................................................................................................   8858
45_cpx.FR.04.04FR_AUK         ---------------------T------....------T-A------.------------------.-------------.------------------.-----.-----------------------------AC-C---G-G----.A--------------   9680
46_BF.BR.07.07BR_FPS625       -----C---------------G------....---------------.------------------.-------------.-------------A----.-----.----------G----------------A-A--------C-G------------------   9113
47_BF.ES.08.P1942             .....................................................................................................................................................................   8926
48_01B.MY.07.07MYKT021        -----C---------------G-A----....---------------.------------------.----------A--.--G---------------..................................................................   8884
49_cpx.GM.03.N26677           .....................................................................................................................................................................   8519
50_A1D.GB.10.12792            -----C---------......................................................................................................................................................   8978
51_01B.SG.11.11SG_HM021       ---------------------G------....------------...........................................................................................................................   8807
52_01B.MY.03.03MYKL018_1      -----C---------------G-AG---....------------...........................................................................................................................   8925
53_01B.MY.11.11FIR164         -----C---------------G-A----....---------------A-----------..........................................................................................................   8954
54_01B.MY.09.09MYSB023        -----C---------------G-AG---....---------------.----------------.....................................................................................................   9069
55_01B.CN.10.HNCS102056       -----C---------------G-A----....--------.......................................................................................................................   8925
56_cpx.FR.10.URF5_patient_A   .....................................................................................................................................................................   8396
57_BC.CN.09.09YNLX19sg        --------------------GT------....----...........................................................................................................................   8813
58_01B.MY.09.09MYPR37         -----C---------------G-AR---....---------------A.....................................................................................................................   8975
59_01B.CN.09.09LNA423         .....................................................................................................................................................................   8380
60_BC.IT.11.BAV499            ---------------------T------....---------------.--------------AAT-.----------T--.------------------.-----.----------------------------A-C---T----T-------------------   9684
61_BC.CN.10.JL100010          --------------------G-------....---------------.------------------.---...............................................................................................   9024
62_BC.CN.10.YNFL13            ---------------------T------....---------------.AG-CTC-A-............................................................................................................   8969
63_02A1.RU.10.10RU6637        ---------------------G------....---------------.------------------.----------T-..................................................................................... .   9187
64_BC.CN.09.YNFL31            ---------------------T------....----------.....................................................................................................................   8969
65_cpx.CN.10.YNFL01           ---------------------CT-----....---------------.AG-CTC-A.............................................................................................................   9038
67_01B.CN.11.ANHUI_HF115      -----C---------------G-A----....A-C---.........................................................................................................................   8918
68_01B.CN.11.ANHUI_WH73       -----C---------------G-A.............................................................................................................................................   8903
70_BF1.BR.10.10BR_PE004       .....................................................................................................................................................................   8524
71_BF1.BR.10.10BR_PE008       .....................................................................................................................................................................   9198
72_BF1.BR.10.10BR_MG002       .....................................................................................................................................................................   9062
O.BE.87.ANT70                 -----C--------C-----CTCTAGCT....------G--CG----.CG----------------.-------AG.........................................................................................   9753
O.CM.98.98CMA104              -----C--------C---C-CTCTAGCT....------G--------.CG...................................................................................................................   9243
O.CM.98.98CMABB141            .....................................................................................................................................................................   9182
O.CM.98.98CMABB212            .....................................................................................................................................................................   9177
O.CM.98.98CMU5337             -----C--------C-----CTCTAGCT....------G--......................................................................................................................   9171
O.CM.99.99CMU4122             -----C--------C-----CTCTAGCT....------.........................................................................................................................   9186
O.FR.92.VAU                   -----C--------C-----CTCTAGCT....------G--------.CG----------------.-------AGAA...-C-.................................................................................   9330
O.GA.11.11Gab6352             -----C---C----C-----CTCTAGCT....------G--------.CG----------------.-------A-AA...-C--------T-A-----.--C...A--C-----GT---------------T-ACG-AGACTGAGTA-A--------------   9072
O.SN.99.99SE_MP1299           -----C--------C-----CTCTAGCT....------G--------.CG----------------.-------AGAA...-C--------T-A-----.--....C---C-----T--------.......................................   9858
O.US.10.LTNP                  .....................................................................................................................................................................   9103
N.CM.02.DJO0131               .....................................................................................................................................................................   8938
N.CM.02.SJGddd                .....................................................................................................................................................................   8615
N.CM.04.04CM_1015_04          .....................................................................................................................................................................   8926
N.CM.06.U14296                .....................................................................................................................................................................   8545
N.CM.06.U14842                .....................................................................................................................................................................   8637
N.FR.11.N1_FR_2011            .....................................................................................................................................................................   8892
P.CM.06.U14788                ------------------TCT-.AGCA.....----------G----AAA...................................................................................................................   9238
P.FR.09.RBF168                -------------------CTTAAGCA.....-------------GGAAG----------------.-------AG-A--.-C--------T-A-----.--CA-..A-------GT----------------G-AC--CTAG-C-A------------------   9868
CPZ.CD.06.BF1167              --A--------T------TAGT-AG--C.AT--GA------------.CG----------------TGA-----T-ATC-....----------A---CA--CC-CGCCGG-CCTGGGG---------------T-TC-GA--G-TAA-C.A-----------C-G   9916
CPZ.CM.05.SIVcpzMT145         -----------------------A----....---------------.------------------.---------G--......................................................................................   9208
CPZ.GA.88.GAB1                -----------------.A-C-GTG-A-....---------------.------------------.-------TA-TTGA-C----------TA---A.--CA-.AC-------------------------TTAAA---TAGTCAAG.---------------   9811
CPZ.TZ.06.SIVcpzTAN13         ------------------TAGTGACT-GTAA--GA--G---------.CG----------------TGA-----T-.........................................................................................   9322
CPZ.US.85.US_Marilyn          -----C---------------GTGTA......---------------.------------------.----------T--.-------------TA--C.--CA-.---------------------------A-TC--A-AAGTAG--.---------------   9781
GOR.CM.07.SIVgor2139_287      ------------------TCTT.AGCA.....---------------.CG----------------A-------AA-T.......................................................................................   9252
GOR.CM.07.SIVgorCP2135con     ------------------TCTT.AGCA.....---------------.CG----------------A-------AA-T.......................................................................................   9246
```

III

HIV-2/SIV Complete Genomes

Contents

III-1	Introduction	**153**
III-2	Annotated features	**154**
III-3	Sequences	**156**
III-4	Alignments	**158**

III-1 Introduction

The HIV-2/SIVsmm alignment contains HIV-2 genomes, plus examples from related species SIVmne, SIVstm, and SIVmac. Compared to HIV-1, fewer HIV-2 genomes have been sequenced, so we are able to include all available HIV-2 genomes, removing only problematic sequences and multiples from the same patient. This year, we have been more selective in removing redundant SIVsmm sequences from the alignment, so it has fewer total sequences than in previous years.

The HIV-2/SIVsmm family is presented together in spite of their different hosts, because their genomic structure is the same—the complete SIV alignment (presented in the next chapter along with a few HIV-1 sequences) is much more stretched out because of the presence of a *vpx* and absence of *vpu* genes in some of these viruses.

The Mac239 sequence (accession M33262) is the master sequence in this alignment. The alignment was generated by a HMMER model, then subsequently codon-aligned using GeneCutter and manual edits to fix obvious misalignments. The alignment presented cannot be considered an "optimal alignment" to any single criterion; it is a compromise between optimal alignment, readability, and codon alignment. Most gaps have been introduced in multiples of 3 bases to maintain open reading frames when the alignment is translated.

III-2 Annotated features

Features of HIV-2 annotated in the alignment that follows.

Feature	Location	Page
5' LTR U3 start	1	158
TCF-1 alpha	309-324	159
NF-κ-B-II	415-424	160
NF-κ-B-I	441-449	161
TATA Box	488-493	161
TAR element start	518	161
5' LTR U3 end	520	161
+1 mRNA start site	521	161
5' LTR R repeat begin	521	161
TAR element end	642	162
Poly-A signal	671-676	162
5' LTR R repeat end	694	162
5' LTR U5 start	695	162
5' LTR U5 end	818	163
Lys tRNA primer binding site	821-839	163
Gag and Gag-Pol start	1053	165
Gag p17 Matrix end	1457	167
Gag p24 Capsid start	1458	167
Gag p24 Capsid end	2144	171
Gag p2 Spacer start	2145	171
Gag p2 Spacer end	2195	171
Gag p8 Nucleocapsid start	2196	171
Gag p8 Nucleocapsid end	2351	172
Gag p1 Spacer start	2352	172
Gag-Pol -1 ribosomal slip site	2354	172
Pol start	2354	172
Gag p1 spacer end	2393	173
Gag p6 start	2394	173
Pol Protease start	2555	174
Gag p6 end	2585	174
Gag end	2585	174
Pol Protease end	2851	176
Pol p66 and p51 RT start	2852	176
Pol RT end	4168	184
Pol p15 RNAse H start	4169	184
Pol RNAse H end	4528	186
Pol p31 Integrase start	4529	186
Vif start	5340	191
Pol, Gag-Pol, and p31 Integrase end	5410	191
Vpx start	5812	194
Vif end	5984	195
Vpx end	6150	196
Vpr start	6151	196
Tat exon 1 start	6302	197
Vpr end	6456	197
Rev exon 1 start	6528	198
Tat Rev exon 1 end	6597	198
Tat Rev intron start	6598	198

Feature	Location	Page
Env start	6604	198
Env signal peptide end	6669	199
Env gp120 start	6670	199
V1 loop start	6940	200
V1 loop end	7104	202
V2 loop start	7108	202
V2 loop end	7236	202
V3 loop start	7534	204
V3 loop end	7635	205
V4 loop start	7807	206
V4 loop end	7899	207
V5 start	8017	207
V5 end	8034	207
Rev Responsive Element (RRE) start	8123	208
Env gp120 end	8178	208
Env gp41 start	8179	208
RRE end	8479	210
Tat Rev intron end	8805	212
Tat Rev exon 2 start	8806	212
Tat end	8902	213
Rev end	9059	214
Nef start	9077	214
Env gp41, gp160 end	9243	215
Premature stop in SMM239	9355	215
3' LTR U3 start	9462	216
TCF-1 alpha binding	9770-9785	218
Nef end	9868	218
NF-κ-B-II	9876-9884	219
NF-κ-B-I	9891-9900	219
TATA box	9949-9955	219
TAR element start	9980	220
3' LTR U3 end	9978	220
3' LTR repeat start	9979	220
TAR element end	10103	220
Poly-A signal	10132-10137	220
3' LTR R repeat end	10155	221
3' LTR U5 start	10156	221
3' LTR U5 end	10279	221

III-3 Sequences

Sequences included in the HIV-2/SIV complete genome alignment.

Name	Accession	Country	Author	Reference
MAC.US.x.239	M33262	United States	Kestler, H.	*Science* **248**(4959):1109-1112 (1990)
A.CI.88.UC2	U38293	Cote D'Ivoire	Barnett, S.W.	*Virology* **222**(1); 257-61 (1996)
A.DE.x.BEN	M30502	Germany	Kirchhoff, F.	*Virology* **177**(1):305-311 (1990)
A.DE.x.PEI2_KR_KRCG	U22047	Germany	Kraus, G.	*ARHR* **14**(1); 65-77 (1998)
A.GH.x.GH1	M30895	Ghana	Hasegawa, A.	*ARHR* **5**(6):593-604 (1989)
A.GM.87.D194	J04542	Gambia	Kuehnel, H.	*Proc Natl Acad Sci USA* **86**(7):2383-2387 (1989)
A.GM.x.ISY_SBL_6669_85	J04498	Gambia	Franchini, G.	*Proc Natl Acad Sci USA* **86**(7):2433-2437 (1989)
A.GM.x.MCN13	AY509259	Gambia	Schmitz, C.	*J Virol* **78**(4):2006-2016 (2004)
A.GM.x.MCR35	AY509260	Gambia	Schmitz, C.	*J Virol* **78**(4):2006-2016 (2004)
A.GW.86.FG_clone_NIHZ	J03654	Guinea-Bissau	Zagury, J.F.	*Proc Natl Acad Sci USA* **85**(16):5941-5945 (1988)
A.GW.87.CAM2CG	D00835	Guinea-Bissau	Tristem, M.	*J Gen Virol* **72**(PT 3):721-724 (1991)
A.GW.x.MDS	Z48731	Guinea-Bissau	Becker, M.	Unpublished
A.IN.07.NNVA	EU980602	India	Gurjar, S.R.	*JAIDS* **52**(3); 329-35 (2009)
A.IN.95.CRIK_147	DQ307022	India	Santhosh, C.V.	*ARHR* **24**(10); 1315-7 (2008)
A.JP.08.NMC786_clone_41	AB731742	Japan	Umeki-Sakamoto, Y.	Unpublished
A.PT.x.ALI	AF082339	Portugal	Reeves, J.D.	*J Virol* **73**(9); 7795-804 (1999)
A.SN.85.ROD	M15390	Senegal	Clavel, F.	*Nature* **324**(6098):691-695 (1986)
A.SN.86.ST_JSP4_27	M31113	Senegal	Kumar, P.	*J Virol* **64**(2):890-901 (1990)
B.CI.88.UC1	L07625	Cote D'Ivoire	Barnett, S.W.	*J Virol* **67**(2):1006-1014 (1993)
B.CI.x.20_56	AB485670	Cote D'Ivoire	Takekawa, N.	Unpublished
B.CI.x.EHO	U27200	Cote D'Ivoire	Rey-Cuille, M.A.	*Virology* **202**(1):471-476 (1994)
B.GH.86.D205_ALT	X61240	Ghana	Dietrich, U.	*Nature* **342**(6252):948-950 (1989)
B.JP.01.IMCJ_KR020_1	AB100245	Japan	Kusagawa, S.	*ARHR* **19**(11):1045-1049 (2003)
G.CI.92.Abt96	AF208027	Cote D'Ivoire	Brennan, C.A.	*ARHR* **13**(5):401-404 (1997)
AB.CM.03.03CM_510_03	EU028345	Cameroon	Yamaguchi, J.	*ARHR* **24**(1):86-91 (2008)
H2_01_AB.CI.90.7312A	L36874	Cote D'Ivoire	Gao, F.	Unpublished
H2_01_AB.JP.04.NMC307_20	AB731738	Japan	Umeki-Sakamoto, Y.	Unpublished
H2_01_AB.JP.07.NMC716_01	AB731740	Japan	Umeki-Sakamoto, Y.	Unpublished
H2_01_AB.JP.08.NMC842_10	AB731744	Japan	Umeki-Sakamoto, Y.	Unpublished
U.CI.07.07IC_TNP3	KC693505	Cote D'Ivoire	Ayouba, A.	*AIDS* **27**(15):2488-2491 (2013)
U.FR.96.12034	AY530889	France	Damond, F.	*ARHR* **20**(6):666-672 (2004)

Name	Accession	Country	Author	Reference
MAC.US.x.17EC1	AY033233	United States	Anderson, M.G.	*Virology* **195**(2):616-626 (1993)
MAC.US.x.251_1A11	M76764	United States	Marthas, M.L.	*J Med Primatol* **18**(3-4):311-9 (1989)
MAC.US.x.251_32H_PJ5	D01065	United States	Rud, E.W.	*(in) Brown,F, Chanock,RM and Ginsberg,HS(Eds); VACCINES* **92**: MODERN APPROACHES TO NEW VACCINES INCLUDING PREVENTION OF AIDS: 229-235; Cold Spring Harbor Laboratory Press, New York (1992)
MAC.US.x.251_BK28	M19499	United States	Hirsch, V.	*Cell* **49**(3):307-319 (1987)
MAC.US.x.MM142_IVMXX	Y00277	United States	Chakrabarti, L.	*Nature* **328**(6130):543-547 (1987)
MNE.US.82.MNE_8	M32741	United States	Kimata, J.T.	*J Virol* **72**(1):245-256 (1998)
MNE.US.x.MNE027	U79412	United States	Kimata, J.T.	*J Virol* **72**(1):245-256 (1998)
SMM.CI.79.SIVsmCI2	JX860430	Cote D'Ivoire	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.LR.89.SIVsmLIB1	JX860431	Liberia	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.SL.92.SIVsmSL92A	JX860432	Sierra Leone	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.SL.92.SL92B	AF334679	Sierra Leone	Chen, Z.	*J Virol* **70**(6):3617-3627 (1996)
SMM.US.04.G078	JX860415	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.G932	JX860416	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M919	JX860417	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M922	JX860418	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M923	JX860419	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M926	JX860420	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M934	JX860421	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M935	JX860422	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M940	JX860423	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M946	JX860424	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M947	JX860425	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M949	JX860426	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M950	JX860427	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M951	JX860428	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M952	JX860429	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.05.D215	JX860413	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.06.FTq	JX860414	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.11.SIVsmE660_FL10	JQ864084	United States	Wu, F.	*J Virol* **86**(16); 8835-47 (2012)
SMM.US.86.CFU212	JX860407	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.x.F236_H4	X14307	United States	Hirsch, V.M.	*Nature* **339**(6223); 389-92 (1989)
SMM.US.x.H9	M80194	United States	Courgnaud, V.	*J Virol* **66**(1):414-419 (1992)
SMM.US.x.PBJA	M31325	United States	Dewhurst, S.	*Nature* **345**(6276):636-640 (1990)
SMM.US.x.PGM53	AF077017	United States	Novembre, F.J.	*J Virol* **72**(11):8841-8851 (1998)
SMM.US.x.SME543	U72748	United States	Hirsch, V.	*J Virol* **71**(2):1608-1620 (1997)
SMM.US.x.pE660.CG7G	JX648292	United States	Lopker, M.J.	Unpublished
STM.US.89.STM_37_16	M83293	United States	Novembre, F.J.	*Virology* **186**(2):783-787 (1992)

5' LTR U3 start

```
MAC.US.x.239          TGGAAGGGATTTATTACAGTGCAAGAAG.ACATAGAATCTTAGACATATACTTAGAAAAGGAAGAAGGCATCATACCAGATTGGCAGGATTACACCTCAGGACCAGGAATTAGATACCCAAAGACATTTGGCTGGCTATGGAAATTAGTCCCTGTAAATGTATCAGATGA 169
A.CI.88.UC2           ----------G-T-------AGGGAT--.----------C-----T-G---C----------G-----GG-A--------------A----T--TCAT--G------G-G--G-------T-TGC--C--G-----G------C----A--A-----CA-G---C-A--  169
A.DE.x.BEN            --C-------G-T-------AGG--G--.----------C----------C-------A--G-----G--A--------------A----T--TCAT--G------G-A--G-------T-TAC--C--G-----G-----GC----AT-A---G-AC-C---C-A--  169
A.DE.x.PEI2_KR_KRCG   ----T-----G----------AG-----G---------------------A-G-------------G--A--T----------A-C--T--TCAT--G------G-A--G--------TTC-----G----------GC----A--A---G-C--CC--C-A-G  170
A.GH.x.GH1            ..........................................................................................................................................  0
A.GM.87.D194          ..........................................................................................................................................  0
A.GM.x.ISY_SBL_6669_85 ..........................................................................................................................................  0
A.GM.x.MCN13          ..........................................................................................................................................  0
A.GM.x.MCR35          ..........................................................................................................................................  0
A.GW.86.FG_clone_NIHZ ..........................................................................................................................................  0
A.GW.87.CAM2CG        --------C-G-T--------A-----.----------C-G--T-------------A-------G--A--TG----------A-C--T--TAGT--G------G-A---------T-TTC-----G----------GC----A--A---G--ACC---C-A--  169
A.GW.x.MDS            ..........................................................................................................................................  0
A.IN.07.NNVA          ---------G----------AG-----.--GAGA----A----TT-----------------G--G--A--T----C------A-C--T---AGT--G-------A-----T----T-TTCCA---G----------GC----A--A--------CC--C-A--  169
A.IN.95.CRIK_147      ---------G-T--------A-----------------------------------------------------------------------------------------------------------  24
A.JP.08.NMC786_clone_41 ---------G-T--------AG---.----------C-----T-G---C------A--G-----G--A----------A----T--TCAT--G------G--A--G--------TGC--C--G-----G-----GC----A--A---G-CA-CA--C-A--  168
A.PT.x.ALI            ---------G-T--------AG-----.-------------------------------G--A--T----------A-C--T--TCAT--G-------A--G-----G-T-TTC-----G----G-----GC----A--A---G----CC--C-A--  169
A.SN.85.ROD           ..........................................................................................................................................  0
A.SN.86.ST_JSP4_27    ..........................................................................................................................................  0
B.CI.88.UC1           ----------C--T---AG-----.---------AC-----C---TC---G--T--G-------TG--T-T-GA----AA-C--T--A-AT--G-----G--A--G--T---G---C----G-----G-----GC----G--A---G--A--C----G--  169
B.CI.x.20_56          .----------C--T---AG-----.---------AC----T-C---TC---G--T--G------G-TG--T-T-GA----AA-C--T--ACAT--G------G-A--G--T--------C-----G-----G-----GC----G--A------A--C----G--  168
B.CI.x.EHO            ----------TC--T---AG-----.G----A---AC------C----------T----------TG-GT-T-GA-----A-C-----ACAT--G-----G-A-----T--C---TTC---------------GC----G--AA-----A-GAT--CA--  169
B.GH.86.D205_ALT      ---------G-----T---AG-----.---C-----A-------C---T--T--G--T----------TG-GT-T-GC-----AA-C--T--TCAT--G-----G--A--GC-T--C--ATAC-----T-----G-----GC-G--A--A---G-G--GC---CA-C  169
B.JP.01.IMCJ_KR020_1  ..........................................................................................................................................  0
G.CI.92.Abt96         ..........................................................................................................................................  0
AB.CM.03.03CM_510_03  ..........................................................................................................................................  0
H2_01_AB.CI.90.7312A  ----------T---T----A------.------G--A-----T-C---T--T--G--T--GAG---T--AG-TGGT-GA------A-C-----A-AT--G-----G--A----------ACAC----------G------C----G--A--GG-A--GG---CA-T  169
H2_01_AB.JP.04.NMC307_20 ----------T---T----AG-----.------G--A-----T-C------T--G--T--G-----T--AG-CT-T-GA-----AA-C-----ACAT--G-----GG-A-----T-----ATAT----------G------C----A--AA--G-A--GG---CA-C  169
H2_01_AB.JP.07.NMC716_01 ----------T---T----AG-----.---------A-----T-C------T--G--T--G-----T--AG-CT-T-GA-----A-C-----ACAT--G-----GG-A-----T-----TAT----------G------C----A--AA--G-A--GG---CA-C  169
H2_01_AB.JP.08.NMC842_10 ----------T---T----AG-----.------G--A-----T-C------T--G--T--G-----T--AG-CT-T-GA-----A-C-----ACAT--G-----GG-A-----T-----TTC----------G------C----A--A---G-A--GG---CA-C  169
U.CI.07.07IC_TNP3     ..........................................................................................................................................  0
U.FR.96.12034         ..........................................................................................................................................  0
MAC.US.x.17EC1        ----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------  169
MAC.US.x.251_1A11     ----------T----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------  169
MAC.US.x.251_32H_PJ5  ----------.----------.----------G----------G----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------C----------.----------  169
MAC.US.x.251_BK28     ----------.----------.----------G----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------.----------  169
MAC.US.x.MM142_IVMXX  ..........................................................................................................................................  0
MNE.US.82.MNE_8       ..........................................................................................................................................  0
MNE.US.x.MNE027       ..........................................................................................................................................  0
SMM.CI.79.SIVsmCI2    ..........................................................................................................................................  0
SMM.LR.89.SIVsmLIB1   ..........................................................................................................................................  0
SMM.SL.92.SIVsmSL92A  ..........................................................................................................................................  0
SMM.SL.92.SL92B       ..........................................................................................................................................  0
SMM.US.04.G078        ..........................................................................................................................................  0
SMM.US.04.G932        ..........................................................................................................................................  0
SMM.US.04.M919        ..........................................................................................................................................  0
SMM.US.04.M922        ..........................................................................................................................................  0
SMM.US.04.M923        ..........................................................................................................................................  0
SMM.US.04.M926        ..........................................................................................................................................  0
SMM.US.04.M934        ..........................................................................................................................................  0
SMM.US.04.M935        ..........................................................................................................................................  0
SMM.US.04.M940        ..........................................................................................................................................  0
SMM.US.04.M946        ..........................................................................................................................................  0
SMM.US.04.M947        ..........................................................................................................................................  0
SMM.US.04.M949        ..........................................................................................................................................  0
SMM.US.04.M950        ..........................................................................................................................................  0
SMM.US.04.M951        ..........................................................................................................................................  0
SMM.US.04.M952        ..........................................................................................................................................  0
SMM.US.05.D215        ..........................................................................................................................................  0
SMM.US.06.FTq         ..........................................................................................................................................  0
SMM.US.11.SIVsmE660_FL10 ----------------A--AG-A---.---------A-----T--G---A-G-------------A--A-------------AA------A-----G-------C--------T-T-TAC-A---G-----C-------------A---G----C--------  169
SMM.US.86.CFU212      ..........................................................................................................................................  0
SMM.US.x.F236_H4      ----------------A--AG-----.---------A-----T--G-----G-------------A--A-------------AA------A-----G---------------T-T-CAC-A---G-----C-------------A---G----C--------  169
SMM.US.x.H9           ..........................................................................................................................................  0
SMM.US.x.PBJA         ..........................................................................................................................................  0
SMM.US.x.PGM53        ..........................................................................----------AA------A-----G-------C--------T-T-TTC-------------------G----A---G---C--------  95
SMM.US.x.SME543       ----------------A--AG-A---.---------A-----T--G---A-G-------------A--A-------------AA------A-T---G-------C--------T-T-TAC-----G-----C-------------A---G----C--------  169
SMM.US.x.pE660.CG7G   ----------------A--AG-A---.---------A-----T--G---A-G-------------A--A-------------AA------A-----G-------C--------T-T-TAC-A---G-----C-------------A---G----C--------  169
STM.US.89.STM_37_16   --------------------AG-----.G-------------------------------------------------------------------------------------------------  30
```

TCF-1 alpha

```
MAC.US.x.239                GGCACAG............GAGGATGAGGAGCATTATTTAATGCATCCAGCTCAAACTTCCCAGTGGGATGACCCTTGGGGAGAGGTTCTAGCATGGAAGTTTGATCCAACTCTGGCCTACACTTATGAGGCATATGTTAGATACCCAGAAGAGTTTGGAAGCAAGTCAG 327
A.CI.88.UC2                 ----G--.............T------ACCA-C-GC--------C-----A-----AAG-AGA-AT-----ATACAT--G---ACA----TG---CG----A-CT-C-TG--------TGAG--CA----C-TCAC-CT-----------A-----GCA-------- 327
A.DE.x.BEN                  ----G--.............A-------CCA-C-GC---G-A--C-----A-----AAG-AGACAT-----TGAGCAT--G---ACAT---TG---C-------CT-C-TG--------T-AC--CA----C-TCAC-CTG----------------GCA-------- 327
A.DE.x.PEI2_KR_KRCG         T.....................--A-----C--C-GC---C-A--C-----A-----AAG-GG--CT---------CAT--G--AACAT--ATG----G------C--T-GG--------TGAG---AC---T-T-AA-C------------A-----GTAT------- 319
A.GH.x.GH1                  ..................................................................................................................................... 0
A.GM.87.D194                ..................................................................................................................................... 0
A.GM.x.ISY_SBL_6669_85      ..................................................................................................................................... 0
A.GM.x.MCN13                ..................................................................................................................................... 0
A.GM.x.MCR35                ..................................................................................................................................... 0
A.GW.86.FG_clone_NIHZ       ..................................................................................................................................... 0
A.GW.87.CAM2CG              --G-G--GACACTGAGACT--CAC----ACT--C-GC---T-A--C-----A-----AAG-AG-CAT------ATGCAT--G---ACA--G-TC----------CT-C-TG------CTT-AG-------C-T-AC-C------------A-----GCA-------- 339
A.GW.x.MDS                  ..................................................................................................................................... 0
A.IN.07.NNVA                --G-G-A.............CAC---AACT--C-GC--------C-----A-----AAGTA---TT------AGACAT--G---ACA----T-----GA--------C--AT-A-----TGA-----TA--C-T-AGGCT-C------G--A-----GCA-------- 327
A.IN.95.CRIK_147            ..................................................................................GA--CAT---C-TCAGATT-----T---G--A------CA-------- 73
A.JP.08.NMC786_clone_41     ----G--.............C------ACCA-C-GCC---CA--C-----A-----AAG-AGA-AT-----TGAGCAT--G--AACA--GATG----------CT-T-TG--A-----T-AC--CAT---C-TCA---AGC-T------------GTA-C------ 326
A.PT.x.ALI                  A-GGG--.............CAC----ACT--C-GCC-GC-A--C----TA-----AAG-AG-CAT------A-GCAT--G---ACAT---TT----GA-----C--T-AG-----TC-TGA---CA-A--C-T-A--CT-C-------G--A-----GTA-------- 327
A.SN.85.ROD                 ..................................................................................................................................... 0
A.SN.86.ST_JSP4_27          ..................................................................................................................................... 0
B.CI.88.UC1                 --A-AG-.............GA-CG--AACCAG--G-C-GG----C-----A--G-TC--TTCA---------ATCCAC--G---ACC--T--C---CG------C--CCTC-----AC-TGAC---TA--C-TCAAC--G------------------GTATC------ 327
B.CI.x.20_56                A-A-G--.............AGA-CA--AA-CAG--G-C-GG----C-----A--G--C---TCA----------CCAC--G--AACC--T-TC---C-------C--CCTC-----AC-TGAC---ATA--C-TCAAC--G-T----------------GTATC------ 326
B.CI.x.EHO                  AC--G--.............C--G--AACC----G-C-GG----------A--G--C---GCA----------CCAC-A----ACC--T-TC---C-------CT-CCTC--A--A--TGAC----T---C-TCAGC--G-T------G-------GTATC------ 327
B.GH.86.D205_ALT            -A-C-GAGAG.............G--AACC----GCC-------C--G--A--G-TC---TCA--------ATCCAT--G---AC---TATC---C--------T-CCTC-----A--TGA-----T---T-TCAA---G-TT----------C--GTA-C------ 330
B.JP.01.IMCJ_KR020_1        ..................................................................................................................................... 0
G.CI.92.Abt96               ..................................................................................................................................... 0
AB.CM.03.03CM_510_03        ..................................................................................................................................... 0
H2_01_AB.CI.90.7312A        AA-C-G-GAG..........--A--G---ACC----G-C--G----C-----A--G--CG--GCA--------.T-CCAT--G---AC---T-TC---C--------T-CCTC-----A--TT-AG-------C-TCAAC--G-T----------------GTATC------ 329
H2_01_AB.JP.04.NMC307_20    AA-C---GAG.............G---ACC----G-C--G----C-----A--G--C--TGCA----------CCAT--G---AC---T-TC---C----C-----CTTC-----A--TT-A--------C-TCAAC-AG-T----------------GTATC------ 330
H2_01_AB.JP.07.NMC716_01    AA-C--AGAG.............G---ACA----G-C--G----C-----A--G--C---GCA----------CCAT--G---AC---T-TC---C----C-----CTTC-----A--TT-A--------C-TCAAC-AG-T----------------GTATC------ 330
H2_01_AB.JP.08.NMC842_10    AA-C---GAG.............G---ACC----G-C--G----C-----A--G--C--TTCA----------CCAT--G---AC---T-TC---C--------T-CCTC-----G--TT-A--------C-TCAAC--G----------------GTATC------ 330
U.CI.07.07IC_TNP3           ..................................................................................................................................... 0
U.FR.96.12034               ..................................................................................................................................... 0
MAC.US.x.17EC1              ------.....------------------------------------------------------------------------------------------------------------------------ 327
MAC.US.x.251_1A11           ------.....------------------------------------------------------------C-------------C-------------------------------------------- 327
MAC.US.x.251_32H_PJ5        ------.....---------AG--------------G--------------A-------------------------------------A-------------------------------------- 327
MAC.US.x.251_BK28           ------.....---------AG--------------G--------------A-------------------------G-------------A----------------------C-------------. 326
MAC.US.x.MM142_IVMXX        ..................................................................................................................................... 0
MNE.US.82.MNE_8             ..................................................................................................................................... 0
MNE.US.x.MNE027             ..................................................................................................................................... 0
SMM.CI.79.SIVsmCI2          ..................................................................................................................................... 0
SMM.LR.89.SIVsmLIB1         ..................................................................................................................................... 0
SMM.SL.92.SIVsmSL92A        ..................................................................................................................................... 0
SMM.SL.92.SL92B             ..................................................................................................................................... 0
SMM.US.04.G078              ..................................................................................................................................... 0
SMM.US.04.G932              ..................................................................................................................................... 0
SMM.US.04.M919              ..................................................................................................................................... 0
SMM.US.04.M922              ..................................................................................................................................... 0
SMM.US.04.M923              ..................................................................................................................................... 0
SMM.US.04.M926              ..................................................................................................................................... 0
SMM.US.04.M934              ..................................................................................................................................... 0
SMM.US.04.M935              ..................................................................................................................................... 0
SMM.US.04.M940              ..................................................................................................................................... 0
SMM.US.04.M946              ..................................................................................................................................... 0
SMM.US.04.M947              ..................................................................................................................................... 0
SMM.US.04.M949              ..................................................................................................................................... 0
SMM.US.04.M950              ..................................................................................................................................... 0
SMM.US.04.M951              ..................................................................................................................................... 0
SMM.US.04.M952              ..................................................................................................................................... 0
SMM.US.05.D215              ..................................................................................................................................... 0
SMM.US.06.FTq               ..................................................................................................................................... 0
SMM.US.11.SIVsmE660_FL10    A--T---.............A--C---ACA----GCC-G-------G--A--G---CAT-------------C----------A--G----------------GAAT-A--T--T-GC---A------T-A---AG----------------T--T------- 327
SMM.US.86.CFU212            ..................................................................................................................................... 0
SMM.US.x.F236_H4            A--T---.............A--C---ACA----GCC-GG------G--A--G----AT-------------C----------A--G----------------GAAT-A--T--T-GC---A------T-A---AG----------------T--T------- 327
SMM.US.x.H9                 ..................................................................................................................................... 0
SMM.US.x.PBJA               ..........................................................................G----------------AGT-A--T--T-AC---A------T----GAGC----------------T--TC------ 82
SMM.US.x.PGM53              A--T--A.............A--C---ACA-----C---G-----------G-TA-------------T--C----------G--------------T--CAAT-A------GA--------T-T-A-----C----------------T--T------- 253
SMM.US.x.SME543             A--C---.............A--C---ACA----GCC-GG----------A--G---CAT-------------C----------A--G----------------GAAT----T--T-GC---A------T-A---AG----------------T--T------- 327
SMM.US.x.pE660.CG7G         A--C---.............A--C---ACA----GCC-GG-------G--A--G---CAT-------------C----------A--G----------------GAAT-A--T--T-GC---A------T-A---AG----------------T--T------- 327
STM.US.89.STM_37_16         ..................................................................................................................................... 30
```

```
                                                                                                                                                  NF-κ-B-II
MAC.US.x.239                GCCTGTCAGAGGAAGAGGTT.AGAAGAAGGCTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAAGGAAACTCGCTGA...........................................................AACAGCAGGGACTTTCCACAA.GGGGATGT  436
A.CI.88.UC2                 -AT--C-----A-----TGG.-AGGC--AA--G-AA-------GA-A-CAT-T-GT-AA..................-A-......CAGGAACAACCATACTTGGTCAAAGCAGGAAGTAGCTACTAAGAAATAGCTGAG--T-------------G--.----C---  471
A.DE.x.BEN                  -AT--C-----A----ATGG.-AGGC--AA--G-AA-------GA-A-CAT-T-GT-AA..................-A-......CAGGAACAACCATACTTGGTCAAGGCAGGAAGTAGCTACTAAGAAACAGCTGAGG-T-------------G--.----C---  471
A.DE.x.PEI2_KR_KRCG         -----C----A------TGG.-AGGC--AA--G-AA-------GA-A-CATTT-GT.....................-A-..AGA.CAGGAACAGCTATATTTGGTCAGAACAGGAAGTAGATGA...........TGA---T-------------G--.----C---  452
A.GH.x.GH1                  ..........................................................................................................................................................................    0
A.GM.87.D194                ..........................................................................................................................................................................    0
A.GM.x.ISY_SBL_6669_85      ..........................................................................................................................................................................    0
A.GM.x.MCN13                ..........................................................................................................................................................................    0
A.GM.x.MCR35                ..........................................................................................................................................................................    0
A.GW.86.FG_clone_NIHZ       ..........................................................................................................................................................................    0
A.GW.87.CAM2CG              ----AC----A--T---TGG.-AGGCG-AA--G-AA-------GA-A-CATTT-GT.....................-A-...GGGCAGGAACAACCATACATGGCCAGGGCAGGAAGTAGCTACT.GAAAACAGCTGAG--T-------------G--.----C---  482
A.GW.x.MDS                  ..........................................................................................................................................................................    0
A.IN.07.NNVA                ----AC-----------TGG.-AGGCG--AT---AG----G--A--A-CATTT-AT.....................-A-AGGAGACAGGAACAACTATACTTGGTCAGGGCAGGAAGTAGCTATT.GAAAACAGCTGAG--T-------------G--.----C---  473
A.IN.95.CRIK_147            --T-AC-----------TGG.-AGGCG--A--G-AA----G--AA-A-CCTTT-GC.....................-A-..GAGACAGGAACAACTATACTTGGTCAGAACAGGAAGTAGCTACT.GAAAACAGTTGAG--T-------------G-T.----C---  217
A.JP.08.NMC786_clone_41     -AT--C----------ATGG.---GC---A--G-AA-------GA-A-CAC-T-GC.....................-A-...CAACAGGAACAACCATACTTGGTCAGGGCAGGAAGTAGCTACTGCA.AACAGCTGAG--T-------------G--.----C---  469
A.PT.x.ALI                  -----C----A--T---TGG.-AGGC---A--G-AA-------GA-A-CATTT-GTAAGA---G---CAGC.....---....................TTTGGTCAGGGCAGGAAGTAACTACT.GAAAACAGCTGAG--T-------------G--.----C---  468
A.SN.85.ROD                 ..........................................................................................................................................................................    0
A.SN.86.ST_JSP4_27          ..........................................................................................................................................................................    0
B.CI.88.UC1                 -AT-AC-----A-G---TGA.-AGGCT--A----AA-------GA-A-C--CAGACTAG-CA-GAG-CAGC-GCA..-A-A..........CAGGAACTAGCTGACACTGCACAAG.AGGGAAACT.......AGCTGAC--C-------------A--.-A--C---  475
B.CI.x.20_56                -G--AC-----A-----TGG.-AGGCT--A----AA-------AA-A-C--CAGACTAG-CAGGA--CAGC-GCT..-A-A..........GAGGAACTAGCTGACAGTGCATAA..ACAGGAACT.......AGCTGAC--T-------------AT-.----CA--  473
B.CI.x.EHO                  -AA--C-----A-----TGG.-AGGCT-AA--G-GA-------AA-A-C--CAGA-TAGACAGGA--TA-C-GCA..CA-A..........GAGGAACTAGCTAACACTGCATAGAGAAGGAAACT.......AGCTGAT--T---.---------G--.----C---  475
B.GH.86.D205_ALT            -G--AC-------G---TGG.-AGGCT--A----AA-------GA-A-C--CAGATTAG-CA-GAG-CAGC-GCA..-A-A..........CAGGAACTAGCTGACACTGCACAAG.AAGGAAACT.......AGCAGAC--T-------------A--.-A--C---  478
B.JP.01.IMCJ_KR020_1        ..........................................................................................................................................................................    0
G.CI.92.Abt96               ..........................................................................................................................................................................    0
AB.CM.03.03CM_510_03        ..........................................................................................................................................................................    0
H2_01_AB.CI.90.7312A        -AT-AC-----A-G--ATGG.-AGGCT--A----AA--------A-A-C--CAGA-TAGACAGGA---C-GCAG-A.-A-A......GAGGAACTAGCTGACGCTGCATAAGAAAGGAAACTGGCT...........GAC--T-------------G--.----C---  479
H2_01_AB.JP.04.NMC307_20    -AT-AC-----A-G--ATGG.-AGGCG--A--G-AA--------A-A-C--CAGA-TAGACAGGA--AC--CAG-A.-A-A...........GGGAACTAGCTGACGCTGCATAAGAAAGGAAACT.......GGCTGAC--T-------------G--.----C---  479
H2_01_AB.JP.07.NMC716_01    -AT-AC-----A-G--ATGG.-AGGCG--A----AA--------A-A-C--CAGA-TAGACAGGA--AC--CAG-A.-A-A...........GGGAACTAGCTGACGCTGCATAAGAAAGGAAACT.......GGCTGAC--T-------------G--.----C---  479
H2_01_AB.JP.08.NMC842_10    -AA--C----AA-G--ATGG.-AGGCT--A--G-AA--------A-A-C--CAGA-TAGACAGGA-GAC-GCAG-A.-A-A..........GGGGAACTAGCTGACACTGCATAAGAAAGGAAACT.......GGCTGAC--T-------------G--.----C---  480
U.CI.07.07IC_TNP3           ..........................................................................................................................................................................    0
U.FR.96.12034               ..........................................................................................................................................................................    0
MAC.US.x.17EC1              --------------------.------------------------------------------------------------...........................................................---------------------.--------  436
MAC.US.x.251_1A11           --------------------.------------------------------------------------------A-A---...........................................................G--------------------.----.---  435
MAC.US.x.251_32H_PJ5        -----C--------------.------------------------------------------------------------...........................................................G--------------------.--------  436
MAC.US.x.251_BK28           --------------------.-------------------------------------------G----------------...........................................................G-T------------------.--------  435
MAC.US.x.MM142_IVMXX        ..........................................................................................................................................................................    0
MNE.US.82.MNE_8             ..........................................................................................................................................................................    0
MNE.US.x.MNE027             ..........................................................................................................................................................................    0
SMM.CI.79.SIVsmCI2          ..........................................................................................................................................................................    0
SMM.LR.89.SIVsmLIB1         ..........................................................................................................................................................................    0
SMM.SL.92.SIVsmSL92A        ..........................................................................................................................................................................    0
SMM.SL.92.SL92B             ..........................................................................................................................................................................    0
SMM.US.04.G078              ..........................................................................................................................................................................    0
SMM.US.04.G932              ..........................................................................................................................................................................    0
SMM.US.04.M919              ..........................................................................................................................................................................    0
SMM.US.04.M922              ..........................................................................................................................................................................    0
SMM.US.04.M923              ..........................................................................................................................................................................    0
SMM.US.04.M926              ..........................................................................................................................................................................    0
SMM.US.04.M934              ..........................................................................................................................................................................    0
SMM.US.04.M935              ..........................................................................................................................................................................    0
SMM.US.04.M940              ..........................................................................................................................................................................    0
SMM.US.04.M946              ..........................................................................................................................................................................    0
SMM.US.04.M947              ..........................................................................................................................................................................    0
SMM.US.04.M949              ..........................................................................................................................................................................    0
SMM.US.04.M950              ..........................................................................................................................................................................    0
SMM.US.04.M951              ..........................................................................................................................................................................    0
SMM.US.04.M952              ..........................................................................................................................................................................    0
SMM.US.05.D215              ..........................................................................................................................................................................    0
SMM.US.06.FTq               ..........................................................................................................................................................................    0
SMM.US.11.SIVsmE660_FL10    --T---------------AA-AG----------------------A----A-------------------AA-----...............................................................G--------------------.----C---  437
SMM.US.86.CFU212            ..........................................................................................................................................................................    0
SMM.US.x.F236_H4            --T---------------A.-AG----------------------.----A-------------------AA-----...............................................................G--------------------.----C---  435
SMM.US.x.H9                 ..........................................................................................................................................................................    0
SMM.US.x.PBJA               --T-----A---------G.CAG----------------------T-A--A---------------A----AA---A-............................GACAGCAGGGACTTTCC.ACAAA........G--------------------A---------  214
SMM.US.x.PGM53              --T------G--------A.-AG----------------------T-A--A-------------------GA-----...............................................................G--------------------A---------  363
SMM.US.x.SME543             --T---------------A.-AG-------------------A--TA---A-------------G------AA-----............................GACAGCAGGGACTTTCC.ACAA.......................----------------.----C---  450
SMM.US.x.pE660.CG7G         A-T---------------A.-AG----------------------A----A-------------------AA-----...............................................................G--------------------.----C---  436
STM.US.89.STM_37_16         .............................-------------------A--G-------------------AA-----...............................................................G--------------------.----C---  109
```

Draft of 30 September 2015, not for public release
hiv2dna_main.tex r8876–10546

```
                                                TAR element end         Poly-A signal  5' LTR R repeat end  5' LTR U5 start
MAC.US.x.239                  ACTCTCACCAGCACTTGGCCGGTGCTGGGCAGAGTGACTCCACGCTTGCTTGCTT..AAAGCCCTCTTCAATAAA.GCTGCCATTTTAGAAGTAAGCTA..GTGTGTGTTCCCATCTCTCCTAGCCGCCGCCTGGTCAACTCGGTACTCAA.TAAT.AAGAA........ 758
A.CI.88.UC2                   -----------TG--C------CA---------.C-G------------------..----A------.------.-------A.-------C---T--A.-----------------------T------------.-T-----GT---TC-G-GT--C--........ 794
A.DE.x.BEN                    -----------T-----------A---------.C-G------------------..----A------.------.-------G.-------C---T--A.-----------------------T------------.-T-----GT---TC-G-GT--C--........ 795
A.DE.x.PEI2_KR_KRCG           -----------T-----------A---------.C-G------------------..----A------.------.-------G.-------C---T--A.-----------------------T------------.-T-----GT---CC---GTG-C--........ 774
A.GH.x.GH1                    -----------TG---------CA---------.C-G------------------..----A------.------.-------A.-------C---T--A.-----------------------T------------.-T-----GT---TC-G-AT--C--........ 239
A.GM.87.D194                  -----------TG---------CA---------.C-G------------------..----A------.------.-------G.-------C---T--A.-----------------------T------------.-T-----GT---TC-G-GT--C--........ 239
A.GM.x.ISY_SBL_6669_85        -----------TG---------C----------.C-G------------------.A----A------.------.-------G.-------C---T--A.-----------------------T------------.-T-----GT--TCC-G-GT--C--........ 240
A.GM.x.MCN13                  ---------G-TG---------CA---------.C-G------------------.A----A------.------.-------G.-------C---T--A.-----------------------T------------.-T-----GT---TC-G-AT--C--........ 240
A.GM.x.MCR35                  ---------G-TG---------CA---------.C-G------------------.A----A------.------.-------G.-------C---T--A.-----------------------T------------.-T-----GT---TC-G-AT--C--........ 240
A.GW.86.FG_clone_NIHZ         -----------TG---------C----------.C-G------------------..----A------.------.-------A.-------C---T--A.-------C---------------T------------.-TC----GT---CC-GTAT--C--........ 238
A.GW.87.CAM2CG                -----------TG---------CA---------.C-G------------------..----A------.------.-------G.-------C---T--A.-----------------------T------------.-T-----GT---CC-G-AT--C--........ 805
A.GW.x.MDS                    -----------TG---------CA---------.C-G------------------..----T------.------.-------G.-------C---T--A.-----------------------T------------.-T-----GT---CC-G-AT--C--........ 239
A.IN.07.NNVA                  -----------TG---------C----------.C-G------------------..----A------.------.-------G.-------C---TC-A.-----------------------T------------.-G-----GT---CC-G-GT--C--........ 789
A.IN.95.CRIK_147              -----------TG---------CA---------.C--------------------..----A------.------.-------G.-------C---T--A.------------------T----T------------.-T-----GT---TC-G-AT--C--........ 540
A.JP.08.NMC786_clone_41       -----------T-----------A---------.C-G------------------..----A------.------.-------A.-------C---T--A.-----------------------T------------.-------GT---TC-G-GT--C--........ 792
A.PT.x.ALI                    -----------TG---------CA---------.C-G------------------..----A------.------.-------A.-------C-G-T--A.AG---------------------T------------.-T-----GT---CC-G-GT--C--........ 790
A.SN.85.ROD                   ---------------------------------.C-G-C----------------..---AA------.------.-------G.-------C---T--A.-------C---------------T------------.-T-----GT---CC-G-GT--C--........ 239
A.SN.86.ST_JSP4_27            -----------TG---------CA---------.C-G------------------.A----A------.------.-------G.-------C---T--A.-------C---------------T------------.-T-----GT---TC---AGT-AC-A....... 241
B.CI.88.UC1                   ---------------------------------.C-G------------------..---A-------.------..------A.-------C---T-GA.-----------------------T------------.-T--AA-G---T-C---AAT-A-......... 795
B.CI.x.20_56                  ---------------------------------.C-G------------------..---A-------.------..------C.-------C---T-GA.-----------------------T------------.-T--A--G---C-C-G-GTTCA-......... 793
B.CI.x.EHO                    ---------------------------------.C-G------------------..---A-------.------.A-----TA.-------C---AC-A.-----------------------T------------.-T-----G---C-C--GA.............. 791
B.GH.86.D205_ALT              ---------------A----A---T--------.C-G------------------..---A-------.------..------A.-------C---.--A.-----------------------T--.-----------------G---C-C-G--A............. 792
B.JP.01.IMCJ_KR020_1          .......................................................................................................................................................................... 0
G.CI.92.Abt96                 ------------G---------------------Y-G------------------AC----A------.------.-------AA------------A-GT-----------------------T-------------C------GT--GCG--GCA-............ 163
AB.CM.03.03CM_510_03          .......................................................................................................................................................................... 0
H2_01_AB.CI.90.7312A          ---------------------------------.C-G------------------..---A-------.------.-------A.-------C---TC-A.-----------------------T------------.-------GT---GC-G-AT--AG......... 801
H2_01_AB.JP.04.NMC307_20      ---------------------------------.C-G------------------..---A-------.------.------TA.-------C---TC-A.-----------------------T------------.-T-----GTC--GC-TTAG............. 797
H2_01_AB.JP.07.NMC716_01      ---------------------------------.C-G------------------..---A-------.------.------TA.-------C---TC-A.-----------------------T------------.-T-----GTC--GC-TTAG............. 797
H2_01_AB.JP.08.NMC842_10      ---------------------------------.C-G------------------..---A-------.------.------TA.-------C---TC-A.-----------------------T------------.-T-----GTC--GC--TAG............. 798
U.CI.07.07IC_TNP3             ------------G--------------------.C-G------------------..----A------.------.-------A.-------C---T--A.-----------------------T------------.-T-----GT---TC-GGCT--C--........ 267
U.FR.96.12034                 -----------T-----------A-------------------------------..---A------.-------.-------G.-------C------A.-----------------------T-------------T.----AGA-TC-CCTTAT-............ 287
MAC.US.x.17EC1                -------------------------------------------------------..------------------.-----------------------..--------------------------------------------------.----.-----........ 758
MAC.US.x.251_1A11             -------------------------------------------------------..------------------.-----------------------..-----------------------T------------------------GG.----.-----........ 757
MAC.US.x.251_32H_PJ5          ------------------------------------G------------------..----A-------------.-----------------------..-----------------------T------------------------GG.----.---G-........ 758
MAC.US.x.251_BK28             --------------------A---------------G------------------..----A-------------.---------------------C-..-----------------------T------------------------GG.----.-----........ 746
MAC.US.x.MM142_IVMXX          ------------------------------------G------------------..----A-------------.--------.--------------..-----------------------T------------------------GG.----.--A--GACCCTGG 247
MNE.US.82.MNE_8               ------------------------------------G------------------..----A-------------.------T-.------------C-..-------C---------------T------------------------G-.----.-----........ 239
MNE.US.x.MNE027               ---------------------------T--------G------------------..----A-------------.--------.------------C-..-C---------------------T--------------------------.----.-----........ 239
SMM.CI.79.SIVsmCI2            .......................................................................................................................................................................... 0
SMM.LR.89.SIVsmLIB1           .......................................................................................................................................................................... 0
SMM.SL.92.SIVsmSL92A          .......................................................................................................................................................................... 0
SMM.SL.92.SL92B               -------A-CAGCAACA-G-CA-TGCT--GTAGAC-G------------------.....GA----.--------G-----A-C.-------C---TC-A.-A--------T------------T-------------TTCG--GT-C-G...G--.--CG......... 232
SMM.US.04.G078                ..........................................................................................................................................T----------TTG-G--AT-AG-A....... 25
SMM.US.04.G932                ..........................................................................................................................................C----------G-TAC................ 16
SMM.US.04.M919                ..........................................................................................................................................T----------G-AAC--.-----........ 22
SMM.US.04.M922                ..........................................................................................................................................T----------G-AAC--.-----........ 23
SMM.US.04.M923                ..........................................................................................................................................T----------G-.AC--.-----........ 22
SMM.US.04.M926                ..........................................................................................................................................T----------GC.-C--.-----........ 22
SMM.US.04.M934                ..........................................................................................................................................T----------GC.-C--.-----........ 22
SMM.US.04.M935                ..........................................................................................................................................T----------G-AAC--.-----........ 23
SMM.US.04.M940                ..........................................................................................................................................T----------TTG-G--AT-AG-A....... 25
SMM.US.04.M946                ..........................................................................................................................................T----------GC.-C--.-----........ 22
SMM.US.04.M947                ..........................................................................................................................................T----------G-.AT--.-----........ 22
SMM.US.04.M949                ..........................................................................................................................................C----------TTGAT--.-----........ 23
SMM.US.04.M950                ..........................................................................................................................................C----------GT.-C--.-----........ 22
SMM.US.04.M951                ..........................................................................................................................................T----------TTGAT--.-----........ 23
SMM.US.04.M952                ..........................................................................................................................................C----------GC.-C--.-----........ 22
SMM.US.05.D215                ..........................................................................................................................................T----------GT..C-G.T-A-......... 20
SMM.US.06.FTq                 ..........................................................................................................................................T-----------C..T-G.C-A.......... 19
SMM.US.11.SIVsmE660_FL10      ------------------------------------G------------------..----A-------------.--------.------------A-..-------C---------------T-------------T----------G-CAC--.-----........ 758
SMM.US.86.CFU212              ..........................................................................................................................................T----------GCTA-CAG--........... 21
SMM.US.x.F236_H4              ------------------------------------G------------------..----A-------------.-------A.------------A-..-----------------------T-------------T----------G-CAC--.-----........ 755
SMM.US.x.H9                   -----------------------------------RG------------------.A----A-------------.-------G.------------C-..-----------------------M-------------T----------G-.AC--.-----........ 241
SMM.US.x.PBJA                 ------------------------------------G------------------..----A-------------.--------.------------C-..-------------------------------------T----------G-.AC--.-----........ 534
SMM.US.x.PGM53                ----------------------------------C-G------------------..----A-------------.------GC.------------A-..-----------------------T-------------T----------G.AAC--A-----........ 684
SMM.US.x.SME543               ------------------------------------G------------------..----A-------------.--------.------------A-..-----------------------T-------------T----------G-CAC--A-----........ 772
SMM.US.x.pE660.CG7G           ------------------------------------G------------------..----A-------------.--------.------------A-..-----------------------T-------------T----------G-CAC--.-----........ 757
STM.US.89.STM_37_16           --------------------A--------T------G------------------..-T-TAT------------.-------A-------------A-..-----------------------T---------------------------.................. 422
```

5' LTR U5 end Lys tRNA primer binding site

```
MAC.US.x.239          .....GACCCTGGT.CTGTTAGGACCCTTTCTGCTTTGGG.AAACCGAAGCA...GGAAAATCCCTAGCAGATTGGCGCCTGAACAGGGACTTG.AAGGAGAGTGAGAGACTCC.TGAGT.ACGGCTGAGTGAAGGCAGTAAGGGCGGCAGGAACCAACCACGACGG  912
A.CI.88.UC2           .....---------.----------C-C--------..T--A----G..----------G-------C----------.---AG--C-----AG-C-T.G--AC.----------A----------  948
A.DE.x.BEN            .....---------.----------CTC--------..T--A-G---...----------G-------C----------.---AG--C-----AG-C-T.G--AC.T--------------------A----------  949
A.DE.x.PEI2_KR_KRCG   .....---------.----------CT---------.G--------G..--------A----------G-------C-----.----------.----------C----------.G--AC.--------------------A----------  917
A.GH.x.GH1            .....---------.----------C.C--------..T--A-G---...----------G-------C.----------.---A-G--C-----AG-C-T.G--AC.------.-------T----.----------A----------  389
A.GM.87.D194          .....---------.----------C-C-----A--..T--A-G---...----------G-------C----------A---AG--C-----AG-C-T.G--AC.----------------------A----------  394
A.GM.x.ISY_SBL_6669_85 .....---------.----------CT---------.------G---...----------G-------C----------.-G-A---C-----AG-CTT.G--AC.----------------------A----------  394
A.GM.x.MCN13          .....---------.----------.------A-----...----------G-------C----------A.G--A---C---A-AG-CTT.G--AC.----------------------A----------  394
A.GM.x.MCR35          .....---------.----------.-----A---------.------A---...----------G-------C----------.---A---C---A-AG-CTT.G--AC.----------------------A----------  394
A.GW.86.FG_clone_NIHZ .....---------.----------CT---------.-------G--GGC----------G-------C-----.----------.---A---C-----AG-CTT.G--ACA-G-----G----------------------A-----GC----  395
A.GW.87.CAM2CG        .....---------.----------CT---------.------A-G---...----------G-------C----------.---A---C---A-AG-CTT.G--AC.----------G----------------------A----------  959
A.GW.x.MDS            .....---------.----------CT---------.------G---...----------G-------C----------.---A---C-----AGTCTT.GA-AC.----------------------------------A----------  393
A.IN.07.NNVA          .....---------.----------CT---------..T--A-G--T..----------G-------C----------.---A-GC----AG-C-T.G--AC.----------------------A----------  943
A.IN.95.CRIK_147      .....---------.----------CT---------A--C----G---...----------T--G-------C----------.--------C-----AGTCTT.G--AC.----------------------A----------  695
A.JP.08.NMC786_clone_41 .....---------.----------.---------.------G---...----------G-------C----------.---AG--C-----AG-C-T.G--AC.----------------------A----------  946
A.PT.x.ALI            .....---------.----------.----G--G---...----------G-------C----------.---A---C-----AGTCTA.G--AC.----------------------A----------  944
A.SN.85.ROD           .....---------.----------CT---------.------G---...----------G-------C----------.---A---C-----AGTCTT.G--AC.----------------------A----------  393
A.SN.86.ST_JSP4_27    .....---------.----------.-----A-G---...----------G-------C----------.---A---C-----AG-CTT.G--AC.----------------------A----------  395
B.CI.88.UC1           .....---------.----------.-----A-G---...----------.-------C----------A-----A---------T--T-.G---C.----------------------T----------  949
B.CI.x.20_56          .....---------.----------.---C---.G--T--G---...A----------.-------C----------.-------A---------T--T-.G---C----T----------------------T----------  947
B.CI.x.EHO            .....A---------.----------C----------.-----A-G---...----------.-------C----------C---.G-A---------T--T-.G-GAC.----------------------T----------  945
B.GH.86.D205_ALT      .....---------.----------CT---------.---------G..----------.-------C----------C---.G-A---------T--T-AA---C--.A----------------------T----------  946
B.JP.01.IMCJ_KR020_1  .....---------.----------.----------.----------...----------.-------C----------.-G-A---------T--T-.G---C.----------G----------------------T----------  91
G.CI.92.Abt96         .....---------.----------.---------R.G---AG---G..----------.-------C----------.-GAAG--------T-CTT.---AC.-----A----------------------T----------  317
AB.CM.03.03CM_510_03  .....---------.----------.----------.----------...----------.-------C----------C---.G-A---------T--T-AG---C--.A----------------------T----------  85
H2_01_AB.CI.90.7312A  .....---------.----------CT---------.-----T-----...----------G-------C----------CC----A---------T--T-.G--AC.--A----------------------T----------  954
H2_01_AB.JP.04.NMC307_20 .....---------.----------CT---------.-----T-----...----------G-------C----------CC----A---------T--T-.G-GAC.--A----------C----------T----------  950
H2_01_AB.JP.07.NMC716_01 .....---------.----------CT---------.-----T-----...----------G-------C----------.---A---------T--T-AG-GAC.--A----------------------T----------  951
H2_01_AB.JP.08.NMC842_10 .....---------.----------CT---------.-----T-----...----------G-------C----------CC----A---------T--T-.G-GAC.--A----------C----------T----------  951
U.CI.07.07IC_TNP3     .....---------.----------T------.--------A-G...----------.-------C----------.----------G--------AC.----------------------A----------  421
U.FR.96.12034         .....A--------A----------CT---------.-----AG----...----------G-------C----------.---.----------T--T-TGA-AC.T----------G----------------------A----------  440
MAC.US.x.17EC1        .....---------.----------.----------.----------...----------.-------.----------.----------.-----.----------.----------.----------  912
MAC.US.x.251_1A11     .....---------.----------.----------.----------...----------.-------C----------.----------.-----.----------.----------.----------  911
MAC.US.x.251_32H_PJ5  .....---------.----------.---------A-----------...----------.-------.----------.----------.-----.----------.----------.----------  912
MAC.US.x.251_BK28     .....---------.----------.---------A-----------...----------.-------C----.-----.----------.-----.----------.----------.----------  899
MAC.US.x.MM142_IVMXX  TCTGTTAG------.----------.----------.----------...----------.-------C----------.----------.-----.----------.----------.----------  409
MNE.US.82.MNE_8       .....---------.----------.----------.----------...----------.-------C----------.----------.-----.----------.----------.---G---  393
MNE.US.x.MNE027       .....---------.----------.----------.----------...----------.-------C----------.----------.-----.----------.----------.---G---  393
SMM.CI.79.SIVsmCI2    ..........................................................................................................................----  4
SMM.LR.89.SIVsmLIB1   ..........................................................................................................................----  4
SMM.SL.92.SIVsmSL92A  ..........................................................................................................................----  4
SMM.SL.92.SL92B       .....AG-------A----C--A-----C--ACTC------.GGCAACTCCTGAGT-A-----T--G---T--------C----------.---AC.----------.GAC.T----------.----.----------.----G-----  375
SMM.US.04.G078        .....---------.----------.----------.------AG----...----------.-------C-G----------.----------.-----.----------.----------.----------  179
SMM.US.04.G932        .....A---------.----------.----------.----------...----------G-------C----------.----------.-----.----------.----------T---------G---  170
SMM.US.04.M919        .....---------.----------.----------.-------C-----G---...----------.-------C----------.----------.-----.----------.----------.----------  175
SMM.US.04.M922        .....---------.----------.----------.-------C-----G---...----------.-------C----------.----------.-----.----------.----------.----------  176
SMM.US.04.M923        .....---------.----------.----------.----------G---...----------G-------C----------.----------.-----.----------C----------.----------  174
SMM.US.04.M926        .....---------.----------.----------.----------G---...----------G-------C----------.----------.-----.----------.----------T----------  176
SMM.US.04.M934        .....---------.----------.----------.----------G---...----------G-------C----------.----------.-----.----------.----------T-------G---  176
SMM.US.04.M935        .....---------.----------.----------.-------C-----G---...----------.-------C----------.----------.-----.----------.----------.----------  176
SMM.US.04.M940        .....---------.----------.----------.------AG----...----------.-------C----------.----------.-----.----------.----------.----------  179
SMM.US.04.M946        .....---------.----------.----------.----------G---...----------G-------C----------.----------.-----.----------.----------T-------G---  176
SMM.US.04.M947        .....---------.----------.----------.----------G---...----------G-------C----------.----------.-----.----------C----------.----------  175
SMM.US.04.M949        .....---------.----------.----------.------G---...----------.-------C----------.----------.-----.----------.----------.----------  167
SMM.US.04.M950        .....---------.----------.----------.----------G---...----------G-------C----------A---------.-----.----------.----------T-------G---  176
SMM.US.04.M951        .....---------.----------.----------.------AG----...----------.-------C----------.---A---G---------.C----.----------.----------A----------  177
SMM.US.04.M952        .....---------.----------.----------.----------G---...----------T-------C----------.----------.-----.----------.----------T-------G---  176
SMM.US.05.D215        .....A---------.----------.----------.------A---...----------A-------C----------CC----------.-----.----------.----------.----------  173
SMM.US.06.FTq         .....---------.----------.----------C-------.-G--T-...----------.-------C----------.----------AA--.----------.----------.----------  174
SMM.US.11.SIVsmE660_FL10 .....---------.----------.----------.----------G---...----------G-------C----------.----------.-----.----------.----------.----------  911
SMM.US.86.CFU212      .....---------.----------.----------.-------C-----G---...----------.-------C----------.----------.-----.----------.----------.----------  175
SMM.US.x.F236_H4      .....---------.----------.----------.---------A-----G---...----------G-------C----------.----------.-----.----------.----------.----------  908
SMM.US.x.H9           .....-S--------.----------Y----------.----------.-----G---...----------G-------C----------.----------.-----.----------C----------.----------  394
SMM.US.x.PBJA         .....-C--------.----------.----------.----------.-----G---...----------G-------C----------.----------.-----.----------C----------.----------  687
SMM.US.x.PGM53        .....---------.----------.----------.-------C-----G---...----------.-------C----------.----------.-----.----------.----------.----------  837
SMM.US.x.SME543       .....---------.----------.----------.----------.-----G---...----------G-------C----------.----------.-----.----------.----------.----------  924
SMM.US.x.pE660.CG7G   .....---------.----------.----------.----------.-----G---...----------G-------C----------.----------.-----.----------.----------.----------  910
STM.US.89.STM_37_16   .....---------.----------.----------.-------CT--T--G-----.A----------.-------C----------.----------.-----.----------.----------.----------  576
```

```
MAC.US.x.239                AGTGCTCCTATAAA.GGCG.....CGGGTC..GGTACCAGA...CGGCGTGAGGAGC.GGGAGAGGAAGAGGCCTCC.GGT.TGCAGGTAAGTGCAACAC....AAA.....AAAGAAATAG....CTGTCTT.TTA.TCC..AGGAAGGGG.......TAATAAGATAG 1044
A.CI.88.UC2                 ----------G---.A---.........----C-AA-------A-AGG-------T-----.-----T-A-----------G---.GAA.---------CTACACC.A---ACTGT-GCC-G.A-AAGGCT---TA-CC--.CTTTT--AC---............--GA-G...-T 1094
A.DE.x.BEN                  ----------G---.----.........----C-GA-------A-GG.-------T-----.-----T-A-----------G---.GAA.-------A-CTACACC.A---ACTGT-GCC-G.A-AAGGCT---TA-CC--.C-TTT--AC---............--GA-G...-T 1094
A.DE.x.PEI2_KR_KRCG         ----------G---.----.........----C-GA-------A-GG.-------T-----.-----T.A-----------G---.GAA.-------A-CTACACC.A---ACTGT-GCC-G-A-AAGGCT---TA-CC--.C-TTT--AC---............--GA-G...-T 1062
A.GH.x.GH1                  ----------G---.A---.........----C-GA-------A-GG.-------T-----.-----T-A-----------G---.GAA.---------CTACACC.G---ACCGT-GCC-G.A-AAGGCTA--TA-CC--TCTTTT--AC---............--GA-G...-T  535
A.GM.87.D194                ----------G---.A---.........----C-GA------GA-G..-------T-----.-----T-A-----------G---.GAA.-------A-CTACACC.G---ACTGT-GCC-G.A-AAGGCT---TA-CC--.C-TTT--AC---............--GA-G...-T  538
A.GM.x.ISY_SBL_6669_85      ----------G---.----.......-A--C-AA-------A-GG.-------T-----.-----T.C-----------G---.GAA.-------A-CTACACC.A---ACTGT-GCCGG.A-AAGGCT---TA-CC--.C-TTT--AC---............--GA-G...-T  538
A.GM.x.MCN13                ----------G---.----.......-A--C-AA-------A-GG.-------T-----.-----T.A-----------G---.GAA.-------A-CTACACCAA---ATTGT-GCC-G.G-AGGGCT---TA-CC--.C-TTT--AC---............--GA-G...-T  539
A.GM.x.MCR35                ----------G---.----.......-A--C-AA-------A-GG.-------T-----.-----T.A-----------G---.GAA.-------A-CTACACCAA---ATTGT-GCC-G.A-AGGGCT---TA-CC--.C-TTT--AC---............--GA-G...-T  539
A.GW.86.FG_clone_NIHZ       ----------G---.----.........----C-GA-------A-G..-------T--.-.-----T.CG----------G---.GAAG-------A-CTACACC.A---ACTGT-GCC--.AGAGGGCT---TA-CC--.C-TTT--AC---............--GA-G...-T  538
A.GW.87.CAM2CG              --------AG---.----.......-A--C-GA-------A-GG.-------T-----.-----T.C-----------G---.GAA.-------A-CTACACC.A---ACTGT-GCCTG.A-AGGGCT---TA-CC--.C-TTT--AC---............--GA-G...-T 1103
A.GW.x.MDS                  ----------G---.A---.........----C-AA-------A-GG.-------T-----.-----T-A-----------G---.GAA.-------A-CTACACC.G---ACTGT-GCC--.A-AGGGCT---TA-CC--.C-TTT--AC---............--GA-G...-T  538
A.IN.07.NNVA                ----------G---.A---.........----C-AA-------A-GG.-------T-----.-----T-A-----------G---.GAA.-------A-CT.CACC.A---ACTGT-GCC-G.A-AAGGCT---TA-CC--.C-TTT--AC---............--GA-GA..-T 1088
A.IN.95.CRIK_147            ----------G---.----.........----C-GA------.A-GG.-------T-----.-----T.C-----------G---.GAA.-------A-CTACACC.A---ACTGT-GCC-G.A-AGGGCT--CTA-CC--.C-TTT--AC---............--GA-GA..-T  839
A.JP.08.NMC786_clone_41     ----------G---.----.........----C-GA-------AGG..-------T-----.-----T-A-----------G---.GAA.-------A-CTACACCAA---CTGT.-GCC-G.A-AGGGCT---T.ATCCTAC-TTTG-AT---............--GA-G...-T 1090
A.PT.x.ALI                  ----------G---.----.......-A--C-AA-------A-GGC-------T-----.-----T.-----------G---.GAA.-------A-CTACACC.A---TTGT.-GCC.G.A-AGGGCT---TA-CC--.C-TTT--AC---............--GA-G...-T 1087
A.SN.85.ROD                 ----------G---.----.........----C-GA-------A-GG.-A-----T-----.-----GA------------G---.GAA.-------A-CTACACCAA---CTGT.-GCC.G.A-AGGGCT--CTA-CC--.C-TTT--AC---............--GA-G...-T  537
A.SN.86.ST_JSP4_27          ----------G---.A---.......-A--C-GA-------AGGG.-------T-----.-----T-A-----------G---.GAA.---------CTACACC.A--TACAGT-GCC-G..-AGGGCT---TA-CC--.C-TTT--ACG--............--GA-G...-T  539
B.CI.88.UC1                 --A------G---.----.......-A--C-..--------G...-A-----------.-----GAA-----------G--AGGAA----G---A-T....CC..---TTCTG--TT.G..............TC-G.CT-TA--A----A-.........--GGGCA..TA 1075
B.CI.x.20_56                --A------G---.----.........----C-..--------G..------------.-----GAA-----------G--AGGAA----G---A-T....CC..---TTTTG--TT.G..............TC-G.CT-TA--A-G--A-.........--GG-CA..TA 1073
B.CI.x.EHO                  ---------G----.AA--.......-A--C-..--------G...-A-----------.-----GA--------TC-GG-AGTAAA--------A-T....CC..---TTTGC--TTCG..............TC--.CT-TAG-AC----..........--CGG-A..TA 1071
B.GH.86.D205_ALT            --A------G----.A---.......-A--C-..--------G...-A-----------.-----GA-------.----G--A.GTA--------A-T....CC..---TTTTGC-TT.G..............TC--.CT-TA-A--G--A-.........--GGGCA..TA 1070
B.JP.01.IMCJ_KR020_1        --A------G----.A---.......-A--C-..--------G...-A-----------.-----GA-----------G--A.GTA---G---A-T....CC..---TTTTA--TT.G..............TC--.CT-TA--A----A-.........--GGGCA..TA  216
G.CI.92.Abt96               ----T-----G---.----.........-----------.............-A--------A-.-----T.-----------G--GC---------------....CCAA--TTTGT-GCC---A--GGCTG--.AA-CC--.C-TTG-AT-GTCA-G........--GA-G.....  457
AB.CM.03.03CM_510_03        --A------G---.----.......-A--CT..--------G...-A-----------.-----GAA-----------G--A.GTAG-------A-T....CC..---TTTTGC-TT.G..............TC-G.CT-TA-A--.--A-.........--GGGCA..TA  209
H2_01_AB.CI.90.7312A        --A------G---.----.......-A-..-..--------G...-A-----------.-----GA-----------G--A.GTA--------A-T....CC..---TTTTG--CT.G..............TC--.CT-TAGA-G--A-..........--GGGCA..TA 1076
H2_01_AB.JP.04.NMC307_20    --C------G---C.A---.........----C-..--------G...-A-----------.-A---GA-----------TG--A.GTA--------A-T....CCAA--TTTTG--CT.G..............TC--.CT-TAG--G--A-..........--GG-CA..TA 1076
H2_01_AB.JP.07.NMC716_01    --C------G---.----.......-A--CG..-------AG...-A-----------.-A---GA-----------TG--A.GTA--------A-T....CC..---TTTTG--TT.G..............TC--.CT-TAGA-G--A-..........--GG-CA..TA 1075
H2_01_AB.JP.08.NMC842_10    -------------C.A---.........----C-..--------G...-A-----------.-A---GA-----------TG--A.GTA--------A-T....CC..---TTTTG--CT.G..............TC--.CT-TAGA-G--A-..........--GG-CA..TA 1075
U.CI.07.07IC_TNP3           ---------..--G-.AAG-.......-C-GGCC--------G..---------------.-----T.------------G---.GAA.-------A-CTTCACC.A---CTTGT-GCC-G.A-AGGGCT-----ATCCTAC-TTT--AC---............--GA-G...-T  562
U.FR.96.12034               ----------G---.----.......-A--CT..-------AG...-A-----------.-----T..G-----T----.---.-----------------A...A---....T-G-ACC-A--GGTTG-C-GCC.......--T.G--A----A-.......AGG--GT---T  574
MAC.US.x.17EC1              --------------.----.........-------..-----------..--------------.------------------.---.-----------------....---......-----------.........-------.---.---.---------.........------------ 1044
MAC.US.x.251_1A11           --------------.----.........-------..-----------..--------------.------------------.---.-----------------....---......-----------.........-------.---.---.---------A.........------------ 1043
MAC.US.x.251_32H_PJ5        --------------.----.........-------..-----------..--------------.------------------.---.-----------------..A---......-----------.........-------.---.---.---------A.........------------ 1045
MAC.US.x.251_BK28           --------------.----.........-------..-----------..--------------.--------G---------.---.-----------------....---......-----------.........-------G---.---.---------A.........------------ 1032
MAC.US.x.MM142_IVMXX        ----------G---.----.........-------..-----------..--------------.---GAGA-----------T---.-----------------....---......G-----------.........-------.---.---.---------A.........------------  542
MNE.US.82.MNE_8             ----------G---.----.........-------..-----------..--------------.------------------.---.-----------------....---......----G----.........-------.---.---.---------A.........------------  524
MNE.US.x.MNE027             ----------G---.----.........-------..-----------..--------------.------------------.---.-----------------....---......-----------.........-------.---.---.---------A.........------------  524
SMM.CI.79.SIVsmCI2          --------AG----TA---.....G---C-..--------G...--C----T-----.-----T.-----------GAC-.GA-C--------GC-GTA.......----..CTTG--CC-AGAAGGCTGT-AG-.C--.C--TT--AC...........GGG--GA-G--  140
SMM.LR.89.SIVsmLIB1         ----------G---.----.........----C-..--------G..-------------.A---TCA------------.---.---G--------GCA--..A-------..-C--GT-ATC............G--.CT-...T---A-AA.......GGG--GATA--  129
SMM.SL.92.SIVsmSL92A        ----------G---.A---GGGAC--A-C-..-------..-G...T-C----T-----.-A---CT-.---------T.---.---------------------A....CTT...GGT-GCCG-A--GGCTG-C-A-CC---.GA-A................GG---GT-ATT  137
SMM.SL.92.SL92B             ---T.....................T-ACC-..---G----TGGG-T-A-C-.----T.ACAG-TAAG--------GTG--A...---.----GT-C..........................---GTG-GAA..TA-CCTAC.C--.....CGG-AT---...........--GA-GT..-A  484
SMM.US.04.G078              --C-------G---.----.........----C-..--------G..-------T-----.-----CC-------------.---.-----------------.............GT--CC-G--GGCTGATAG--.C-G.CT-..TT--G-AAA.......G-GAT--..-A  310
SMM.US.04.G932              ----------G---.----.........-------..-----------..--------------.-----C.------------.---.-----------------....G-----..............--------------.---AG--A-----A.........--G--------  302
SMM.US.04.M919              ----------G---.----.........-------..-----------..--------------.-----TC-G----------.---.-----------------....-----........G--TC----GACTAA--TCC.C--.CTT..TTA-G-AAA.......G-G--G-..--  307
SMM.US.04.M922              ----------G---.----.........----C-..--------G..--------------.-----TC-G----------.---.-----------------....-----........G--TC----GACTGA--TCC.C--.CTT..TT--G-AAA.......G-G--G-..--  308
SMM.US.04.M923              --------------.----.........----C-..--------G..--------------.-----TC-G----------.---.-----------------....-----........G--TC----GACTGA--TCC.C--.CTT..TT--G-AAA.......G-G--G-..--  307
SMM.US.04.M926              ---------G----.----.........-------..-----------..--------------.-----GT-G----------.---.-----------------....-G-.........--AGG----...........-------.---.---.--A-----A.........--G--------  307
SMM.US.04.M934              ---------G----.----.........-------..-----------..--------------.-----GT------------.---.-----------------....-G-.........--AG-G---...........-------C--.-.-.---------A.........--G--------  306
SMM.US.04.M935              ----------G---.----.........----C-..--------G..--------------.-----TC-G----------.---.-----------------....-----........G--TC----GACTGA--TCC.C--.CTT..TT--G-AAA.......G-G--G-..--  308
SMM.US.04.M940              ----------G---.----.........-------..--------G..-------T-----.-----CC-------------.---.-----------------.............GT--CC-A--GGCTGATAG--.C--.CT-..TT--G-AAA.......G-GAT--..-A  310
SMM.US.04.M946              ---------G----.----.........-------..-----------..--------------.-----GT------------.---.-----------------....-G-.........--AG-G---...........-------C--.-.-.---A-G--A.........--G--------  306
SMM.US.04.M947              --------------.----.........----C-..--------G..--------------.-----TC-G----------.---.-----------------....-----........G--TC----GACTGA--TCC.C--.CTT..GT--G-AAA.......G-G--G-..--  307
SMM.US.04.M949              ----------G---.----.........----C-..--------G..--------------.-----CC-------------.---.-----------------.............GT--CC-G--GGCTGATAG--.C--.CT-..TT--G-AAA.......G-GAT--..-A  298
SMM.US.04.M950              ---------G----.----.........-------..-----------..--------------.-----GT------------.---.-----------------....-G-.........--AG-G---...........-------C--.-.-.-------A--A.........--G--------  306
SMM.US.04.M951              ----------G---.----.........-------..-----------..--------------.-----CC-------------.---.-----------------.............GT--CC-G--GGCTGATAG--.C--.CT-..TT--G-AAA.......G-GAT--..-A  307
SMM.US.04.M952              ---------G----.----.........-------..-----------..--------------.-----GT------------.---.-----------------....-G-.........--AG-G---...........-------C--.-.-.---------A.........--G--------  306
SMM.US.05.D215              ----------G---.----.........----C-..--------G..-------T-----.-----CTA-----------.---.-----------------.............-T--CCG---GGCAAGCAGCA.-C-.CT-TG----GAA-A.......G-G-G....--  304
SMM.US.06.FTq               ----------G---.----.........----C-..--------G..-------T-----.-----GC------T----.---.-----------------....-----........G--CC-G--GGCTG-AGT-..C--.CT-TCC-----A-AAGAAGTGAGGAG--TA--  317
SMM.US.11.SIVsmE660_FL10    --A-------G---.----.........----C-..--------G..--------------.-----TC-G---------.---.-----------------.....G-..........---TC----GACTGA--TCC.C--.CTT..TT--G-AAA.......G-G--G-..--- 1043
SMM.US.86.CFU212            ----------G---.----.......-A--C-..--------G...-A----T-----.-----TC-G--------T.---.-----------------.................--CT-G---TTGA-A--AC-.C.........T---A-AA.......GGG--G...-A  296
SMM.US.x.F236_H4            --A-------G---.----.........----C-..--------G..--------------.-----TC-G----------.---.-----------------.....G-..........---TC----GACTGA--TCC.C--.CTT..TT--G-AAA.......G-G--G-..--- 1040
SMM.US.x.H9                 --------------.----.........----C-..--------G..--------------.-----TC-G----------.---.-----------------....-----........G--TC----GACTGA--TCC.C--.CTT..TT--GAAAA.......G-G--G-..--  526
SMM.US.x.PBJA               --------------.----.........----C-..--------G..--------------.-----TC-G----------.---.-----------------....-----........G--TC----GACTGA--TCC.C--.CTT..TT--GAAAA.......G-G--G-..--  819
SMM.US.x.PGM53              --------------.----.........----C-..--------G..--------------.-----TC-G----------.---.---G-------------....-G-.....-G--TC----GACTGA--TCC.C--.CTT..TT--G-AAA.......G-G--G-..--  970
SMM.US.x.SME543             --A-------G---.----.........----C-..--------G..------------A----TC-G----------.---.-----------------.............---TC----GACTGA--TCC.C--.CTT..TT--G-AAA.......G-G--G-..--- 1056
SMM.US.x.pE660.CG7G         --A-------G---.----.........----C-..--------G..--------------.-----TC-G----------.---.-----------------.....-G-..........---TC----GACTGA--TCC.C--.CTT..TT--G-AAA.......G-G--G-..--- 1042
STM.US.89.STM_37_16         ---------G----.----.....----........-G-AG..-------------.-----GC------A-TC-..---.----A--G---------.........-.....GT--CC-G--GGCTGAT-G--.C--.CT-..TT--G--AA.......G-GGT--..-A  700
```

Gag and Gag-Pol start

```
MAC.US.x.239                  AGTGGGAGATGGGCGTGAGAAACTCCGTCTTGTCAGGGAAGAAAGCAGATGAATTAGAAAAAATTAGGCTACGACCCAACGGAAAGAAAAAGTACATGTTGAAGCATGTAGTATGGGCAGCAAATGAATTAGATAGATTTGGATTAGCAGAAAGCCTGTTGGAGAACAAA 1214
Gag                                   _M__G__V__R__N__S__V__L__S__G__K__K__A__D__E__L__E__K__I__R__L__R__P__N__G__K__K__K__Y__M__L__K__H__V__V__W__A__A__N__E__L__D__R__F__G__L__A__E__S__L__L__E__N__K_
A.CI.88.UC2                   T--------------C-------------C--AG------A--G-----C------------------T----G---GG-----------------G---A--A---A-T--G--------G--C-----G--C-----CA------------A----------TCA--- 1264
A.DE.x.BEN                    T--------------C----------------AG------A--------C------------G-----T----G---GG-----------------G---A--A---A-T--G--------G--------G----A---C-----G-----G------------TCA--- 1264
A.DE.x.PEI2_KR_KRCG           T--------------C-----G----------AG------A----T----------------------T----G---GG---G--------A--T-GAC-A------A-T--G--------G--------G-GC-A---C-----G------------------TCA--- 1232
A.GH.x.GH1                    T--------------C----------------AG------A--------------------G------T----G----G-----------------G-C-A--A---A-T--G--------G--------G--C-A---C-----G-----G------------TCA---  705
A.GM.87.D194                  T--------------C----------------AG------A--------C------------G-----T----G----------------GA----G---A--A-----T--G--------G-----------C-----C-----G-----G-----------ATCA---  708
A.GM.x.ISY_SBL_6669_85        T--------------C--A-------------AG------A--G------------------------T----G---GG---G--------A----GA--A--A---A-T--G--------G--------G--C-----C------A----G------------TCA---  708
A.GM.x.MCN13                  T--------------C----------------AA------A--------C----------C-------T----G---GG------------A----G-C-A------A-T--G--------G--------G--C-----C-----------G------------TCA---  709
A.GM.x.MCR35                  T--------------C----------------AA------A--------C----------C-------T----G---GG------------A----G-C-A------A-T--G--------G--------G--C-----C-----------G------------TCA---  709
A.GW.86.FG_clone_NIHZ         T--------------C----------------AG------A---------------------------T----G---GG-----------------A---A--A---A-T--G--------G--------G--C-----C-----------G------------TCA---  708
A.GW.87.CAM2CG                T--------------C----------------AG------A--------C------------G-----T----G---GG------------A----A---A--A---A-T--G--------G--------G--C-----C-----G-----G-----------ATCA--- 1273
A.GW.x.MDS                    T--------------C-----G----------AG------A--------C------------G-----T----G---GG------------A----G---A--A---A-T--G--------G--------G--C-----C-----------G-----------ATCA---  708
A.IN.07.NNVA                  T--------------C--AG---------C--AG------A-----T-------------G-------T----G---GG------------A----GAG-A--A-----T--G--------G--C-----G--C-----C------A----------------ATCA--- 1258
A.IN.95.CRIK_147              T--------------C--AG-G-------C--AG------A-----T-------------G-------T----G---GG------------A----GAC-A--A-----T--G--------G--C-----G--C-A---C------A-----------------TCA--- 1009
A.JP.08.NMC786_clone_41       T--------------C----------------CG------A---------------------------T----G---GG-----------------A---A--A---A-T--G--------G--C-----G--C-----C-----------G------------TCA--- 1260
A.PT.x.ALI                    T--------------C----------------AG-----GA--------C----------G-------T----G---GG------------A--TCA-C-A--A---A-T--G--------G--------G--C-----C-----G------------------TCA--- 1257
A.SN.85.ROD                   T--------------C----------------AG------A-------------------G---C---T----G---GG-----------------G-C-A--A---A-T--G--------G---A----G--C-----C-----------G------------TCA---  707
A.SN.86.ST_JSP4_27            T--------------C----------------AG------A--------C-----------G------T----G---GG------------A--T-G-C-A--A---A-T--G--------G--------G--C-----C-----G-----G------------TCA---  709
B.CI.88.UC1                   ---------------C-----G----------------------A-----------------G-----T----G---GG-----------G---TTGTC-A--A---A--A---------TG--C-----G--C-----------------G-----------ATCA--- 1245
B.CI.x.20_56                  ---------------C----GG----------------------A-----------------G-----T----G---GG---------------TTGTC-A-------------------TG--C-----G-------A------------G-C----------TCA--- 1243
B.CI.x.EHO                    ---------------C----GG---------------------GA----------------GG-----T----G---GG-----------GA------C-A--A---A------------TG-----------G-----------G-----G----G-----GATCA--- 1241
B.GH.86.D205_ALT              ---------------C----GG-------C--------------A-----------------G-----T----G---GG---------------T---C-A--A----------------TG--C--------C-----------------G-----------ATCA--- 1240
B.JP.01.IMCJ_KR020_1          ---------------C----GG----------------------A-----------------G-----T----G---GG------------A--TTGTC-A------A------------TG--C-----G--C----A------------G------------TCA---  386
G.CI.92.Abt96                 T--------------C-----G-------C----------A-----T---------------G-----T----G---GG------------A--------A--A-----T--G--------G--C-----G--C-----C-----------G-----------ATCA---  627
AB.CM.03.03CM_510_03          ---------------C-----G-------C--------------A-----------------G-----T----G---GG-----------G---T--------A------A---------TG--C--------------------------G-C---------ATCA---  379
H2_01_AB.CI.90.7312A          ---------------C----GG-------C----------A--GA-----------------G-----T----G---GG---------------------A--A------A----------------------C-----C-----G-------A---------ATCA--- 1246
H2_01_AB.JP.04.NMC307_20      ---------------C----GG-------C----------A--GA-----------------G-----T----G---GG---------------------A--A------A----------------------C-----------------------------A-CA--- 1246
H2_01_AB.JP.07.NMC716_01      ---------------C-----G-------C-----------G--A-----------------G-----T----G---GG-----------G------------A------A----------------------C-----------------G-----------ATCA--- 1245
H2_01_AB.JP.08.NMC842_10      ---------------C----GG------------------A--GA-----------------G-----T----G---GG----------------G----A--A---A--A----------------------C-----------------------------A-CA--- 1245
U.CI.07.07IC_TNP3             T--------------C----C--------C--------------A-T--------------GG-----T----G----------------G-------C----A-----T--R--------G--C-----G--------------------G-----------ATCA---  732
U.FR.96.12034                 T--------------C--------.--------------------T----------------G-----T----G-----------------A--------A--A-----T--G--------G--C-----G--C------------------------------GC----  743
MAC.US.x.17EC1                -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1214
MAC.US.x.251_1A11             ---------------C---------------------------------------------------------------------------------------------------------------------------------------------------------- 1213
MAC.US.x.251_32H_PJ5          ---------------C-------------------------------------------------------------GG------------------------------------------------------------------------------------------- 1215
MAC.US.x.251_BK28             ---------------C-------------------------------------------------------------GG------------------------------------------------------------------------------------------- 1202
MAC.US.x.MM142_IVMXX          ---------------C---------------------------------------------------A---------GG-------------------------------------------------------------------------------------------  712
MNE.US.82.MNE_8               ---------------C-------------------------------------------------------------GG---G---------------------------------------------------------------------------------------  694
MNE.US.x.MNE027               ---------------C-------------------------------------------------------------GG---G--------------------------------------------------------------------C------------------  694
SMM.CI.79.SIVsmCI2            T--------------C---G------------------------------------------G-----T----G---GG------A--G-----------A--A---C-T--G--------G--C-----G--------------------T------------TCA---  310
SMM.LR.89.SIVsmLIB1           T--------------C----------------------------------C-----------G-----T----G----------G---------------A-----------G-----------C-----G--C-----------------T------------G-T---  299
SMM.SL.92.SIVsmSL92A          T------A-------C---G--------------------A---------------------G-----T----G---GG---------G-----T--------A-----GA-T--------G--------G--C-----C-----------------------ATCA---  307
SMM.SL.92.SL92B               --C------------C----GG------------------A---------------------G-----T----G---GG-----G-------------C-T------A-TA----------C-GA--------C----------CG-----------------A-G----  654
SMM.US.04.G078                G--------------C-----G------------------A--------C---------------T-TT----G-----------A--G-GA--T--------A--------G-----------------G--------------------------------T------  480
SMM.US.04.G932                ---------------C------------------------A---------------------G---------------------GA--G-GA--------A--------------------G--------G--------------------------------A-C----  472
SMM.US.04.M919                ---------------C----C-------------------------G------------------G--T----G--A--T--------------------------------G-----------C-----G----------------------------------C----  477
SMM.US.04.M922                ---------------C----------------------------------------------G-----T----G-----------------------------A--------G-----------C-----G--C-------------------------------C----  478
SMM.US.04.M923                ---------------C----------------------------------------------------T----G---GG-------------------C-------------------------------G--C------------------------------------  477
SMM.US.04.M926                ---------------C----------------------------------------------G-----T----G----------G---------T---C-------------------------------G--------------G-----------------A-GT---  477
SMM.US.04.M934                ---------------C-------------------.-------------------------G------T----G--------------G-----T---C-------------------------------G--C------------------G-------------T---  475
SMM.US.04.M935                ---------------C----------------------------------------------------T----G----------G---------------------------G-----------C-----G--C-------------------------------C----  478
SMM.US.04.M940                G--------------C-----G------------------A---------------------G-----T----G----------GA--G--A--T-----------------G-----------------G---------------------------------------  480
SMM.US.04.M946                ---------------C-----G----------------------------------------G-----T----G--------------G-----T---C-------------------------------G--C--------------------------------T---  476
SMM.US.04.M947                ---------------C------------------------A---------------------------T----G---GG-------------------C-------------G-----------C-----G--C-------------------------------C----  477
SMM.US.04.M949                G--------------C-----G----------------------------------------G-----T----G----------GA--G-----T-----------------------------------G---------------------------------------  468
SMM.US.04.M950                ---------------C----------------------------A-----------------G-----T----G---GG---------G------------------A----------------C-----G--C-------------------------------C----  476
SMM.US.04.M951                G--------------C------------------------A---------------------G---A-T----G---GG-----GA----G---T------------A----------------------G---------------------------------------  477
SMM.US.04.M952                ---------------C----------------------------A-----------------G-----T----G--------------G-----T---C-------------------------------G--C-------------------------------C----  476
SMM.US.05.D215                ---------------C---GTG------------------A---------------------G-----T----G---GG------A--G--A--T--------A-----G--G-----------------G--------------------G-----------A--T---  474
SMM.US.06.FTq                 ---------------C-----G------------------A--GA----------------GG-----T----G--------------------------------------G--------G--C-----G--------------G------------------------  487
SMM.US.11.SIVsmE660_FL10      ---------------C---------------------------------------------G------T----G--------------------T-----------------------------------G--C-----------------------------T------ 1213
SMM.US.86.CFU212              -------A-------C-----G------------------A--------------------GG-----T----G---GG---------------------------------G-----------C-----G--C-----C-----G-----G-----------A-R----  466
SMM.US.x.F236_H4              ---------------C-------------------------G--------------------G-----T----G--------------------T-----------------------------------G--C-----------------------------T------ 1210
SMM.US.x.H9                   -------------T------------------------------------------------------T----G---GG-------------------C--------A----------------------G--C------------------------------------  696
SMM.US.x.PBJA                 ---------------C----------------------------------------------------T----G---GG-----------G----CA-C--------A----------------------G--C------------------------------------  989
SMM.US.x.PGM53                ---------------C----------------------------------------------------T----G--------------------------------------------------C-----G--C-------------------------------C---- 1140
SMM.US.x.SME543               ---------------C----------------------------------------------------T----G--------------------T-----------------------------------G--------------------------------T------ 1226
SMM.US.x.pE660.CG7G           ---------------C----------------------------------------------------T----G--------------------T------------A----------------------G--C-----------------------------T------ 1212
STM.US.89.STM_37_16           G--------------C-----G------------------A---------------------G-----T----G---GG---G--------------------A--------G-----------C-----G--C-----------G------------------TCA---  870
```

```
MAC.US.x.239                   GAAGGATGTCAAAAAATACTTTCGGTCTTAGCTCCATTAGTGCCAACAGGCTCAGAAAATTTAAAAAGCCTTTATAATACTGTCTGCGTCATCTGGTGCATTCACGCAGAAGAGAAAGTGAAACACACTGAGGAAGCAAAACAGATAGTGCAGAGACACCTAGTGGTGGA 1384
Gag                            _E__G__C__Q__K__I__L__S__V__L__A__P__L__V__P__T__G__S__E__N__L__K__S__L__Y__N__T__V__C__V__I__W__C__I__H__A__E__E__K__V__K__H__T__E__E__A__K__Q__I__V__Q__R__H__L__V__V__E
A.CI.88.UC2                    -----T--C-----G--T--CA-A--T----A---------A--------G----------------------T------C-----------T------T-G---------------------G-T-----A--G------A-AC----A--------T-------CA-- 1434
A.DE.x.BEN                     -----T--C-----G--T--CAGA--T----A---------A--------G----------------------T------C-----------T------T-G---------------------G-T--------------GA-AC---CA--------T-------CA-- 1434
A.DE.x.PEI2_KR_KRCG            -----T--C--------TA--A-A--T----A------------------A----------------------T------------------T------T-G---------------------G-T-----A-G---------A-----A--------T-------CA-- 1402
A.GH.x.GH1                     -----T--C-----G--T--CA-A--T----A---------A--------G----------------------T---C--A-----------T------T-G---------------------G-T-----A---------A-AC----A--------T----G--CA--  875
A.GM.87.D194                   -----T--C-----G--T---AAA--T----AA--------A--------G----------------------T------C-----------T------T-G---------------------G-T-----A--------GA-AC---CA--------T-------CA--  878
A.GM.x.ISY_SBL_6669_85         -----T--C--------TA----A--T----AA--------A--------G------------------------------AC---------T------T-G---------------------G-T-----A---------AGA-----AGG------T-------CA--  878
A.GM.x.MCN13                   -----T--C----G---T---A-A--T----G---------A--G-----T----------------------T------------------T--------A---------------------G-T-----A-G---------A-----A--------T----C--CA--  879
A.GM.x.MCR35                   -----T--C----G---T---A-A--T----A---------A--G-----T----------------------T------------------T--------A---------------------G-T-----A-G---------A-----A--------T----C--CA--  879
A.GW.86.FG_clone_NIHZ          -----T--C--------T---A-A--T----A---------A--G-----T----------------------T---------------------------A---------------------G-T-----A-G---------A-----A--------T-------CA--  878
A.GW.87.CAM2CG                 -----T--C----G---T---AAA--T----A---------A--G-----T----------------------T------------------A--------A---------------------G-T-----A---------AGA----CA-T------T----C--CA-- 1443
A.GW.x.MDS                     --G--T--C--------T---A-A-------A---------A--G-----T----------------------T------------------T-TT-----A--------G------------G-T---------------A--T-C-CA-G------T-------CA--  878
A.IN.07.NNVA                   --G--T--C--------T---A-A--T----A------------G-----T----------------------T------------------T--------A---------------------G-T-----A---------A--T----A--------T-------CA-- 1428
A.IN.95.CRIK_147               --G--T--C--------T---A-A--T----A------------G-----T----------------------T------------------T--------A---------------------G-T-----A---------A-------A--------T-------CA-- 1179
A.JP.08.NMC786_clone_41        -----T--C----G---T---A-A--T---AA-------A-A--------G----------------------T------A-----------T------T-G---------------------G-T-----A---------A--C----A--------T-------CA-- 1430
A.PT.x.ALI                     -----T--C----G---T---AAA--T----AA-----------------A----------------------T---------------AG-T------G-G---------------------G-T-----A-G---------A---A-A--------T----C--CA-- 1427
A.SN.85.ROD                    --G--T-----------T---A-A--T----A----A-G--A--G-----T-----------------T----T------------------T--------A---------------------G-T-----A-G---------A-------G------T-------CA--  877
A.SN.86.ST_JSP4_27             --G--T--C--------T---A-A--T----A---------A--G-----G----------------------T------------------T-----T--A----------------C----G-T-----A-----------A-AG--A--------T-------CA--  879
B.CI.88.UC1                    -----------C-----T---A-------------------A--------A----------------------T------------------T-AC---T-G---------------------G-T---------------A------CA--------T----C--C--- 1415
B.CI.x.20_56                   C----------C-----T---A--------------A-G--A--------A----------------------T------------------T-AC---T-G---------------------G-T---------------A------CA-----C--T----C------ 1413
B.CI.x.EHO                     --------C-GG-----T-GGAAA-------GA--------A--------T-----------------T-----------C-------------TT---C-G---------------------G-T---------------A------CA---C----T----CA-C--- 1411
B.GH.86.D205_ALT               -----------------T---AAA-----------------A--------G----------------------T-----TC-----------T-TT---C-G---------------------G-T--A------------A------CA--------T----C--C--- 1410
B.JP.01.IMCJ_KR020_1           -----------C-----T---A---------AA--------A--------A----------------------T------------------T-AC--TC-G----------T----------G-T--------T--T---A------CA-----C--T----C--C-A-  556
G.CI.92.Abt96                  -----T--C--------T-----A--A--------------A--------T--------------G--------------------------T-----TT-G-----C---C-------A-----T-----A-----------AG-------------T--------A--  797
AB.CM.03.03CM_510_03           --------C--------T---AAA-----------------A--------A----------------------T------------------T-TC---C-G---------------------G-T--A-----G------A------CA---C--------...-C---  546
H2_01_AB.CI.90.7312A           -----------G-----T---A-----------------A----------G----------------------T------------------T------T-A---------A-----------G-T--C------------A--C---CA-----C--T--------A-- 1416
H2_01_AB.JP.04.NMC307_20       -----------G-----T---A---C-------------A----------G---------G------------T------------------T------T-G---------A-----------G-T--C------------A--T---CA-----C--T----C--CA-C 1416
H2_01_AB.JP.07.NMC716_01       -----------G-----T---AAA------------------------------------G------------T------------------T------T-G---------C-----------G-T---------------A--C---CA-----C--T----C--CA-C 1415
H2_01_AB.JP.08.NMC842_10       -----------G-----T---A---C-C-----------A----------A---------G------------T------------------T------T-G---------A-----------G-T--C------------A--T---CA-----C--T----C--CA-C 1415
U.CI.07.07IC_TNP3              -----T-----------T-----A--T-----------------G-----G--------------G---------------AC---------T------T-G-----------------------T-----------------AC----A--------T--------A--  902
U.FR.96.12034                  -----------------TT-A-----G--------------T--------T--------------G--------------------------T-----TT-G-----------------------T--------------G--------A--------T----C------  913
MAC.US.x.17EC1                 ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1384
MAC.US.x.251_1A11              -----------------------------------------------------------------------------------------------------------------------------------------------------A-------------------- 1383
MAC.US.x.251_32H_PJ5           ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1385
MAC.US.x.251_BK28              ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 1372
MAC.US.x.MM142_IVMXX           ---------------------------------------------------------------------------------------------------------------------------------------------------------------------A----  882
MNE.US.82.MNE_8                -----------------------------------------------------------------G--------------------------------------------------------------------------------------------------------  864
MNE.US.x.MNE027                -------------------------C---------------------------------------G--------------------------------------------------------------------------------------------------------  864
SMM.CI.79.SIVsmCI2             -----------C-----TT-G--A--T--------------A--G-----T----------------------T---------G--------T--------A--T-----------------------C--------------A-----------------------A--  480
SMM.LR.89.SIVsmLIB1            -----T-----------T-------------T---------T--------T--------------G--------C-----------------T--------A-----------------------T--A--------------A--------------T-----------  469
SMM.SL.92.SIVsmSL92A           -----T-------G---TT-G--A--G----T---------C-----------------------G--T-----------------------T--------A---------AT----------A-T--------------G-------------A--------C---A--  477
SMM.SL.92.SL92B                -------------G----T-AG-A--AC-----------A-------------G---------------T-G-T--G------------AG-T-----TC-G---------AT----------G----A--------T--GA-A-C---A-----C--T--------A--  824
SMM.US.04.G078                 --------C--------T--C-----------C--------A--G-----T--------------G---T-G-----------------T--T--------------------------------T-------------G---A--------------T-----------  650
SMM.US.04.G932                 ------------------A----------------------A-----------------------G-----C--------------------T-----TC-------------------------T-----A--G--------A-----A--------------------  642
SMM.US.04.M919                 -----C-------G---TA----A--T------------A-T--G-----T--------------G--------------------------T--------------------------------T-------------T---A------A-------T-----------  647
SMM.US.04.M922                 -----T-----------TA----A--T--------------T--G-----T------G-------G--------C-----------------T--------------------------------T-----------------A-C------------T-------C---  648
SMM.US.04.M923                 -----T-----------T-----A--T--------------T--G-----T--------------G--------------------------T--------------------------------T--C--A-----------A--------------T-----------  647
SMM.US.04.M926                 -----C--------------------T-----------------------T--------------G--------------------------T------T-G-----------------------T-----------------A-------------CT-----------  647
SMM.US.04.M934                 -----T--C----G---T-----A--T--------------T--------T--------------G--------C-----------------T-----TG-G-----------------------T-----------------AG-------------T-----------  645
SMM.US.04.M935                 -----T-----------TA--G-A--T--------------T--G-----T--------------G--------C-----------------T--------------------------------T-----------------A--------------T-----------  648
SMM.US.04.M940                 --------C--------T--C-----------C--------A--G-----T--------------G---T-G--C-----------------T--------------------------------T-------------G---A----------A---T-----------  650
SMM.US.04.M946                 -----T--C----G---T-----A--T--------------T--G-----T--------------G--------C-----------------T-----TT-G-----------A-----------T--------------G--AG-------------T-----------  646
SMM.US.04.M947                 -----T-----G-----T-----A--T---AA---------T--G-----T--------------G--------------------------T--------------------------------T-----A-----------A--------------T-----------  647
SMM.US.04.M949                 -----------------T--C-----------C--------A--G-----T--------------G---T-G--------------------T--------------------------------T-----------------A--------------T-----------  638
SMM.US.04.M950                 -----T-------G---T-----A--T--------------T--G-----T--------------G--------C-----------------T------T-G-----------------------T-----------------AG-------------T-----------  646
SMM.US.04.M951                 --------C--------T--C-----------C--G--------G-----T--------------G---T-G--------------------T--------------------------------T-----------------A--------------T-----------  647
SMM.US.04.M952                 --C--T-------G---T-----A--T--------------T--G-----T--------------G--------------------------T------T-G-------------------G---T-----------------AG-------------T-----------  646
SMM.US.05.D215                 -----T-------G---TA----------------------A--G-----T--------------G--------C-----------------T--------------------------------T-----------------A--------------T--------A--  644
SMM.US.06.FTq                  -----T-----------T--C--------------------A--------T--------------G--------------------------T--------C---------C-------------T--A--------------A--------------T-----------  657
SMM.US.11.SIVsmE660_FL10       -----T-----------T--------T-----------------------T--------------G--------------------------T--------------------------------T-----------------A--------------T----------- 1383
SMM.US.86.CFU212               -----T--------R--T-----A--T------------ACA--------G--------------G-----C--C-----------------Y--------------------------------T--Y--------------A--------------T--------A--  636
SMM.US.x.F236_H4               -----T-----------T-----A--T--------------T--G-----T--------C-----G--------------------------T--------------------------------T-----------------A--------------T----------- 1380
SMM.US.x.H9                    -----T-----------T-----A--T--------------T--G-----T--------------G-----------C------C-----C-T--------------------------------T-----A-----------A--------------T-----------  866
SMM.US.x.PBJA                  -----T-----------T-----A--T--------------T--G-----T--------------G-----------C------------C-T--------------------------------------A-----------A--------------T----------- 1159
SMM.US.x.PGM53                 -----T-----------TA----A--T--------------T--GG----T--------------G--------------------------T--------------------------------T-----------------A--------------T----------- 1310
SMM.US.x.SME543                -----T-----------T--------T--------------T--G-----T--------------G--------------------------T--------------------------------T-----------------A--------------T----------- 1396
SMM.US.x.pE660.CG7G            -----T-----------T--------T--------------T--G-----T--------------G--------------------------T--------------------------------T-----------------A--------------T----------- 1382
STM.US.89.STM_37_16            -----G-----------TA-AA---------AA-----------------T--------------G--T----T------------------T-----T--C-----------------------T--------------G--AG-G--AA-------T--T-----A-- 1040
```

Gag p17 Matrix end_Gag p24 Capsid start

```
MAC.US.x.239              AACAGGAACAACAGAAACTATGCCAAAAACAAGTAGACCAACAGCACCATCTAGCGGCAGAGGAGGAAATTACCCAGTACAACAAATA...GGTGGTAACTATGTCCACCTGCCATTAAGCCCGAGAACATTAAATGCCTGGGTAAAATTGATAGAGGAAAAGAAATTTG 1551
Gag                       __T__G__T__T__E__T__M__P__K__T__S__R__P__T__A__P__S__S__G__R__G__G__N__Y__P__V__Q__Q__I__.__G__G__N__Y__V__H__L__P__L__S__P__R__T__L__N__A__W__V__K__L__I__E__E__K__K__F__
A.CI.88.UC2               ---------TG----G-AA--------T--------------------C----T--G-A---------C-T---C--------GGC-...------C-----------TG-----C-G-----CC----TC-------T-----------AG------------G--C- 1601
A.DE.x.BEN                ---T-----TG----G-AA--------T--------------------C----T--G-A-A-------C-----C--G------GCG...------C-----------TG-----C-G-----CC----TC-------A-----------AG-G----------G--C- 1601
A.DE.x.PEI2_KR_KRCG       ---------TG----C-AA-------GC-----------G--------C-------G-----G-----------C--G----------...-C---C------TC---TG----G--G-----CC----CC-------T--------G--AG-G--A--------G--C- 1569
A.GH.x.GH1                ---------TG------AA-------GT--------------------C-------G------A----C-T---T--------G-C-GGC----C------A-----G-----C-G-----CC----TC-------T-----------AG-------C-----G--C- 1045
A.GM.87.D194              ---------TG----G-AA--------T-T------------------C----T--G-A---------C-T---C--G-----GGC-...-C--C------A----TG----GC-G-----CC----TC-------T-----------AG------------G--C- 1045
A.GM.x.ISY_SBL_6669_85    -----A---TG----G-AA--------T-T------------------C-------G-A---G-----C-T---C--G----------...-C--C-----------T-----GC-G--T--CC----CC-------T--------G--AG-------G--A--G--C- 1045
A.GM.x.MCN13              ---------TG----G-AA--------T--------------------C-------G-AG--------C-T---C--------G--...-C--C--T---AC---TG----GC-G--T--TC----CC-------T-----------AG------------G--C- 1046
A.GM.x.MCR35              ---------TG----G-AA--------T--------------------C-------G-A---------C-T---C--------G--...-C--C--T---AC---TG----GC-G--T--TC----CC-------T-----------AG------------G--C- 1046
A.GW.86.FG_clone_NIHZ     ---------TG----G-AA--------T--------------------C------.....-GAA---C-T---C--G--G---G--...-C---C------AC---TA-A---C-G--T--CG----TC-------T-----------AG-G-------------C- 1039
A.GW.87.CAM2CG            ---------TG----G-AA------G-T--------------------C----T--G-AG--------------C--G---AGC---...-C--C------ACT--TG----GC----T--CC----TC-------T-----------AG------------G--C- 1610
A.GW.x.MDS                ---------TG----G-AA--------T--------------------C-------G-----------C-----C--G--GA-T---...-C--C------AC---TG----G--G--T--CC----TC-------T-----------AG----A--------G--C- 1045
A.IN.07.NNVA              ---------TG----G-AA--------T-------------------CC------AGG----G-----C-TT--C-------GC---...-C--C-----CAC---G-----GC-G-----C-----TC------TT-----------AG----A--G-----G--C- 1595
A.IN.95.CRIK_147          ---T-----TG----G-AA--------T--------------------C-------GG----------C-T---C-------C---...-C--C--T---AC---TG----GC-G--T--C-----TC-----------------------AG------------G--C- 1346
A.JP.08.NMC786_clone_41   ---------TG----G-AA------G----------------------C----T--G-A---------C-T---C--------G--...-C--C------------TG----GC-G-----CC----CC-------T-----------AG-G----------G--C- 1597
A.PT.x.ALI                --T--A----G----G-AA-------GC--------------------C----T-AACAG--G-----C-T---C--------G--...CC--C-----CAC---TG----GC-G-----CC----C---G---T-----------AG----A--------G--C- 1594
A.SN.85.ROD               ---------TG----G-AA-------GC--------------------------AG-AG---------------G-----TG--...-C--C-----CAC---TA-A--GC-G--T--CC----CC--------------------AG----------A--G--C- 1044
A.SN.86.ST_JSP4_27        ----AA---T-------AA-------GT--------------------C-------G-AC--------C-T---C--------G-G...CC--C------AC---TG-----C----T--CC----CC-------T----------C-AG------------G--C- 1046
B.CI.88.UC1               C.........-----G-AA------GCT--------------------C-----......--C------------G--G--G----...-C---C-----C------A----GC-------TC----C-----------------G--AG----A--------G--C- 1567
B.CI.x.20_56              C.........-AG----AA------GCT--------------------C-----......--C------------G--G--G----...CC-----T---------G----GC----T--CC----CC----C-------------AG-G-----C-----G--C- 1565
B.CI.x.EHO                C.........-------AA------GCT-TG----A-----GTAA----A-----......C-CCT-GC----T-----G--G-------...-C---C--T---TC---------GC-------CC----C-------------------AG----A--G-----G--C- 1563
B.GH.86.D205_ALT          C.........-------AA------GCT----A--A--------------C-----......--C--------T-----G--G---C-G...-C---C-----C-----------GC-------CC----C-------T--------G--AG----A--------G--C- 1562
B.JP.01.IMCJ_KR020_1      C.........--G----AAGC----GCT-------A-------------C----T.......--CA-------T-----G--G--GG-G...CC-----T--CAC---------GC-------CC----C-----------------------AG-----------------C-  708
G.CI.92.Abt96             ---T-----T-------AAG---TGCC-----C-------T-------C----T---------------C--------G--G---G--...-C--C-------------T--T-----------T--TC----T--------A-----------G----A--------G--C-  964
AB.CM.03.03CM_510_03      C.........G------AA---A--GCT----A--A------------C-----......--C--------T-----G--G--------...-C---C-------------G----GC-------CC----C-------T-----------AG----A--------G--C-  698
H2_01_AB.CI.90.7312A      C.........--T----AA-CA--T-CTG--GAC-A----G----GA-------T.......--CA------C------G------G--...-C---C--T----------G-----------------CC----CC-----------------------AG-G-------------A- 1568
H2_01_AB.JP.04.NMC307_20  C.........--T----AAGCA----CT......GC-GAC-A-C--G--G-G-CATCT--T--C-----C--------G------G--...-C---C--T----------G-----C----T--CC----CC-------------G-----AG-G----------G--C- 1568
H2_01_AB.JP.07.NMC716_01  C.........--T----AAGCA---GCT......GC-GAT-A-C--G--G-G-CATCT--T--C-----C--------G------G--...-C---C--T----------G-----C----T--CC----CC-----------------------AG-G----------G--C- 1567
H2_01_AB.JP.08.NMC842_10  C.........--T----AA-CA----C-C---C-GC-GAT-A-C--G--G-G-CATCT--T--C--------------G------G--...-C---C--T--------TG----G-----T--CC----C-------------------------AG-G----------G--C- 1573
U.CI.07.07IC_TNP3         ---T-----TG----C-AA------GCC--------G-----------C----T-----A-----------------G---G-G...--A--C--T-----A--T-----TC-G--T--TC-G--TC-C-----T-----G-----AG-G--A------------- 1069
U.FR.96.12034             ---T-A---TG------AA------GCC-------------------YT---C----T---G-CA-----------G--G--G--G---...------C--T----C------A--TC----T--T------C-C---------------C--G-------G--A------- 1080
MAC.US.x.17EC1            ------------------------------------------------------------------------------------------...------------------------------------------------------------------------------------ 1551
MAC.US.x.251_1A11         ----------G------T--------------C--------T-------------------------------------------...------------------------------------------------------------------------------------ 1550
MAC.US.x.251_32H_PJ5      ----------G-----------------------------------------------------------------------------...------------------------------------------------------------------------------------ 1552
MAC.US.x.251_BK28         ----------G-----------------------------------------------------------------------------...------------------------------------------------------------------------------------ 1539
MAC.US.x.MM142_IVMXX      ----------G----------------------------------------T-------------------------------------...---------------AC------A-------------------------------------A-------G--------- 1049
MNE.US.82.MNE_8           ----------G---------------------------------------------T------------------------G--...---------------AC------A----------------------------------------------G--------- 1031
MNE.US.x.MNE027           ----------G---------------------------------------------T---A-A-------------------------...---------------AC------A----------------------------------------------G--------- 1031
SMM.CI.79.SIVsmCI2        ---------TG----C-AA------GC---------------------C--G-T----A-------------------G---G-G...--A-----T--CAC----T-A-------T--TC----C-------------------A-------------G--C-  647
SMM.LR.89.SIVsmLIB1       ---------TG----C--AG----------------------------C----T-A---GA------------------G-----G...-C--C--T--------T----GC-------C--G---C-------A-A------GC-AG-G--A--------G----  636
SMM.SL.92.SIVsmSL92A      --AT-A---TG----CCAAT----TGCC--------------------C----T----AGA-------------T-----G---...-A--C--T-----T-----A--T--------CC-C--G--G-----T-----T--GC-AG-G--A--------G----  644
SMM.SL.92.SL92B           --GT-----TG------AAT-----GCTCA---C-----------T---C----T......--------C--T----G--G---G--...--AAA---T-----G---ACT---C-GTC---AC-G---C-C-----G-----C---C--G---A-------------  985
SMM.US.04.G078            ---------TG----C-AA-------C----------------------C----T----A--------------------G---G-C...------C--T--CAC----T----GC----T--A-----T--G-----A--------G---G---A-------GG----  817
SMM.US.04.G932            ----------G-----G-----------T--------------------C----T-----------------------------G--...-C-----T--AC------A--T-----T--T--G-----------T-----------G-G--A--G--A-------  809
SMM.US.04.M919            ----------G----C-AA-------T------C---------------C----T-------------------G--G---G--...------C--T--CAC------A--C------T--A--------------------------C--G---A--G--A-------  814
SMM.US.04.M922            ----------G----T-AA------GC------C----T-GT-------C----T----A-------------G--G--G---G--...------C--T--CAC------AT-C-----T--A--------------T----------C--G---A--G--A-------  815
SMM.US.04.M923            ----------G----C-AA------GC----------------------C----T-------------------G--G---G--...------C--T-----------A--T------T--A--------------T-------------G---A--G--A-------  814
SMM.US.04.M926            ---------------T-ACC-------G-------C-------------C----T-----------------------G---G--...-C--C-----------TT----G------T--A--------------------------G-------G---------  814
SMM.US.04.M934            ----------G----C-AC-----G----------G-------------C----T-----------------------G---G--...-C--C-----------T-----G------T--A--G---C------------------G-G----------------  812
SMM.US.04.M935            ----------G----T-AA------GC------C---------------C----T-------------------G--G---G--...------C--T--CAC------A--C------T--A--------------T----------G-G--A--G--A-------  815
SMM.US.04.M940            ----------G----C--C------------------------------C----T---GC-A-------------G-----G-T...--A--C--T--CAC----T----GC-----T--A-----C--------A--------G---G-G--A-------G-----  817
SMM.US.04.M946            ----------G----C-AC-----G----------G-------------C----T-----------------------G---G--...-C--C--T---------T----GC-----T--A--G-------------------------G-G----------------  813
SMM.US.04.M947            ----------G----T-AA------GC-----------------------------T-------------------G--G---G--...------C--T---ACT----A--C------T--A--------------------------G---A--G--A-------  814
SMM.US.04.M949            ----------G----C--C------------------------------C----T---GC-A------------------G---G-T...--A--C--T--C------T----G------T--A-----C--G--------------G---G---A-------GG----  805
SMM.US.04.M950            ----------G----C-AC-----G----------G-------------C----T-------------------G--G---G--...-C--C-----------T-----G------T--A--G-------------------------G-G--A----------------  813
SMM.US.04.M951            ----------G----C--C-------------------G----------C----T--T-T-A------------------G---G-T...--A--C--T--CAC----T----G------T--A-----C--G--------------G---G---A-----A-GG----  814
SMM.US.04.M952            ----------G----C-AC-----G----------G-------------C----T--T----------------------G---G--...-C--C-----------T-----G------T--A--G-------C---------------G-G----------------  813
SMM.US.05.D215            ---T-A---TG---GG--C-----G-----T-----------------C----T----TGA-------C--------G--G--GG--...--A------T---A------AA--TC-------A--G--T--------T--------G--AG-T--A-------------  811
SMM.US.06.FTq             ---------TG----T--A------GCC---------------------C----T------------------T--G--G---------...-C--C--T-----------A--T------T--A------C-G-----A-----------G-G--A--G------------  824
SMM.US.11.SIVsmE660_FL10  ----------G----C-AA------GC------C---------------C----T-----------------------G--G---------...------C--T-----------A--C------T--A---------------T----G------G---A--G--A------- 1550
SMM.US.86.CFU212          ---------TG----TC-C-----G----GY------------------C----T----TGAAG--------------G--G---G-R...------C--T-----Y----A-------Y--C--G--------C-------------AG-G--R-----------------  803
SMM.US.x.F236_H4          ----------G----C-GA------GC------C---------------C----T-----------------------G--G---G--...------C--T-----------A--C------T--A--------------T-----------AG---A--G--A------- 1547
SMM.US.x.H9               ----------G----C-AA------GC----------------------C----T-----------------------G--G---G--...------C--T---AC------A--T------T--A--------------T-------------G---A--G--A------- 1033
SMM.US.x.PBJA             ----------G----C-AA------GC----------------------C----T----A------------------G--G--------...------C--T---AC------A--T------T--A--------------T-----------------A--G--A------- 1326
SMM.US.x.PGM53            ----------G----C-AA-------T------C---------------C----T-----------------------G--G--------...------C--T--CAC------A--T------T--A--------------T-------------G---A--G--A------- 1477
SMM.US.x.SME543           ----------G----C-AA------GC------C---------------C----T-----------------------G--G---G--...------C--T-----------A--C------T--A--------------T-------------G---A--G--A------- 1563
SMM.US.x.pE660.CG7G       ----------G----C-AA------GC------C---------------C----T-----------------------G--G--------...------C--T-----------A--C------T--A--------------T-----------AG---A--G--A------- 1549
STM.US.89.STM_37_16       G--T-----TG--A-C-AA-----TGCC---------------------C----T--------------------------G---G--...--C--C--T-----------A-----------A-----C--G-----A--------G--AG-G--------------- 1207
```

```
MAC.US.x.239              GAGCAGAAGTAGTGCCAGGATTTCAGGCACTGTCAGAAGGTTGCACCCCCTATGACATTAATCAGATGTTAAATTGTGTGGGAGACCATCAAGCGGCTATGCAGATTATCAGAGATATTATAAACGAGGAGGCTGCAGATTGGGACTTGCAGCACCCA......CAACCA 1715
Gag                       G__A__E__V__V__P__G__F__Q__A__L__S__E__G__C__T__P__Y__D__I__N__Q__M__L__N__C__V__G__D__H__Q__A__A__M__Q__I__I__R__D__I__I__N__E__E__A__A__D__W__D__L__Q__H__P__.__.__Q__P_
A.CI.88.UC2               -G---------------------------C--------C-----G--------T--C-----A---C-T-----------C--T--A-----A--C---------A--------A-----C--T--A--A--A--G--------TGC-----------......AT---- 1765
A.DE.x.BEN                -G---------------------------C--------C-----G--------T--------A---C-T-----------C--T--C-----A--------A--A--------G-----T--T--A--A--A-----C-----T-C-----------......AT---- 1765
A.DE.x.PEI2_KR_KRCG       -G----------------G----------C--------C-----G--------T--------A---C-T-----------C-----------A--------A--A--------G-----T--T--A--A--A-----------TG----A-------......AT---- 1733
A.GH.x.GH1                -G-----------------------A-----C--------C-----G--------T--C--C--A---C-T-----------C--T--C-----A--------A--A--------G-----C--T--C--A--A-----------TGCA----------......AT---- 1209
A.GM.87.D194              -G---------------------------C--------C-----G--------T--C-----A---C-T-----------C--T--C-----A--------A--A--------A-----T--T-----A--A-----------TGC-----------......AT---- 1209
A.GM.x.ISY_SBL_6669_85    -G-------------G-------------C--------C-----G--------T--------A---C-T-----------C-----------A--G-----A--A--------A-----T--T--A--A--A-----C-----TG-A--A--T---......AT---- 1209
A.GM.x.MCN13              -G-----G---------------------C--------C-----G--------T--C--C--A---C-T-----------C-----------A--------A--A----G--G--CG-T--T--A--A--A-----------TG---A--T---......AT---- 1210
A.GM.x.MCR35              -G-----G---------------------C--------C-----G--------T--C--C--A---C-T-----------C-----------A--------A--A----G--G--C--T--T--A--A--A-----------TG---A--T---......AT---- 1210
A.GW.86.FG_clone_NIHZ     -G---------------------------C--------C-----G--------T-----------C-T-----------C-----------A--------A--A----G--A-----T--T-----A--A-----C-----TG-TGCA--T--C......AT---- 1203
A.GW.87.CAM2CG            -G---------------------------C--------C-----G-----------------A---C-T-----------C-----------A--------A--A----G--G-----T--T--A--A--A-----C-----TGCCA-T--T---......AT---- 1774
A.GW.x.MDS                -G---------------------------C--------C-----G--------T--------A---C-T--C--------C--T--------A--------A--A----G--A--C--T--T--A--A--A-----------TGC------T---......AT---- 1209
A.IN.07.NNVA              -G-------------------------G--C--------C-----G--------T--------A---C-T-----------C-----------A--C-----A--A----G--A-----T--T--A--A--A-----------TG---A--T---......AT---- 1759
A.IN.95.CRIK_147          -G--------------A------------C--------C-----G--------T--C-----A---C-T-----------C-----------A--C-----A--A----G--G-----T--T--A--A--A-----------TGC--A--T---......AT---- 1510
A.JP.08.NMC786_clone_41   -G---------------------------C--------C-----G--------T--C-----A---C-T-----------C--T--C-----A--------A--A--------A-----C--T-----A--A-----A-----TGCA----------......AT---- 1761
A.PT.x.ALI                -G---------------------------C--------C-----G--------T--------A---C-T-----------C-----------A--C-----A--A----G--G-----C--T--A--A--A-----C-----TG-TGCA--T--C......AT---- 1758
A.SN.85.ROD               -G---------------------------C--------C-----G--------T--C--C--A---C-T-----------C-----------A--C--------A----G--G-----C--T-----A--A-----A-----TG---A--T---......AT---- 1208
A.SN.86.ST_JSP4_27        -G---------------------------C--------C-----G--------T--------A---C-T-----------C-----------A--------A--A----G--A-----T--T--A--A--A--------------GCA--A-------......AT---- 1210
B.CI.88.UC1               -G----------A--------C--------A--------A-----T--T-----T--A---------C----------A-----------G--A--C-----A-----T--G--A--A--C--T--------G-----C------CA---------G......AT---- 1731
B.CI.x.20_56              -G--------G----------C--A-----A--G-----G--------T--C--T------------C----------A-----T--C--G--A--C-----A-----T--G--A--A--C--T--------G-----C------CA---A-----G......AT---- 1729
B.CI.x.EHO                -G----------A-----------------A--------G--------T-----T--------------------------A-----A-----G--A--C-----A-----T--G--G--A--T--T-----A--A-----C------CAA-----T--......TCG--- 1727
B.GH.86.D205_ALT          -G----------A-----------------A--------A--------T-----T--A---------C----------A-----A-----G--A--C-----A-----T-----A--A--C--T-----A--A-----C------CA---A-----G......TC---- 1726
B.JP.01.IMCJ_KR020_1      -G----------A--------C--------A--------A--------T-----T------------C----------A-----A--------A--C-----A-----T--G--A--A--T--T--T--A--A-----C------CA---------G......TC---G  872
G.CI.92.Abt96             -G--------G-----------------G------------C---CTA--A-----T--C-----A---C----------A-----A--------A--C-----A-----T-----A--A--T--T--------A-----K-----TC-T--A--T---..........---A- 1128
AB.CM.03.03CM_510_03      -G-----G-----A--------C--A-----A--------A-----T--T-----T--A---------C----------A--G--------G--A--C-----A-----T-----A--A--T--T--T--A--A-----C------CAA--A-----G......TC---- 862
H2_01_AB.CI.90.7312A      -G----------------------------A-----G--A--T-----T-----T------------C----------A-----A-----G--A--C-----AG-C--T--G--A--A--C--T--------G-----C------CA---------G......GTG--- 1732
H2_01_AB.JP.04.NMC307_20  -G------A----A----------------A-----G--A--T-----T-----T------------C----------A-----A--C--G--A--C-----AG-C--T-----A--A--C--T--A-----G-----C------CA---A-----G......GT---- 1732
H2_01_AB.JP.07.NMC716_01  -G------A----A----------------A-----G--A--T-----T-----T--C---------C-------C--A-----A-----G--A--------AG-C--T-----A--A--C--T--A-----G-----C------CA---------G......AT---- 1731
H2_01_AB.JP.08.NMC842_10  -G------A----A----------------A-----G--A--T-----T-----T--C-----A---C----------------A--C--G--A--C-----AG-C--T-----A--A--C--T--A-----G-----C------CA---A-----G......AT---- 1737
U.CI.07.07IC_TNP3         ----------------------------T-A--------C-----A-----C--T--C-----A---C----C-----A--R--A-----G--A--C-----A--C--T-----A--A--T--T-----A--A-------------A-A---A-T--T..........---T 1233
U.FR.96.12034             -----------------------------C-----G--C-----G--------T--------A----------C--------A--------T--------A--------G--A--A--T--T--T--A--A-------------ACA-----------......A-C-A- 1244
MAC.US.x.17EC1            ---------------------------------------------------------------------------------------------------------------------------------------------------------......------ 1715
MAC.US.x.251_1A11         ------------------------------------------------------------------------------------------------------C--C--------------------------------------------......------ 1714
MAC.US.x.251_32H_PJ5      -----------------------A--------------C--------------------------------------------------------------C---------T------------------------------------......------ 1716
MAC.US.x.251_BK28         --------------------------------------C-----------------------------------------------------------------------T------------------------------------......------ 1703
MAC.US.x.MM142_IVMXX      -------------T------------------------C---CT-----------------------------------------C----------------------T------------------------------------......----A- 1213
MNE.US.82.MNE_8           --------------------------------------C---------------------------A--------A--------------A--------------------------------------------------......----A- 1195
MNE.US.x.MNE027           --------------------------------------C---T-----------------------A--------A--------------A-----------------------------C-----------------......----A- 1195
SMM.CI.79.SIVsmCI2        -G-----------------------A-----A--------C-----T--------T--C-----A--------------------A--------A--A--------A--T-AG--A--C--T--T-----A--A--------------------T--T......A-TGT-  811
SMM.LR.89.SIVsmLIB1       -G-----------------------------A--------C-----T--G--------C-----------------C-----------A-----G--A--C-----A-----A-----A--A--T--T--A--A--A--------A--T--------T---............  800
SMM.SL.92.SIVsmSL92A      -------------C-----T--------GT----------A--------------------A--C-T-------T--G--G--C-----A--A-----A--A--T-----A--C--C--T--A-----A-----C---------------T--G.......-G---  808
SMM.SL.92.SL92B           -------------------------------A--------C--T--A------------------C-T-------A--T--A-----G--A--C-----A-----------A--A--T--T-----A--A-----C------C-A--A------AGAGGA----AG 1155
SMM.US.04.G078            --------------------------------------C--------T--------C--C--A---C----------A--G--A--------A--C-----A-----A-----A-----------T--A--A--A-----C-----T-----A--T---.......-G---  981
SMM.US.04.G932            -------G--------------------G---------C-----A------------------C----------A-----A--------A--------A-----------A--C-----T--------------C------C-A--------G.......-G---  973
SMM.US.04.M919            -G-----G-----------------A--G--T-------C-------------------C-----A---C----------A-----A-----G--A--C-----A-----T-----A--------T--A--A-----T--------T--A--A-----G.......-G---  978
SMM.US.04.M922            -G-----G-----------------A--G--A-------C-----T--------T-----------A---C----------A-----A-----G--A--C-----A--G-T-----A--------T--A--A-----T-----------A--A-----G.......-G---  979
SMM.US.04.M923            -G-----G-----A-----------A--G--A-------C-----T-----------C-----A---C----------A-----A-----G--A--C-----A-----T-----A--------T-----------T--------T--A--A-----G.......-G---  978
SMM.US.04.M926            -G----------------T-----------A-------C--T-------------C-----A---C-G----------G--A--------A--C--------G--------A--C-----T-----A-----G----------------T---.......-G---  978
SMM.US.04.M934            ------------------C-----------------C--T-----T-----T--C-----A------------------G--A--------A--C-----A--G--------A--------T-----A----------------------T---.............  976
SMM.US.04.M935            -G-----G-----------------A--G--A-------C-----T-------------------A---C----------A-----A-----G--A--C-----A--G-T-----A--------T--A--A-----T--------T--A--A-----G.......-G---  979
SMM.US.04.M940            -------G---------------------T---------C--T-----T--------C--C--A---C----------A--G--A--------A--C-----A-----------A--A-----T--A--A--A-----C-----T--------T---.......-G-A-  981
SMM.US.04.M946            ------------------C-----------A-------C--T-----T--------C-----A-------------------G--A--------A--C-----A--G--------A--------T-----A-----------------------------.......-G---  977
SMM.US.04.M947            -G-----------------------A-------------C-----T-----------C-----A---C----------A--G--A-----G--A--C-----A-----T-----A--------T--A--A-----T--------------A--A--------.......-G---  978
SMM.US.04.M949            -------G---------------------T-A-------C--T-----G--------C--C--A---C----------A--G--A--------A--C-----A--G-A-----A--A-----T--A--A--A--------------T--------T---.......-G---  969
SMM.US.04.M950            ------------------C-----------A--G-----C--T-----T--------C-----A------------------G--A--------A--C--------------------A--------T-----A-----G----------------T---.......-G---  977
SMM.US.04.M951            -------G-----------------------------C--T-----T--------C--C--A---C----------A--G--A-----G--A--C-----------A-----A--A-----T--A--A--A-----C-----T--------T---.......-GGT-  978
SMM.US.04.M952            ------------------C-----------A-------C--T-----T--------C-----A-------------------G--A-----G--A--C-----A--G--------A--------T-----A-----G----------------T---.......-G---  977
SMM.US.05.D215            -G---------------------------T-A-------C--------G------G-C--C-----C-----------A-----A-----G--A--C-----A--C--T--G--A--A--C--T--A--A-----T--------TC---A--T---.......-G---  975
SMM.US.06.FTq             ----G--G-----------C--------A--------C--T-----A-----T-----------A------------------G--T-----G--A--C-----A-----A-----A--C-----T-----A--G--T--------C-A--A-------...CAG-CGATG  991
SMM.US.11.SIVsmE660_FL10  -G-----G-----------------A--G--A-------C-----T-----------T-C-----A--C----------A-----A-----G--A--C-----A-----T-----G--------T--A--A-----C--------T--A--A--------.......-CG---  1714
SMM.US.86.CFU212          -G-----G----------G-----------R-------C--T--T--T-----T--C-----------C-------------A--C--R--A--C--------C--A--G--A--C--C--T--A-----A--T--------C---A--T---............. 967
SMM.US.x.F236_H4          -G-----G-----------------A--G--A-------C-----T-----------T-----------A---C----------A-----A-----G--A--C-----A-----T-----G--------T--A--A-----C--------T--A--A-----G............ 1711
SMM.US.x.H9               -G-----G-----------------A--G--A-------C-----T-----Y--T--C-----A---C----------A-----A-----G--A--C-----A-----T-----A--------T--A--------T--------T--A--A-----R.......-G--- 1197
SMM.US.x.PBJA             -G-----G-----------------A--G--A-------C-----T-----------C-----A---C----------A-----G-----G--A--C-----A-----T-----A--------T--A--A-----T--------T--A--A-----G.......-G--- 1490
SMM.US.x.PGM53            -G-----G-----------------A--G--A-------C-----T-----------C-----A---C----------A-----A-----G--A--C-----A-----T-----A--------T--A--A-----T--------T-----A-----G.......-G--- 1641
SMM.US.x.SME543           -G-----G-----------------A--G--A-------C-----T-----------C-----A---C----------A-----A-----G--A--C-----A-----T-----G--------T--A--A-----C--------T--A--A-----G.......-G--- 1727
SMM.US.x.pE660.CG7G       -G-----G-----------------A--G--A-------C-----T--------T--C-----A---C----------A-----A-----G--A--C-----A-----T-----G--------T--A--A-----C--------T--A--A-----G............ 1713
STM.US.89.STM_37_16       -T--------G-------------------A--------C--T-------------C-----AC---------------G--G--------A--C-----A-----A-A---G--C-----T-----A-----T----------A--------T---...CAA-C---G 1374
```

```
MAC.US.x.239                 GCTCCA...CAACAAGGACAACTTAGGGAGCCGTCAGGATCAGATATTGCAGGAACAACTAGTTCAGTAGATGAACAAATCCAGTGGATGTACAGACAACAGAACCCCATACCAGTAGGCAACATTTACAGGAGATGGATCCAACTGGGGTTGCAAAAATGTGTCAGAAT 1882
Gag                          _A__P__.__Q__Q__G__Q__L__R__E__P__S__G__S__D__I__A__G__T__T__S__S__V__D__E__Q__I__Q__W__M__Y__R__Q__Q__N__P__I__P__V__G__N__I__Y__R__R__W__I__Q__L__G__L__Q__K__C__V__R__M
A.CI.88.UC2                  -GC--CTTA-C-GC---G--G-----A--C--AAGG--G--T--C--A----------A--CA----------G-G-------------T--G--G--A--T---G-G--G-----G-----C-----A----------GA-A---C----G--G---------- 1935
A.DE.x.BEN                   -GC--CTTA-C-GC-----G--C--A--C--AAG---G--T--C--A----------A--CA------------G-------------T--G-C---A--T---G----G-----G-----C-----A-----------A-A---C-------G---------A 1935
A.DE.x.PEI2_KR_KRCG          -GC--CTTG-C-GCG--G--G-----A--A--AAG---G--T-----A-----G-----A--CA-------A-----G------------TT---GC---A--T--T-----------G-----C--T------------GA-A--AC----G--G--C-----G-- 1903
A.GH.x.GH1                   -GC--CTTA-C-GC---G--G-----A--C--AAGG--G--T--C--A----------A--CA-------A-----G-------------T--G-C---A--T---G-G--G-----G-----C-----A----------GA-A---C-A--G--G--------G-- 1379
A.GM.87.D194                 -GC--CTTA-C-GC---G--G-----A--C--AAGG--G--T--C--A----------A--CA-------------G-------------T--G----CA--T---G-G--G-----G-----C-----------------GA-A---C-A--G-----------G-- 1379
A.GM.x.ISY_SBL_6669_85       -GC--CTTA-C-GCG--G--G--C--A--T--ACG------T--C--A-----G-----A--CA-------G-----G---G-A--------T--G---G-A--T--TG-----------A-----C--T------------GA-A--AC----G--G---------- 1379
A.GM.x.MCN13                 -G---CTTA-C-GCG--G--G-----A--A--AAG---G--T--C--A-----G-----A--CA-------------G------------TT--G-C---A--T---G--------G--A-----C--T------------GA-A--AC----G--G--C-----G-- 1380
A.GM.x.MCR35                 -G---CTTA-C-GCG--G--G-----A--A--AAG---G--T--C--A-----G-----A--CA-------------G------------TT--G-C---A--T---G--------G--A-----C--T------------GA-A--AC----G--G--C-----G-- 1380
A.GW.86.FG_clone_NIHZ        -GC--CTTA-C-GCG--G--G-----A-----AAG---G--T--C--A-----G-----A--CA-------A--G--G------------TT----C---A--T--TG-------------------C--T----------T--GA-A--AC-------G--C-----G-- 1373
A.GW.87.CAM2CG               -GC--CTTA-C-GCG--G--G--C--A--C--AAG---G--T--C--A----------A--CA-------A-----G-----A-------TT--GGC---A--T---G----------G-----C--T--A----------GA-A---C----G--G--------G-- 1944
A.GW.x.MDS                   -GC--CTTA-C-GCG--G--G-----A--A--AAG---G--T--C--A-----G-----A--CA--------------------------T---GGC-----T--TG-------G--A-----C--T--A----------GA-A---C----G--G--------G-- 1379
A.IN.07.NNVA                 -GC--CTTA-C-GC---G--G--C--A--A--AAG---G--T--C--A----------A--CA-------A-----G---G---------TT---GC---A--T--TG-------G--A-----C--T--A----------GA-A---C-------G--------G-- 1929
A.IN.95.CRIK_147             -GC--CTTA-C-GC-----G------A-----AAG------T--C--A-----G-----A--CA-------G-----G-------------TT---GC---A--T--TG-------G--A-----C--T--A----------GA-A--AC----G--G--------G-- 1680
A.JP.08.NMC786_clone_41      -GC--CTTA-C-GC---G--G--C--A--C--AAGG--G--T--C--A----------A--CA-------A-----GC-------------T--G--G-CA--T---G-G--G-----G-----C--------------A--GA-A---C-A---------------G-- 1931
A.PT.x.ALI                   -GC--CTTA-C-GC---G--G-----A--A--AAG---G--T--C--A----------A--CA-------A-----G-------------T---G-C--G---T--TG-G--------G-----C--T--A----------GA-A---C-A--G--G--------G-- 1928
A.SN.85.ROD                  -GC--CTTA-C-GCG--G--G-----A-----AAGG-----T--C--A-----G-----A--CA-------A-----G------------TT--G-C---A--T--TG-----------A-----C--T--A----------GA-A--A-----G--G--------G-- 1378
A.SN.86.ST_JSP4_27           -GC--CTTA-C-GCG--G--G--C--------AAGG-----T--C--A-----G-----A--CA-------A--G--G------------TT--G-C---A--T--TG-----------A-G---C--T--A----------GA-A---C-A--G--G--C-----G-- 1380
B.CI.88.UC1                  -G---CTG-C-GC--------------C--CAG---G--------A--------C--C--CA-------G---------A--------------GGC---A--T--TG-T--G-----G-----G-----------A---------T---A-A---C----G----------C-G-- 1901
B.CI.x.20_56                 -G---CTG-C-GC--------------C--CAG---G--------A--------C--C--CA-------G---------A--------------GGC---A--T--TG-T--G-----G-----G-----------A---------T---A-A---C----G----------C-G-- 1899
B.CI.x.EHO                   -GC--ATG-C-GC---G-----C-----A--AAG---G-----C--A--G-----C--C--CA-------A-----G--A----------------CC--A--T--TG-C-----G--A-----G--A-----C--T--A--G-----T--GT-A---C-A--G----------C-G-- 1897
B.GH.86.D205_ALT             -GC--ATG-CGGC--------------C--AAG---G--------A--------C--C--CA-------G-----G--A--------------GGCC--A--T--TG-C-----G--A-----G--A--------T--A---------T---T-A--A----G----------C---- 1896
B.JP.01.IMCJ_KR020_1         -GC--ATG-C-GC------G--------C--AAG---G-----C--A--------C-----CA----G--G--G--G--A--------------G-CC--A-----T--C---------A-----------A--------T--A---------T--GT-A---C----G----------C-G-- 1042
G.CI.92.Abt96                -GG--GCCG-C-GC-------M-C-----A---A--------------A-----G--C-------A--A----A---------T-------CAC----G--G--A-----T-------------A--T--Y--T--A-----------RT-A--AC-T----------------G-- 1298
AB.CM.03.03CM_510_03         -GC--CTG-CGGC------G-----A--C--AAGG-----------A--G-----C--C--CA-------G-----G--A--------------GGGC--A--T--T--C-----G--A--------G--A--------T--A---------T---A-A---------G----------C-G-- 1032
H2_01_AB.CI.90.7312A         -G---CTG-C-GC--------------C--CAG------------A--G-----C--C--CA-------G-----G--A--------------GGGT--A-GT--T--C--------G-----------G--------T--A---------T--G--A--A-----G--G-------C-G-- 1902
H2_01_AB.JP.04.NMC307_20     -G---CTG-C-GC-----------A--C--CAG------------A--G-----C--C--CA-------G---------A--------------GGGT--A--TT----C-----G--G--------G--G--------T--A---------T---T-A--A-----G-------------G-- 1902
H2_01_AB.JP.07.NMC716_01     -G---CTG-C-GC---------C--A--C--CAG------------A--G-----C--C--CA-------G---------A--------------GGGT--A--TT--G-C--------G--------------G--------T--A---------T---T-A--A-----G-------------G-- 1901
H2_01_AB.JP.08.NMC842_10     -G---CTG-C-GC-----------A--C--CAG------------A--G-----C--C--CA-------G---------A--------------GGGT--A--TT--G-C-----G--G--------G--G--------T--A---------T---T-A--A-----G-------------G-- 1907
U.CI.07.07IC_TNP3            -GC--CTA-C-GC---G--G-----A--A--TAGG--G-----C--A--------C--C--CA-G--G--A-----G----A--------T--G-----A-----T-----G-----A-----G-----A-----C-----------------T---T-A--AC-C-------------G-- 1403
U.FR.96.12034                -GC--CCA-C-GC---G---------A-----AAG---T--T-------------------CA----G--A--G--G--A---------------G-CG--A--T--A-----G--G--T-----G--G--T-----C--TC-----------------T-----CC-A-----------T----- 1414
MAC.US.x.17EC1               ------...------------------------------------------------------------------------------------------------------------------------------------------------------------------ 1882
MAC.US.x.251_1A11            ------...----------G-------------------------------------------------------------------------------------------------------------------------------------------------------- 1881
MAC.US.x.251_32H_PJ5         ------...----------G-------------------------------------------------------------------------------------------------------------------------------------------------------- 1883
MAC.US.x.251_BK28            ------...----------G-------------------------------------------------------------------------------------------------------------------------------------------------------- 1870
MAC.US.x.MM142_IVMXX         ------...----------G-----------------------------------------A-------A-------------------------------------------------------------------------------------------------- 1380
MNE.US.82.MNE_8              ------...----------G----------------------C-----------------A--------------------------------------------------------------------------------------------------------------- 1362
MNE.US.x.MNE027              ------...----------G----------------------C-----------------A-------------------------------------------------------------T----------------------------------------A--- 1362
SMM.CI.79.SIVsmCI2           -GG---TTG-CGGC------GT-G--A--A---AGT--C-----C-------------A---A---G-----G-------A----------G-----A-----TG-------T--G-----C--T-----------T--------C--A--G-----------G--  981
SMM.LR.89.SIVsmLIB1          AGC---ATA-C-GC---G-----------A---CG---T--C-----A--T--C--C------A---G-----G-----A-----------G-----A-----T-------T--A--------T------------------AA--C--------------------  970
SMM.SL.92.SIVsmSL92A         -G----CTC-C-GC---G-----C--------AAG------T-----A-------------CA-------A-------A---------------GGCTGCC---------C--G--G-------C---C--C------------T-A--T--A-----------T-----  978
SMM.SL.92.SL92B              C-GG-GCAA-CGGCT---GGG--A-----A--A-----------C--A-----G-------CA-CCC-AG------G--AG-A--------T--GGC---A-----AG-C-----G--AG----C--T--A--G--------G-----AC-C---------C-----G-- 1325
SMM.US.04.G078               -GC---CTA-C-GC---G-----C--A--A--AAGG--C--T-----A--------C--C---A----------G-----A------------------G--A----------G--T--A-----G--T--A-----------T---T----AC----G-----------G-- 1151
SMM.US.04.G932               -G----ATA-C-GC-----G--C-----C--AAG------------------G--C------A---G--A--------------------T-----G----T----------------T-----------T--C-----A--------------T--A--------------T----- 1143
SMM.US.04.M919               -G----ATA-C-GC---------C--A--C--AAG--------------------T------A------A--G-----------------T--G--G-------T--------G--A-----------------------------------T-----A-----G----------- 1148
SMM.US.04.M922               -G----CTA-C-GC-------------A-----AAG--------------------T------A---G--A--G----T------------T--G----------------------G--A-----C--------------------------T-----A-----G--------G-- 1149
SMM.US.04.M923               -G---GCTA-C-GC------G--G--A-----AAG------C--------------C-----CA-------A-------T------------T--G--G--A-----T------------T-----T-----C-----A--G--------GT-A---C----G--------A--G-- 1148
SMM.US.04.M926               -GC---TTG-C-GC---G--G--C--A--A--AAG---------C------------T-----CA-------C--G----T------------T--G---------------------G--A-----C-----A--G-----T-----T--C--------G------------ 1148
SMM.US.04.M934               -G----TTA-C-GC---------C-----A--AAG---G------------------C------A---G--A-----------A--------T--G--------T--T-------G--A-----------------------G--T---C-A-----G------------ 1146
SMM.US.04.M935               -G----ATA-C-GC------------A--T--AAG--------------------T------A---G--A--G----T------------T--G----------------------G--A--T--C-------------------T-----A-----G--------G-- 1149
SMM.US.04.M940               -GC--GTTA-C-GC------G--C-----A--AAGG--C--T--C--A--------C--C---A----------------A-----------T--G--G--A-----------G--T--G--------T--A-----------T-----C----G--G--------G-- 1151
SMM.US.04.M946               -G----TTG-C-GC-----------------A--AAG---------------------C------A---G--A--G----------------T--G-----------T-------G--A--------------------G--T---C-A-----G--------G-- 1147
SMM.US.04.M947               -G----ATA-C-GC------------A--T--AAG--------C--------------T-----CA----------------T--A--------T--G--G--A--T--T---------------T-----------A--G----------T-A---C--------------A----- 1148
SMM.US.04.M949               -GC---CTA-C-GC---G--G--C-----A--AAGG--C--T--C--A--------C--C---A----------------A-----------T--G--G--A-----------G--T--G--------T--A-----------T-----C----G--G-----T--G-- 1139
SMM.US.04.M950               -G----TTG-C-GC---------C-----A--AAG--------C--------------C------A---G--A--G-----------------T--G-----------T-------G--G--------------------G--T-----A--G--G----------- 1147
SMM.US.04.M951               -GC---CTA-C-GC---G--G-----A--A--AAG---C--T--C--A--------C--C---A----------------A-----------T-----G--A----TG----G--T--G--------T--A-----------T----AC----G--G--------G-- 1148
SMM.US.04.M952               -G----TTG-C-GC---------C-----A--AAG--------------------C------A---G--A-------T------------T--G-----------T-------G--G--------------------------T-----A-----G----------- 1147
SMM.US.05.D215               -GC--GCTA-C-GC------G--C-----A--AA---G--T--C------------C--C---A---G-----G-------A----------G--G--------G-----------A--------A--G--------T------A-----------T----- 1145
SMM.US.06.FTq                -GC---ATA-C-GC---G--G--C--A--C---AGG--------C------------T-----CA------G--G--G--AG-A--------T--G--G--A--T--T-----------T-----C-----A-----------G--T---C-A-----------T----- 1161
SMM.US.11.SIVsmE660_FL10     -G----CCA-C-GC---G---------A-----AAG--------C------------T-----A--------------------A----------G-----A--------------------------------T--A--G--------T-A---C----G--------A----- 1884
SMM.US.86.CFU212             -G----ATA-C-GC------G--------C--AAG---R-------A--------C--C--CA---G--A-------T--------------G--G-----T--G-----G-----A-----C-----------T-----T-----------------G--G-- 1137
SMM.US.x.F236_H4             -G----CTA-C-GC---G---------A-----AAG--------C------------T-----A--------------------A----------G-----A--------------------------------T--A--G--------T-A---C----G--------A----- 1881
SMM.US.x.H9                  -G----ATA-C-GC-------------A-----AAGR-----C--C------------T-----CA--S---C-------T--A--------Y--G--G--A-----T------------------C-----R--G--------T-A---C----G-----C--A----- 1367
SMM.US.x.PBJA                -G---GATA-C--C-------------A-----AAG------C--C------------T-----CA-------C-------T--A--------T--G--G--A-----T-----------T-----C-----A--G-----------A---C----G--------A----- 1660
SMM.US.x.PGM53               -G----ATA-C-GC---------C--A--T--AAG--------------------T------A---G--A--G----T------------T--G--G----------------------G--A-----------------------------T-----A-----G----------- 1811
SMM.US.x.SME543              -G----CTA-C-GC---G---------A-----AAG--------C------------T-----A------G---------A----------G-----A-------------------------------T-----G--------T-A---C----G--------A----- 1897
SMM.US.x.pE660.CG7G          -G----CTA-C-GT---G---------A-----AAG--------C------------T-----A------------------A----------G-----A--------------------------------T--A--G--------T-A--------G--------A----- 1883
STM.US.89.STM_37_16          -G---GCTG-C-GC---G---------A--A--AAGC--G--------A-----G--T-----C---CC---G--------T--A---------C-G--G--A--T--T-----G-----A-----C-------------T-----A---C-A---------------- 1544
```

```
MAC.US.x.239                GTATAACCCAACAAACATTCTAGATGTAAAACAAGGGCCAAAAGAGCCATTTCAGAGCTATGTAGACAGGTTCTACAAAAGTTTAAGAGCAGAACAGACAGATGCAGCAGTAAAGAATTGGATGACTCAAACACTGCTGATTCAAAATGCTAACCCAGATTGCAAGCTAG 2052
Gag                         __Y__N__P__T__N__I__L__D__V__K__Q__G__P__K__E__P__F__Q__S__Y__V__D__R__F__Y__K__S__L__R__A__E__Q__T__D__A__A__V__K__N__W__M__T__Q__T__L__L__I__Q__N__A__N__P__D__C__K__L__
A.CI.88.UC2                 ---C--------C---G-CT----C-----G--G--A--------AT----C----------------------T-----C-----G-----G--AG----CC-G---------A-----------C-----------G-A--G-----C--------C--T---T--- 2105
A.DE.x.BEN                  ---C--------T-----CT----CA-------G--A--------A--G--C--A------------------------C--G--G--------A-----CC---------A-----------C-----G-----A--A--G-----C--------C-----T--- 2105
A.DE.x.PEI2_KR_KRCG         ---C--T-----C-----C-----C--------G--A-----G-----G--C--A-----------T--A----------CC----G--------A-----CC---------A-----------C-----------G-A--G-----C--------C--T--AT--- 2073
A.GH.x.GH1                  ---C--------T-----CT----C-----G--G--A-----G--A--G--C-----------G-----------T-----C--G--G--------A------C-G---------C--------C-----G-----A--A--G-----C--------C-----T--- 1549
A.GM.87.D194                ---C--------T-----CT-------G--G--G--A--------AT-G--C----------------A-----------CC----G--------A-----CC-G--T-----A-----------C-----G-----A--A--G-----C--------C-----T--- 1549
A.GM.x.ISY_SBL_6669_85      ---C--T-----C-----------CA---------A--------T-G--C--A--------G--T--A-----------C-----G--------------------G-------------C--G--G----AG-G---...T-G--------C--T---T--- 1546
A.GM.x.MCN13                ---C-----G--C-----C-----CA---------A-----G--A-----C--A--T--------T--A-----------C--G--G--------A------C-------G-------------C--G-----A--AG-A--G-----C--------C--T--AT--- 1550
A.GM.x.MCR35                ---C-----G--C-----C-----CA---------A--------A-----C--A--T--------T--A-----------C--G--G--------A------C--A--------------------C--G-----A--AG-A--G-----C--------C--T--AT--- 1550
A.GW.86.FG_clone_NIHZ       ---C--T-----C-----C-----CA----T-----A-----G--A-----C--A-----------T-------------C--G--G--------A------C---------A-----------C-----------A--A--G-----C-----------T--AT--- 1543
A.GW.87.CAM2CG              ---C-----G--C-----C-----CA-------G--A-----G--T-----C--A-----------T--A-----------C--G--G--------A------C---------A-----------C-----------AG-G--G-----C--------C--T--AT--- 2114
A.GW.x.MDS                  ---C--T-----C-----C-----C--------G--A-----G-----G--C--A-----------T--A-----------C--G--G--------A---------------G-A--------------G--------A--A--------C--------C--T--AT--- 1549
A.IN.07.NNVA                ---C--T--G--C-----C-----CA-------G--A--------A--G--C--A-----------T--A--------G--C--G-----------------C---------------------C------T-A-----A--------C-----------T--AA--A 2099
A.IN.95.CRIK_147            ---C--T--G--T-----C-----CA-------G--A--------A--G-----A-----------T--A--------G--C--G-----------------CC--------------------C------T-A---G-A--------C-----------T--AT-G- 1850
A.JP.08.NMC786_clone_41     ---C--------T-----C-----CA-------G--A-----G-----G--C----------------------T-----C--G--G--------A-----CC-G-----G-----C--------C-----G-----A--A--G----------------C-----T--- 2101
A.PT.x.ALI                  ---C--------C-----C-----CA---------A-----G--------C--A-----------T--A-----------C-----G--------A------C---------------------------------G-A--G-----C--------C-----AT--- 2098
A.SN.85.ROD                 ---C-----G--C-----C-----CA-------G--A-----G-----G--C--A-----------T--A-----------C--G--G--------A------C-------G-------------C-----------AG-A--------C--------C--T--AT--- 1548
A.SN.86.ST_JSP4_27          ---C--------C-----C-----CA-------G--A-----G--------C-----T--------T--A--------G--C--G--G--------A------C---------A-----------C-----------AG-G--G-----C--------C--T---T--- 1550
B.CI.88.UC1                 ---C--T--T--C-----AT----CA----G--G--A--------A--T--C--A----C-----T--A-----------C---C-G-----G--A----CC-------G--A-----------A-----------A-----G-----------------------T--- 2071
B.CI.x.20_56                ---C--T--C--C--T--AT-------------G-----G--G--T-C--C--A-----------T---------------CC--C-G-----G--A--G--CC-------G--A-----------A--G--------A-----G-----C-----------------T--- 2069
B.CI.x.EHO                  ---C--T--T--T--T--A--G--CA----G--------------A--C--C--A-----------T--A-----------C---C-G-----G--A----CC---------A-----------A-----------------G----------------------AT--- 2067
B.GH.86.D205_ALT            ---C--T--T--C-----AT----CA----G--G--A-----G-----C--C--A-----------T--A-----------C---C-G--------A----CC-------G--A-----------A-----------------G-----------------------T--- 2066
B.JP.01.IMCJ_KR020_1        ---------T--C--T--A--G--CA----G--G--A-----G--A--C--C--A------------------A-------------C-G--------A----CC-G-------A-----G-----G-----------------G--------------C-----AT--- 1212
G.CI.92.Abt96               ---Y--T-----------C---------------------------A--C-----TCT-------T--A--------G---C-R-----------A--T---C---------A-----------A--G-----A--------------C-----------T--AT--- 1468
AB.CM.03.03CM_510_03        ---C--T--T--C-----AT-G--CA----G--G--A--------A--C--C--A--T-------T--A-----------C---C-G-----G--A----CC-------G--A-G---------A--------A--G---G-----------------------T--- 1202
H2_01_AB.CI.90.7312A        ---C--T--T--T--T--A-----------------------------A--C--C--A-------G--T--A--------G--C---C-G-------AG-G--CA-G--C--G-GAGCG--------AG-G-----A---G-C--G-------------------------- 2072
H2_01_AB.JP.04.NMC307_20    ---C--T--T--T--T--A-----------------------------A--C--C--A----------T--A-----------C---C-G-------AG----CA-G--C--G-GAGCA--------AG-G-----A-----C--G-------------------------- 2072
H2_01_AB.JP.07.NMC716_01    ---C-----T--T--T--A-----------------------------A--C-----A----------T--A-----------C---C-G---------------CCAG--C--G-GAGCA--------AG-G------T----C--G-------------------------- 2071
H2_01_AB.JP.08.NMC842_10    ---C--T--T--T--T--A-----------------------------A--C--C--A----------T--A----------CC--C-G-------AG----CA----C--G-GAGCA--------AG-------A---G-C--G-------------------------- 2077
U.CI.07.07IC_TNP3           ------T-----------C--G--C----------------------T-T----------------T-----------G--C--G--G-----G-------CC-G-----G--------------A--------------------------A--------C--T------- 1573
U.FR.96.12034               ---C--T-----T--T--CT-----A----G-----------G-----------A--T-------T--A---------------G----------A------C-------G--A-----------------G--G-----------G--------------C-----A--C- 1584
MAC.US.x.17EC1              ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2052
MAC.US.x.251_1A11           ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2051
MAC.US.x.251_32H_PJ5        ------------------------------------A--------------------------------------------C-------------A----------------------------------------------------------------------T--- 2053
MAC.US.x.251_BK28           ---------------------------------------------------------------------------------C-------------A------------------------------------------------------------------------------ 2040
MAC.US.x.MM142_IVMXX        -----------------------------------------------------------------------------------------------A-----C------------------------------------------------------------------------- 1550
MNE.US.82.MNE_8             -------------G-------------------------------------------------------------------C-------------A-----C------------------------------------------------------------------------- 1532
MNE.US.x.MNE027             -----------------------A--G------------------------------------------------------C-------------A-----C-----------------------------------------------------------------------G- 1532
SMM.CI.79.SIVsmCI2          ------T-----C-----C--------G--------C--C--G-----T------------------------------G--C--G-----T--G--A-----C--G-----G--A-----------A--G---------------------------C--T------- 1151
SMM.LR.89.SIVsmLIB1         ------T-----T--------G-----G--G-----------G-----C------------------------------------G-----G--A------C-G---------A----------------------------------C--T-----C--T--AT-G- 1140
SMM.SL.92.SIVsmSL92A        ---C--T---GTC-----C--G---A-C--G--------C--------T--------T--------T--A--------G--C--G-----------A------C-----------------------A--G--G--------C--------------T-----T-----G- 1148
SMM.SL.92.SL92B             ------------------C--G-----T--------T-----G-----C-----A--T--------T------------------G-------G-----------C-G-----G--------------A---------------G--C--C-----T--------A---- 1495
SMM.US.04.G078              ------T----TG-----C-----C--G--------------------------A--T--C---------------T----G-----------A-----CC-G-----G--A-----------------------------------------------T--AT-G- 1321
SMM.US.04.G932              ---C--T--------------------G--------------G-------C--A----C----------------------C------------A------C-------G------------------G--------A-----------------------T---T--- 1313
SMM.US.04.M919              ---C--T---GT-----------CA------------G-------G--C---------------------------------C----------G--A------C---G--G-----------------------------------------------T--AT-G- 1318
SMM.US.04.M922              ---C--T-----------C--G--------------------------G--C---------------------------------C---------------------C-----------------C------------------------------------T--AT-G- 1319
SMM.US.04.M923              ---C--T------------T-G-----G--------A--------T-----C--A-----------T--A--------G---C------------A-----CC-------G--A-----------------------------------C-----G--C--T--AT--- 1318
SMM.US.04.M926              ---C--T---GT---------G--CA-----------G--G------------------------------------G-C---------------------C-------G-----------------------------C--G-----C-----------T---T-G- 1318
SMM.US.04.M934              ------T--------------G--CA-----------G-----A---C-------------------------------G-C------------A------C-------G------------------------------------C-----------T--AT-G- 1316
SMM.US.04.M935              ---C--T--------------G--CA-------------------G--C----------------------------------C--G--------------C-G---------------C--------------------------------------T--AT-G- 1319
SMM.US.04.M940              ------T---GTG-----C-----C--G--------------------------A-----C----------------GT---G-----------A-----CC-G-----G-------------------G---------------C-----------T--AT--- 1321
SMM.US.04.M946              ------T--------------G--CA-----------------AA---C-----------------------------G-CC-----------A------C-------G----------------------------------------G-----C-----------T---T-G- 1317
SMM.US.04.M947              ------T------------T-------G--------A-----------C-----------T--A--------G---C------------A------C-------G--A----------------------------------------------AT-G- 1318
SMM.US.04.M949              ------T-----G-----C-----C--G--------------------------A-----------------------G-----G---------A------C-G---------A-------------G--------------------------T--AT--- 1309
SMM.US.04.M950              ------T--------------G--CA-----------G----------C------------T------------------C-------------A------C-------------------------------------------G-----C-----------T-----G- 1317
SMM.US.04.M951              ------T----TG-----C-----CA----------------------------A-----C-----------------GT-C--G----------A-----CC-G-----G--A-------------G------------------------------T--AT--- 1318
SMM.US.04.M952              ------T--------------G--CA-----------G----------C-----A----------------------C-----G---------A------C-------G-------------------------------------G--C--C-----------T---T-G- 1317
SMM.US.05.D215              ------------T------------A----G--------T---A----T--------------------A--------G-----G---------A-----CC------------A-----------A--G-------------------C--------C--T---T-G- 1315
SMM.US.06.FTq               ------T-----------------------G--G--A--------A--T-----GC--------T--A-----------C-----G-----G--A------C-G----------------------------------G--------------------A-----------T--AG-GA 1331
SMM.US.11.SIVsmE660_FL10    -----------------T-------G--------A-----------------A-----C-----T--A-----T-----C-----------G--A------C-G-------------------C-----------------------------------------T--A--G- 2054
SMM.US.86.CFU212            ------T--------------------G-----G--A--C--------------ATCT-------------------------G--R-----R--A------C-R-----W--------------------G-----------------C--C----------Y--T--AT-G- 1307
SMM.US.x.F236_H4            -----------------T-------G--------A-----------------A-----C-----T--A-----------C-----------G--A------C-C-------------------------C-------------------------------------T--A---- 2051
SMM.US.x.H9                 ------T------------T-G-----G--------A--------T-----A-----------T--A--------G---C-------------A-----CC-G---S-G--A-----------------------------------C-----G--C--T--AT-G- 1537
SMM.US.x.PBJA               ------T------------T-G-----G--------A-----------------A-----------T--A--------G---C-------------A-----CC-G-----G--A-----------------------------------C-----G--C--T--AT-G- 1830
SMM.US.x.PGM53              ---C--T--------------G--C-------------G--------G--C-------C-----T------------------C-------------A------C-------G-------------------------------------------------T--AT-G- 1981
SMM.US.x.SME543             -----------------T-------G--------A-----------------A-----C-----T--A-----------C-----------G--A-----CC-G-----------------------C-------------------------------------T--AT-G- 2067
SMM.US.x.pE660.CG7G         -----------------T-------G--------A-----------------A-----C-----T--A-----T-----C-----------G--A------C-G-----------------------C-------------------------------------T--AT-G- 2053
STM.US.89.STM_37_16         ---C--T---GTT------T----C-----G-------------AA-G--C--A--T-------------------G-----G-----------AG-----C--T----G--A------------GG--------A--------------T-----C-----AT-G- 1714
```

Draft of 30 September 2015, not for public release
hiv2dna_main.tex r8876–10546

```
                                                                                                                                                         Gag-Pol -1 ribosomal slip site
                                                                                                                                                                   Pol start
                                                                                                                                         Gag p8 Nucleocapsid end  Gag p1 Spacer start
MAC.US.x.239                ............CCAATTAAGTGTTGGAATTGTGGGAAAGAGGGACACTCTGCAAGGCAATGCAGAGCCCCAAGAAGACAGGGATGCTGGAAATGTGGAAAAATGGACCATGTTATGGCCAAATGCCCAGACAGACAGGCGGGTTTTTTAGGCCTTGGTCCATGGGGAAA 2377
Pol                                                                                                                                                                F  F  R  P  W  S  M  G
Gag                         _.__.__.__.__P__I__K__C__W__N__C__G__K__E__G__H__S__A__R__Q__C__R__A__P__R__R__Q__G__C__W__K__C__G__K__M__D__H__V__M__A__K__C__P__D__R__Q__A__G__F__L__G__L__G__P__W__G__K
A.CI.88.UC2                 ............A------G---------C--------G-----------G--G-AA--G---C----A--T-----------C--------------C--GCCA-GA--CA-C-----A--T-----G---------------------GA-G--CT--C-------- 2421
A.DE.x.BEN                  ............G------G--A------C-----A--G-----------G-----A--G---C----A--T-----------C--------G-----C--GCCA-GA--CA-C-----A--C-----G--A---------A-----------GT-G--C---C-------- 2421
A.DE.x.PEI2_KR_KRCG         ............A----------C-----------A--G--T--G-----G-----A------C----A--T-----------C--------------C---TCA-GA-----C-----A--C---------A---------T-----------GA----C----------- 2389
A.GH.x.GH1                  ............GT-----G---------C--C--A--G-----------G--G--A--G---C----G--T-----------C--------------C--G-CA-GA--C--C-----A-------------A---------A-----------GA-G------------- 1865
A.GM.87.D194                ............G------G---------------A--G-----------G--G-AA--G---C----A--C-----------C--------G-----C--GTCA-GA--CA-C-----A--C-----G--A---------A-----------GA-G--C---C-------- 1865
A.GM.x.ISY_SBL_6669_85      ............G----------------------A--G--A--G-----G-----A------C----G--T-----------C--------------C--GTCA-GA--CA-C-----A--C---------T---------T-----------G-----A----------- 1862
A.GM.x.MCNI3                ............A-------A--C-----------A--G--A--------G-----A------C----A--T-----------C--------G-----T---CCA-GA-----C--A--A--T---------T--------T------------GA-G--C--CC-------- 1866
A.GM.x.MCR35                ............A-------A--C-----------A--G--A--------G-----A------C----A--T-----------C--------G-----T---CCA-GA-----C--A--A--T---------T--------T------------GA-G--C--CC-------- 1866
A.GW.86.FG_clone_NIHZ       ............A--T-C-----C-----------A--G--A--------G-----A----G--C--G--T-----------C--------G-----T--GTCA-GA-----C-----A--T---------T---------T-----------GA----C-AC--------- 1859
A.GW.87.CAM2CG              ............A-G-----A--C--------------G-----G-----G-----A------C----A--T-----G-----C--------G-----C--GCCA-GA--CA-C---A-A--T-----T--T---------------------GA----C--C--------- 2430
A.GW.x.MDS                  ............A----------C-----------A--------G-----G-----A------C----A--T-----------C--------------C--G-CA-GA--A-C-----A--G------------------------------GA----C--C--------- 1865
A.IN.07.NNVA                ............A------G---C-----------A--G--A--------A-----A--G---C----A--T----A------C--T-----G-----C--GCCA-GA--A-C-----A--C--T-----T---------------------GA-G--C--CC-------- 2415
A.IN.95.CRIK_147            ............A------G---C-----------A--G--A--------A-----A------C----A--T--G--------C--T-----G-----C--G-CA-GA-----C-----A--C---------A---------------------GA-AA-C--CC-------- 2166
A.JP.08.NMC786_clone_41     ............G--T-------------C-----A-----A--------G--G--A--G---C----A--C-----G-----C--------------C--GCAA-GA--CA-C---T-A--C---------A---------A-----------GA-AA-----C-------- 2417
A.PT.x.ALI                  ............A----------C-----------A--G--A--G-----G-----A------C----A--C----A-----C--------------C--GCCA-GA---C-------A--C---------A---------T-----------G-----C----------- 2414
A.SN.85.ROD                 ............G--T----A--C-----C-----A--G--A--G-----G-----A------C----A--T-----G-----C--------G-----T--GCCA-GA--CA-C---A-A--C---------T---------A-----------A--G--C--T-------- 1864
A.SN.86.ST_JSP4_27          ............A----------C---------C--A--G--A--G-----G-----A------C----A--T--------A--C--------------C--GGCA-GA--CA-C-----A-------------A---------------------GT-G--C----------- 1866
B.CI.88.UC1                 GGG.........A-GG-G-CA--C-----C-----C--G-TA------A-A--C-A---------G-----T--------------------------CAA-GA--CA-C---T-A------------A---------T-----------GT-A--A--C-------- 2399
B.CI.x.20_56                ............A----A-CA--C-----C-----C----GA--C---A-A--C----------AG-----T------------------------G---CCA-GA--C--C---T-AC-----------T--------T-----------AT-C--A--C-------- 2394
B.CI.x.EHO                  ............A--G-G-CA--C-----C--C--C--G-C------TA-A--C-----G----AG-----T-----G-----------------------C-GCAA-GA--CA-C---T-A------------A---------------------GT-C--C--C-------- 2392
B.GH.86.D205_ALT            GGG.........A--G-G-CA--C-----C-----C---C--------A-A--C-----------G-----T-------------------------------CA-GA--CA-C---T-A------------A---------------------GT-A--A--C-------- 2394
B.JP.01.IMCJ_KR020_1        GGA.........G--G-G-CA--C-----C-----C---ATA------A-A--C--------T-A------T-----G-----------------C------CAA-GA--CA-C---T-A------------A---------------------AT----C--C-------- 1540
G.CI.92.Abt96               AGC.........A----A--A--------------C--G--------TA-A-TT--A--G-----------C-----------G--------G-----G--GCCA-G---CA-------------T-----G---------A-----------GT----A--T-------- 1799
AB.CM.03.03CM_510_03        GGG.........A--G-G-CA--C-----C-----C----G-------A-A--C--A--G----AG-----T--------------------------CAA-GA--CA-C---T-A------------A--------TT-----------GT-A--A--C-------- 1530
H2_01_AB.CI.90.7312A        GGG.........A--G-G-CA--C-----C-----C----G------TA-A--C-A---------G--T--T--------------------------CCA-GA--CAAC-----A--T--T-----A---------T-----------G-----A--C-------- 2400
H2_01_AB.JP.04.NMC307_20    GGG.........A--T-G-CA--C-----C-----C----G------TA-A--C-----------G-----T-----------G---------------------CA-GA--CAAC-----A--T----------------------------G-----A--C-------- 2400
H2_01_AB.JP.07.NMC716_01    GGG.........A--T-G-CA--C-----C-----C----G------TA-AK-C-----------G-----T--G--------G---------------------CA-GA--CAAC-----A--T------------R---------------GT----A--C-------- 2399
H2_01_AB.JP.08.NMC842_10    GGG.........A--T-G-CA--C-----C-----C----T------TA-A--C-A---G-----G-----T-----------G--------G----------C--GA--CAAC-----A--T----------------------------G-----A--C-------- 2405
U.CI.07.07IC_TNP3           AGC.........ATG--A--A--C-----C-----C--G--A--------A--------G---C-C--A--C-----------G--T-----G-----C--GCCA-GA--CA-C-----G------------A---------------------GT----C----------- 1904
U.FR.96.12034               ...............G-C-GA--------C-----C--G--A------A----------G--T-A------T-----------------------A---GCCA-GA--CAAC-----T-GC--------A-CCA---T-----------G-----C--T----A--- 1912
MAC.US.x.17EC1              ............------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2377
MAC.US.x.251_1A11           ............------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2376
MAC.US.x.251_32H_PJ5        ............------------------------------A------------------------------------------------------------------------------------------------------------------------------------------ 2378
MAC.US.x.251_BK28           ............--------------------------G---------------------------------------------------------------------------------------------------------------------------------------------- 2365
MAC.US.x.MM142_IVMXX        ............--------------------------G--A------------------------------------------------------------------------------------------A-------------------------------C--T-------- 1875
MNE.US.82.MNE_8             ............--------------------------------------------------------------A-----------C------------C-----G-----------------------------------A-----------T---C----------- 1857
MNE.US.x.MNE027             ............----------------------------------------------------------------------------C------------C-----G-----------------------------------A-----------T---C----------- 1857
SMM.CI.79.SIVsmCI2          AGG.............C--A----------------G--A-----AG--T-------T-AG--A--T----G-----C--T-----------C--TCA-G-----------T-----T----A--------T-----------AT----C----------- 1482
SMM.LR.89.SIVsmLIB1         ............A--G---G---------C----A--G--A-----T-------------T-AG--A--C--G----------------G-----C----CA-G--CA-C-----A-------------G---------T------------T-------T-------- 1468
SMM.SL.92.SIVsmSL92A        ............AA-G-C-GA--C-----C-----C-----A--G----A--G-----G--T-A---A--C-----------C--------------T------GA--CA-C-----------------A---------T-----------AT-------T-------- 1470
SMM.SL.92.SL92B             GGGAATAAGCCTATC--C-G-----TT--C-----A--GACA-----------G--A--G--------T--C----GA------------------GG--GAA-GA-G-A--CA------C------A--CAGA--------------------------C----------- 1829
SMM.US.04.G078              ............A--G-G--A--C-----C-----------A--G--TA------A--G--T-AG----------G-----T-----------C------CA-G-----C------------------T---------------------GT----C----------- 1649
SMM.US.04.G932              ..............G--------C-----C-----A-----------T----G--------------------G-------C------------------C--G------------------------------G---------------T----C----------- 1638
SMM.US.04.M919              ............A-------------------------G-----C----------A-----------T-----------C------------------CCA-G-----C-------------A-C--A-------------------G-------------------- 1643
SMM.US.04.M922              ............A----------------C-----A--G--A--C----------A--G----------T-----------C--------------------CA-G----------------------C--A---------------------G-----C----------- 1644
SMM.US.04.M923              ............A----------------C-----A--G--A-------------AA------A-----T-----------C--------G----------CA-G-----C-----T--------C--A---------------------G-----C--G-------- 1643
SMM.US.04.M926              ............A-----C----------C--A--G--A--G-----------A--------G----------G-----C-------------------GT-----C-----T-----T----A---------------------GT----C----------- 1643
SMM.US.04.M934              ............A------C---------------A--G--A--G----------AA-------------------C-G-----C------------------CA-G-----C-----T-----------A---------------------GT----C--T-------- 1641
SMM.US.04.M935              ............A----------------C-----A--G--A--C----------A--G----------T-----------C------------------CC-G-------------------C--A---------------------G-----C----------- 1644
SMM.US.04.M940              ............A--G-A--A--C-----C-----------C--G--TA---------G--T-AG-----------G-----C--------G--C------CA-G-----C-------------T---------------------------GT----C----------- 1649
SMM.US.04.M946              ............A------C---------------A--G--A--G----------A-----------------C-G-----C------------------CA-G-----C-----T-----------A---------------------GT----C--T-------- 1642
SMM.US.04.M947              ............A-------A--C-----------A--G-----------------AA--------------T-----------C--------G--------GCA-G-----------------T--C--A---------------------G-----C----------- 1643
SMM.US.04.M949              ............A--G-A--A--------C-----------A--G--TA---------G--T-AG-----------G-----C--------G-----G---CA-G-----C-------------T--A--------A--------------GT----C----------- 1637
SMM.US.04.M950              ............A----C-C---------------A--G--A-------------A-----------G-----G-----C-----------------G-CA-G-----C-----T-----------A---------------------AT----C--T-------- 1642
SMM.US.04.M951              ............A----A--A--C-----C-----A-----A--G--TA---------G--T-AG-----------G-----C--T----GG------C---CA-G---A-C-------------T--A---------A--------------GT----C----------- 1646
SMM.US.04.M952              ............A------C---------------A--G--A--G--A-------A----------------G-----C--T------------------CA-G-----C-----T-----------A---------------------GT----C--T-------- 1642
SMM.US.05.D215              ............ATG-----A--C-----C-----C--G--A-------------A--G-----G---------------------------------------GT--CAC-----A--C-----T------------------------AT----C--T-------- 1643
SMM.US.06.FTq               CAT.........A--G-A-----C-----------A--G----------------AA--------------T--G-----A--C-----------------------G---A-C-----T-----T-----G-------T------------T----------------- 1659
SMM.US.11.SIVsmE660_FL10    ............A----C-----------------A--G----------------A--------G-----T-----------C-----GGG--C------C--GT-------------------T--A---------------------G-----C----------- 2379
SMM.US.86.CFU212            AGA.........A--G-C--A--------------A-----T------A-------A--G--T-A---T---------R--C-----------------------G---C--C-----T------A---A-------A---------------RT----C----------- 1635
SMM.US.x.F236_H4            ............A----C-----------------A--G----------------AA--------G-----T-----------C--------G--------C--GT--------------------T--A---------------------G-----C----------- 2376
SMM.US.x.H9                 ............AT----W------R---C---R-A--G--A-------------A----T---------T--------R--C--------G--------GCA-G-------------T--------C--A---------------------G-----C----------- 1862
SMM.US.x.PBJA               ............AT---------------C-----A--G--A-------------A---------------T-----------C--------G--------GCA-G-------------T--------C--A-----A---------------G-----C----------- 2155
SMM.US.x.PGM53              ............A----------C-----------A--G-----C----------AA--------------T-----------C------------------CA-G-----C-----------------C--A---------------------G-----C----------- 2306
SMM.US.x.SME543             ............A----C-----------------A--G----------------A--------G-----T-----------C-----GGG--C------C--GT-------------------T--A---------------------GT----C----------- 2392
SMM.US.x.pE660.CG7G         ............A----C-----------------A--G----------------A--------G-----T-----------C-----GGG--C------C--GT-------------------T--A---------------------G-----C----------- 2378
STM.US.89.STM_37_16         ............A--G-A--A--C---------------CA--G---A------AA--G--T-A--G---G--------A--T--T-----------C--CA--G----CAG----------------A-------T-------------T---C----------- 2036
```

Draft of 30 September 2015, not for public release
hiv2dna_main.tex r8876–10546

```
                                                                                                                                          Gag end
                                                                                                       Pol Protease start        Gag p6 end
MAC.US.x.239                  GAGAAAAGCAGAGAGAAAGCAGAGAGAAGCCTTACAAGGAG......................................................GTGACAGAGGATTTGCTGCACCTCAATTCTCTCTTTGGAGGAGACCAGTAGTCACTGCTCATATTGAAGGACAGC 2609
Pol                            E  R  K  A  E  R  K  Q  R  E  A  L  Q  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  G  G  D  R  G  F  A  A  P  Q  F  S  L  W  R  R  P  V  V  T  A  H  I  E  G
Gag                           R  E  K  Q  R  E  S  R  E  K  P  Y  K  E  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  V  T  E  D  L  L  H  L  N  S  L  F  G  G  D  Q  *
A.CI.88.UC2                   -G--G............CAG--G----GA--A---------GTGACGGAAGACTTACTGCACCTCGAGCAGGGAGAGACACCTCGCAGAGAG...-C---------C---------------------------AA--------------A--AT-C--C--G-AT---- 2692
A.DE.x.BEN                    -G--G............CAG--G----GA--A-----A---GTGACGGAGGACTTGCTGCACCTCGAGCAGAGAGAGACACCTCACAGAGAG...-A---------C---------------------------AA--------------A--GT-C--C--G-AT---- 2692
A.DE.x.PEI2_KR_KRCG           A---G............CAG-A-AT--GA--A---------GTGACAGAGGACTTACTGCACCTCGAACAAGGAGAGACACCACACAAAGAG...-C-------------------------------------AA--------------A--AT--G----G--T---- 2660
A.GH.x.GH1                    -G--G............CAG--G----GA--A-----A---GTGACGGAAGACTTACTGCACCTCGAGCAGGGAAAAGCACCTCACAGAGAG...-C---------C---------------------------AA--------------A--AT-C--C--G-T----- 2136
A.GM.87.D194                  ----G............CAG--G----GA--A---------GTGACGGAGGACTTACTGCACCTCGAGCAGGGAGAGACGCCCCACAGAGGG...-C---------C-----A---------------------AA--------------A--ATTC--C--G-AT---- 2136
A.GM.x.ISY_SBL_6669_85        ----G............CAG---C-A-GA--A-----A--AGTGACAGAGGACTTGCTGCATCTCGAGCAAGGAGAGACACCACACAGAGAG...AC---------C---------------------------AAC-------------A--AT-C-----G-AT---- 2133
A.GM.x.MCN13                  -----............CAG-------GA--A-----A--AGTGACAGAGGACTTACTGCACCTCGAGCAGGGAGAGGCACCATGCAGAGAG...AC---------C------------------------T--AA--------------G--AT-CG-C--G--C---- 2137
A.GM.x.MCR35                  -----............CAG-------GA--A-----A--AGTGACAGAGGACTTACTGCACCTCGAGCAGGGAGAAGCACCATACAGAGAG...AC---------C------------------------T--AA--------------G--AT-CG----G--C---- 2137
A.GW.86.FG_clone_NIHZ         -G--G............CAG-------GA--G--------AGTGACAGAGGACTTACTGCGTTTCGAGCAGGCAGAGACACCATGCAGGGAG...AC---------C---------------------------AA--------------A--A--C-----G--T---- 2130
A.GW.87.CAM2CG                -G--G............CAG---C---GA--A-----A---GTGACAGAGGACTTACTGCACCTCGAACAAGGAGAGGCCCCATGCAGGGAG...ACA--------C---------------------------AA--------------A--A--C-----G--T---- 2701
A.GW.x.MDS                    -G--G............CAG-------GA--A-----A---GTGACAGAGGACTTGCTGCACCTCGAGCAGGGAAAGACACCATGCAAGGAG...ACA------------------------------------AC-----------T--A--A--C-----G------- 2136
A.IN.07.NNVA                  -G--G............CAG-------GA--A---------GTGACAGAGGACCTACTGCACCTCGAGCAGGGAGAAGCACCCTGCGGGGAA...AC------A--C---------------------------AA-----T--------A--AT-C--C--G--T---- 2686
A.IN.95.CRIK_147              -G--G............CAG-------GA--A---------GTGACAGAGGACCTACTGCACCTCGAGCAGGGAGGGGCACCCTGCGGGGAA...A----------C---------------------------AA--------------A--AT-C-----G--T---- 2437
A.JP.08.NMC786_clone_41       -G--G............CAG--G----GA--A-----A---GTGACAGAGGACTTACTGCACCTCGAGCAAGGAGAGACACCTCACAGAGAG...-C----A----C----------------------------A--------------A--AT-------G-AT---- 2688
A.PT.x.ALI                    ----G............CAG-------G---A-----A---GTGACGGAGGACTTACTGCACCTCGAGCAGGGAGAGACACCACACAAAGAG...-----------C---------------------------AA--------------A--CT-C-----G--C---- 2685
A.SN.85.ROD                   ----G............CAG-------GA--A--------AGTGACAGAGGACTTACTGCACCTCGAGCAGGGGGAGACACCATACAGGGAGCCACCA--------C---------------------------AA--------------A--AT-C-----G--T---- 2138
A.SN.86.ST_JSP4_27            ----G............CAG-------G---A-----A--AGTGACAGAGGACTTCCTGCAGCTCGAGAAACAAGAGACACCATGCAGAGAG...AC---------C---------------------------AA--------------A--A---G----G--C---- 2137
B.CI.88.UC1                   ----G...............C------GA--C---------......................................................----------------------------------------A-------------GA--ATG---C--G--T---- 2673
B.CI.x.20_56                  ----G...............C------GA--C---------......................................................-----------C----------------------------A-------------AA--AAG------G--T---- 2668
B.CI.x.EHO                    AG--G.............A--------GA--C---------......................................................----------------------------------------A-------------AA--AAC------G--T--AT 2663
B.GH.86.D205_ALT              ----G...............C------GA--C---------......................................................----------------------------------------A-------------AA--ATG---C--G--T---T 2668
B.JP.01.IMCJ_KR020_1          A--GG............G----G-----A--C---------......................................................----------------------------------------A-----------T-AA--AT-------G--T--AT 1814
G.CI.92.Abt96                 AG--G............CAG---A-C--A--C---------......................................................--------R--H------------GC--------------AC----------------AT-C--A---A-T---- 2019
AB.CM.03.03CM_510_03          ----G...............C------GA--C---------......................................................-----------C----------------------------A-------------GR--ATG---C--G--T---- 1804
H2_01_AB.CI.90.7312A          ----G.............AGCAG-G--G---C---------......................................................-----------C----------------------------A------------CAA--AT----C--G--T---- 2674
H2_01_AB.JP.04.NMC307_20      ----G.............AGCA--G--GA--C---------......................................................-----G-----C----------------------------A------------CAA--AT----C--G--T---- 2674
H2_01_AB.JP.07.NMC716_01      ----G.............AGCA--G--GA--C---------......................................................-----G-----C----------------------------A------------CAA--AT----C--G--T---- 2673
H2_01_AB.JP.08.NMC842_10      ----G.............AGCA--G--GA--C---------......................................................-----G-----C----------------------------A------------CAA--A-----C--G--T---- 2679
U.CI.07.07IC_TNP3             AG--G.................G--A-GA--A---------......................................................-----G----------------------------------AC----------------A--C--A--G--T---- 2124
U.FR.96.12034                 AG--G.............A--------GA--A---------......................................................-----G---------------------------------AAC-----A----------AT-C--A--G--G--A- 2174
MAC.US.x.17EC1                -----------------------------------------......................................................--------------------------------------------------------------------------- 2609
MAC.US.x.251_1A11             -------A---------------------------------......................................................A-------------------------------------------------------------------------- 2608
MAC.US.x.251_32H_PJ5          -----------------------------------------......................................................--------------------------------------------------------------------------- 2610
MAC.US.x.251_BK28             -----............------------------------......................................................--------------------------------------------------------------------------- 2585
MAC.US.x.MM142_IVMXX          -----............-------G----------------......................................................--------------------------------------------------------------------------- 2095
MNE.US.82.MNE_8               -----............----A-AG----------------......................................................----------------------------------------A--------------------------G------- 2077
MNE.US.x.MNE027               -----............----A-AG----------------......................................................---G------------------------------------A--------------------------G------- 2077
SMM.CI.79.SIVsmCI2            ----G..............G-A---A-G---------A---......................................................--A--G--A--C----------------------------AC----------------AT-C--A--G--T---- 1729
SMM.LR.89.SIVsmLIB1           -G--G.............A------A-GA--C---------......................................................----------------------------------------AC-------------A--C--C-----G--T---- 1688
SMM.SL.92.SIVsmSL92A          A---G......................GA--A---------......................................................-----G-----C----------------------------A--------------A--------A---------- 1690
SMM.SL.92.SL92B               AGAC-..............G--GAGC-G-------------......................................................-----G-----------------------------C----A-------------GA--A-TA--A--G--T---- 2040
SMM.US.04.G078                -G--G.............A--------GA--C---------......................................................-----G------------------G---------------A--------------A---T-C-----G--T---- 1869
SMM.US.04.G932                -G--G......................GA---CA--C---------.................................................-----------C----------------------------A-----------T------T-C--A---------- 1858
SMM.US.04.M919                -G--G.............A--------GA------------......................................................-----G----------------------------------A-----------------CT-C-------A----- 1863
SMM.US.04.M922                -G--G.............A--------GA------------......................................................-----G----------------------------------A-----------------CT-C-------A----- 1864
SMM.US.04.M923                -G--G.............A--------GA------------......................................................-----G-----C----------------------------AC----------------CT-C-------A----- 1863
SMM.US.04.M926                ----G.............A---GAG---A------------......................................................--------A--C----------------------------A------------------T-------G------- 1863
SMM.US.04.M934                A---G.............A---GAG----------------......................................................--A-------------------------------------AC-----------------T-C-----G------- 1861
SMM.US.04.M935                -G--G.............A--------GA------------......................................................-----G----------------------------------A-----------------CT-C-------A----- 1864
SMM.US.04.M940                ----G......................GA--C---------......................................................----------------------------------------A--------------A--CT-C-----G--T---- 1869
SMM.US.04.M946                A---G......................GAG--G-------------.................................................-----------C----------------------------AC----------T------T-C-----G------- 1862
SMM.US.04.M947                -G--G.............A--------GA------------......................................................-----G----------------------------------A-----------------CT-C-------A----- 1863
SMM.US.04.M949                -G--G......................GA--C---------......................................................-----G----------------------------------AC-------------A--CT-CG----G--T---- 1857
SMM.US.04.M950                ----G......................GAG----------------.................................................----------------------------------------AC-----------------T-C-----G------- 1862
SMM.US.04.M951                -G--G.............A--------GA--C---------......................................................-----G----------------------------------A--------------A--CT-C-----G--T---- 1866
SMM.US.04.M952                ----G.............A---GAG--G-------------......................................................-----------C----------------------------AC-----------------T-C-----G------- 1862
SMM.US.05.D215                -G--G.............A--------GA------------......................................................----------------------------------------A------A-------A--CT----------T---- 1863
SMM.US.06.FTq                 -G--G......................GA--A---------......................................................----GG------------------G---------------A--------------A--C-----------T---- 1891
SMM.US.11.SIVsmE660_FL10      -G--G.............A--------GA------------......................................................----------------------------------------A-----------------CT-C-------A----- 2599
SMM.US.86.CFU212              ----G.............A--------GA------------......................................................-----G----------------------------------AC----------G--A--G--GG-C--G--T---- 1855
SMM.US.x.F236_H4              -G--G.............A--------GA------------......................................................----------------------------------------A-----------------CT-C-------A----- 2596
SMM.US.x.H9                   -G--G.............A--------GA------------......................................................--A--G----------------------------------A-----------------CT-C-------A----- 2082
SMM.US.x.PBJA                 -G--G.............A--------GA------------......................................................--A--G----------------------------------A------A-------G--CT-C-------A----- 2375
SMM.US.x.PGM53                -G--G.............A--------GA------------......................................................-----G-----------------------------------------------G----CT-C-------A----- 2526
SMM.US.x.SME543               -G--G.............A--------GA------------......................................................----------------------------------------A-----------------CT-C-------A----- 2612
SMM.US.x.pE660.CG7G           -G--G.............A--------GA------------......................................................----------------------------------------A-----------------CT-C-------A----- 2598
STM.US.89.STM_37_16           ----G......................GA---CA--C---------.................................................----------------------------------------A--------------A--C--------G--T---- 2256
```

HIV-2 Genomes

```
                                                                                   Pol Protease end Pol p66 and p51 RT start
MAC.US.x.239              ATGACAGGGGACACCCCGATTAACATTTTTGGTAGAAATTTGCTAACAGCTCTGGGGATGTCTCTAAATTTTCCCATAGCTAAAGTAGAGCCTGTAAAAGTCGCCTTAAAGCCAGGAAAGGATGGACCAAAATTGAAGCAGTGGCCATTATCAAAAGAAAAGATAGTTGC 2949
Pol                       _I__M__T__G__D__T__P__I__N__I__F__G__R__N__L__L__T__A__L__G__M__S__L__N__F__P__I__A__K__V__E__P__V__K__V__A__L__K__P__G__K__D__G__P__K__L__K__Q__W__P__L__S__K__E__K__I__V
A.CI.88.UC2               --------A--T-----A--C-----C-----C------A-TT-G-----CT-A--C-----AT-------A--AG-C----GGA-------AA-----A-AA-A--G-----------A--C-------GGC----A--A-----CC--A----------AG---AA-- 3032
A.DE.x.BEN                --------A--T-----A--C-----C-----C------A-T--G-----CT-A--C-----AT-------A--AG-T--C--GA-------AA-------AA-A--G--------G--A----------GGC----A--A-----CC--A-------G--A----AA-- 3032
A.DE.x.PEI2_KR_KRCG       --------T--T-----A--C-----------C-----CA-T--G-----CT-A--C-----AT-----C-A--AG-C--C--G-----C--GA-------AATAC-G--A--------A------------G-A-GA--A-----TC--A---------------AG-- 3000
A.GH.x.GH1                --------A--T-----A--C-----C-----C------A-C--G-----CT-A--C-----AT-------A--A--C--C--GA----A--AA-------AA-G--G--------G--A----------G-C---GA--A-----CC--A-------G--A----AA-- 2476
A.GM.87.D194              --------A--T-----A--C-----------C------A-T--GG--A-CT-A--C-----AT----CC-A--AG-C--C--GT----C--AA-------AA-A--G--------G--A----------GGC----A--A-----CC--A----------A----AA-- 2476
A.GM.x.ISY_SBL_6669_85    --------T--T-----A--C-----------C------A-C--G-----CT----C-----AT-------A--AG-C--C---A----A--A--------AA-A--------------A-----G------CAA-GA--A-----C---A--G-------A----AA-- 2473
A.GM.x.MCN13              --------T--------A--C-----------C------A-T--G-----CT-A--C-----AT-------A--AG-C--C---A----A--AA-----A-AATG--------------A------------C---G---A-----C---A----------A----AG-- 2477
A.GM.x.MCR35              --------T--------A--C-----------C------A-T--G-----CT-A--C-----AT-------A--AG-C--C---A----A--AA-----A-AATG--------------A------------C---G---A-----C---A----------A----AG-- 2477
A.GW.86.FG_clone_NIHZ     --------T--T--------C-----------C------G-T--G------T-A--C-----AT----CC-G--AG-T--C--GA----A--AA-----A-AATGC----------G--A----------G-C-A--A--A-----T---A----------A----AA-- 2470
A.GW.87.CAM2CG            --------C--------A--C-----------C------A-T--G-----CT-A--C-----AT-----C-A--AG-C--C--GA----A--AA-----A-AATGC-------------A----------G-----GA--A-----C--GA----------A----AA-- 3041
A.GW.x.MDS                --------T--T-----A--C-----------C------A-TT-G-----CT-A--C-----AT-----C-A--AG-C--C--GA----A--AA-----A-AATGC----------G---------------C---GA--------T---A----------A----AA-- 2476
A.IN.07.NNVA              --------T--T-----A--C-----------C-----CA-TT-G-----CT-A--C------T----CC-A--AG-C--C--GA----A--AA-------AAAGC----A---------------------C---GA--------C---A----------A----AG-- 3026
A.IN.95.CRIK_147          --------T--T-----A--C-----------C-----CA-TT-G----TCT-A--C-----AT-----C-A---G-C--C--GA----A--AA-----A-AATGC----A---------------------C---GA--------CC--A-C--------A----AA-- 2777
A.JP.08.NMC786_clone_41   --------A--T-----A--C-----C-----C------A-C--GG-G--CT-A--C-----AT-------A--A--C--C--GA----A--AA-C-----AA-A--G--------G--A-----T----GGC-A--A--A-----CC--A----------A----AG-- 3028
A.PT.x.ALI                --------T--------A--C--T--------C-----CA-TT-G-----CT-A--C-----AT----CC-A--AG-T--C--GA-------AA--G-G--AAGA--------------A--C--G--------A-GA--A-----C---A----------A----AG-- 3025
A.SN.85.ROD               --------C--------A--C-----------C------A-T--G-----CT-A--C-----AT-----C-A--AG-C--C-----------AA-----A-AATGC----------G--A------------C---GA--A-----C---A----------A----AA-- 2478
A.SN.86.ST_JSP4_27        --------T--T-----A--C-----------C-----CA-T--G-----CT-A--C-----AT-----C-A--AG-C--C--GA----A--AA-----A-AATGC-G------------------------C---GA--A-----C---A----------A----AG-- 2477
B.CI.88.UC1               --------A--------A--A-----------C------A-TT---ATA-AT-A--C---A----------C--AG----A--GA----A--A--------AAAC--------T--G--A------------A-C-GA--A-----TC----C------------T-G-- 3013
B.CI.x.20_56              --------A--------A--A-----------C------A-CT---ATA-CT----C---A----------C--AG----A---A----A--A--------AAAC--G-----T--G--A------------A-C-GA--A-----CC----C------------C-G-- 3008
B.CI.x.EHO                -----------------A--A-----------C------A-TT---ATAGCT-A--C---A----------C--AG----A-GGA----A--A---------CAG--------T-A---A-----G------A-C-GA--A-----CC----C-----G--A---C-A-- 3003
B.GH.86.D205_ALT          --------A--T-----A--A-----------C------A-TT---ATA-CT----C---A--T-------C--AG-G--A--G-----A--A--------T-AG-----A--T-----A-----G-----GA-C-GA--A-----TC----C-GG---------C-A-- 3008
B.JP.01.IMCJ_KR020_1      --------A--T-----A--A-----------C------A-T----ATA-CT----C---A----------C--AG----A---A----A--A-----G--AAAC-----A--T-----A------------A-C-GA--A-----TC----T------------C-G-- 2154
G.CI.92.Abt96             -----------T-----T--A--T--C-----A--G--CA-TT--GTTAAG--A--C-----------------A--T--A--G--------A--------AAGGC----A--T--C-T------G------A-C-G---A-----CC----T--G-----A---CAA-- 2359
AB.CM.03.03CM_510_03      --------A--------A--A-----------C------A-TT---ATA--T-R--C---A----------C--AG-G--A-RGR----A--A--------T-AG--R--A--T-----R-----R-----GA-C-GA--A-----TC----C-GG--R------C-A-- 2144
H2_01_AB.CI.90.7312A      --------A--------A--A--T--------C------A-CT---ACA-CT-A--C---A----------C--AG-G--A---A-------A--G--G---CAG--------T------------------A-C-GA--A-----TC----T------------C-G-- 3014
H2_01_AB.JP.04.NMC307_20  --------A--------A--A--T--------C------A-CT---GC--CT-A--C---AG---------C--A--G--A---T-------A--G--G---CAG--------T------------------A-C-GA--A-----TC----T------------C-G-- 3014
H2_01_AB.JP.07.NMC716_01  --------A--------A--A--T---------------A-CT---GT--CT-A--C---AG-T-------C--A--G--A---T-------A--G--G---CAG--------T-----A------------A-T-GA--A-----TC-G--T--G--G------C-G-- 3013
H2_01_AB.JP.08.NMC842_10  --------A--------A--A--T--------C------A-CT---GC--CT-A--C---AG---------C--A--G--A---T-------A-----G---CAA--------T-----A------------A-C-GA--------T-----T------------T-G-- 3019
U.CI.07.07IC_TNP3         -----------T-----T--A-----------A------A-T--GGTCAAAT-A--A--------------C--AG-G--A--G-----AA-A--------AA-AC----A-------T-------------A-T-GA--------RC----T-----G--A----AA-- 2464
U.FR.96.12034             --------A--------A--C--T-----------G---C-TT-G----AA-----C-----------C-----A----GA----A---CA-TA--C-TT-GG-GAC---A-----C-----------------A-GA--------T-----T--G-----AT---AA-- 2514
MAC.US.x.17EC1            -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 2949
MAC.US.x.251_1A11         -----------------------------------------A---------------------------C-------------------------------------------------A-T------------------------------------------------ 2948
MAC.US.x.251_32H_PJ5      -----------------------------------------A---------------------T----CC---------------------------------------------------------------------------------------------------- 2950
MAC.US.x.251_BK28         --------------T--------------------G---------------------------------C-------------G------------------A------------------T------------------------------------------------ 2925
MAC.US.x.MM142_IVMXX      -----------------------------------------A---------------------------C-------------G--------------GTCGC------------------------------------------------------------------- 2435
MNE.US.82.MNE_8           -------------------------------------------------------A---------------C-----------G------------------A----------------A----------------G--------------------------------- 2417
MNE.US.x.MNE027           --------------------------------------------------------A--------------C-----------G------------------A----------------A----------------G--------------------------------- 2417
SMM.CI.79.SIVsmCI2        --------A--------A--A--T--------C------A-TT-G-----AT-A--C-----AT----C-GC--AT-G--A--G-----A--A-----G--AA-T--G-----------A--A-----------A--A--A-----GC-G----------------AG-- 2069
SMM.LR.89.SIVsmLIB1       --------A--T-----A-----T--------C------C-CT-G----AAT-A--A-----------------AG-------------A--A-----G--AAAAC----A-----G--A----------------GA--A------C-T--------G-------AG-- 2028
SMM.SL.92.SIVsmSL92A      -----------T--A--T--------C-----A------A-TT-------AT-A--A--T----------A---AG-G--A------C-A--A-----G--AA-AC----A---------------------CAA-GA--------CC-G--T-------------AA-- 2030
SMM.SL.92.SL92B           --------A--T--------A--T--------C------A-T---G----AT-A---G-C---T------AC--AG-G--A--G-----ATA-AC------AAAG--G---GAG----T------G-----G--A--A--A-----C-----T--G---------CAA-- 2380
SMM.US.04.G078            -----T--A--------A-----T--------C-----CC-AT-G-----AA-----------------C-A--AG-G-----G--------A-----G--GAAG-----A-----G--A--A--------G--A--A--------GC----T--------A----AG-- 2209
SMM.US.04.G932            -----------------A-----T--------C--------A--G---A-------------------C-----A-----------G---------------A-AC----A--------A----------------GA--A------------------------A---- 2198
SMM.US.04.M919            --------A--T-----A--------------C------------------A----C------T-----C-C---G-------G--G--------------AA-AC----A-----------C-----------A-GA--------GC-G-----------A-------- 2203
SMM.US.04.M922            --------A--T-----A--------------C------C-A---------A----C------T----CC-C---G-------G-----------------AA-AC----A--------A--C-------------GA--------GC----------G----------- 2204
SMM.US.04.M923            --------A--T-----A--------------C------C-----------A----C------------C-C-----------G--G------A-------AA-AC----A--------A--------------A-GA--------GC----------G------A---- 2203
SMM.US.04.M926            --------A--------------T--------C---------T-G---A--A----C--------------C--------------G--A------------A--C----A--------A-----------G----GA--A--------------------A---A---- 2203
SMM.US.04.M934            --------A--------A-----T--------C---------T-G---A--A----C-----------C-GC---G--------------------------A--C----A--------A-----------G----GA--A--------------------A---A-C-- 2201
SMM.US.04.M935            --------A--T-----A--------------C------C-----------A----C------T-----C-C-----------G---------A-------AA-AC----A--------A--------------A-GA--------GC----------G----------- 2204
SMM.US.04.M940            -----C--A--------------T--------C--G--CC-AT-G-----AT----A------------C-A--AG-G----GG--G--A--A-----G--GAAG-----A-----C--A--G--------G-----A--------GC----C-------------AG-- 2209
SMM.US.04.M946            --------A--------A-----T--------C---------T-G---A--A----C--------G----GC--------------------------G---A--C----A--------A----------------GA--A--------------------A---A---- 2202
SMM.US.04.M947            --------A--T-----A--------------C------C----GG-----A----C--------------C-----------G--G--A---A-T-----AA-A-----A--------A--------------A-GA--A-----GC-----------------A---- 2203
SMM.US.04.M949            -----C--A--------------T--------C------C-AT-G-----AT----A------------C-A--AG-G-----------A--A-----G--GAAG-----A--------A--G--------G--A--A--A-----GC----C-------------AG-- 2197
SMM.US.04.M950            --------A--------A-----T--------C--------AT-G---A--A----C-------------GC-----------------A--------G---A----------------A-----------G----GA--A--------------------A---A---- 2202
SMM.US.04.M951            -----C--A--------A-----T--------C--G--CC-AT-G-----AT-------------G---C-A--AG-G----GG-----A--A-----G--GAGG-----A--------A--G-----------A--A--------GC----C-------------AG-- 2206
SMM.US.04.M952            --------AA-------A-----T--------C---------T-G---A--A----C-------------GC-----------G--T--A--------G---A--C-G--A--------A-----------G----GA--A--------------------A---A---- 2202
SMM.US.05.D215            --------A--------------T--------------CC-AT-G--TA-AA----------AT----------AG-------G--------A-----G--GA-A--G--A---------------------C-A-GA--A-------------------------AA-- 2203
SMM.US.06.FTq             -----G-----T-----A--C-----------C------C--T-----AGAT-A---------T----------AG------GG-----A--A--------GA-AC----A-----GC----------------A-GA--------------T--------A----AG-- 2231
SMM.US.11.SIVsmE660_FL10  -----G--A--T-----A-----T--------C------------------A----C-----CT-----C-C-----------G--G------A-------AA-AC----A-----G--A--A-------------GA--------GC-----------------A---- 2939
SMM.US.86.CFU212          --------A--T-----A--C--T--------C--------A---------T----T-----CT----C--C--A--G--CC----G-----G--------AA-AC----A--------A------------G-A-G---A------C-----------------CAG-- 2195
SMM.US.x.F236_H4          -----G--A--T-----A-----T--------C------------------A----C------T-----C-C-----------G--G------A-------AA-AC----A--------A--A-----------A-GA--------GC-----------------A---- 2936
SMM.US.x.H9               --------A--T-----A--------------C------C-----------A----C------------C-C-----------G--G--R---A-------AA-AC----A--------A--------------A-GA--------GC----G------------A---- 2422
SMM.US.x.PBJA             --------A--T-----A-----T--------C------C-----------A----T------------C-C-----------G---------A-------AA-AC----A--------A--------------A-GA--------GC-----------------A---- 2715
SMM.US.x.PGM53            --------A--T-----A--------------C------C--T--------A----C-----CT-----C-C---G-------G--G------A-------AA-AC----A--------A--C-------------GA--------GC-------------A-------- 2866
SMM.US.x.SME543           -----G--A--T-----A-----T--------C------------------A----C------T-------C-----------G--G------A-------AA-AC----A-----G--A--A-------------GA--------GC-----------------A---- 2952
SMM.US.x.pE660.CG7G       -----G--A--T-----A-----T--------C------------------A----C------T-----C-C-----------G--G------A-------AA-AC----A-----G--A--A-------------GA--------GC-----------------A---- 2938
STM.US.89.STM_37_16       -----T--A--------C--------------C--G--CC-AT-G------T----------AT----------AG-------------A--A-----G--GA-A-----A--------A------------A-A--A--A-------------------------AG-- 2596
```

Draft of 30 September 2015, not for public release
hiv2dna_main.tex r8876–10546

```
MAC.US.x.239                  GTCCAATTAGGAATACCACACCCTGCAGGACTAGCAAAAAGGAAAAGAATTACAGTACTGGATATAGGTGATGCATATTTCTCCATACCTCTAGATGAAGAATTTAGGCAGTACACTGCCTTTACTTTACCATCAGTAAATAATGCAGAGCCAGGAAAACGATACATTTA 3288
Pol                           _E__V__Q__L__G__I__P__H__P__A__G__L__A__K__R__K__R__I__T__V__L__D__I__G__D__A__Y__F__S__I__P__L__D__E__E__F__R__Q__Y__T__A__F__T__L__P__S__V__N__N__A__E__P__G__K__R__Y__I
A.CI.88.UC2                   A-T--GC-------T--------A------T----C----A------G-----T-----A---G----G-----C--C--T--------A---T-------T-------C---T-----A-----CC-------------C----T---A---------A-------A-- 3371
A.DE.x.BEN                    A-T--GC-------T--------G-----------C----A------G--CT-TA--T-A---G----G-----C-----T--------A---C-------T-----------T-----A-----CC-----G-------C---ATG--A---------A----T--A-- 3371
A.DE.x.PEI2_KR_KRCG           A-T--G--------T--------A------T----C--G-AA-G---------T-----A--------G-----C--C--T--------A---C----G--C-----A--A--T-----A------C-----A----G--C--------A--------GA----T--A-- 3339
A.GH.x.GH1                    A-T--GC-------T--------G------T----C----A------G--C--T-----A---G----G-----C--C--T--------G---C-------T-----------T-----A-----CC-------------C--------A---------A----T--A-- 2815
A.GM.87.D194                  A-T--GC-------T--------G------T----C----A------G-----T-----A---G----G-----C--C--T--------A---C-------T--------A--T-----A-----CC-------------C-------------A----A----TG-A-- 2815
A.GM.x.ISY_SBL_6669_85        --T--G--------T--------A------T----C--G-AA-G---------T--GT-A---G----A-----C--C--T--------C---T----G--T-----A-----T-----A------C-G----------C--------A---------A----T--A-- 2812
A.GM.x.MCN13                  A-T--G--------T--------A------T-G--C--G-AA-----------T-----A--------G-----T--C--T--------A---C-------C-----A-----T-----A----------------A---C--------A---------A----T--A-- 2816
A.GM.x.MCR35                  A-T--G--------T--------A------T-G--C--G-AA-----------T-----A--------G-----T--C--T--------A---C-------C-----A-----T-----A----------------A---C--------A---------A----T--A-- 2816
A.GW.86.FG_clone_NIHZ         A-T--G--------T--------A------T-G--C----A--GG--------T-----A---G----G-----T--C--T--------A---C----G--C-----A-----T-----A--C---C-------------C--------A---------A----T--A-- 2809
A.GW.87.CAM2CG                A----G--------T--------A------T----C----A--G---------T-----A---G----G-----T-----T--------AT--C-------C-----A-----------A------C----------G--C--------A---------A-------A-- 3380
A.GW.x.MDS                    A-T--G--------T--------A--G--------C--G-A--G---------T-----A---G----G-----T--C--T--------A--GC-------C-----A-----------A-----CC----------G--C--------A---------A----T--A-- 2815
A.IN.07.NNVA                  A-T--GC-------T--------G------T-G--C--G-AA-G------C--T-----A---G----G-----T--C--T--------A--GC----G--C-----A-----------A-----C--G--G-------C--------A---------A----T--A-- 3365
A.IN.95.CRIK_147              A-T--G--------C--------A------T-G--C--G-A--G------C--T-----A---G----G-----T--C--T--------A--GC-------C-----A-----T-----A-----------G-------G--C--------A---------A----T--A-- 3116
A.JP.08.NMC786_clone_41       A-T--GC-------T--------G-----GT----C----A------G--C--T-----A---G----G-----C-----T--------A--GT-------T-----------T-----A-----CC--------AC---C--------A---------A----T--A-- 3367
A.PT.x.ALI                    A-T--G-----G--T-----T--A------T----C--G-AA-G------C--T--G------G----G-----T--C--T--------A--GC----GAGC-----A-----T-----A------C-------------C--------A---------A----T--A-- 3364
A.SN.85.ROD                   A-T--G--------T--------A-----GT-G--C--G-A--G---------T-----A---G----G-----T--C--T--------A---C----G--C-----A-CA--T-----A------C----------G--C--------A---------A-------A-- 2817
A.SN.86.ST_JSP4_27            A----G--------T--------A-----------C--G-A---C--------T--C--A---G----G-----T--C--T--------A---C----G--T-----A-----T-----A------C--------A---C-----T--A---------A-------A-- 2816
B.CI.88.UC1                   ------C-G--T--T--C-----G-----GT-G---G---AA-GG----A--------A---G-G--A-----C--C---AGT--C--A------CC-A-C--C--A-----T--A--A--C--C--G------A---C----------------GA---------- 3352
B.CI.x.20_56                  -----GC-G--T--T--C-----G--T--G--G---G---AA-G-----A------T-A---G----A-----C--C---AGTG-C--A-----CCC---T--C----------A--A-----C--G---C----G------T---GA---------A---------- 3347
B.CI.x.EHO                    -----GC-G--T--T--T-----A--------G---TC--A--------A--------A---G----A-----C--C---AGTG-C--A------CC---C--C--A--A--T--A--A--------G---G----------------A---------GA----TC---- 3342
B.GH.86.D205_ALT              --.---C-G--T--T--C-----G-----..-G---G---AA-GG----A------A-A---G-G--A-----C--C---AGT--C--A------CC-A-C--C--A-----T--A--A--C--C--G-----------C------------------GA---------- 3344
B.JP.01.IMCJ_KR020_1          -----GC-G--C--T--C-----G-----G--G---G----A-------A--------A--------A-----C--C---AGT--C--A------CC---T--C--A--A-----A--A-----C--G-------G----------A---------A-------C-- 2494
G.CI.92.Abt96                 -----G--G--C-----T----------TT-G---G---AA-GG----A------------------G-----C--C---AG---T----------T---T--C-----------A--T--C-------------------C--------A---------A-------C-- 2698
AB.CM.03.03CM_510_03          -----R--G--T--C--C-----G-----GT-G---G----A-GG----A------A-A---G-G--A--C--C--C---AGT--C--A------CC-R-Y--C--A-----T--A--A--C--CY-G------AC---C--YA------------GA---------- 2483
H2_01_AB.CI.90.7312A          ------C----T--T--C-----G-----G--G---G---AA-GG----A-----GA-A---G----A-----C--C---AGTG-C--A-----CCC---C--C-----A--T--A--A--C-----G------A-------------------------A-G-------- 3353
H2_01_AB.JP.04.NMC307_20      ------C----T--T--C-----A------T-G---G---AA-GG----AT-------A---G----A-----C--C---AGTG-C--A------CC---C--C--A--A--T--A--A--C--C--G------A-------------------------A---------- 3353
H2_01_AB.JP.07.NMC716_01      ------C----T--T--C-----A------T-G---G----A-GG----AT----G--A---G----A-----C--C---AGTG-C--A------CC---C--C--A--A--T--A--A--C-----G------A-------------------------A-G-------- 3352
H2_01_AB.JP.08.NMC842_10      -----GC----T--T--C-----G--------G---G----A-G-----AT-------A---G----A-----C--C---AGTG-C--A-----CCC---C--C--A--A--T--A--A--C-----G------A-------------------------A---------- 3358
U.CI.07.07IC_TNP3             --G--------T-----T------------------G-G-AA-G-----A----------CG---A-----C-----AG---T-----------C----C---CA----A--T--C---C-------------C--------A--------GA-------C-- 2803
U.FR.96.12034                 -----G-----T-----C-----A-----------TG-G-AA--G--G--A-----T--A---G----G-----C--C---AG---C--G--GTG------C--C-----A--T-----T-------------CA-T--------------------A-------C-- 2853
MAC.US.x.17EC1                --------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 3288
MAC.US.x.251_1A11             -----------------------------------------G-------------------C-------------------------------------------------------------------------------------------------- 3287
MAC.US.x.251_32H_PJ5          -----------------------------------------G----------------------------------------------------------------------------------------------------------G----------- 3289
MAC.US.x.251_BK28             -----------------------------------------G-------------------------------------------------------------------------------------------------------------------- 3264
MAC.US.x.MM142_IVMXX          -----------------------------------------G-------------------C----------T------------------------------------------------------------------------------------------- 2774
MNE.US.82.MNE_8               --------G-----------------------------G--G--C-----------G-------C-------------------------------------------T------------------------A----------------------- 2756
MNE.US.x.MNE027               --------------------------------------G--G--C-----------G-------C-------------------------------------------T------------------------A----------------------- 2756
SMM.CI.79.SIVsmCI2            --T--GC----------------A------T-G---G---A--G-----A------T-A--C-----A-----C-----TAGT--T--AT--TGG-----G-----A--A--T--A--A--CG--C-C--T-----C--C--------A-----T---A-G--T----- 2408
SMM.LR.89.SIVsmLIB1           A-T--G--G-----C-------------------------G---G-----------GT-A---G----A---------------T--T--CT-G------------------T----------C--G-----G-----C--------A--------GA----TG-G-- 2367
SMM.SL.92.SIVsmSL92A          A-T--G--------C--T-----A-----G---CAGC------GG----C------T-A---G-G--A-----T-----T-----C--AT-G---CCG--C--------A--T--A--T--------G--T--T-------G---T--A---------A-------C-- 2369
SMM.SL.92.SL92B               --A--GC-G--T--C-----T--G-----C---AAGG---T-G----G--A-----G--A--------G-----C-----T--AG----C-----CCCG--------A--A-----A--A--C---A----------G--C--CCA---A--------GA-G--T----- 2719
SMM.US.04.G078                A-T--G--------T--G-----C------T-----------A-G------C------T-A---G----G----------T--------AT-G-----G-----C--A------------------CC----------T--C--C-----A--------GA-G--TG-G-- 2548
SMM.US.04.G932                ------------------------------------T-----A--G-----C------T-----G----------------------------GC----G------------------A---------------------C--------A--------G--------C-- 2537
SMM.US.04.M919                --A--G-----G------------------------G----G------C------T-----G------------------------------------------C-----A--------T-----------------------C------------------A-------C-- 2542
SMM.US.04.M922                --A--G------------------------------G----G------C------T-----G---------------T----------G----------C-----A------------------------------G--C--------A---------A-------C-- 2543
SMM.US.04.M923                --A--G----------------------------------A-G---G--C------T-----G-------------------------T------------C-----A------------------------------------C--------A---------A-------C-- 2542
SMM.US.04.M926                -----G-----------T-----A-----------T--G----G---G--C-------------G-G-----C--C-----T--------CT-G--------C--------A-----A--T--------------------------------------------C-- 2542
SMM.US.04.M934                -----G-----------------G-----------T--G--A-G------C-------------G-------C--C-----T--------CT-G----------------------A-----A--------------------------A-----------G-----C-- 2540
SMM.US.04.M935                --A--G------------------------------G----G------C------T-----G-----------------T------------G-----------C-----A------------------------------G--C--------A---------A-------C-- 2543
SMM.US.04.M940                A-T--G--------T--G-----C------T-----------A-G------C------T-A---G----G----------T--------AT-G-----G--G--------------------C--CC----------T--C--------A--------GA-G--TG-G-- 2548
SMM.US.04.M946                -----G-----------------A-----------T--G----G------C-------------G-------C--C--C--T--------CT-G--------C--------A-----A--T-------------------------A-----G--G--G-----C-- 2541
SMM.US.04.M947                --A--G-----G------------------------G--A-G---G--C------T-----G-----------------------T------G--G--C-----A------------------------------------C--------A---------A-----C-C-- 2542
SMM.US.04.M949                A-T--G--------T--G-----C------T-----------A-G------C------T-A---G----G----------T--------AT-G-----G--T-----------------------C--C----------T--C--------A--------GA-G--TG-G-- 2536
SMM.US.04.M950                -----G-----------------A-----------T--G----G------C-------------G-------C--C-----T--------CT-G--G-----------------T--A--T------C-----------------------A--------G--G--T--C-- 2541
SMM.US.04.M951                A-T--G--------T--G-----C-----CT-----------A-G------C--------A--------G----------T--------AT-G-----G-----------------------C--CC----------T--C--------A--------GA-G--TG-G-- 2545
SMM.US.04.M952                -----G-----------------A-----------T--G--A-G------C-------------G-------C--C--C--T--------CT-G------A-C--------A--T--A--T-------------------------A--------G--G-----C-- 2541
SMM.US.05.D215                --G--G--G-----T--------A--------------G--A-G---G---------T-A---G-G--A--------C-----T-----AT-G-----GA-C-----------T-----T-----CC-G----------C--------A---------A---------- 2542
SMM.US.06.FTq                 --T---C---------------------------------A-G------------------G----G--------C--T--------A--G--CCC------C--A--A--T------------CC----------T--C--------------GA-------C-- 2570
SMM.US.11.SIVsmE660_FL10      --A--G------------------------------G--A-GG--G--C------T-----G------------------------------------------C-----A-----------------------T--A---------A-------C-- 3278
SMM.US.86.CFU212              --G--G--G-----T--------A------T-G-----G--A-----G--------C-----G----G-----------T----CT-------G--C--C-------T-----T--C--CC-------G----------------------A----T--C-- 2534
SMM.US.x.F236_H4              --A--G------------------------------G--A-GG--G--C------T-----G---------------------------------G-----C---------------------------------------A---------A-------C-- 3275
SMM.US.x.H9                   --A--G-----G-----------R------------G--A-G---G--C-----GT-----G---------------------T------------C-----A-------------C-------G-----C--------A---------A-------C-- 2761
SMM.US.x.PBJA                 --A--G-----G------------------------G--A-G---G--C-----GT-----G---------------------T------------C-----A-------------C-------G-----C--------A---------A-------C-- 3054
SMM.US.x.PGM53                --A--G-----G------------------------G----G------C------T-----G------------------------------------------C-----A--------T-----------------------C--------A---------A-------C-- 3205
SMM.US.x.SME543               --A--G------------------------------G--A-GG--G--C------T-----G-----------------T----------------C---------------------------------------A---------A-------C-- 3291
SMM.US.x.pE660.CG7G           --A--G------------------------------G--A-GG--G--C------T-----G------------------------------------------C---------------------------------------A---------A-------C-- 3277
STM.US.89.STM_37_16           A-T--GC----T------------------------G----G------C------T-----G---G----------T--------C--------G-GC----------T-----T--C--C-------------C--------A---------A-------C-- 2935
```

Draft of 30 September 2015, not for public release
hiv2dna_main.tex r8876–10546

```
MAC.US.x.239               TACAGTCAAAGGAACTCTTGAATAGCATAGGGTTTTCTACCCCAGAAGAGAAATTCCAAAAAGATCCCCCATTTCAATGGATGGGGTACGAATTGTGGCCAACAAAATGGAAGTTGCAAAAGATAGAGTTGCCACAAAGAGAGACCTGGACAGTGAATGATATACAGAAG 3628
Pol                        V__L__Q__S__K__E__L__L__N__S__I__G__F__S__T__P__E__E__K__F__Q__K__D__P__P__F__Q__W__M__G__Y__E__L__W__P__T__K__W__K__L__Q__K__I__E__L__P__Q__R__E__T__W__T__V__N__D__I__Q_
A.CI.88.UC2                -G---CT---A------C-A---G--C-------C-----------G-----G-----------C--T--G--------------C--T---C----------T--G------C----G--A---C-AC----C--G-A---A--A--------C-----C--C--A--A 3711
A.DE.x.BEN                 -G---CT---A-----TC-A---G--C-------------T-----T-----G-----------C--T-----------------C-GT---C-A--------T---------C----G--AC--C-AC----C--G-A---C-TA--------C-----C--C--A--- 3711
A.DE.x.PEI2_KR_KRCG        -G----T---A-----T--A---G--C----A--C--C--------T-----G-----------C-------ACA----------C--T---C-A--------C---------C-------A---C-A-----C----A---AGTA--------C-----C--C--A--- 3679
A.GH.x.GH1                 -G---CT---A-----TC-----G--C-------C-----------T-----G-----------C--T--GC-------------C--T---C-A--------T---------C----G--AT--C-AC----C--G-A---A-TA--------C--C--C--C------ 3155
A.GM.87.D194               -G--ACT---A-----TC-----G--C-------C-----------T-----G--------G--C--T--G--------------C--T--------------T--------AC----G--A---C-A--A--T--G-A---A-TA--------C-----C--C--A--A 3155
A.GM.x.ISY_SBL_6669_85     -----CT---------TC-A---G--C-G--A-----C--------C--A--G--------G--C--T----AC-----------C--T---C----------C--------A--------A---C-A-----C----AG--AGTA--------T-----C--C-----A 3152
A.GM.x.MCN13               -G---CT---------TC-A----A-C----A--------------T-----G--------G--C--T----AC--C--------C--T---C----------T--G------C----G------C-------C----A---TGTA--------A-----C--C--A--- 3156
A.GM.x.MCR35               -G---CT---------TC-A----A-C----A--------------T-----G--------G--C--T----AC--C--------C--T---C----------T--G------C----G------C-------C----A---TGTA--------A-----C--C--A--- 3156
A.GW.86.FG_clone_NIHZ      -G---CT---A-----TC-A---G-AC----A-----C--------T-----------G-----C--T----A--GC--------C--T-----A--------T------------------A---C--C----C----A---AGTA--------C-----C--C--A--- 3149
A.GW.87.CAM2CG             -G---CT---------TC-A----A-C----A--C-----------T-----G-----------C--T----AC-GC--------C--T---C----------T--------------G--A---C-A-----C----AG---GTG--------C-----C--C--A--- 3720
A.GW.x.MDS                 -G---CT----------C-A------C-G-----C-----------T--A--G-----------C--T----AC-GC--------C--T---C----------T---------C----G--A---C-------C----A---AGTA--------C-----C--C--A--A 3155
A.IN.07.NNVA               -----CT---------TC-A---G--C-------------------T-----G--------------T----ACT-C--------C--T-----A--------T--------------G--A---C-------C--G-AG---GTA--------C-----C--C--A--- 3705
A.IN.95.CRIK_147           -----CT---------TC-A---G--C-------------------T-----G--------------T----AC-----------T--T---C-A--------T--------------G--A---C-------C----AG---GTA--------C-----C--C--A--- 3456
A.JP.08.NMC786_clone_41    -----CT---A--G--TC-----G--C-------C-----------T--A--G-----------C--T--G--------------C--T---C-A------AT----------C-A--G--A---C-A-----T--G-A---AGTA--------C-----C--T--A--- 3707
A.PT.x.ALI                 -G----T----------C-A---G--C----A-----C--------T-----G-----------C--T----ACA----------C--T-G-C----------T---------C-------A---C-A-----C--G-A---AGTA--------C-----C--C--A--A 3704
A.SN.85.ROD                -G---CTC--------TC-A---G--C----A--------------T-----G-----------C--T----AC--C--------C--T---C-A--------T--------------G--A---C-------C----A---A-TA--------C-----C--C------ 3157
A.SN.86.ST_JSP4_27         -G---CT---------TC-A---G--C-G--A-----C--------T-----G-----------C--T----AC-----------C--T---C----------T---------C------GA---C-A-----C----AG--AGTA--------C-----C--C--A--A 3156
B.CI.88.UC1                CC--ACT---A--G--A--A---GA---G--A--C--------------A--G-----------C--T--G--CA----------T--T--GC-C-------A--GG-----AC-------A---C-AC----G----AG--AGTT--------------C--T--A--A 3692
B.CI.x.20_56               CT--ACT---AC----A--A---GA---G--A--C-----------------G-----------C--------CA-------------T--GC-C-------A---G-----AC-------A---C--C-A--G----A---AGT---------A-----C--C-----A 3687
B.CI.x.EHO                 CT--ACT---A--G--A--A----A---G--A--C-----T--------A--G-----------C--T-----CA-------------T--GC-C-------AG--------AC-------A---C--C-A--G----A----GTT--------A-----C--T------ 3682
B.GH.86.D205_ALT           CC--ACT---A--GT-A--A---GA---G--A--C--------------A--G-----------C--T--G--CA----------T--T--GC-C-------A---G-----AC-------A---C-AC----G----A---AGTT-------------CA--T--A--A 3684
B.JP.01.IMCJ_KR020_1       -T--ACTG--A--G--A--A---GA------A--C--------------A--G--------G--C--------CA-------------T--GC-C-------A--G------AC-------A---C--C----G--------A-GA--------A--------T-----A 2834
G.CI.92.Abt96              -G---CT------------A----AA-CR-----C--------T--T--A--G-----------C--T----A------------A--T-----R-------A--G------AC-A--------TA-CC----MG---A------T--------A--------T--A--A 3038
AB.CM.03.03CM_510_03       CC--ACT---A--G--A--A----A---G--A--C--------------A--R-----R--R--C--T--G--C-G---------T--T--GC-Y-------A--G------AC-------A---C-AC----G----AG--AGTT--------------C--T--A--A 2823
H2_01_AB.CI.90.7312A       CT--ACT---A--G--AC-A---GA---G--A-----C-----------A-----------G--C--T-----C--G--------T--T--GC-C-------AG--G------C-------A---C-AA-A--G----G----GTT--------A-----C--T-----A 3693
H2_01_AB.JP.04.NMC307_20   CT--ACT---A--GT-AC-A---GAT--G--A-----C-----------A--G--------G--C--------C--G--------T--T--GC-C-------A--G------AC-------A---C-AA-A--G----G----GTT--------A-----C--T-----A 3693
H2_01_AB.JP.07.NMC716_01   CT--ACT---A--GT-AC-A---GAA--G--A-----C-----------A--G--------G--C--------C--G--------C--T--GC-C-------A--G------A--------A---C-AA-A--G----G----GTT--------A-----C--T-----A 3692
H2_01_AB.JP.08.NMC842_10   CT--ACT---A--GT-GC-A---GA---G--A-----C-----------A-----------G--C--------C--G--------C--T--GC-C-------A--G-------C-------A---C-AA-A--G----G----GTT--------A-----C--T-----A 3698
U.CI.07.07IC_TNP3          ------T------T--TC-A---CAGC-G-----C-----------T-----G-----G-----C-------A------------T--T-----A-----C-AG--G------C----------CA-T-----G----A----GTA--------T-----C--T--A--- 3143
U.FR.96.12034              C----CTC-----G--AC------AG-C---A---AGC--------C--A-----T--G-----C--A---CACA----------C--T---C-A-----T-A--G------A-----G--A--T---C-C---G-------TGTT--------A-----C--T------ 3193
MAC.US.x.17EC1             ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 3628
MAC.US.x.251_1A11          ------T------------A------------------------------------------------------------------------------------------------------------------------------------------------------ 3627
MAC.US.x.251_32H_PJ5       ------T------------------------------------------A---------------------------------------------------------------------------------------------------------------------- 3629
MAC.US.x.251_BK28          -----CT------------A--------------C-----------------------------------------------------------------G------------------------------------------------------------------ 3604
MAC.US.x.MM142_IVMXX       ------T--A---------A-----------------AT-------------------------------------------------------------G------------------------------------------------------------------ 3114
MNE.US.82.MNE_8            ------T------------A----------------------------------------------------------------------------------------------------------------------------------------------------- 3096
MNE.US.x.MNE027            ------T------------A--------------------------------G-----------------------------------T-----A------------------------------------------A------------------------------- 3096
SMM.CI.79.SIVsmCI2         C----CTC-----GT-G--A---G-AT-G--A--CAGC--A-----T--C-----T--G-----C--A--CCA-A----------C--------A-------AG--G-----A--A--G--A--CA-A------G----G--C-GA--------A-----C--T--A--A 2748
SMM.LR.89.SIVsmLIB1        ------T---A-----T--A---G--T----A--C--A--T--------A--G--T--G-----C-------A-----------A---TA-G--A--------------A--A-----G--A-----------------A--A-GA--------A-----C--T--A--A 2707
SMM.SL.92.SIVsmSL92A       C----CT---A---T-G--A---GATT-G-----C--C-----C-----A-----T-----------G--T--C-----------A--T---C-A-AT--T-A----------------------ACCC----G----A---AGT---------A--------T------ 2709
SMM.SL.92.SL92B            C-----T--GA--CA-GC-C--C-ATC-G-----CAG------------C--G--T-----GA----A-----A--G--------A--TTTG--A-AC--C-AG--G-----A--A--G--A-----AC-C--TG---AG---CGA--------A-----C-----A--A 3059
SMM.US.04.G078             -G--ACT---------TC-A---G-GC----A--C-----------T-----G-----G-----C--------A--G-----------T--GC-------C--C---------C-A--G--A-----A----------A---A-AT--------A-----C--T-----A 2888
SMM.US.04.G932             -G--A-T---------T--A-----------A-----C-----------A--G-----G--G--------------G--------A--T--G--------------------------G--A----------------A---A-T---------A-----C--------A 2877
SMM.US.04.M919             ------T---------TC---C--T------A--C--C------------------G-----G--------------C--G--------A--T--------------C--------AC-A--G--A-----C--------------A-----------A-----C-----A--A 2882
SMM.US.04.M922             -----CT------G--TC-A--C--T--G--A-----C------------------G-----G--------------C--G--------A--T--------------C--------A-----G--A--------------------A-----------A-----C-----A--A 2883
SMM.US.04.M923             ----A-T---------TC---C--T------A-----C------------------G-----G--------------C--G--------A--T-----------C--C--------AC-------A-----------G--------A-----------A-----------A--A 2882
SMM.US.04.M926             -G----T---------TC-A--C----C---A--------T-----T-----G--------G-----------C--------------T------------G-----------C-------A-----C----------A---A--T--------A-----C--------- 2882
SMM.US.04.M934             -----CT---------TC-A--C--------A-----C--T-----T-----G-T---------A--------CA-------------T------------------------C-------A-----A----------A---A-----------A-----C--------- 2880
SMM.US.04.M935             ------T------G--TC-A--C--T-----A-----C------------------G-----G--------------C--G--------A--T--------------C--------A-----G--A--------------------A-----------A-----C-----A--A 2883
SMM.US.04.M940             -G--ACT---------TC----CG-GC----A--------------T-----G-----G-----C--------A--G-----------T--GC-------C--C--G------C-A--G--------AC---------A---A-AT--------A-----C--T-----A 2888
SMM.US.04.M946             -----CT---------TC-A--C-A---G--A--------T-----T-----G-----------A--------CA-------------T------------------------C-------A-----A----------A---------------A-----C--------- 2881
SMM.US.04.M947             ------T--G------TC-A--C--T--G--A-----C------------------G-----G-----C--------------------A--T--------------C--------AC-------A--------------------A--A--------A-----------A--A 2882
SMM.US.04.M949             ----A-T------G--TC-A---G-GC-G--A--------------T-----G-----G-----C--------A--C-----------T--C--------C--C--G----G-C-A-----A-----AC-T-------A---A-G---------A-----C--T--A--A 2876
SMM.US.04.M950             -----CT---------TC-A--------G--A--------T-----T-----G-T------------------CA-------------T--G---------------------C-------A-----A----------A---------------------C--------- 2881
SMM.US.04.M951             ----A-T---------TC-A---G-GC----A--------------T-----G-----G--------------A--GG----------T--GC-------C--C--G----G-C-A--G--A-----A----------A---A-G---------A--------T-----A 2885
SMM.US.04.M952             -----CT---------TC-A--C----C---A--------T-----T-----G-T---------C--------CA-------------T--------------------------------A-----A--------------A-----------------C--------- 2881
SMM.US.05.D215             ----A-TG--A--G-----A---G-------A--C-----------T--------T--G-----C--T----AC--T--------A--T--G--------------------A--------A--------A-----G--G--------------A-----C--C-----A 2882
SMM.US.06.FTq              -G----T---A-----A------G--C-------C--C--------T--------T--G-----C--T----A---T--------A--T--G--A-----T------------C-------------C-A------------A-G---------A-----C--T------ 2910
SMM.US.11.SIVsmE660_FL10   ------T---------TC----CG-------A--C---------------------G-----G--------------C--G--------A--T--------------C--------AC----G--A--------------------------------A-----C-----A--A 3618
SMM.US.86.CFU212           -G----T---A--------A--CG-A--G--C--C-----T--G--T-----G-----------C--T----A------------A----------------A-------------------A-----A-----T--G-A---C-AT--------A-----C--T-----A 2874
SMM.US.x.F236_H4           ------T---------TC----CG-------A--C---------------------G--------------------C--G--------A--T--------------C--------AC----G--A--------------------------------A-----C-----A--A 3615
SMM.US.x.H9                ----A-T---------TC---C--T------A-----C------------------G-----G--------------C--G--------A--T-----------C--C--------AC-------A-----------G--------A-----------A---A-------A--A 3101
SMM.US.x.PBJA              ------TG--------TC---C--T------A-----C------------------G-----G--------------C--G--------A--T-----------C--C--------AC-------A--------------------A-----------A-----------A--A 3394
SMM.US.x.PGM53             ------T---------TC---C---------A-----C------------------G-----G--------------C--G--------A--T--------------C--------AC----G--A--------------------A-----------A-----C-----A--A 3545
SMM.US.x.SME543            ------T---------TC----CG-------A--C--C------------------G-----G--------------C--G--------A--T--------------C--------AC----G--A--------------------------------A-----C-----A--A 3631
SMM.US.x.pE660.CG7G        ------T---------TC-A--CG-------A--C--C------------------G-----G--------------C--G--------A--T--------------C--------AC----G--A--------------------------------A-----C-----A--A 3617
STM.US.89.STM_37_16        ------TG---------C-A----ATT----A-----C-----T-----------T--G--G-----T-----------------A--T--G--A-----------G-----AC-------A-----A--A--T--G-----CGTT-----------------C--A--A 3275
```

Draft of 30 September 2015, not for public release
hiv2dna_main.tex r8876–10546

```
MAC.US.x.239              ATGTTATTACCAAGAAGGCAAGCCATTAGAAGCCACGGTAATAAAGAGTCAGGACAATCAGTGGTCTTATAAAATTCACCAAGAAGACAAAATA.CTGAAAGTAGGAAAATTTGCAAAGATAAAGAATACACATACCAATGGAGTGAGACTATTAGCACATGTAATACAG 3967
Pol                       __G__C__Y__Y__Q__E__G__K__P__L__E__A__T__V__I__K__S__Q__D__N__Q__W__S__Y__K__I__H__Q__E__D__K#_I__L__K__V__G__K__F__A__K__I__K__N__T__H__T__N__G__V__R__L__L__A__H__V__I_
A.CI.88.UC2               -CAC------------AG--AGA---G--G--A---A-CCA---A-----A--------A---A-A--------A-----G-----G-----C.--A-------------A------A-----A--C--C-----------G--C---T--C-------G---G-T--- 4050
A.DE.x.BEN                --A-------------AA--AGA------G--A--AA-CCA---A--C--A-GAC----A---A-A--C-----A-----G-----G-----C.--A-----------G-A-----------A-----C-----------G--C---T--C-------G---G-T--- 4050
A.DE.x.PEI2_KR_KRCG       ---C------------AA---GA---------A--A--CCA----GA---A------------A-A--------A-------G---G-----C.--A-------------A-----------A-----C-----------G--C---T-G-----------G-T--A 4018
A.GH.x.GH1                --AC------------AA--AGA------G--A--AA-CCA---A-A---A------------A-A--C-----A-----G-----G-----C.--A--------G---A-----------A-----C-----------G--C---T--C-------G---G-T--- 3494
A.GM.87.D194              --CC--C--T--G----AAG-AGA-C-------A--A--C--C--A--C--A------------G-A--C-----A-----G-G---G-GGG-T.--A-----------G-A---G--------A-----T-----------G--C------C-------A---G-C--A 3494
A.GM.x.ISY_SBL_6669_85    -CAC----------G-AA--AGAG--------A--A--TCA----GA---A------------A-A------G-A-----G-G---A-----T.--A-------------A-----------A-----C--------C--G--C--GT-G--G-----G---G-T--- 3491
A.GM.x.MCN13              -CAC------------AA--AGAG--------A--A--CCA----GA---A------------A-A------G-A-----G-G---G-----T.--A--------G---A-----------A-----C-----------G--C---T-G--------A---G-T--A 3495
A.GM.x.MCR35              -CAC------------AA--AGGG--------A--A--CCA----GA---A------------A-A------G-A-----G-G---G-----T.--A--------G---A-----------A-----C-----------G--C---T-G--------A---G-T--A 3495
A.GW.86.FG_clone_NIHZ     -CAC--------G----AA--AAAG--------A--A--CCAG--AGA---A------------A-A-----GG-A-----G-GG--G-----TCTCA--G--G-------A-----------A-----C--------C--.--C---T-G--------G---G-T--A 3488
A.GW.87.CAM2CG            -CAC--C-----G----AA---GAG--------A--A--TCA----GA---A------------A-A-----G--A-----G-----A-----T.--A-------------A------A-----AC----C---------------C-A-T-----------G---G-T--A 4059
A.GW.x.MDS                -CAC----------G-AA--AGA---------A--A--CCA----GA---A------------A-A------G-A-----G-G---A----CT.--A-------------A-------G-----A--C--C--------C-----C---T-G--------G---G-T--- 3494
A.IN.07.NNVA              -CAC------A----G-AA--AGA-C----G--A--A--CCA---AGAC--A------------A-G--C--GG-A--T--G-G---A-----C.--A-------------A-------G-----A---G-C--------C--G--C---T--C--------G---G-T--- 4044
A.IN.95.CRIK_147          -CAC----------G-AA---GAGC----------A--CCA---AGAC--A------------A-A------G-A-----G-G---A-----C.--A-------------A-------G-G--A---G-C-----------G--C---T-GC-------G---G-T--- 3795
A.JP.08.NMC786_clone_41   G--C------------AA--AGA------G--A--A--CCA---------A------------A-A--C-----A-----G-----G-----C.--A--------G---A-----------A-----C-----------G--C---T--C-------G---G-T--A 4046
A.PT.x.ALI                GCAC------------AA---GAG--------A--A--CCA---AGA---A---------A---A-A--------A-----G-G---A-----T.--A-----G-A---G-A------A--G--A-----C--------C--G--C---T-G--------G---G-T--A 4043
A.SN.85.ROD               -CAC------------AA--AGAGC-------A--A--CCA----GA---A--G---------A-A--------A-----G-----A-----T.--A-------------A-------G-G--A--C--C------------A-C---T-G--------G---G-T--- 3496
A.SN.86.ST_JSP4_27        ---C----------G-AA---GAGC-------A--A--CCA---AGA---A------------A-A------G--A-----G-G--GA-----T.--A-------------A-------G-----A-----C--C-----C-----C-----CC-------A---G-T--A 3495
B.CI.88.UC1               --CC--C---A-G-----GGTA---C--------A--A---CAG--A-AC-TA-CA---------A-A--C--------T--G-G-A-T-G---T.T-A-------------A-------G-T--A--C--C--C-----C--G--A------C-G----------G-T--- 4031
B.CI.x.20_56              --CC--C---A-------AGTA--TC-----G--A--A--GCAG--A-A--TA-CA--C------A-A--C--------T--G-GGA-T-GG-CC.--------------A-------G-T--A--C--C--------------------C-G--G-----GG-C--- 4026
B.CI.x.EHO                --CC------A-------GGTA--T--------A--A--GCAG--A-A--TA-CA---------A-A--C--G-----T--G-G---T-----C.--A-------------A-------G-T--A--C--T--C-----------A--------G--T------G-C--A 4021
B.GH.86.D205_ALT          --CC--C---A-G---A-GGTA--GC-------A--A---CAG--A-AC-TA-CA---------A-A--C--------T--G-G-A-T---G-C.--A-------------A-------G-T--A--C--G--C-----C--G--A------C-G----------G-T--- 4023
B.JP.01.IMCJ_KR020_1      --CC--C--TA----G--A-TA--TC-------A--A---CAG--A-AC-TA-CA---------A-A--C-----C--T--G-G-----G---C.--A-G-----------A-------G-T--A--C--C--C-----C-----A------C-G-------------C--- 3173
G.CI.92.Abt96             CGCC------A------RAG----T--R--------T---CAG--A-A-TTA------------A-A--C---G----T----G-A---G---C.--A--------T---A-------------A-----C-----A--------A---T----G--T-----GG-C--A 3377
AB.CM.03.03CM_510_03      --CC--C---A-G-----GGTA-----------A--A---CAG--A-AC-TA-CA---------A-A--C--G-----T--G-G-A-Y-G---G.T-A-------------A-------G-T--A--C--C--C-----C--G--A--RA-RC-G----------G-T--- 3162
H2_01_AB.CI.90.7312A      --C---C---A-------GGTA--TC-------A--A--GCAG--A-ACTTA-CA--C-------A-A--C--------T--G-G-A-T---G-C.--A-------------A-------G-T--A--C--C--C------------A---A--C------------G-C--A 4032
H2_01_AB.JP.04.NMC307_20  --CC--C---A-------GGTA--TC-------A--A--GCAG--A-ACTTA-CA--C-------A-A--C--------T--G-G---T---G-C.--A-------------A-------G-C--A--C--C--C---------G--A---A--C---------------C--A 4032
H2_01_AB.JP.07.NMC716_01  --C---C---A-------AGTA--TC-------A--A---CAG--A-ACTTA-CA--C-------A-A--C--------T--G-G---T---G-C.--A-------------A-------G-C--A--C--C--C------------A---A--C---------------C--A 4031
H2_01_AB.JP.08.NMC842_10  --CC--C---A-------GGTA--TC-------A--A--GCAG--A-ACTTA-CA--C-------A-A--C--------T--G-GG--T---G-C.--A--G---------A-------G-C--A-----C--C------------A---A----G--------------C--A 4037
U.CI.07.07IC_TNP3         CACA--C---AG------GG-A--T--------A--A--GCAG----A--TA-----C------A-A----G--C--T--G-GCA-T-----C.T-A------------------------------C--C--A-----G--A--GT----G--T------G--T--A 3482
U.FR.96.12034             GGC--------AGG----AAG-AAATC-------A--A--GC-T--A-A---A--T-----A---A-C--C--G--C--T--G-GT----GG-C-.T-A-----------------C---G----------T-----A---------TC-C--GC-----------GG----- 3532
MAC.US.x.17EC1            ---------------------------------------------------------------------------------------------.------------------------------------------------------------------------------------ 3967
MAC.US.x.251_1A11         ---------------------------------------------------A------------------------------------------.-----------------------------------------------------T----------------------------- 3966
MAC.US.x.251_32H_PJ5      ---------------G-----------------------------------------------------------------------------.-------------G---------------------------------------T----------------------------- 3968
MAC.US.x.251_BK28         ---------------------------------------------------------------------------------------------.-----------------------------------------------------T----------------------------- 3943
MAC.US.x.MM142_IVMXX      ----------------A------------------------G---------------------------------------------------.-----------------------------------------------------T----------------------------- 3453
MNE.US.82.MNE_8           ----------------------------G---------------A-------------------G----------------------------.--A---------------------------A-----------------------T----------------------------- 3435
MNE.US.x.MNE027           ----------------------------G---------------A-------------------G----------------------------.-------------------------------A-----------------------T----------------------------- 3435
SMM.CI.79.SIVsmCI2        -GCC-T----AGG----AAG-AAAT--------A--A--GCA-----A---A--T-----A---A-A--C--G--A--T----GTAGT-GGG--.--A-----------G-----C--------------C-------------TC-T--CC-G----------G-G--- 3087
SMM.LR.89.SIVsmLIB1       C--C-------------AT-----T--------A--A--GG---A-A---A----GC--A------------G--C--T--G-----T-----T.--A-G-------C-------T--AG-G--A--C----------------A--GT-GC-G----------G-T--A 3046
SMM.SL.92.SIVsmSL92A      C-C------T--G--G--AG-------------A--T-C--------A---A--A--C--A------A--C--G--A-------G---A--G--T.T-A--------T---A---------C--A-----T--C--A----------C-----C--G--C--G---------A 3048
SMM.SL.92.SL92B           -A-A------AG-----ATG-A--GC-------A-----TC-----A---A------------AG-----G-----T----GT----GG--T.--C--G-----G--G---------A------A-----------A------A-A--GT--C-G--TA-----G-T--A 3398
SMM.US.04.G078            G--C-------------AT--A--CC-------A--CA--G-G----------------C--A------------------T----G---A--GG--.--A---------C----A--------------------C--A---------C--GT-GC-----------G-C--- 3227
SMM.US.04.G932            ------------G----------------------------G--A-A-------------------C------------------------.T-A-----------G------G-A--------C--T---------------C--GT---------------G----- 3216
SMM.US.04.M919            ------C--------G--A------C-------A--A--------A--------T-----A-----A----------------------G---.T-----------C-----C--------C--A---------C--A---------C---T-----G---------G-G--- 3221
SMM.US.04.M922            ------C-----------G------C----G--A--A--------A--------T-----A-----A----------------------G---.---------G--C-----C--------T--A--------------A--C-----C---T-----------C--GG-G--- 3222
SMM.US.04.M923            ---C--C--T--G--G--A--A---C----G--A--A-----------------T-----A-----A--------------------------.-----------T--------------T--A--------------A-------C--GT--C-----------G-G--- 3221
SMM.US.04.M926            ---------------------A----G------------------A------T------------------G--C--------------G-C.T-A-----------------G--A----A--C--------A--------T-------G--------G------ 3221
SMM.US.04.M934            ---C-----------------A----G------------------A------T-----A-----------G--C-----------------C.T-A-------------------------A--C--------A--------T--GT----G-------------- 3219
SMM.US.04.M935            ------C-----------A-----C----G--A--A--------------------T-----A-----A----------------------G--.-----------C-----C--------T--A--------------A--C-----C---T----------------G-G--- 3222
SMM.US.04.M940            G----------------AT--A--CC----G--A--TA----G----------------A----------G-----T----G---A-GGG--.--A--G-----C----A--------------------A--C-----C--A--------C--GT-G------------G-C--- 3227
SMM.US.04.M946            ---C-----------------A----G------------------A------T-----A----------G--C------------------C.T-A-------------------------A--C--------A--------T--GT-G--G--------G------ 3220
SMM.US.04.M947            G--------T-----G--A------C-G--G--A--A-----------------T-----A-----A--------------------------.--A--G-----C---------------T--A--------------A-------C---T-----------C---G-G--- 3221
SMM.US.04.M949            G----------------AT--A--CC----G--A--CA--G-G--------------C--A----------G-----T----G---A--G---.--A--G-----C----A---C--AG----A---------C--A---------C--GT-G------------G-C--A 3215
SMM.US.04.M950            ---C-----------------A----G------------------A------T-----A----------G--C------------------C.T-A----------G--------------A--C--------A--------T--GT-G---------------G------ 3220
SMM.US.04.M951            G--C----------C--AT--A--CC----G--A--CA--G-G-------------T-----A---------------T----G---A--GG-G.--A--G-----C----A---C---G-G--A---------C--A---------C--GT-GC-----------G-C--- 3224
SMM.US.04.M952            ---C-----------------A---------------------A-------------A----------G--C------------------C.T-A-------------------------A--C--------A--------T--GT-G--G--------G------ 3220
SMM.US.05.D215            G--------T--G----A-----A---------A--T---G-G----A---------------------G--C--T--G----GA--G-CC.T-A--G-------------A-----A--C--------------------T--GT----G-----C----------- 3221
SMM.US.06.FTq             G-----C--------G-AT-----C--------A--CA--G-G--A--C--A--T----------------G--C--T-----A-T------.-----------T---A------A--------------------A--------C--GT----G----------G---A 3249
SMM.US.11.SIVsmE660_FL10  ------C--------G--A--A---A----G--A--A-----------------T-----A-----A-----------------------G--.-----------T-----------G-T--A--------------A-------C---T-----------C---G-G--- 3957
SMM.US.86.CFU212          ------C--------G--A-----T----G--A------G-G--A-A-----T---------------C-----C-----G-----T--G--.T-----------C---------C--A--------------------A--------T--G-----G---------GG---A 3213
SMM.US.x.F236_H4          ------C-----------A--A---A----G--A--A-----------------T-----A-----A-----------------------G--.-----------T-----------G-T--A--------------A-------C---T--C--------C---G-G--- 3954
SMM.US.x.H9               ------C--T--G--G--A------C----G--AC-A-----------------T-----A-----A-------------G---------.--A-------C---------------T--A--------------A-------C--CT-----------------G-G--- 3440
SMM.US.x.PBJA             ------C--T--G--G--A------C----G--A--A-----------------T-----A-----A--------------------------.-----------C---------------T--A--------------A-------C---T-----------------G-G--- 3733
SMM.US.x.PGM53            ------C--------G--A-------------G--A--A-----------------T-----A-----A-------C--------------G--.-----------C---------------T--A--------------A-------C---T-----------------G-CG-- 3884
SMM.US.x.SME543           ------C--------G--A--A---A----G--A--A-----------------T-----------A-----------------------G--.--A-----G--C-----------G-T--A--------------A-------C---T--C--------------G-G--- 3970
SMM.US.x.pE660.CG7G       ------C--------G--A--A---A----G--A--A-----------------T-----A-----A-----------------------G--.-----------T-----------G-T--A--------------A-------C---T-----------C---G-G--- 3956
STM.US.89.STM_37_16       GCAA--C----G-----A------TC----G--A--A---G-G--AGA------------A-----C-----G--A-----------------.T-----------C--G---------A-----A--C---------A--------T---T----G--------G--T--- 3614
```

Draft of 30 September 2015, not for public release
hiv2dna_main.tex r8876–10546


```
MAC.US.x.239                 AAAATAGGAAAGGAAGCAATAGTGATCTGGGGACAGGTCCCAAAATTCCACTTACCAGTTGAGAAGGATGTATGGGAACAGTGGTGGACAGACTATTGGCAGGTAACCTGGATACCGGAATGGGATTTTATCTCAACACCACCGCTAGTAAGATTAGTCTTCAATCTAGT 4137
Pol                          _Q__K__I__G__K__E__A__I__V__I__W__G__Q__V__P__K__F__H__L__P__V__E__K__D__V__W__E__Q__W__W__T__D__Y__W__Q__V__T__W__I__P__E__W__D__F__I__S__T__P__P__L__V__R__L__V__F__N__L
A.CI.88.UC2                  ------------------C----C--T-------G-A-A--------T---C----G--G----GA--GACC-----G---------GATA----C--------G--A-----C--A--------C--CG-A--T--C-----A--G--C--G--GACA--T--C----- 4220
A.DE.x.BEN                   -----------A--G---C----C--T-------G-A-A--------T---C-G-----G----GA--GACC-----G---------GATA----C-----A--G--A-----C--A--G-----C---G-A--T--C-----A--G--C--G---ACA--T--C----- 4220
A.DE.x.PEI2_KR_KRCG          -----------A------C----C--T-------GAA-A--------T---C-------A--A-GA--AACC-----G---------GATA----------A--G--A-----C--A--C-----C--CG-A--T--T-----A--G--C--G----CA--T--C----- 4188
A.GH.x.GH1                   -----------A------C----C----------G-A-A-----------------G--G----GA--GACC-----G---------GATA----C-----A--G--A-----C--A--G-----C--CG-A--T--C-----A-----C--G---ACA--T--C----- 3664
A.GM.87.D194                 ------------------C-G--C--T-------GA--G--------T---C----G--A----GA--CACC-----G--A------GATA----C-----A-----A---G-C--A--G-----C--CG-A--T--C-----A--G--C--G--GACA--T--CT-G-- 3664
A.GM.x.ISY_SBL_6669_85       -----------A------C----C--T-------GAA-A--------T---C-------A--A-GA--GACC---------------GATA----------A--G--A-----C--A--C-----C--CG-A--C--C------T-G--C--G----CA--T--C--G-- 3661
A.GM.x.MCN13                 --G--------A------C--A-C--T-------GAA-A--------T---C-------A--A-GA--GAC----------------GAT-----C--------G--A-----C--T--C-----C--CG-A--T--C--G-----G--C---C---CA--T--C--G-- 3665
A.GM.x.MCR35                 --G--------A------C----C--T-------GAA-A--------T---C-------A--A-GA--GACC---------------GAT-----C--------G--A-----C--T--C-----C--CG-A--T--C--G-----G--C---C---CA--T--C--G-- 3665
A.GW.86.FG_clone_NIHZ        -----------A------C----C--T-------G-A-A--------T---C-------A----GA---ACC-----G---------GATA----C-----A-----A-----C--A--C-----C---G-A--T--C-----A--G--C--GC---CA--T--C--G-- 3658
A.GW.87.CAM2CG               -----------A------C----C--T--A----GAA-A--T-----T---C-------A----GA--G--C-----G---------GATA----C--------G--A-----C--A--C-----C--CG-A--T--C-----A--G--C---C---CA--T--CT---- 4229
A.GW.x.MDS                   -----------A------C----C--T-------GAA-A--------T---C----G--A----GA--AA-C-----G---------GAT-----C--------G--A-----C--A--C-----C---G-A--T--C-----A--G--C--G-----A--T--C----- 3664
A.IN.07.NNVA                 ------------------C----C--------G-GAA-A--------T---C-------A--A-GA--AACC---------------GAT--T--C--------G--A-----C--A--C-----C--CG-A--T--C-----A--G--C---C---CA--T--C----- 4214
A.IN.95.CRIK_147             --G--------A------C----C----------GAA-A--------T---C-------A----GA--AACT-----G---------GAT-----C-----------A-----C--A--C-----C---G-A--C--C-----A--G--C---C---CA--T--C----- 3965
A.JP.08.NMC786_clone_41      -----------A------C----C----------GAA-A------------C----G--G----GA--GACC-----G---------GATA----C-----A-----A-----C--A--G-----C---G-A--T--C--G--A--G--C--GC--ACA-----C--G-- 4216
A.PT.x.ALI                   -----------A------C-G--C--T-------GAA-A----G---T---C-------A--A-GA--AACC---------------GAT-----C-----A--G--A-----C--A--C-----C---G-A--T--C-----A--G--C--GC---CA--T--C----- 4213
A.SN.85.ROD                  -----------A------C----C--T-------GAA-A--------T---C-------A----GA--AA-C-----G---------GATA----C-----A--G--A-----C--A--C-----C--CG-G--T--C-----A--G--C--G----CG--T--C--G-- 3666
A.SN.86.ST_JSP4_27           -----------A------C----C--T-------GAA-A--------T---C-------A--A-GA---ACC---------------GATA----C-----A--G--A-----C--A--C-----C--C--A--T--C--G--A--G--C--------A--T--C--G-- 3665
B.CI.88.UC1                  -----------A-----CC----C---------G--A-A---GT------TC-G-----A--A-GA--GAC------C--------------T--C-----A-----T-----C--A------------G-------C-----AT--------C---C--A---C----- 4201
B.CI.x.20_56                 -----------A------T----T---------G-AA-A----TG-------C-G--G--A----GA--GACG-----T-----------------C-----------T-----C--A--------------T-----C-----AT--------C---C--A---C----- 4196
B.CI.x.EHO                   ------------------T-G--C---------G--A-A----TG------TC-------A--A-GA--GAC------T--------------T--C-----A-----------C--A------------G-------C-----AT--A----G---C---AT--C--G-- 4191
B.GH.86.D205_ALT             --------C--A-----CC----C---------G--A-A---GTG-----TC-G-----A--A-GA--GAC------C--------------T--C-----A-----------C--A--G-----C---G----G--C-----AT--A-----C---C--A---C----- 4193
B.JP.01.IMCJ_KR020_1         ------------------T-G--C-C-------G-AA-A----TG-------C-G-----A--A-GA--GAC------T--A-----------T--C-----A-----------C--A------------G-------------TT-GA-----C---C--AT--C----- 3343
G.CI.92.Abt96                --R---------------C----A--A------A-AT-A--CTTC--T--TC-G-----A--A-G---CACC--------A--------------------A-----T-----C--A--T-----C---G-G--T-----G-AT--------G-------A---C----- 3547
AB.CM.03.03CM_510_03         -----------A-----CC----C---------G--A-A----T------TC-G-----A--R-GA--AAC------C--------------T--Y-----A--G--------C--A--G-----C---G----G--C-----AT--A-----C---C--A---C----- 3332
H2_01_AB.CI.90.7312A         -----------A------T-G--T--T--------AA-A----TG-------C-G-----A----GA--GAC------C--A-----------T--C-----A--G--T-----C--A--------C---G-------C-----AT-G-----GC---C--A---C----- 4202
H2_01_AB.JP.04.NMC307_20     -----------A------T-G--T--T--------AA-A----TG------T--G-----A--A-GA--AAC------C--A-----------T--C-----A--G--T-----C--A--------C---G-------C-----AT-G------C---C--A---C--G-- 4202
H2_01_AB.JP.07.NMC716_01     -----------A------T-G--T--T--------AA-A----TG------T--G-----A--A-GA--AAC------C--A-----------T--C-----A-----T-----C--A--------C---G-------C-----AT-G-----GC---C--A---C--G-- 4201
H2_01_AB.JP.08.NMC842_10     -----------A------C-G--T--T--------AA-A----TG------T--G-----A----GA--AAC------C--A-----------T--C-----A-----T-----C--A--------C---G----G--C-----A---------C---C--A---C--G-- 4207
U.CI.07.07IC_TNP3            --G--------A-----------A--------GA-A--G----TC-------C-G-----G----GA--AAC------------------------------A--G--------T--A--G---------G-A--T--------TT-G-----G----CA-AT--C--T-- 3652
U.FR.96.12034                --G--------A-----C-----A--A------A-AA-T----TG------TC-G-----A--A-GA--G--C--------A--------T--------------C------G-C-TA--C---------G-A--C--------C--G--G---C-G-CA-AT-------- 3702
MAC.US.x.17EC1               ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 4137
MAC.US.x.251_1A11            --------------------------------------------------------------------------------------------------------------------A------A-------------------A-----G-------------------- 4136
MAC.US.x.251_32H_PJ5         --------------------------------------T-----------------------------------------------------------------------------A--------------------------A-------------------------- 4138
MAC.US.x.251_BK28            ----------------------------------------------------------------G------------------------------------------------------G-----------------G-----A-------------------------- 4113
MAC.US.x.MM142_IVMXX         --------------------------------------------------------------------------------------------------------------------------------C--------------AT------------------------- 3623
MNE.US.82.MNE_8              -----------------------------------------------------------------A---------------------------------------------------A-------------------------A-----------------------G-- 3605
MNE.US.x.MNE027              -----------------------------------------------------------C--------------------------------------------------------A--------------------G--G--A-------------------------- 3605
SMM.CI.79.SIVsmCI2           -----------A------T-G--T---------A--A-A--GTTC-----TC-G-----A--A-GA--A--------G--A--------------------A-----------C--A--T---------G-A--C--------AT-G--C--------A--------G-- 3257
SMM.LR.89.SIVsmLIB1          -----------A------T----A----AA--G---A-T-----G--------------A--A-GA--AA-C-A------A-----------T-----------G--A--------A---A----------A--C-----C--TT-G--T--G-----------C-CG-- 3216
SMM.SL.92.SIVsmSL92A         --G--T-----------------A--T------A----G--------------G-----A--A-GA--GC-------G--A--------------------A-----T---G-T--T--T-----C---G-A--C--C--C--A---A----GA-G-CT--T---T-G-- 3218
SMM.SL.92.SL92B              --G--------A---AGTC-------A------A--ACA---TTT--T---C-T-----A--A-G---G--T-----T--A-----------T--------A-CT--T-----T--A--G-----C--C-----C---------T--A-C--G-----G--T---T---- 3568
SMM.US.04.G078               -----------A-----------A--T--------A--G--T-----T--TA-G-----A----GA--AACT--------A-----------T--C-----A--G--A-----------------C--CG-G--G--C--T--T-----T--------G------T-G-- 3397
SMM.US.04.G932               -----------A--------------------------T-----------------------A-GA--GA----------A-----------T--------A--G-----------C------------G-T-----------AT---G---C-G---A--T--CT-G-- 3386
SMM.US.04.M919               --------G--A--G--------A--T-----------G--------T--T--G-----A----GA--AA-T--------A-----------T--------A--G-----------A--------C---G-G--------T--CT----C--------------CT-G-- 3391
SMM.US.04.M922               --------G--A--------------T-----------G--------T--T--G-----A--A-G---AA-T--------A-----------T--------A--G-----------A--------C---G-G--------T--CT----C-------------------- 3392
SMM.US.04.M923               -----------A-----------A--T-----------G----G---T--T--G-----A--A-GA--AA-T--------A-----------T--------A--------------A--------C---G-G--------T--CT----C---C----T----------- 3391
SMM.US.04.M926               -----------------------A------------A--------------C-G-----C----GA-----T-----G--A--------------------A--G-----------T--------C---G----G-----C--AT--------------------T---- 3391
SMM.US.04.M934               ------------------------------------A----------------------C--A-GA-----G-----G--A--------------------A--G--T--------T--------C---G----G-----C--AT-------------------CT---- 3389
SMM.US.04.M935               --------G--A--------------T-----------G--------T--T--------A--A-GA--GA-T--------A-----------T--------A--G-----------A--------C---G-G--------T--CT----C-------------------- 3392
SMM.US.04.M940               -----------A-----------A--T--------A--G--T-----T--TA-G-----G----GA--AACT--------------------T--C-----A--G--A-----------------C--CG-G-----C--T--T-----C--------G-----CT---- 3397
SMM.US.04.M946               -----------A------------------------A----------------G-----C--A-GA-----G-----G--A--------------------A--G--T--------T--------C---G----G-----C--AT-------------------CT---- 3390
SMM.US.04.M947               -----------A-----------A--T-----------G--------T--T--G-----A--A-GA--AA-T--------A-----------T--------A--------------A--------C---G-A--------T--CT----------------T--C----- 3391
SMM.US.04.M949               -----------A-----------A--T--------A--G--T-----T--TA-G-----A----GA--AA-C--------------------T--C-----A--G--A--------A-----------CG-G--G--C--T--T-----C--------G------T-G-- 3385
SMM.US.04.M950               ------------------------------------A----------------G-----C--A-GA-----G-----G--A--------------------A-----T--------T--------C---G----G-----C--AT-------------------CT---- 3390
SMM.US.04.M951               ---G-------A-----------A--T--------A--G--T-----T--TA-G-----A----GA--AA-C--------------------T--C-----A--G--A--------------------CG-G-----C--T--T-----C--------A------T-G-- 3394
SMM.US.04.M952               -----------------------A------------A----------------G-----C--A-GA-----G-----G-----------------------A-----T--------A--------C---G----G-----C--AT-------------T-----CT---- 3390
SMM.US.05.D215               --G-----G--A-----------A--T--------A--G-----------TC-G--G--A----GA--A--T-----------------------C-------C---T--------A--G-----C---G-G--C-----C--A--------------------C----- 3391
SMM.US.06.FTq                --G--------A-----------A--T------A-A--G--T-----T--TA-G-----A--A-GA--AA-T-----G-----------------------------A--------A-----------C--A--------C--TT----T--GC----G--T-----G-- 3419
SMM.US.11.SIVsmE660_FL10     -----------A------C----A--T------G----G-----------T--G-----A--A-GA--AA-T--------A-----------T--------A--G-----------A--T-----C---G-G--------T--CT----C--------------C----- 4127
SMM.US.86.CFU212             --G-----------G--T----------------G-A-------------TC-------A----GA--AA-C-----G--A--------------C--------G-----------A--C-----C---G-A--------C--TT-G--C--------T-----C----- 3383
SMM.US.x.F236_H4             -----------A------C----A--T------G----G-----------T--G-----A--A-GA--AA-T--------A-----------T--------A--------------A--T-----C---G-G--------T--CT----C--------------C----- 4124
SMM.US.x.H9                  -----------A---------A-A--T-----------G----G---T--T--G-----A----RA--AA-T--------A-----------T--------A--------------A--------C---G-G--------T--CT----C---C----T-----C----- 3610
SMM.US.x.PBJA                -----------A-----------A--T-----------G--G-G---T--T--G-----A----GA--AA-T--------A-----------T--------A-----T--------A--------C---G-G--------T--CT----C---C----T-----C----- 3903
SMM.US.x.PGM53               --------G--A-----------A--T-----------G--------T--T--G-----A----G---AA-T--------A-----------T--------A--------------A--------C---G-G--------T--CT----C--------------C----- 4054
SMM.US.x.SME543              --------------G---C----A--T------G----G-----------T--G-----A--A-GA--AA-T--------A-----------T--------A--------------A--T-----C---G-G--------T--CT----C--------------C----- 4140
SMM.US.x.pE660.CG7G          -----------A------C----A--T------G----G-----------T--G-----A--A-GA--AA-T--------A-----------T--------A--G-----------A--T-----C---G-G--------T--CT----C--------------C----- 4126
STM.US.89.STM_37_16          --G-----G--A------------------------A----------------G-----A----GA--G--C--------A------G----T--------------A--------A------------G-A--C-----T--TT----------------T---T---- 3784
```

```
                                     Pol RT end  Pol p15 RNAse H start
MAC.US.x.239                GAAGGACCCTATAGAGGGAGAAGAAACCTATTATACAGATGGATCATGTAATAAACAGTCAAAAGAAGGGAAAGCAGGATATATCACAGATAGGGGCAAAGACAAAGTAAAAGTGTTAGAACAGACTACTAATCAACAAGCAGAATTGGAAGCATTTCTCATGGCATTGA 4307
Pol                         __V__K__D__P__I__E__G__E__E__T__Y__Y__T__D__G__S__C__N__K__Q__S__K__E__G__K__A__G__Y__I__T__D__R__G__K__D__K__V__K__V__L__E__Q__T__T__N__Q__Q__A__E__L__E__A__F__L__M__A__
A.CI.88.UC2                 AGGA--T------CCA--CAC---G----TC--C--------------C----G---------G----A-----------G-A-------A--A-G-----G---G--A-------A----C----G---------C-A--G--C---GCG------C--G 4390
A.DE.x.BEN                  AGGA--T------CCA--C-C---G----TC--C--------------C----G------------G--A-----------G-A-------A--A-----------------A-------A----C----G--G-------A----TC----GG------C--G 4390
A.DE.x.PEI2_KR_KRCG         A--A--T--C---CCA--T-----G----TC--C-----------C-------GG--A--------G--A-------------A-------A--G-G-----G---GGA-A--G--G--A----C----G-----------A-----C--CGCA--------A- 4358
A.GH.x.GH1                  AGGA--T------CCA--C-C---G----TC--C--------------C----G---------G----A------C-----G-G-------A--A-G-----G---G--AC------GA----C----G-----------A-----C---GCG---A--C--- 3834
A.GM.87.D194                AGGA--T------CCA--CAC---G----T---C--------------C----G---------------A-----------G-A-------A--G-G----GG---G--A-----G--A--AT-C-----G---------C-A-----C---GCG------C--G 3834
A.GM.x.ISY_SBL_6669_85      A--A--T------CCA--C-C---G----TC--C--G--------T--C----GG--A--------G--A---------------A-------A--A---------------GGA-A-----G--A----C--C--------------A-----C--GCA------G-A- 3831
A.GM.x.MCN13                A--A--T------CCAA--AC---G--T-TC--C-----------C--C---GG--A-----G----A-----------G-A-------A--G-G-----G---GGA--C-------A----C----G-----------A-----C--GCA------C-A- 3835
A.GM.x.MCR35                A--A--T------CCAA--AC---G--T-TC--C-----------C--C---GG--A-----G----A-----------G-A-------A--G-G-----G---GGA--C-------A----C----G-----------A-----C--GCA------C-A- 3835
A.GW.86.FG_clone_NIHZ       AGGA--G---G--CCA--C-C-------T-TC--C-----------C--C---GG---------------------------A-------A--G-G----GG-------A-----G--A----C----G-----------A-----C---GCA------C-A- 3828
A.GW.87.CAM2CG              AGG---T------CCA--CAC---G----TC--C-----------C--C---GG--A-----------A------------G-A-------A--A-G-----G-----A-AC----G--A----C----G---------G--A-----C---GCG------C-A- 4399
A.GW.x.MDS                  A--A--T------CCA--T-C--------TC--C-----------C--C---GG--A-----------A------------G-A------AA--G-G-----G-----G--AC---G--A----C----G-----------A-----C---GCA------C-A- 3834
A.IN.07.NNVA                A--A--T--C---CCA--C-C---G---ATC--C-----C-----C--C-----G--------------A------------G-A-------A--G-G----GG-----G--AC---G--A----C----G-----------A-----C---GCA------C-A- 4384
A.IN.95.CRIK_147            A--A--T--C---CCA--CAC--------TC--C-----------C--C---GG-------------A--G---------G-A-----C--A--G-G-----G-----G--AC---G--A----C----G--T------A-----C---GCA------C-A- 4135
A.JP.08.NMC786_clone_41     AGGA--T------CCA-AT-C--------TC--C--------------C--C-GG--------------A-------------G-G------AA--G-G----GG---------------------A----C----G-----------A-----C---GCA-----G---G 4386
A.PT.x.ALI                  A--A--T------CTA--C-C---G----TC--C-----C--GC-C-------GG--A------------A---------------A-------A--G-G-----G--G--G--AC-------A----C----G-----------A-----C--CGCGC-----G-A- 4383
A.SN.85.ROD                 AGG---T------CCA--T-C---G----TC--C-----------C--C---GG--A-----------A------------G-A-------A--G--------G----GAAAC---G--A----C----G---------C-A-----C---GCG------C-A- 3836
A.SN.86.ST_JSP4_27          ---A--T--C---CTA--C-C--------TC--C-----------C--C-----G--A---G-----A-----------C--A-------A--A-G-----G--G-GGC-A-----G--A--C--C----G-----------A-----C---GCG------G-A- 3835
B.CI.88.UC1                 C--A-----CC----AAAG-TG--------C--C-----------C--C--C-G-GCC----G--G--A--G---------G-------C-----A----T--G--T------C----G-----A--A---------------C-T---------GCAC-A---C-AC 4371
B.CI.x.20_56                C--------CC----AAAG-TG--------C--C--------G--C--C--C-G-GCC----G----A--G---------G-------C-----A----T-GG--G------C----G--A--A--A---------------C-T---------GCAC-A---C-AC 4366
B.CI.x.EHO                  C--A-----CC----A----T------T--C--C-----------C-----C--GCC------------------------G-------C-----A--G--T-----T--CCA--------A--A--A----G--------GC-T---------GCAC-A---C-AC 4361
B.GH.86.D205_ALT            C--A-----CC----A--GAG---------C--C--------G--C--C----G-ACC----G-----A-------------G---T--C-----A----T--G--T--------------------A--A--C-------------C-T---------GCAT-A-----A- 4363
B.JP.01.IMCJ_KR020_1        C--A-----CC-G-------T---G-----C--C--------------C-----C-GGACC--G---------A--G---------G-------C-----A--G--T-----T----CA-----G--A--A--A-----G---------C-T---------GCAC-A-----AC 3513
G.CI.92.Abt96               C--A--A---C-----CAGAC---G-------------C--G--C--C------ACC--C--G--R--A---------C--CG-------C--A--A--------G-----R--C-----G-----A--A--C--G--G--R---C-T--G------GCA------C---C 3717
AB.CM.03.03CM_510_03        C--A-----C---------RA----------C--C--------G--C--C------GCC------------A-----------G-------C-----A----T--G--Y------------------A--A---------------C-T---------GCAT-A-----RC 3502
H2_01_AB.CI.90.7312A        C--------CC----AAAG-T---------C--C--------G--C--C--C-G-ACC----G-----A--G-----G---G-------C-----A-GG--------T---CCA-----G-----A--A---------------C-T---------GCAC-A---C-AC 4372
H2_01_AB.JP.04.NMC307_20    C--------CC----AAAG-T--------TC--C--------G--C--C--C-G-GTC----G---------G-----G---G-------C--A--A-G-----G--T---CCA-----G-----A--A--C-------------C-T--G--G--GCAC-----C-AC 4372
H2_01_AB.JP.07.NMC716_01    C--------CC----AAAG-T--------TC--C--------G--C--C--C-G-GTC--------------G-----G---G-------C-----A-G-----G--T---CCA-----G-----A--A--C-------------C-T---------GCAC-----C-AC 4371
H2_01_AB.JP.08.NMC842_10    C--------CC----AAA--T------T-TC--C--------G--C--C--C-G-GTC----G---------G-----G---G---G--C-----A-G--------T---CCA-----G-----A--A--C-------------C-T--G--G--GCAC-----C-AC 4377
U.CI.07.07IC_TNP3           C--A--G---C----AC---TG-----T--C--C-----------C-----C-G-A-C--T--G--G----G---------G-------C--A--A--G--T--G--C--G-----G--G--A--A--A--C--G--------GC-T--G--------G-----C---C 3822
U.FR.96.12034               A--A--A--AT--CTA...-----G--A-------------C--T-----C-GGGCT------C-G-----------C---G-G--------A--A-------G-------CCT------------C--C--C---------------A-----C---GCT-----C--AG 3869
MAC.US.x.17EC1              ----------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4307
MAC.US.x.251_1A11           -----------------------------------------------------------------------------------------------------------------------G-----------------C-----------A- 4306
MAC.US.x.251_32H_PJ5        ------------A---------------------------------------------------------------------------------------------------------------------------------------------------------- 4308
MAC.US.x.251_BK28           -------------------------------------------------------------------------------------------------------------------------------------A------------------------ 4283
MAC.US.x.MM142_IVMXX        -------------------------------GT-------------G-------------------A-------------------------C-------------G-------------------------------------------------------- 3793
MNE.US.82.MNE_8             A-----------A------------------GT-----------------------------------------------------------------------------C------------------------------------------------G 3775
MNE.US.x.MNE027             -------------------------------GT------G----------------------------------------------------------------G-G---C--------------------------------------------------- 3775
SMM.CI.79.SIVsmCI2          C--A--T--C---CCA--G-------G-T---C--------------C--C-G-GTC--T--G-------------G--G-A-------A---GG--A---------CCTC---------C--------AG---------A--G--TA--GCTCCA--TC-T- 3427
SMM.LR.89.SIVsmLIB1         A-----G--CC---------T-------TC--GT---A------C---G-G-A--T--GAT--A--G----G--C--A--A-C--A--A-G---------G-CT-------A----C----GA-------GC-A-----------T-----A--C 3386
SMM.SL.92.SIVsmSL92A        A-----T--C---CCA--------G--A--C---GTG-------C--C----------------C--A-----------G----T--C--A--A-----T--G-----G--C---------------C--C--G---------C-A--G--C---GCA-----CC-AG 3388
SMM.SL.92.SL92B             A--A--T--A-----AAAG--G---GTA-----C-TT-----C--C--C----G-A-CAGT--------------------G-------C-----A-----A--G--CCTGCCA---------G----C-----G--------GC--C-G--CC-GT-AT----C---- 3738
SMM.US.04.G078              A-----T-----T--A----C------A-T----GTG-----C--T--C-------------------A-----------C--A-----C--A--G-GGA-----A----GTCC---------------C----G-----------A-----C-----T------G-A- 3567
SMM.US.04.G932              A--A-----C-----A-----G-----A------GT----------------G---------G-T---------------------------C--A----G------G--------C--------A-----------G-----------A-----C--C--T---------- 3556
SMM.US.04.M919              A--A--G------C------C------A-T----GT---------C--C----GG--A--------------G-----C---G-A--------A----G-------------CCTC-G-----------C--------------G--A-----C--CTGTC-A--C--AG 3561
SMM.US.04.M922              A--A--G------C----G-C------A-T----GT------T--C-------GG--A-----------A--G-----C---G-G--------A----G-------------CCT--G-----------C-----------------------C--CTATC-A--C--AG 3562
SMM.US.04.M923              A--A--G------C------C------G-T----GT---------C-------GG-------G------A--G--G--C---G-G--------A----G------G-C-----C--C-G-----------C--------------G---------C--CTATC-A--C--AG 3561
SMM.US.04.M926              A--A-----------------------A--C---GT-----------------C-G------------T-----------------C---------C--A--A--G-----G-------A-C-------------------------------------------C-A- 3561
SMM.US.04.M934              A--A-------------A------------A------GTG---------------C-G-------------T------G---------C---------C--A--A-----T--G-------A-C------------------------------A--G--C-------------C-AC 3559
SMM.US.04.M935              A--A--G------C------C------A-T----GT------T--C-------GG---------------A--G-----C---G-G--------A----G-------------CCT---------------C-----------------------C--CTATC----C--AG 3562
SMM.US.04.M940              A-----T--C--T--AA-G-C------A-T----GTG-----C--T--C-----G------------------------------C--A--------A--G-GGAG----A----GTCC--------------------G-----------A-----C--C---------G-A- 3567
SMM.US.04.M946              A--A------------A------------A------GTG---------------C-G-------------T------G--------C---------C--A--A-----T--G-------A-C-------------------------------A-----C-------------C-AC 3560
SMM.US.04.M947              A--A--G------C-A----C------A-T----GT-----------T-------GG-------G---------G-----C--CG-G--------A----G---------C-G--C--C-G--G---------C--------------G---------C--CTATC-A--CC-AG 3561
SMM.US.04.M949              A--A--T--C--C--AA-G-C------A-T----GTG-----C--T--C------------------G--A------------C--A-----C--A--G-GGAG----A--G-GTCC---------------C-----G-----------A--G--C--C---------G-A- 3555
SMM.US.04.M950              A--A-----------------------A------GTG---------------C-G-------------T-----------------C---------C-----A-----T--G--G---A-C--------------------G-----------A-----C-------------C-AC 3560
SMM.US.04.M951              A-----T--C--C--A--G-C------A-TC---GTG-----T--T--C--------------------A-----------C--A-----C--A--A-GGAG----A----GTCT---------------C-----G-----------A-----C--C---------G-A- 3564
SMM.US.04.M952              A--A----------A--------------A------GTG---------------C-G-------------T---A--G----------T-----C--A--A-----T-GG-------A-C-----G-------------------------A--G--C---------------AC 3560
SMM.US.05.D215              T--A------------A--G---------A-T----GT-------------C----GG--A-------------------------C--C--A-----------A-G-----GG------ACC---------------C--C--G---------GC-A-----C--C--T-----TG-A- 3561
SMM.US.06.FTq               A--A------------A--G-C------A-TC---GTT-----T--T-------GG---A-G--G-T---A-----------C--A-----C--A--A-----T--G-------C-C---------------C-----------------A-----T-------------G-A- 3589
SMM.US.11.SIVsmE660_FL10    A--A--G------C----G-C------A-TC---GT---------C-------G---------G------A-------C---G-G--G-----A----G---------C----C-T---------------C--C-----------G--A-----C---TATC-A--C--AG 4297
SMM.US.86.CFU212            A-----T---C-T--A-----------A-TC--C--------------C--C-GG-----T-----------------T---G-A------G-A--A------GG-----G--------------------C--------G--------GC-A--G--C---------------A- 3553
SMM.US.x.F236_H4            A--A--G------C------C------A-T----GT---------C-------GG-------G------A-------C---G-G--G-----------G---------C----C-T---------------C--C-----------G---------C--CTATC-A--C--AG 4294
SMM.US.x.H9                 A--A--G------C------C------A-T----GT----A----C-------GG-------G------A-------C---G-G--------A----G--G---GAC----C--C-G-----------C--------------G---------C---TATC-A--C--AG 3780
SMM.US.x.PBJA               A--A--G------C------C------A-T----GT---------C-------GG-------G------A--G-----C---G-G--------A----G-------G-C-----C--C-G-----------C--------------G---------C--CTA-C-A--C--AG 4073
SMM.US.x.PGM53              A--A--A------C------C---G--A-T----GT---------C-------GG--A----G------A--G-----C---G-A--------A----G-------G-C----CCT--G-----------C--------------G---------C--CTATC-A--C--AG 4224
SMM.US.x.SME543             A--A--G------C------C------A-T---CGT---------C-------GG-------G------A--------C---G-G--G-----------G---------C----C-T---------------C--C-----------G---------C--CTATC-A--C--AG 4310
SMM.US.x.pE660.CG7G         A--A--G------C-A--G-C------A-TC---GT---------C-------GG-------G------A-------C---G-G--G-----------G---------C----C-T---------------C--C-----------G---------C--CTATC-A--C--AG 4296
STM.US.89.STM_37_16         A--A--A--CT-G--A---AC------A-T----GTG-----C--------C-GG--A--T------------------C--C--A--T-----A--A---A------------CT--------------------------------A---------GC------C---G 3954
```

```
MAC.US.x.239               CAGACTCAGGGCCAAAGGCAAATATTATAGTAGATTCACAATATGTTATGGGAATAATAACAGGATGCCCTACAGAATCAGAGAGCAGGCTAGTTAATCAAATAATAGAAGAAATGATTAAAAAGTCAGAAATTTATGTAGCATGGGTACCAGCACACAAAGGTATAGGA 4477
Pol                        L__T__D__S__G__P__K__A__N__I__I__V__D__S__Q__Y__V__M__G__I__I__T__G__C__P__T__E__S__E__S__R__L__V__N__Q__I__I__E__E__M__I__K__K__S__E__I__Y__V__A__W__V__P__A__H__K__G__I_
A.CI.88.UC2                ----------T-------TT--C--C--------C-----G-----A---------G--G--A-CCAA--A-----G--------T--AA----A--C-----C--T-----C-----A--G--AGA--C-G-C-----T------A-C-----C--T-----C------ 4560
A.DE.x.BEN                 ----------C-------TT-----C--------------G-----A-----G---G--G----CCAG--A-----G-----A-AT--AA----G--C--G--C--------------A--G---GA--C-G-C-----T--------C-----C--T-----C------ 4560
A.DE.x.PEI2_KR_KRCG        ----------T-----A--C--------------C-----G-----A---------G--G-G--CCAG--A--------------T-AA-----A--C-----C--------------A------GA-AC-C-C-----T--------C-----C--------C------ 4528
A.GH.x.GH1                 -------G--T-----A-TT-----C--------C-----G-----A---------G--GT---CCAA--A-----G-----A--T--AA----A--C-----C--T--G--C-----A--G--AGA--C-G-C-----C--G-----C-----C--T-----C------ 4004
A.GM.87.D194               ----------T--C----TT-----C--------C-----G-----A-----G---G--G----CCAA--A-----G-----A-AT--AA----A--C-----C--T--G--C-----A--G--AGA--C-G-C-----T--------C-----C--T-----C------ 4004
A.GM.x.ISY_SBL_6669_85     ----------T-----A-TC------G-------C-----G-----A---------G-------CCAA--GG-T-----------T--AA----A---A----T--------G-----A------GA--C---C-----T--------C--G--C--------C------ 4001
A.GM.x.MCN13               ----------T-----A--C--------------C-----G-----A-----G---G--G----CCAG--A--------------T--AA----A--------C-----G--G-----A------GA--C---C-----T--------C-----C--T-----C------ 4005
A.GM.x.MCR35               ----------T-----A--C--------------C-----G-----A-----G---G--G----CCAG--A--------------T--AA----A--------C-----G--G-----A------GA--C---C-----T--------C-----C--T-----C------ 4005
A.GW.86.FG_clone_NIHZ      ----------T-----A--T--------------C-----------A-----G---G--G----CCAA--A-----G-------AT--AA----A-----G--C--------------A--G--AGA--C---C-----T--G-----C-----C--------C------ 3998
A.GW.87.CAM2CG             ----------T-----A--T--------------C-----G-----A---------G-GG----CCAA--A-----G-----A-AT--AA----A--------C--------------A------GA--C---C-----T--------C-----C--T-----C-----G 4569
A.GW.x.MDS                 ----------C-----A--T--------------C-----G-----A---------G-GGT---CCAA--A-----G-----A-AT---A----A-----G--C-----G--------A------GA--C---C-----T--------C-----C--------C------ 4004
A.IN.07.NNVA               ----------T-----A-TC--------------C-----G-----A---------G-GG----CCAA--A-----G-----A-AT--A-----A-----G--C--------------A------GA--C---C-----T--------C-----C--------C------ 4554
A.IN.95.CRIK_147           ----------T-----A--C----------G---------------A---------G-GG----CCAA--A-----G-----A-AT-A-A---G-----G--C--------------A------GA--CC--C-----T--------C-----C--------C------ 4305
A.JP.08.NMC786_clone_41    ----------TA------TT-----C--------C-----------A---------G-GG--A-CCAA--A-----------A--T--AA----A--C-----C--T--G--C-----A------GA--C-G-C--C--T--------C-----C--T-----C------ 4556
A.PT.x.ALI                 ----------T-----A--C--------------------G-----A---------G--G----CCAG--A-----------A-AT--AA----A--------C--------------A------GA--CC--C-----T--G-----C-----C--------C-----G 4553
A.SN.85.ROD                -------G--T-----A-TT--------------C-----G-----A-----G--C-GTG--A-CCAA--A-----G-----A--T-AAA---G--C--G--C--------------A------GA--C---C-----T--------C-----C--------C-----G 4006
A.SN.86.ST_JSP4_27         ----------T--------C--C-----------C-----------A---------G--G----CCAA--A-----G--------T-AAA---A--------C--------------A------GA--C---C-----T--------C-----C--T-----C------ 4005
B.CI.88.UC1                AG--T-----A---C-A-TC--C--C--------C-----------C---------G--G-T---CAG--A------A-G---TCACCA----A-----------T-----------C-----AGA--C---A-------G-------G-----T----G---AC----T 4541
B.CI.x.20_56               AG--T-----A---C-A-TC--C--C--------C-----------C------------G-T---CAA--A------A---CTCACCAA---A---A-G-----T-----------C------GA--C---A-------G-------G-----T--------AC----T 4536
B.CI.x.EHO                 AG--------A---C---TC-----C--------------------C---------G--G-T-C-CAA--A------A---ATCACC-A---A-GAG------T-----------C------GA-A---A-------G-----------T-----G--AC-G--T 4531
B.GH.86.D205_ALT           ---------AA---C-A-TT--C--C--------------------C------------G-T-C-CAG--A------A---ATCACCAA---AGCAA------T-----------C-----AGAG-C-G-A-------G-----------T-----G--AC-G--T 4533
B.JP.01.IMCJ_KR020_1       AG--T-----A---C---TC-----C-----G--C-----------C------------G-T-C-CAG--A------A---ATCACCAA---A--GAC----G-T-----------C------GA--C---A--C----G-------G-----T--T-----AC-G--T 3683
G.CI.92.Abt96              AG--------T-----A-TC-----C-----R-----C--G-----C-----R--C---G----RCAA--A--G---A---ATC-CC-T---A---A-------T-------R--------AGA-RCTC-A-------G-----------T--T-----A------ 3887
AB.CM.03.03CM_510_03       AG--T--R--A---C-A-TT--C--C--------------------C------------G-T-C-CAR--A------A---ATCACCAA---R-CAAR------T-----------C-----RGAG-C---A-------G-----------C-----G--AC----T 3672
H2_01_AB.CI.90.7312A       AG--T-----A---C-A-TC--C--C--------C-----------C---------G--G-T---CAG--A------A----TCACC-A---A---A-G-----T-----------C------GA--C---A-------G-------G-----T--------AC----T 4542
H2_01_AB.JP.04.NMC307_20   AG--T-----A---C-A-TC-----C--------------------C-----G------G-T---CAG--A------A----TCATCAA---A---AGG-----T-----------C------GA--C-G-A-------G-------G-----T--------AC----T 4542
H2_01_AB.JP.07.NMC716_01   AG--T-----A---C-A-TC-----C--------------G-----C-----G------G-T---CAG--A------A----TCATCAA---A--CAGG-----T-----------C-----AGA--C-G-A-----G-G-------G-----C--------AC----T 4541
H2_01_AB.JP.08.NMC842_10   AG--T-----A---CGA-TC-----C--------------G-----C------------G-T---CAA--A------A----TCATCAG---A---AGG-----T-----------C------GA--C-G-A-------G-------G-----T--------AC----T 4547
U.CI.07.07IC_TNP3          AG--------TT--C-A--T------G----------C--------C------------G-----CAG--C------A---ATCACCA--G--A--CA----C--T-----------C-TG---GA--CTT-A---R---GC-----------C--------A------ 3992
U.FR.96.12034              A---T-----A-----A-T---C---G-GAC------------A-A------------G-----CAG--A------A----TC-CCA----A-G-A----C---------C-T--A-GG--AGA--C-G-G-----T-GC--------T--T--T--G-----T--T 4039
MAC.US.x.17EC1             ------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4477
MAC.US.x.251_1A11          --------------------------------------------------------------------------------A-------C--------A---------------------------------------------------------------------- 4476
MAC.US.x.251_32H_PJ5       ----------------------------------------------------------------------------------------C-------------------------------------------------------------------------------- 4478
MAC.US.x.251_BK28          ----------------A---------------------------------------------------------------------C-------------------------------------------------------------------------------- 4453
MAC.US.x.MM142_IVMXX       ----------------------------C---------------------------------------------------------C--------------C-----A----------G----------------------------------------------- 3963
MNE.US.82.MNE_8            ----------------------------------------------------------------------------------------C-----------------------A-------------------G------------------------------- 3945
MNE.US.x.MNE027            ----------------------------------------------------------------------------------------C-----------------------A-------------------G------------------------------- 3945
SMM.CI.79.SIVsmCI2         G---TG----A-----------G-GG-GAC---C----------------T--TG--G-----CA------G--GA---CTCTCC------A---G-G--C-----G--GC-A--A--G--AGA--CC--C-----T-G-----------C---------T----C 3597
SMM.LR.89.SIVsmLIB1        A---------T---G-A-------------------C-----C-------G--TG--G----GCAG-----T--------A-AT-AC----A---------T--------G-----C-----A---CT--A---------A---------T--T-----A------ 3556
SMM.SL.92.SIVsmSL92A       -------------G--A--T------G-CAC---C--T--G-----GT-A--G--TG-GTT----CAA--A---------ATCTCCAA---A---G-G-----A-CC-T-----GTT---AGA---GC-G-----G--C-----------T--T-----AT-G--- 3558
SMM.SL.92.SL92B            A---T--TCCAAGT----TC---G-AG--ACG--C-----G-----CT--AAC------------CAG--AT-------G--TTCAGATA---GGCA-----T--T---C--C-AG--C-G--AGA--C-G-A--CA---G------T-----T--T-----C------ 3908
SMM.US.04.G078             ----T-----AA--G-A------------------T--G-----G-----G--G-------GCAA--C--T---------A---AAA---A-----G------------------------A---C---C-----G---------------C--T-----CC----- 3737
SMM.US.04.G932             T---T-----------A--------G--------------------------------------T----------------T--A--G--A--C------------------C-----C------AT---------------------------------C------ 3726
SMM.US.04.M919             ----T-----A--------C-----C--------C--------C-------C-----G----TCAG--C--T---------A--T--AT----G-----T-------G------------------GA--C---------G-----------T-----T-----A------ 3731
SMM.US.04.M922             ----T-----A--------C---G-C--------T--------------C-----G----TCAA--C--T---------A-AT--AT----A--C--G-------G--G--------------GA--C---------G-----------T-----T-----A------ 3732
SMM.US.04.M923             ----T-----A-----A------------------C--------------C--G--G----TCAG--C--T---------A--T--AT----A--C--G-------G------------------GA--C---------------------T-----T-----A------ 3731
SMM.US.04.M926             T---T-----A--T--A--------------------------A--------G-G-----------C-----G------A--T-AAT----A--C--G----------G-----C-----C------GA---G-------G--C---------------T-----A------ 3731
SMM.US.04.M934             A---T-----A--T--A--------------------------C--------G-G-------------G--------A--T--AT----A--C--------------------------------AGA--G---C-----G--------------G--T-----A------ 3729
SMM.US.04.M935             ----T-----A-----A--C---G-C--------T-------------C-----G----TCAA--C--T---------A--T--AT----A--C--G----------------------------GA--C---------G-----------T-----T-----G------ 3732
SMM.US.04.M940             ----T-----AA--G-A------------------T-------G-----G--G-------GCAA--C--T---------A---AAA----A-----G----------G-----C-----C------A---C-G-C-----GA-------------T--T------C----- 3737
SMM.US.04.M946             A---T-----A--T--A------------------G-------C--------G-G------------C-----------A-AT--AT----A--C---------------G--------------AGA--GG--C-----G---------------G--T-----A------ 3730
SMM.US.04.M947             ----------A-----A----------------C--C--------------C-----G----TCAG--C--T---------A-----AT----A--C--T-------G------------------GA--C------------G--------T-----T-----A------ 3731
SMM.US.04.M949             ----T------G--G---------C--------T-------G-----G--G-G----GCAA--C--T---------A---AAA----A-----G---------------------------A---C-G-A----G---------------T--T--G---C------ 3725
SMM.US.04.M950             A---T--------T--A--------G-------------------C-----G--G-G-----------------------A--T---T-G--A-GC-------------G-------G--AGG---G--C----G-------------------T-----A----G 3730
SMM.US.04.M951             ----T-----AA--G-A-----C---G--------T-------G-----G--G-------GCAA--C--T---------A---AAA----A-----G-----------------------------A---C---------G-------G----T--T------C------ 3734
SMM.US.04.M952             A---T-----A--T--A------------------C--------G-G----------------------------T---T----A-GC-------------G-------G--AGA---G--C-----G---------G-----T-----A------ 3730
SMM.US.05.D215             A---T-----A--T--A--C------G----G-----T--------A-----G--G--G-----CAG--C--C---------A-ATTCTA----A-----G--T-------------C------A---C---------G-------G---------------C-----G 3731
SMM.US.06.FTq              ----T-----A---G-A-----C-----------C--T--G----C---------G--G-----CAA-----T---------A-----G----A--C-----------G--G-----C------A---C---C-----G---------------C--T-----C------ 3759
SMM.US.11.SIVsmE660_FL10   ----T--GA-A--G--A------------------C-------------C-----G----TCAA--C--T---------A--T---T----A--C--G-------G--G--------------GA--C---------------------T-----T-----A------ 4467
SMM.US.86.CFU212           ----------A-----A-------C--------------------A---------G--G----TCAA-----T---------A-AT-A-A----A-GC-----C-------------C------A-G-C---------CA-------------T--------GC----- 3723
SMM.US.x.F236_H4           ----T--G--A-----A------------------C--------------C-----G----TCAG--C--T---------A--T--AT----A--C--G-------G------------------GA--C---------------------T-----T-----A------ 4464
SMM.US.x.H9                ----T-----A-----A------------------C--------------C--G--G----TCAA--C--T---------A--T--AT----A--C--G-------GA-----------------GA--C---------------------T-----T---RRA------ 3950
SMM.US.x.PBJA              ----T-----A-----A------------------C--------------C--G--G----TCAA--C--T---------A--T--AT----A--C--G--------------------------GA--C---------------------T-----T-----A------ 4243
SMM.US.x.PGM53             ----T-----A--------C-----C--------C--------------C-----G----TCAA--C--T---------A--T--AT----A--C--G-------G------------------GA--C---------G-----------T--T--T-----A------ 4394
SMM.US.x.SME543            ----T--G--A-----A------------------C--------------C--G--G----TCAG--C--T---------A--T---T----A--C--G----------G---------------GA--C---------------------T-----T-----A------ 4480
SMM.US.x.pE660.CG7G        ----T--G--A-----A------------------C--------------C-----G----TCAA--C--T---------A--T---T----A--C--G-------G--G-------------AGA--C------------------G--T-----T-----A------ 4466
STM.US.89.STM_37_16        -G--------T-----A---------G-G-----C-----------G--------C---------CAG--C--------G-------A-T----A-----G------------------------GA--C-------------------------T-----A------ 4124
```

```
                                                     Pol RNAse H end Pol p31 Integrase start
MAC.US.x.239                GGAAACCAAGAAATAGACCACCTAGTTAGTCAAGGGATTAGACAAGTTCTCTTC.TTGGAAAAGATAGAGCCAGCACAAGAAGAACATGATAAATACCATAGTAATGTAAAAGAATTGGTATTCAAATTTGGATTACCCAGAATAGTGGCCAGACAGATAGTAGACACCT 4646
Pol                         _G__G__N__Q__E__I__D__H__L__V__S__Q__G__I__R__Q__V__L_#_F__L__E__K__I__E__P__A__Q__E__E__H__D__K__Y__H__S__N__V__K__E__L__V__F__K__F__G__L__P__R__I__V__A__R__Q__I__V__D__
A.CI.88.UC2                 --------G---G-------TT----A-----G--C--C--------A--T---.C----G--A--------T--C-----G--------A-----T-----C---A-------GC--A-CCAT---------A-----CA-C----A--A-----------A----A- 4729
A.DE.x.BEN                  --------G---G-------TT----A--------C--C--------AT-A---.C----------------C--T-----G--------A-----T-----C-T-A--------C-AACCCAT---------A-----CTTC----A--A-----------A--T-A- 4729
A.DE.x.PEI2_KR_KRCG         -----T--G---G----T--TT----A-----G--C-----------AT-A---.C-A-----A--------C--T--G-----------G-----T-----C----------------ATCCCAT---------C-G----A-C-------A-----A-----A----A- 4697
A.GH.x.GH1                  --------G---G-------TT----A-----G--C--C--------AT-G---.C-A----G---------C--C---------------A-----T-----C---A-G------C-AACCCAT---------A-----CA-C----A--A-----------A----A- 4173
A.GM.87.D194                --------G---G-------TT----A-----G--C--C--------AT-A---.C----------------C--T-----------C--A-----T-----C---A--------C-AACCCAT---------A-----CA-C-------A-----------A----A- 4173
A.GM.x.ISY_SBL_6669_85      -----T---------T------T---A-----G--C--C--------AT-A---.C-A--G-GA--------C--T--G----------GA-----T-----C-------------C-A-CCCAT--G-------------ACC-G-----A-----A------A----A- 4170
A.GM.x.MCN13                --G--T--G--GG----T--GT----A-----G--C--C--------GT-G---.C-------A--------C--T--G-----------G----------C-------------C-ATCCCAT----------G----A-T----A--A-----A------A----A- 4174
A.GM.x.MCR35                --G--T--G--GG----T--GT----A-----G--C--C--------GT-G---.C-A-----A--------C--T--G-----------G----------C-------------C-ATCCCAT----------G----A-T----A--A-----A------A----A- 4174
A.GW.86.FG_clone_NIHZ       --------G---G----T--TT----A-----G--C-----------AT-A---.C-A--G--A-----A--C--G--G-----------A-----T-----C---A--------C--TCCCAT---------A-----AGC-------A-----A------A----A- 4167
A.GW.87.CAM2CG              --------G---G----T--TT----G--------C--C--------A--G---.C-A--------------T--T--G-----------A-----T----CC-----------GC-ATGCCAT-------ATA-----CA-C-------A-----A------A----A- 4738
A.GW.x.MDS                  -----T--G---G----T--T-----A-----G--C--C--------AT-G---.C-A-----A--------C--T---------------A-----T-----C-------------C-ATCCCAT----------------CA-C-------A-----A------A----A- 4173
A.IN.07.NNVA                -----T--G--GG----T--TT----A-----G--C--C--------AT-G---.C-A-----A--------C--T---------------A-----T----AC-----G------C-ATCCCA------------------A-C-------A-----A------A----A- 4723
A.IN.95.CRIK_147            -----T--G--GG----T--TT----G-----G--C--C--------AT-G---.C-A-----A--------C--T--------G-----A-----T-----C-------------C-ATCCCAT-----A---------A-C-------A-----------A----A- 4474
A.JP.08.NMC786_clone_41     --------G---G-------TT----G-----G--A--------GA-AT-A---.A-A--------------G--T---------------A-----T-----C--CA--------C-AACCCAT---------A-----CAGC----A--A-----------A----A- 4725
A.PT.x.ALI                  -----T--G---G-------TT----A-----G--C--C--------AT-G--T.C-A-----A-----------T--G-----------A--G--T-----C-----G------C-ATCCCAT----------------ATC-G-----A-----------A----A- 4722
A.SN.85.ROD                 --------G---G----T--TT----G-----G--T--C--------GT-G---.C-------A--------C--T--G-----------A-----T-----C-------------C--TCTCAT---------A-----ATT-------A--G--A------A--T-A- 4175
A.SN.86.ST_JSP4_27          -----T--G--GG----T---T----A-----G--C--C--------AT-A---.C-A--G--A-----A--C--T--G--G--------A-----T-----C-------------C-ATCCCAT---------C-G----A-T-------A-----A------A----A- 4174
B.CI.88.UC1                 --T--T--G---G-------------A--------A-----G--G-----A---.C-A-----------A------------------A--G------G-C---------------A--TCAT---------C-C--ACAGT-------A-A---A------A--T--- 4710
B.CI.x.20_56                --T--T--G---G-------------A-----G--A--C--G--G-----A---.C-A-----------A------------------A----------G-C--------------C-A--TCAT---------A-C--ACAGT-------A-A---A------A-TT--- 4705
B.CI.x.EHO                  --T--T--G---G-------------G--C-----A---------A-C--A--T.C-A-----A-----A-----T-------------A-----T---A----------------C-A--CCAT-------GA-T--ACA-T-------A-----A------A-TT--- 4700
B.GH.86.D205_ALT            --T--T--G---G-------------A--------A--C-----G--CT-G---.C-A-----A-----A-----C--G-----G-----A-----T--G-C--------------C---CCAT-----C---A-T--ACA-T-------A-A-----------A-TT--- 4702
B.JP.01.IMCJ_KR020_1        --T--T------G-------------G--------A--C-----CA-C--G--T.C-A-----T-----A-----C--------G-----A-----T--G-------------------A--TCAT---------A-T--ACA-T-------A-A---A------A--T--- 3852
G.CI.92.Abt96               --T--T-----GG------RT-----A--C--G--A-----------A--A---.---R-----A-----A--T--T--------------A--R-TT-------------------G--AACTCAT--G-----CA-T--TCA-T----A--A-A-----------A-TT-A- 4056
AB.CM.03.03CM_510_03        --T---------G-------------A--------A--C-----G--CT-G---.C-A-----A-----A-----T--G-----G-----R-----T--G-C------------MC----CCAT-----Y---M-C--ACA-T----R--A-A-----------A-TT-Y- 3841
H2_01_AB.CI.90.7312A        --T--T--G---G-------------A--------A-----G--G-----A---.C-A-----------A-----C--G--G--------A-----T--G-C------------GC----TCAT---------A-T--ACA-C-------A-A---A------A--T--- 4711
H2_01_AB.JP.04.NMC307_20    --T--T--G---G-------------A--------A-----G--G-----A---.C-A-----------A------------------A-----T--G-C--------------C----TCAT---------C-T--ACA-C-------A-A---A------A--T--- 4711
H2_01_AB.JP.07.NMC716_01    --T--T--G---G-------------A--------A-----G--G-----A--T.C-A-----------A------------------C--G--T--G-C------------GC-A--TCAT---------A-T--ACA---------A-A---A------A--T-T- 4710
H2_01_AB.JP.08.NMC842_10    --T--T------G-------------G--C-----A-----G--G-----A--T.C-A-----------A------------------A-----T--G-C------------GC----TCAT---------C-T--ACACC-------A-A---A------A--T--- 4716
U.CI.07.07IC_TNP3           --C--TTT---GG----T--TT-G--------G--A--C-----G---T-A---.C-A-----A-----A-----T--G------------C---TT----------------GC-A--C-----------CA----A---T-------A-A---A-----G--T--T- 4161
U.FR.96.12034               --C--T------G----T--TT-G--G-----G--A-----G-----A--G--T.C-A-----A-----A-----T--G------------G----------C-----T-----GC-ACAGCAT--------TA----ACA-T----A-----A---A------A-T--T- 4208
MAC.US.x.17EC1              ----------------------------------------------------------.--------------------------------------------------------------------------------------------------------------------- 4646
MAC.US.x.251_1A11           -----------G-----------------G----------------------------.-------------------------G------------------------------------------------------------------------------------------ 4645
MAC.US.x.251_32H_PJ5        ----------------------------------------------------------.---------------------G---------G-------------------------------------------------------------------------------A- 4647
MAC.US.x.251_BK28           -----------------------------G----------------------------.--------------------------------------------------------------------------------------------------------------------- 4622
MAC.US.x.MM142_IVMXX        ----------------------------------------------------------.---------------------------AG-------------CA--------------------------------------------C--------A-------------A- 4132
MNE.US.82.MNE_8             ----------------------------------------------------------.-------------------------------------------------------------------------------------C--------A-------------A- 4114
MNE.US.x.MNE027             ----------------------------------------------C-----------.-------------------------------------------------------------------------------------C--------A-------------A- 4114
SMM.CI.79.SIVsmCI2          --------G---G----T--T-----G---------A-----G--C--G--T.C-T---C-A-----A-------G-----G-----A-----T-----C-----C--GC-G--A--GCA---------CA-----CA-T-------T-A---A-----GA------ 3766
SMM.LR.89.SIVsmLIB1         -----T-GG---G----T--TT-G--------A-A------------A-CT..-A--G--A-----A-----C------------G--G--T-----C---A------------G--T----A---CA----A---T-G--A--A-A---------------- 3724
SMM.SL.92.SIVsmSL92A        ------------G----T--T-----C--------A--C-----G--A--A---.--A-----C-----A---------G-------------T--C-----A--------C----C-AT--------TA-C----.-T----A--T-A---A------A-T---- 3726
SMM.SL.92.SL92B             --G---A-T--GG------GTT-------------A--A-----G--C--G---.C-A----GT-------------------C-----C-----T-----C--------G--G--A-CTCAG----A-AACA-C--ACA-T----A--T-A-----------A-TG-T- 4077
SMM.US.04.G078              -----T-------------T--G--C-------------G-----CT-G---.C-A-----------------T-------G-----A-----T--C--C-------------------C--T--------GA----A---T-G--A--A-A---------G--T---- 3906
SMM.US.04.G932              -----T----------T--T-----C------------------G---------.C-A-----A-----------------G-----A----------------------------------T--------C----T---C----A--T-A---A--------T--A- 3895
SMM.US.04.M919              -----T------G----T-----G-----------A------------C--A---.--A-----A-----A-------------G--C--A--G-------C---------------A------------TA----T--GC----A--A-A---------------A- 3900
SMM.US.04.M922              -----T------G----T-----G-----------A---------AG---A---.--A-----A-----A-------------G-----A--G-----C--C----------------------------TA----T--GC----A--A-A-------A-------A- 3901
SMM.US.04.M923              -----T------G----T--T--G-----------A------------C--A---.--A-----A-----A--------G-----G-----A---------C---------------A------------T-----T--GC----A--A-A---------------A- 3900
SMM.US.04.M926              -----T--------------------G--A-----G----------------.---------A-------------------G-----A-G--------C-----------------------A----A----T---T----A----A--------------T--A- 3900
SMM.US.04.M934              -----T--------------------G--A-----G----------------.------G--A-----A--------------G--C--A-G------C--C---------------A-----------AC--GA----T---T-G--A----A---------------A- 3898
SMM.US.04.M935              -----T------G----T-----G-----------A------------C--A---.--A-----A-----A-------------G-----A--G----C--C----------------------------TA----T--GC----A--A-A-------A-------A- 3901
SMM.US.04.M940              --G--T-------------T--G--C-------------G-----CT-G---.C-------A-----------T-------G-----A-----T--C--C----------------C--T--------A----A---T----A--A-A---------G--T---- 3906
SMM.US.04.M946              -----T-----------------------G----------------------------.--A--G--A-------------------G--C--A-G------C--C--------------A---------A--GA----T---T----A----A---------------A- 3899
SMM.US.04.M947              -----T------G-G--T--T--G--A--------A------------C--A---.C-A-----A-----A--G-------------------A--G-------C-----------G------------------C-----T--GC----A--A-A---------------A- 3900
SMM.US.04.M949              --G--T-------------TT-G---------------G-----CT-G---.C-A-----------A-----C-------G--C--A----------C--------------------C--T--------GA----A---C----A--A-A-------------T---- 3894
SMM.US.04.M950              -----T--------------------G--A-----G----------------.-----G--A----------------------------G-----G-G------C--C-------------A---------A--GA----T---T----A----A---A---A---T---- 3899
SMM.US.04.M951              -----T-------------T--G--C-------------G-----CT-G---.C-------------A-----C-------G--C--A----T--C--------------------C--T--------GA----A---T----A--A-A-------------T---- 3903
SMM.US.04.M952              -----T--------------------G--A-----G----------------.-----G--A-----------G-------G-----A-G------C--C-------------A---------A--GA----T---T----A--A-A---------G--T--A- 3899
SMM.US.05.D215              -----T------G--------T-G--C--------A--C-----------A--T.C-A-----A-----A-----T--G--G--T-----A--------C-----------G--GC----G--T----A---CA----A---C----A--A-AG--A-----------A- 3900
SMM.US.06.FTq               -----T--G--------T--TT-------------A--C--------C--G--T.C-A--G--A-------------------------G--G-------------------------C----G--------A--A--------T----A--A-A-------A-T--T--T- 3928
SMM.US.11.SIVsmE660_FL10    -----T------G----T----------C--G--A-----G-----C--A---.--A-----A-----A-------------G-----A--G-------C------------------------------T-----T--GC----A--A-A---------------A- 4636
SMM.US.86.CFU212            -----T------G-----------C-----G--A--C-------AT-G--T.C-A-----------A-----T-------G-----A-------C--C------------G-------------A---CA----A---T-G--A--A-A---A---------T---- 3892
SMM.US.x.F236_H4            --C--T------G----T-----G-----C--G--A----------C--A---T-AAA---A-----A-------------G-----A--G-------C-------------------------------T-----T--GC----A--A-A---------------A- 4634
SMM.US.x.H9                 -A---T------G----T--T-------------AA----------C--A---.--A-----A-----A----G--G----G-----A-----------C--------A------------------R-T-----T--GC----A--A-A---A-----------A- 4119
SMM.US.x.PBJA               -----T------G----T--T--G-----------A------------C--A---.--A-----A-----A----G--G----G-----A-----------C-----------------------------T-----T--GC----A--A-A---------------A- 4412
SMM.US.x.PGM53              -----T------G----T-----G-----------A------------C--A---.--A-----A-----A-------------G-----A--G-------C-----------------------------TA----T--GC----A--A-A---------------A- 4563
SMM.US.x.SME543             -----T------G----T-----G-----C--G--A------------C--A---.--A-----A-----A-------------G-----A--G-------C-----------------------------TA----T--GC----A--A-A---------------A- 4649
SMM.US.x.pE660.CG7G         -----T------G----T-----G-----C--G--A------------C--A---.--A-----A-----A-------------G-----A--G----------------------------------A-T-----T--GC----A--A-A---------------A- 4635
STM.US.89.STM_37_16         -----T--G--GG----T--TT-G--A-----G--A--------G--C--A---.C-A-----A-----A-----T--------G-----A-----T-----C-----G-----GC-A--C-----G-----CA----A--GC----A--A-AG--A-----G--T--A- 4293
```

```
MAC.US.x.239               GTGATAAATGTCATCAGAAAGGAGAGGCTATACATGGGCAGGCAAATTCAGATCTAGGGACTTGGCAAATGGATTGTACCCATCTAGAGGGAAAAATAATCATAGTTGCAGTACATGTAGCTAGTGGATTCATAGAAGCAGAGGTAATTCCACAAGAGACAGGAAGACAG 4816
Pol                        T__C__D__K__C__H__Q__K__G__E__A__I__H__G__Q__A__N__S__D__L__G__T__W__Q__M__D__C__T__H__L__E__G__K__I__I__I__V__A__V__H__V__A__S__G__F__I__E__A__E__V__I__P__Q__E__T__G__R_
A.CI.88.UC2                ---CCC----C--A------------C-----------A-T----G----AA----CGT-----------C--C--A--CT----A--------C--T-----A----------T-----------------------C--C-----G--AT-------G--- 4899
A.DE.x.BEN                 ---CCC----C--A----------A--C----------A-T----G----AA----CGT-----------C-AC--A--CT----A--------C--T-----A----------T--A-------------------A--C--C-----G--AT-------G--- 4899
A.DE.x.PEI2_KR_KRCG        ---CCC-------A-------G--A-------------A-T-G--G----A--G--C-------------C--C--A--CT-----------------C--T-----A----------T--A--C--G--T-----------A--T--C-----G--A--G-----G--A 4867
A.GH.x.GH1                 ---CCC----C--A------------C-------------T----G----AA----TGTC----------C--C--A--CT----A-----G--C--T-----A----------T--A-------------------A--T--C-----G--AT-------G--- 4343
A.GM.87.D194               ---CCC----C--A----------A--C----------A-T----G----AA----CGT-----------C--C--A--CT----A--------C--T-----A-----G----T--A-------------------A--C--C-----G--AT-------G--- 4343
A.GM.x.ISY_SBL_6669_85     ---CCC-G--C--A--A-----G--A------------A-T----G----A-----C--C----------C--C--A--CT----A--------C--T-----A----------T--A--------T----------A--C--C-----G--AT-------G--A 4340
A.GM.x.MCN13               ---CCC-------A-------G----------------A-T-G--G----AT----C-------------C--C--A--CT----A-----G--C--T-----A----------T--A-------------------A--C--C-----G--AT-------G--- 4344
A.GM.x.MCR35               ---CCC-------A-------G----------------A-T-G--G----AT----C-------------C--C--A--CT----A-----G--C--T-----A----------T--A-------------------A--C--C-----G--AT-------G--- 4344
A.GW.86.FG_clone_NIHZ      ---CCC-TGTG--A-------G----------------A-T----G----A-----C-------------C--C--A---T----A------G-C--T-----A----------T--A--------T-----G-----A--T--C-----G--AT-------G--A 4337
A.GW.87.CAM2CG             ---CCC---A---A-------G----------------A-T----G----AG-G--C--C----------C--C--A------------------C--------A----------T--A--------T----------A--C--C-----G--AT-------G--A 4908
A.GW.x.MDS                 ---CCC-------A------------------------A-T----G----CT----C-------------C--C--T---T----A------G-C--T-----A----------T--A--------T-----G-----A--C--C-----G--A--------G--- 4343
A.IN.07.NNVA               ---CCC-------A-------G-----------A--A-T----G----AT----C--G-----------C--C--A---T----A------G-C--------G----------T--A--------T-----G-----A--C--C--------AT-------G--A 4893
A.IN.95.CRIK_147           ---CCC-------A------------------------A-T----G----A-----T-------------C--C--A---T----A------G-C--T-----A-----G----T--A--------T-----G-----A--C--C-----G--A--------G--A 4644
A.JP.08.NMC786_clone_41    --CCCC-------A--A---------C-----------A-T----G----AA----TGTC-------------C-----CT----A--------C--T-----G----------T--A-------------------A--C--C-----G---T-------G--- 4895
A.PT.x.ALI                 ---CCC-------G----G-------------------A-TG---G----A-----C-------------C--C--A--CT----A------G-C------A-A-----G----T--C-----------------------C--C--G--G--AT----------A 4892
A.SN.85.ROD                ---CCC-------A-------G--A-------------A-T----G----A-----C-------------C--C--A---T----A-----G--C--T-----A----------T--A--------T----------A--C--C-----G--AT----------A 4345
A.SN.86.ST_JSP4_27         -CACCC-------G-------G----------------A-T----G----AT----C-------------C--C--A--CT----A--------C--T-----A----------T--A--------T----------A--C--C-----G--AT-------G--A 4344
B.CI.88.UC1                ----------C--A--A-----G--A---G-T-----A----T----G----A-----A--A-----G-----C-----A--CT----A-----GG-C--A-----G-----C--------C--------T-----G-----------A--C-----A-----------A 4880
B.CI.x.20_56               ----------C--A--A-----G--A-----T-----A----T----G----A-----A--A-----G-----C-----A--CT----A-----GG-C--A-----G-----C--------C-----G--------G-----------A--C-----A--G--------A 4875
B.CI.x.EHO                 ----------C--A--A-----G--A-----T-----A----T---------A----------A-----------C-----A---T-----------GG-T--A-----G-----T--------C-------------------A-----A--C-----A------------ 4870
B.GH.86.D205_ALT           ----------C--A--A-----G--A-----T-----A----T----G----C----------A-----G-----C-----A---T----A--------T--A-----G-----C--------C-----G--T----------------A--C------------------ 4872
B.JP.01.IMCJ_KR020_1       ----C--------A-----------A-----C-----A----TG---G----A----------A-------------------A--CT--------------T--A-----G-----C--------C--------T----------A-----A--C-----A------------ 4022
G.CI.92.Abt96              ACC-C-GT--C--A-----------A--C--T-----A----T----G----AT-G--T--------------C--C------T----A------G-----------A-----T-----G--C-----C--TG---------A-----C--------A---------G--A 4226
AB.CM.03.03CM_510_03       ----C-----C--A--R-----R--A-----T-----A----T----G----A----------A-----G-----C--C--A---T----A-----G--T--A-----A-----C--------C-----G--Y----------AR----A--C-----AT------------ 4011
H2_01_AB.CI.90.7312A       ------------A--A--------A---G-T-----A----T----G----A----------A-----------C-----A---T----A--G---G-C--A-----A-----C--------C-----G--T-----------------A--C-----A-----------A 4881
H2_01_AB.JP.04.NMC307_20   ------------A--A--G-----A---G-T-----A----T----G----A----------A-----------C-----A---T----A---------C--A-----A-----C--------C-----------------A-----A--C-----A-----------A 4881
H2_01_AB.JP.07.NMC716_01   ------------A--A--------A---G-T-----A----T----G----A----------A-----------C-----A---T----A---------C--A-----A-----C--------C-----G-----------A--G--A--C-----A------------ 4880
H2_01_AB.JP.08.NMC842_10   -------G--C--A--A--G-----A---G-T-----A----T----G----A----------A-----G-----C-----A---T-G--A---------C--A-----A-----C--------C-----------------A--G--A--C-----A-----------A 4886
U.CI.07.07IC_TNP3          ------------C--A--G-----A--C----------A-T----A----A----------A--------------C--A---T----A------G-C--------------C--------C--------T-----G-----A--C--------G--A-----G------ 4331
U.FR.96.12034              --C--GTT--C--G----------A--C--G-----A----T----A----AG---------A-----------C--------CT-----------G-TG----C--G-----------G--C-----C--T-----------A--G--C--G--G--A--T--T--G--A 4378
MAC.US.x.17EC1             ------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 4816
MAC.US.x.251_1A11          -----------------------A--------------T---------------------------------------------------------------------------------------------------------------------------------------- 4815
MAC.US.x.251_32H_PJ5       ----------------A-------A--------------T------------------------------------------------------------------------------------------------------------A------------------------ 4817
MAC.US.x.251_BK28          -----------------------A--------------T----------------------C--------------A--------G-------------------------------------------------------------A------------------------ 4792
MAC.US.x.MM142_IVMXX       ----------------A-------A--------------T--------C-------------------------------------G-------------------------------------------------------------A---------------A--------- 4302
MNE.US.82.MNE_8            -----------------------A--------------T-------------------------------------------------------------------------------------------------------------A------------------------ 4284
MNE.US.x.MNE027            -----------------------A--------------T-------------------------------------------------------------------------------------------------------------A------------------------ 4284
SMM.CI.79.SIVsmCI2         -CAC-C-G-----G-----------A-----------A----T----G----A--------A-----------C---------T-G--A-----G-C--------A----G-----T--C-----G--------G-----A-----C-----G------------A 3936
SMM.LR.89.SIVsmLIB1        ----C-----C-----A--------A--A--C--C-----A-T----G----AT----A--C--------A--C-----A----------C-------T-----G--T-----------C--T-----C-------A--G-----G-------------G--- 3894
SMM.SL.92.SIVsmSL92A       ----C-----C--G-----G--------A--T-----A--A-TT---G----CA----A--A--------------C--T--C-----A--G-----T--T-----A-----C-----G---------------G-----A-----------C--G--A--G--------A 3896
SMM.SL.92.SL92B            --A-C-----C--A-----G-----A--G--------T---A-----G----GG----C--A-----G-----C--C-----CT----A-----GG-T-----------A-----------G--------C--T---------------A--TAGG--A------------ 4247
SMM.US.04.G078             ----------C--G-----------A-----------A--A-T---C-----AT-G--AGTC--------------C-----C--------C-----TG-T-----A-----G-----G-----------------G-----A-----C--------A-----------A 4076
SMM.US.04.G932             ----------C--C-----------A--C--------A--A-T---------C------------------------------C--G--A----------------G--------------C--------TG---------A--G--C--------A------------ 4065
SMM.US.04.M919             -------------A-----------A--C-----------A-T----G----A------------------C-----A--C--------C--G----------------G-----G-----------------G-----A-----C-----G--A-----------A 4070
SMM.US.04.M922             ----------C--C-----------A--C-----------A-T----G----A------------------C-----G--C--------C--------------------------G-----------------G-----A-----C--G--G--A-----------A 4071
SMM.US.04.M923             ----------C--C-----------A--C--------A--A-T----G----A------------------C-----A--C--------C--------T----------------G-----------------------A-----C--G-----A-----------A 4070
SMM.US.04.M926             -------G--C-----A-----G--A-----------A----T---------T----A----------------------C-----A--C---G-G-------------------------------------------A--G-----------------G--- 4070
SMM.US.04.M934             ----------C--C--A-----G--A--------------A-T---------C-----A----------------------C-----A--C---G-----------------------------------T---------A-----------------------------A 4068
SMM.US.04.M935             ----------C--C-----------A--C--------A--A-T----G----A------------G-----C-----G--C--------C--------T----------------G-----------------G-----A-----C--G--G--A-----------A 4071
SMM.US.04.M940             ----------C--G-----------A----------------T---CA----GT----AGT------------C--C--A--C--------C---G-TG-T-----G-----G-----G-----------T-----G-----A-----C------------------------ 4076
SMM.US.04.M946             ----------C--C--A-----G--A--------------A-T---------C-----A----------------------C-----A--C---G--------C------------------------C--T-----------A-----------------------------A 4069
SMM.US.04.M947             -------------C-----------A--C--------A--A-T----G----A------------------C-----A--C--------C--------T----------------G--C--------------G-----A-----C--G--G--A--------G--A 4070
SMM.US.04.M949             --A----------G-----------A-----------A----T---CA----GT----AGT------------C-----A--C--------C-----TG-T-----A-----G-----G-----------------G-----A-----C------------------------ 4064
SMM.US.04.M950             ----------C--C--A-----G--A-----------A--A-T---------C-----A----------------------C-----A--C---G---------------------------------T-----------A-----------------------------A 4069
SMM.US.04.M951             ----------C--G--A--------A--G-------------T---CA----AT----AGT------------C--C--A--C--------C---G-C--T-----A-----G-----G--------A--------G-----A-----C-----G--A-----G------ 4073
SMM.US.04.M952             ----------C--A--A-----G-----------------A--A-T---------CT----A--C-----------------------C-----A--C---G---------------------------------T-----------A-----------------------------A 4069
SMM.US.05.D215             ----C-----C--C--A-----G--A-----------A--A-TC---G----AT----A--------------------C--A---T----A--T---G-C--A-----A-----C-----------------------------A-----A-----G--A------------ 4070
SMM.US.06.FTq              ----C-----C--C--A-----G-----A--------A--A-T----G----GT------------------C--C--A--CT----A-----GG-----------A-----G-----------------------G-----A-----A-----G--A------------ 4098
SMM.US.11.SIVsmE660_FL10   ----------C--C-----------A--C-----------A-T----G----A------------------C-----G--C-----A--C--------T----------------G-----------------G-----A-----C--G--G--A-----------A 4806
SMM.US.86.CFU212           ----C-----C--C--A-----G--A-----------A--A-TT---G----------A--C----------C--A--CT----A-----------G-----G----------G---------------------------A--------------------G------ 4062
SMM.US.x.F236_H4           ----------C--C-----------A--C-----------A-T----G----A------------------C-----G--C-----A--C--------T----------------G-----------T-----G-----A-----C--G--G--A-----------A 4804
SMM.US.x.H9                ----------C--C-T---------A--Y-----------A-T----G-R-RA------------------C-----A--C--------C--------T----------------G-G---------------G-----A-----C--G-RG--A-----------A 4289
SMM.US.x.PBJA              ----------C--C-----------A--C--------A--A-T----G----A------------------C-----A--C--------C--------T----------------G-----------------G-----A--------G--G--A-----------A 4582
SMM.US.x.PGM53             --C-------C--C-----------A--C-----------A-T----G----A------------------C-----A--C--------C------------------A--------G-----------------G-----A--------G--G--A--------G--A 4733
SMM.US.x.SME543            ------G---C--C-----------A--C-----------A-T----G----A------------------C-----G--C-----A--C--------T----------------G-----------T-----G-----A-----C--G--G--A-----------A 4819
SMM.US.x.pE660.CG7G        ----------C--C-----------A--C-----------A-T----G----A------------------C-----G--C-----A--C--------T-------------C--G-----------T-----G-----A-----C--G--G--A-----------A 4805
STM.US.89.STM_37_16        ----C--G--C--C-----------A-------------A-T----G----GT------------G-----C-----A--------A--C--G-----T-----A------------------G--T-----G-----A-----C--------A-----G------ 4463
```

```
MAC.US.x.239               ACAGCACTATTTCTGTTAAAATTGGCAGGCAGATGGCCTATTACACATCTACACACAGATAATGGTGCTAACTTTGCTTCGCAAGAAGTAAAGATGGTTGCATGGTGGGCAGGGATAGAGCACACCTTTGGGGTACCATACAATCCACAGAGTCAGGGAGTAGTGGAAGC 4986
Pol                        _Q__T__A__L__F__L__L__K__L__A__G__R__W__P__I__T__H__L__H__T__D__N__G__A__N__F__A__S__Q__E__V__K__M__V__A__W__W__A__G__I__E__H__T__F__G__V__P__Y__N__P__Q__S__Q__G__V__V__E
A.CI.88.UC2                --------C--C--A------C----CA-T--G-----A--A-----CT-G--------C-----CC-C-----CA----A--G--------A-----G----------T---T--------AT-------A-----T-----C-----A--C-----------A----- 5069
A.DE.x.BEN                 --------C--C--A------C----CA-T--G-----A--A--G--CT-G--------C-----CC-C-----CA----A--G-----G--------G----------T---T-----A--AT-------A-----T-----C-----A--C-----------A----- 5069
A.DE.x.PEI2_KR_KRCG        --------C--C--A------C----CA-T--G-----A--A-----C--G-----------------C-----CA-C--A--G--------------A--G------A----T-----A--AT-------A-----T-----------A--C--A--------A----- 5037
A.GH.x.GH1                 -----G--C--C--A------C----CA-T--G-----A--A-----CT----------C-----CT-C-----CA----A--G-----G--------G---------AT---T--------AT-------A-----T-----C-----A--C-----------A----- 4513
A.GM.87.D194               --------C--C--A------C----CA-T--G-----A--A-----CT-G--------C-----CC-C-----CA----A--G-----G--------G---------AT---T--------AT-------A-----T-----------A--C-----------A----- 4513
A.GM.x.ISY_SBL_6669_85     --------C--C--A------C----CA-T--G-----A--A-----CT-G-----------------C-----CA----A--G--G-----------A----------T---C-----A--AT-------A-----T-----------A--C-----------A----- 4510
A.GM.x.MCN13               --------C--C--A------C----CA-T--G-----A--A-----CT-G-----------------C-----CA----A--G--------A-----A----------T---T-----A--AT-T--C--A-----T-----------A--C--A--------A----- 4514
A.GM.x.MCR35               --------C--C--A------C----CA-T--G-----A--A-----CT-G-----------------C-----CA----A--G--------A-----A----------T---T-----A--AT-T--C--A-----T-----------A--C--A--------A----- 4514
A.GW.86.FG_clone_NIHZ      --------G--CT-AC-G---C----CA-T--G-----A--A-----C--G-----------------C-----CA----A--G-----G--A-----A----------T---T-----A--A--------A--G--T-----------A--C--A--------A----- 4507
A.GW.87.CAM2CG             --------C--C--A-----------TA----G-----A--A-----CT-------------------C-----CA----A--G--------------A----------T---C-----A--A--------A-----T-----C-----A--C--A--------A----- 5078
A.GW.x.MDS                 --------C--C---------C----TA-T--G-----A--A-----CT-G-----------------C-----CA----A--G-----G--------A---------AT---T-----A--AT-------A-----T-----------A--C--A--------A----- 4513
A.IN.07.NNVA               --------C--C--A-----GC-A--TA-T--------A--A-----C--G--------------A--------CA----A--------G--------A----------T---T--------AT-------A-----T-----------A--C-----G-----A----- 5063
A.IN.95.CRIK_147           --------C--C---------C-A--TA-T--------A--A-----C--G-----------------------CA----A--G--------------G---------AT---T--------AT-------A-----T-----------A--C-----------A----- 4814
A.JP.08.NMC786_clone_41    --------C--C--A------C----CA-T--G-----A--A-----CT-G--------C-----CT-C-----CA----A--G-----G--------G----------T---T--------AT-------A-----C-----C--T--A--C-----------A----- 5065
A.PT.x.ALI                 --------C--C--A------C----TA-T--------A--A-----CT-G----------G-----TC-----CA-C--A--G--------------A----------TT--T--------AT-------A-----T-----------A--C--A--------A----- 5062
A.SN.85.ROD                --------C--C--A--G---C-----A-T--G-----A--A-----CT-G--T--------------C-----CA----A--G--G--G--------A---------AT---T-----A--AT-------A-----T--------------C--A--------A----- 4515
A.SN.86.ST_JSP4_27         --G-----C--C--AC-----C----CA-T--G-----A--A------T-G--------C--------C-----CA----A--G--------------G---------AT---T-----A--AT----C--A-----T-----------A--C--A-------------- 4514
B.CI.88.UC1                -----T--C--C--------------CA-------------C-----C--G--------C--C--C--C-----CA-C--A-----C--G-------CG--C------AT---------A--A-----C--------C--T--C-----A--------------A----- 5050
B.CI.x.20_56               -----T--C--C--------------CA-------------C-----C--G--------C-----C--C-----CA-C--A-----T--G--------G--C-------T------------AT----C--------C--T--C-----A--------------A----- 5045
B.CI.x.EHO                 -----T--C--C--------GC----CA-------------C-----C--G--------C--C-----C-----CA----A-----T--G--A----CA--C------AT---------A--A--A--C--A--G--C--T-----G--A--------------A----- 5040
B.GH.86.D205_ALT           -----T--C--C--AC----G-----CA-------------C-----C-----------C--C-----C-----CA-C--A-C-AGT-----------A--C-------T---A-----A--A--T-----A-----C--T--C-----A-----A-------------- 5042
B.JP.01.IMCJ_KR020_1       -----T--C--C---------C----CA-------------C-----C-----------C--------C-----CA----A-----C--G--A----CA--C------AT---AG----A--G-----C--A-----T--T--------A--------------A----- 4192
G.CI.92.Abt96              ------T----C-----------A-----G-----------C-----C-----T-----C--------C-----CA-C--------G--G--------A--C-------T---C--C-----AG-A-----A-----------C--C--------A--G-----A---T- 4396
AB.CM.03.03CM_510_03       -----T--C--C--------G-----CA----R--------C-----C--G--------C--C-----C-----CA-C--A-----T--G--------A--C-------T---A-----A--G--------A-----C--T--C-----A-----A--------A----- 4181
H2_01_AB.CI.90.7312A       -----T--C--C--------GC-----A----------C--------CT-G--------C--------C------A----A-----T-----------G--T-------T---A-----A--A--------A-----C-----C-----A--------------A----- 5051
H2_01_AB.JP.04.NMC307_20   -----C--C--C-----G--GC-----A-------------C-----C--G--------C--------C------A-C--A-----T-----------G--C-------T---A-----A--AT-------A-----C--------------------------A----- 5051
H2_01_AB.JP.07.NMC716_01   -----C--C--C-----G--GC-A---A-------------C-----C--G--------C--------C------A-C--A-----T--G--A-----G--C-------T---A-----A--AT-------A-----C--------------C-----------A----- 5050
H2_01_AB.JP.08.NMC842_10   -----C--C--C-----G---C----GA-------------C-----C--G--------C--------C------A-C--A-----T-----------G--C-------T---A-----A--AT-------A-----C--------------------------A----- 5056
U.CI.07.07IC_TNP3          --G--C--C--C---------------ACT-----------CT-----T-G--T--------------C-----CA-C--T--GA----G--------A--T-------T---C--C-----AT-A--------G--C--T--Y-----------A-------------- 4501
U.FR.96.12034              -----C--G--C--AC----GC-A-----A--G--------C--T---T-G-----------------C-----CA----A--G--------------A----------T---T--------T--------A-----C--T--C--G--A--C--A-----G--A----- 4548
MAC.US.x.17EC1             -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4986
MAC.US.x.251_1A11          -------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 4985
MAC.US.x.251_32H_PJ5       ---------------------------------------G-------------------------------------C-------------------------------------------------------------------------------------------- 4987
MAC.US.x.251_BK28          -----------------------------------------------------------------------------C-------------------------------------------------------------------------------------------- 4962
MAC.US.x.MM142_IVMXX       ---------------------------A----------------------G----------------------------------------------------------------------------------------------------------------------- 4472
MNE.US.82.MNE_8            ---------------------------A----------------G-----------------------C----------------------------------------------------------------------------------------------------- 4454
MNE.US.x.MNE027            ---------------------------A----------------G-----------------------C------A----------------------------------------------------------------------------------------------- 4454
SMM.CI.79.SIVsmCI2         ------------T-------------TA-------------C--C--CT----------C-----A--C-----CA----C--------G-------------------T---C-----A--A--------------C--T--------A-----A--G--G--A--GT- 4106
SMM.LR.89.SIVsmLIB1        -----T--G------C-------A--T--A--------------C---T-G-----------------C-----CA-C--A--------G--------A------------AAC-----A--G--T---A-A-----------------------A--------A--G-- 4064
SMM.SL.92.SIVsmSL92A       -----CT-G--CT----------A--T-CA---------G-A-AG--C--G--T--T--------A--C--T--CA----T--G--G--G-----------------------A-----A--AT----------------T--------A--------T--T--A----- 4066
SMM.SL.92.SL92B            ------T-G--C--C------A-A---TCA-----------A-AG-----G--T-----C--C-----C--T--CA-C--C--G--G--G--------A---------TT---AG-G--A--GT-------A-----G-----C-----A--------------A----- 4417
SMM.US.04.G078             ------T-G-----C--G-----A--CA----G-----C--C-----C--G-----------------------------------------------A--C-----------T-----A--------------G-----T--------A-----A--------A----- 4246
SMM.US.04.G932             --------G------C-------A--A---------------------T----T--------------C------A----A--G-----G--------A--------------------------------A--G--C-----------A--------G-----A----- 4235
SMM.US.04.M919             --------C------C----------CA----------C--C--------G--T--T-----------C--T--CA-A--A--------G--A--------C--------------T-----G-----C--------T--T-----------C--A-----G-------- 4240
SMM.US.04.M922             --------G------C-------A--TA----------C--C--------G--T--T-----------C--T--CA-A--A--------G--A--------C--------------T--A--G-----------G--T--T--------A--C--A-------------- 4241
SMM.US.04.M923             --------G------C-------A--CA----------C--C-T------G--T--T-----------C--T--CA-A--A-----G--G--A--------T--------T-----T--A--G--------------T--T--------------A--G----------- 4240
SMM.US.04.M926             -----------C-----------------------------------C--C-----------------C-----CA----A--------G--------C--C--------------------G--------A-----C-----------A--C--A--G-----A----- 4240
SMM.US.04.M934             --------G--CT--------------A----G--------------C--T--T--------------C------A----A--------------------C-----------------A--G--------------------------A--C--A--G-----A----- 4238
SMM.US.04.M935             --------G------C-G-----AA-T--T--------C--C--------G--T--T-----------C--T--CA-A--A-----G--G--A--------C--------------T--A--T--------A--G--T--T--------A--C--A-------------- 4241
SMM.US.04.M940             ------T-G--CT----------A--TA----------C--C-----C--G-----G-----------------------------------------A--C-----------T-----A--------------------T--------A--C--A--------A----- 4246
SMM.US.04.M946             --------G--CT--------------A----G--------C-----C--C-----G-----------C------A----A--------G-----------C-----------------A--G-----------------------------C--A--G-----A----- 4239
SMM.US.04.M947             --------G------C-------A--CA----------C--C--------G--T--T-----------C--T--CA-A--------------A--------T--------------T--A--G-----------G--T--T-----------C--A-------------- 4240
SMM.US.04.M949             ------T-G--C-----------A--TA----G-----C--C-----C--G-----------------------------A-----------------A--C-----------T-----A-----T-----------------------A-----A--------A----- 4234
SMM.US.04.M950             --------G--CT-------------------G--------------C--C-----------------C------A----A--------G-----------C-----------------A--G-----------------------------C--A--G-----A----- 4239
SMM.US.04.M951             ------T-G--C-----------A--TA----------C--C-----C--G-----------------------------A-----------------A--C-----------T--------------------------T--C-----A-----A--------A----- 4243
SMM.US.04.M952             --------G--CT-------------------G--------------C--C-----------------C------A----A--------G-----------C-----------------A--G-----------------------------C--A--G-----A----- 4239
SMM.US.05.D215             -----------C-----G---------A-T--------C--C-----C--G--T-----C--------C--------C--------G-----A-----A--C-----------T-----A--G--A-----------T--T--C-----A--C--A-------------- 4240
SMM.US.06.FTq              ------T-G--C-----G---------A----------C--C------T-G--T-----C--------C--T--CA-C--------G--G--------A--C-----------C-----A--G-----------G--T--T-----------C--A--G-----A----- 4268
SMM.US.11.SIVsmE660_FL10   --------G-----------G--A--T--------------C--------G--T--T-----------C--T--CA-A--A-----------A--------C--------------T--A--G-----------G--T--T-----------C--A-------------- 4976
SMM.US.86.CFU212           ------T----C---------------A-T-----------C------T-G--T-----C--------C------A----A-----------A-----A--T------A--AAT-----------------A-----C--T-----------C--A-------------- 4232
SMM.US.x.F236_H4           --------G-----------G--A--T-----------C--C--------G--T--T-----------C--T--CA-A--A--------G--A--------C--------------T--A--G-----------G--T--T-----------C--A-------------- 4974
SMM.US.x.H9                --------G------C-------A--CA----------C--C--------G--T--T-----------C--T--CA-A--A--------G--A--------T-----------R--T--A--G------R-------T--T-----------C--AR-G----------- 4459
SMM.US.x.PBJA              --------G------C-------A--CA----------C--C--------G--T--T-----------C--T--CA-A--A--------G--A--------T--------------T--A--G--------------T--T-----------C--A--G----------- 4752
SMM.US.x.PGM53             --------G------C-------A--TA----------C--C--------G--T--T-----------C--T--CA-A--A--------G--A--------C-----------A--T-----G-----C--------T--T-----------C--A-------------- 4903
SMM.US.x.SME543            --------G-----------GC-A--T-----------C--C--------G--T--T-----------C--T--CA-A--A--------G--A--------C--------------T--A--G--------A--G--T--T-----------C--A-------------- 4989
SMM.US.x.pE660.CG7G        --------G-----------G--A--T--------------C--------G--T--T-----------C--T--CA-A--A--------G--A--------C--------------T--A--G-----------G--T--T-----------C--A-------------- 4975
STM.US.89.STM_37_16        ------T-G--------G--G--A---A-----------G-C-----C--G-----G--------C--------CA-C-----G--G-----------A--C--------G--A-----A--A--------A-----T--------------C--A-------------- 4633
```

Draft of 30 September 2015, not for public release

hiv2dna_main.tex r8876–10546

```
MAC.US.x.239              AATGAATCACCACCTGAAAAATCAAATAGATAGAATCAGGGAACAAGCAAATTCAGTAGAAACCATAGTATTAATGGCAGTTCATTGCATGAATTTTAAAAGAAGGGGAGGAATAGGGGATATGACTCCAGCAGAAAGATTAATTAACATGATCACTACAGAACAAGAGA 5156
Pol                       __A__M__N__H__H__L__K__N__Q__I__D__R__I__R__E__Q__A__N__S__V__E__T__I__V__L__M__A__V__H__C__M__N__F__K__R__R__G__G__I__G__D__M__T__P__A__E__R__L__I__N__M__I__T__T__E__Q__
A.CI.88.UC2               ---------------A--G-----G---AG------T--A-----G--T---A----G-----A-----C--------------------------------------------------------C--G---------C----C--------T--C-----------A- 5239
A.DE.x.BEN                ---------------A--G-----G---AG------T--A-----G------A--A-------A-----C-G------------------------------------------------------C------------C----C--------T--C-----------A- 5239
A.DE.x.PEI2_KR_KRCG       -------------T-A-----C--G---AGC-----T--A--G--G------A--A-G-----A--------------------------------------------------------------C------------C----C--T--------C-----------A- 5207
A.GH.x.GH1                -------------T-A--G-----G---AG------T--A-----G------A--A-------------C--------------------------------------------------------C------------C----C--------T--C-----------A- 4683
A.GM.87.D194              ---------------A--------G---AG------T--A-----G------A--A-------A-----C--------------------------------------------------------C------------C----------------C-----------A- 4683
A.GM.x.ISY_SBL_6669_85    ------------------------G-----A-----T--A--G--G------A--A-G-----A-----C--------------------------------------------------------C----T-------C--G-C--T--------C-----------A- 4680
A.GM.x.MCN13              ---------------A--------G---AG------T--A-----G------A----------A-----C-G-------ACA--C-----------------------------------------C------------C----C--T--------C-----------A- 4684
A.GM.x.MCR35              ---------------A--------G---AG------T--A-----G------A----------A-----C-G-------ACA--C-----------------------------------------C------------C----C--T--------C-----------A- 4684
A.GW.86.FG_clone_NIHZ     ---------T--T--A--------G-----C-----T--A--G--G------A----------A--------------------------------------------------------------C------------A----C--T--------C-----------A- 4677
A.GW.87.CAM2CG            ---------------A--------G---AG------T--A-----G------A----------A-----C------------------------------------------------------------T--------C----C--T--------C------------- 5248
A.GW.x.MDS                C--------------A--------G---AG---------A--T--G------A----------A-----C--------------------------------------------------------C---T--------C----C--T--------C-----------A- 4683
A.IN.07.NNVA              ------------T--A-----C--------C--------A--G--G------A----------A--------------------------------------------------------------C---T----G---C----C-----------CG------------ 5233
A.IN.95.CRIK_147          ------------T--A-----C------AG---------A--G--G------A----------A-----C--------------------------------------------------------C---T----G---C----C--T--------C-----------A- 4984
A.JP.08.NMC786_clone_41   ------C--T-----A--G-----G---AG------T--A-----G-----CA----------A-----C--------------------------------------------------------C------------C----------------C-----------A- 5235
A.PT.x.ALI                ---------------A--------G---AGC-----T--A--G--G------A----G-----A-----C--------------------------------------------------------C------------C-C--C--T-------TC-----------A- 5232
A.SN.85.ROD               ------------T--A-----C------AG---------A-----G------A--A-------A-----C---------A----------------------------G---------------------T-------------C--T--------C------------- 4685
A.SN.86.ST_JSP4_27        ---------------A--------G---AGC-----T--A--G--G-----CA----------A-----C--------------------------------G-----------------------C------------C----C--T---G-----G-------G--A- 4684
B.CI.88.UC1               ------C--T-----A-----C--------------------C--G---GTA---A----G--AG-T--C----------C---C-----------------------------------------C--T-----G---C--G----------A--C-----------A- 5220
B.CI.x.20_56              ------C--T-----A--------------------------C------GTA---A----G--AG-T-GC----------C---C-----------------------------------------C--T-----G---C-------------AG-C-----------A- 5215
B.CI.x.EHO                ------C--T--T-----------G-----C-----T--A--T--G---GTA---A----G--AG-T-G----------AC---C-----------------------------------------C--T---------A--G-C--------A--------------A- 5210
B.GH.86.D205_ALT          ------C--T--------------------C---C----A--C------GTA---A----G--AG-T--C---------AC---C-----------------------------------------C--T---------C--G----------A--C-----G-----A- 5212
B.JP.01.IMCJ_KR020_1      ------C--T--T--A--------------C--------A--C--G---GTA---A----G--AG-T--C----------C---------------------------------------------C--T---------A--G----------A---G-------G---- 4362
G.CI.92.Abt96             ---------T--T--A--G--A----------A---T--A--T------------A-------------G-----------A--------------------------------------------C--T---------A-------------T--------------A- 4566
AB.CM.03.03CM_510_03      ------C--T-----A--------G---AGC---C----A--Y------GTR---A----G--AG-T--C-R-------AC---C-----------------------------------------C--T---------C--G----------A--C-----G-----A- 4351
H2_01_AB.CI.90.7312A      ---------T--T--------------------G-----A--T--G---GTA---A-------AG-T-GC----------C---------------------------------------------C--T-----G---C--G----------A--C-----------A- 5221
H2_01_AB.JP.04.NMC307_20  ------C--T--T--A-----------------G-----A--T--G---GTA---T-G-----AG-T--C----------C---------------------------------------------C--T-----G---C--G----------A--C-----------A- 5221
H2_01_AB.JP.07.NMC716_01  G-----C--T--T--A-----------------G-----A--T--G---GTA---T-G-----AG-T--C----------C---------------------------------------------C--T-----G---C--G----------A--C-----------A- 5220
H2_01_AB.JP.08.NMC842_10  ------C--T--T--A-----------------G-----A--C--G---GTA---T----G--AG-T--C----------C---------------------------------------------C--T-----G---C--G----------A--C-----------A- 5226
U.CI.07.07IC_TNP3         C-----------T--------C-----------------A--T--G-----C---A-G--G--A-----G---------AC---------------------------------------------C--T-----G--------C--T-----A--C-----------A- 4671
U.FR.96.12034             C-----------T--A--G--C--------G--------A-----G------A--A-G--G-------CC----------C---C--------------------------------------------T---------A-C-----------AT-C-----------A- 4718
MAC.US.x.17EC1            ------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5156
MAC.US.x.251_1A11         --------------------------------------------------------------------------------------CG------------------------C-------------------------------------------------------A- 5155
MAC.US.x.251_32H_PJ5      ----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------A- 5157
MAC.US.x.251_BK28         ----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------A- 5132
MAC.US.x.MM142_IVMXX      ----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------A- 4642
MNE.US.82.MNE_8           ------------T--A---------------------------------------A-G------------------------------------------------------------------------------------C-------------------------A- 4624
MNE.US.x.MNE027           ------------T--A---------------------------------------A-G--------------------------------------------------------------------------------G---C-------------------------A- 4624
SMM.CI.79.SIVsmCI2        C--------T--T-----G--C--------C-----T--A--G------------A-------------C-G-------TG---C-----------------------------------------C--T---------A--G----------AT-C-----------A- 4276
SMM.LR.89.SIVsmLIB1       ------C--T--T-----G-CC--------C-A---T--A---------------A-------------C-C-------AC------------------------------------------------------------GG-C--T--------C-----------A- 4234
SMM.SL.92.SIVsmSL92A      ------C-----T--C--------G-----C-----T--A--T------GTA--CC----G--AG----C-G------TAC------------------------------------------------T--------------C--------A--AG----------A- 4236
SMM.SL.92.SL92B           C---G-CTTA--T--T--G--AA-C-------A---A------------G-G--------G---C----C----------CA-----------------------------------------------T---------A--G-A--T--------C-------T---A- 4587
SMM.US.04.G078            ---------T-----A--------G-----G-----T--A--G--G-----C-----------A----GC-C------------------------------------------------------C-----------------C--T-----T--C------------- 4416
SMM.US.04.G932            ------------TT-A-----------------------A-----G-----C-----G--------------------------------------------------------------------------------------------------C-----------A- 4405
SMM.US.04.M919            ------C--T--T--A----CC--G-----------T--A-----G---------A----G--T-----C------------------------------------------------------------------------G-C--T--------C-----------A- 4410
SMM.US.04.M922            ------C--T--T--A----CC--G-----------T--A-----G---------A----G--T-----C-----------------------------------------------------------------G------G-C--T--------C-----------A- 4411
SMM.US.04.M923            ------C--T--T--A----CC--------------T--A-----G---------A-------T-----C--------------------------------------------------------C---------------G----T--------C-----------A- 4410
SMM.US.04.M926            ---------T--T--A--G--------------------A--G--G---------A-G--G-----------------------------------------------------------------------------------C--T-----T--C-----------A- 4410
SMM.US.04.M934            ---------T--T--A--G--------------------A-----G---------A-G--------------------------------------------------------------------------------------C--T--------C-----------A- 4408
SMM.US.04.M935            ------C--T--T--A----CC--G--------------A-----G---------A----G--T-----C-----------------------------------------------------------------G------G-C--T--------C-----------A- 4411
SMM.US.04.M940            ---------T-----A--------G-----G--G--T--A-----G--------------G--A----GC-C------------------------------------------------------C--G--------------C--T--------C------------- 4416
SMM.US.04.M946            ---------T--T--A--G--------------------A--G--G----G----A-------------G--------------C------------------------------------------------------C----C--T--------C-----------A- 4409
SMM.US.04.M947            ------C--T--T--A-----C--G---------G-T--------G---------A----------T--C-----------------------------------------------------------------G------G-C--T--------C-----------A- 4410
SMM.US.04.M949            ---------T--------------G-----G-----T--A--G--G-----C-----------A----GC-C------------------------------------------------------C--------------------T--------C--G--------A- 4404
SMM.US.04.M950            ---------T--T--A--G-----G------G-------A--G--G---------A-G--------------------------------------------------------------------------------------C--T--------------------A- 4409
SMM.US.04.M951            ---------T-----A--G-----G-----A-----T--A--G--G-----C---A-------A-----C-C------------------------------------------------------C--------------------T--------C-----------A- 4413
SMM.US.04.M952            ---------T--T--T--G--------------------A--G--G-----------G-----------------------------------------------------------------------------G--------C--T--------C-----------A- 4409
SMM.US.05.D215            T--------------A--G-CC--C--------G-----A--T--G---------A-------TG----C-C-------AC---------------------------------------------C--------G---C-GG-C--------A--C-----------A- 4410
SMM.US.06.FTq             ---------T---T-A-----C--------C--------A-----------C-----------A-----C-C------------------------------------------------------C-----------------C--T-----T--C-----------A- 4438
SMM.US.11.SIVsmE660_FL10  ------C--T--TT-A----CC--G-----------T--A-----------C---A----G--T-----C------------------------------------------------------------------------G-C--T--------C-----------A- 5146
SMM.US.86.CFU212          T-----------T-------CA--------C-----T--A--G--------------------------C--------------------------------------------------------C------------C-------T-----T--C-----------A- 4402
SMM.US.x.F236_H4          ------C--T--TT-A----CC--G-----------T--A-----------C---A----G--T-----C-----------------T------------------------------------------------------G-C--T--------C-----------A- 5144
SMM.US.x.H9               G-----C--T--T--A----CC--------------T--A-----G---------A-------T-----C---------A----------C---------------------------------------------------G-C--T-----T--C-----------A- 4629
SMM.US.x.PBJA             G-----C--T--T--A----CC--------------T--A-----G---------A-------T-----C------------------------------------------------------------------------G-C--T--------C-----------A- 4922
SMM.US.x.PGM53            ------C--T--T--A----CC--G-----------T--A-----G---------A----G--T-----C------------------------------------------------------------------------G-C--T--------C-----------A- 5073
SMM.US.x.SME543           ------C--T--TT-A----CC--G-----------T--A-----------C--------G--T-----C------------------------------------------------------------------------G-C--T--------C-----------A- 5159
SMM.US.x.pE660.CG7G       ------C--T--TT-A----CC--G-----------T--A-----------C---A----G--T-----C------------------------------------------------------------------------G-C--T--------C-----------A- 5145
STM.US.89.STM_37_16       C--------T--TT-A--G-CC--G-----C-----T--A--T--------CA----------TG----C-C------------C--------------G-A---------C--------------C---------------G----T--------C-----------A- 4803
```

```
MAC.US.x.239                TACAATTTCAACAATCAAAAAACTCAAAATTTAAAAATTTTCGGGTCTATTACAGAGAAGGCAGAGATCAACTGTGGAAGGGACCCGGTGAGCTATTGTGGAAAGGGGAAGGAGCAGTCATCTTAAAGGTAGGGACAGACATTAAGGTAGTACCCAGAAGAAAGGCTAAA 5326
Pol                         E__I__Q__F__Q__Q__S__K__N__S__K__F__K__N__F__R__V__Y__Y__R__E__G__R__D__Q__L__W__K__G__P__G__E__L__L__W__K__G__E__G__A__V__I__L__K__V__G__T__D__I__K__V__V__P__R__R__K__A_
A.CI.88.UC2                 -------C-TC---AG------T-----T--C-----A--C-A---------------------------G--------A--C--T--A--A---C-------G--A--------G-----AG-C---------G-------A--A--------A-----G-----C--G 5409
A.DE.x.BEN                  -------C-TC---AG------T-----T-----------C-A---------------------------G--------A-----T-----A---C-------G--A--------------AG-C-----------------A--A--------A-----G-----C--G 5409
A.DE.x.PEI2_KR_KRCG         -------C-TC--CG-------T--------A--------C----------T------------------G--------A-----T--G--A---C-------G--A--T-----------AG-C-----------------A--AA----G--A-----G--A-----G 5377
A.GH.x.GH1                  -------C-TC---AG------T-----T--------C--C-A---------------------------G--------A-----T--A--A---C-------G--A--C-----------AG-C---------G-------A--A---A----A-----G-----C--G 4853
A.GM.87.D194                -------C-TC---AG------T-----T--------A--C-A---------------------------G--------A-----T--A------C-------G--A--C-----------AG-C---------G-G-----A--A--------A-----G-----C--G 4853
A.GM.x.ISY_SBL_6669_85      -------C-TC---G-------T--------A-------------------T-------------A-----------C-A-----T--G------C-------------C-----------AG-C--------------T--A--A---A----A-----------C--G 4850
A.GM.x.MCN13                -------C-TC--CG-C-----T--------A-------------------T------------------G--------A--------G--A---C-------G--A--C-----------AG-C-----------------A--A--------A--G--G--A--C--G 4854
A.GM.x.MCR35                -------C-TC--CG-C-----T--------A-------------------T------------------G--------A--------G--A---C-------G--A--C-----------AG-C-----------------A--A--------A--G--G--A--C--G 4854
A.GW.86.FG_clone_NIHZ       -------C-TC---G-------T--------A-------------------T------------------G--------A-----T--G--A--GC-------G--A--C-----------AG-C--------------A--A--A--------A-----G--A--C--G 4847
A.GW.87.CAM2CG              -------C-TC---G-C-----T--------A-------------------T------------------GT-------A-----T--G--A---C-------G--A--C-----------AG-C--------A-----T--A--AA--A----A--G--G--A--C--G 5418
A.GW.x.MDS                  -------C-TC---G-C-----T--------A-------------------T------------------GT-------A-----T--G------C-------G--A--C-----------AG-C--------A--------A--AA-------A-----G-----C--G 4853
A.IN.07.NNVA                -------C-TC---G-C-----T--------A-------------------T------------------GT-------A-----T--G--A---C-------G--T--C-----G-----AG-C--------A--------A--AA--A----A-----G-----C--G 5403
A.IN.95.CRIK_147            -------C-TC---G-C-----T--------A-------------------T------------------GT-------A-----T--G--A---C-------G--A--C-----------AG----------A--------A--AA--A----A-----G--A--C--G 5154
A.JP.08.NMC786_clone_41     -------C-TC---AG------T-----T-----------C-A------------G-----------C--G--------A-----T--A------C-------G--A--------------AG-C-----------------A--A---A----A-----G-----C--G 5405
A.PT.x.ALI                  -------C-TC---A-------T-TG--------------C-C------------G--------------G--------A-----T--G------C-------------C-----------AG-T-----------------A--A--------A-----G-----C--G 5402
A.SN.85.ROD                 -------C-TC---G-C-----T--------A---G---------------T------------------GT-------A-----T--G--A---C----------A------------C-AG-C--------A--------A--AA--A----A-----G--A--C--G 4855
A.SN.86.ST_JSP4_27          -------C-TC---G-------T--------AC------------------T------------------G--------A-----T--G--A---C-------G-----C-----------AG-C---------G-T-----A--AA--A----A-----G--A-----G 4854
B.CI.88.UC1                 ----G--CTTC---G-------T-T-------C-------C-A------------------------------C-----A-----T---------------------------------C--A----------------G--C--A---A----------G--A--A--G 5390
B.CI.x.20_56                -------CTTC---G-------T-T-------C-------C-A------------------------------C-----------T---------------------------------C--G----------------G--C--A---A-------------A--A--G 5385
B.CI.x.EHO                  -------C-TC---A-------T-T------CC-------C--------------------------------C-----------T-----T------------------------------A----------------A--C--A---A-------------A--A--G 5380
B.GH.86.D205_ALT            ----G--CTTC---G-------T-T-------C-------C-A------------------------------C-----------T-----A------------------------------A----------------A--C--A-----------G-----A--A--- 5382
B.JP.01.IMCJ_KR020_1        -------C-TC---A-------G-T-------C-------C----------T---------------------C-----------T------------------------------------A----------------A--C--A---A-------------A--A--G 4532
G.CI.92.Abt96               --------------A-------T---------------------------------------------C--G--C-----A-----T--------T--------------------T------A----------------A--C--A--T--------------A--A--- 4736
AB.CM.03.03CM_510_03        ----G--CTTC---G-------T-T-------C-------C-A---Y--------------------------C-----------T-----A------------------------------R----------------A--C--A-----------G--------A--- 4521
H2_01_AB.CI.90.7312A        -------CTTC---G-------T-T-------C-------C-A------------------------------C-----------T-----A-----------G------------------A----------------G--C--A---A-------------A--A--G 5391
H2_01_AB.JP.04.NMC307_20    -------CTTC---G-------T-T-------C-------C-A------------------------------C-----------T------------------------------------G----------------G--C-----------------------A--G 5391
H2_01_AB.JP.07.NMC716_01    -------CTTC---G-------T-T-------C-------C-A------------------------------C-----------T------------------------------------G----------------G--C--A--------------------A--G 5390
H2_01_AB.JP.08.NMC842_10    ------ACTTC---A-------T-T-------C-------C-A------------------------------C-----------T------------------------------------G----------------G--C-----------------------A--G 5396
U.CI.07.07IC_TNP3           -C-----C------A-------T--------C-----------------------------------C-----C--------G---------------------------------------A-------------------A--A---A-------------A------ 4841
U.FR.96.12034               -------C-TC-----------T---------------------------------------------C---T-------A--C--A-CA--CT--C-------G--A--------G---C-TA-T------------------------A-------G-----A--A--G 4888
MAC.US.x.17EC1              ------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 5326
MAC.US.x.251_1A11           ------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 5325
MAC.US.x.251_32H_PJ5        ------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 5327
MAC.US.x.251_BK28           ------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 5302
MAC.US.x.MM142_IVMXX        ----------------------------------------------------------------------G--------------------------------------------------------------A-----------------------G------------ 4812
MNE.US.82.MNE_8             -------C--------------------------------------------------------------G--------A-----T---------------------------------G-------------------------------------------------G 4794
MNE.US.x.MNE027             -------C-------------------------------------------------------------------------------------------------------------------------------------------------------------------G 4794
SMM.CI.79.SIVsmCI2          -------C-TC---A-C--------------C---------------------------------------T----------C--G-C-A-C--TC-T--------A--------------TA-T-----------T------------A----T-----------A--G 4446
SMM.LR.89.SIVsmLIB1         -------C--------------T--------------------------------------------------C-----A--------------G---------------------------C----------------T-----A--------T--G--------A--- 4404
SMM.SL.92.SIVsmSL92A        ------AC--T---A----------------C-----A-----------------------------C-----C--------T--T--------C--------------------------AA----------AT----T--A--A---A----T-----G--A--A--G 4406
SMM.SL.92.SL92B             CC----ACTT-A-T---C----T---------C-------------T--------------A-----C-----------------A-CA-----CC-C--------A-----G--T--TG-AA-T--------------A--A--A--G-----A--------A--A--- 4757
SMM.US.04.G078              ----------------------T----------------------------T---------------C--G--------A-----------------------------------------------------------G--C-----------------G--A------ 4586
SMM.US.04.G932              -------C-----------------------------------------------------------------------------------------------------------------------------------------A-----------------A------ 4575
SMM.US.04.M919              -------C--------------T--------------------------------------------C--G--------A-----------------------------------------------------------G--C-----------A--G--G--A------ 4580
SMM.US.04.M922              -------C--------------T--------------------------------------------C--G--------A-----------------------------------------------------------G--C-----------A--G--G--A------ 4581
SMM.US.04.M923              -------C--------------T--------------------------------------------C--G--------A------------------------------------------C----------------T--C-----------A--G--G--A-T---- 4580
SMM.US.04.M926              -------C-----------------------------------------------------------C-----------------------------------------------------------------------------A--G-----A-----G--------- 4580
SMM.US.04.M934              -------C-----------------------------------------------------------------------------------------------------------------------------------------A--------A-----G--A------ 4578
SMM.US.04.M935              -------C--------------T--------------------------------------------C--G--------A-----------------------------------------------------------A--C-----------A--G--G--A------ 4581
SMM.US.04.M940              ----------------------T--------C-------------------T---------------C-----------------------AT--C--------------------------C-C--------------T-----A--------------G--A------ 4586
SMM.US.04.M946              -------C-----------------------------------------------------------C-----------------------------------------------------------------------------A--------A-----G--A------ 4579
SMM.US.04.M947              -------C--------------T-------------------------------------------A-C--G--------A------------------------------------------C----------------A--C-----------A--G--G--------- 4580
SMM.US.04.M949              -------C--------------T----------------------------T---------------C-----------------------AT--C--------------------------C-C--------------T--------------------G--A------ 4574
SMM.US.04.M950              -------C-----------------------------------------------------------C-----------------------------------------------------------------------------A--------A-----G--A------ 4579
SMM.US.04.M951              -------C--------------T----------------------------T---------------C-----------------------AT--C--------------------------C-C--------------T--------------------G--A------ 4583
SMM.US.04.M952              -------C-----------------------------------------------------------C-----------------------------------------------------------------------------A--------A-----G--------- 4579
SMM.US.05.D215              -------C--------------T---------------------------------------A--------G--T-----A--------------------------------------------------------------A--------------------A-TA--G 4580
SMM.US.06.FTq               -------C--------------T-----------------------------------G--------C-----------A------------------------------------------A----------------------A--------A--G-----A--A--- 4608
SMM.US.11.SIVsmE660_FL10    -------C--------------T--------------------------------------------C--G--------A------------------------------------------C----------------G--C-----------A--G--G--A------ 5316
SMM.US.86.CFU212            -------C--G-----C-----T-----------------------------------------------G--------A-----------A-----------------------------------------------T--A--------G-----------A-----G 4572
SMM.US.x.F236_H4            -------C--------------T--------------------------------------------C--G--------A------------------------------------------C----------------G--C-----------A--G--G--A------ 5314
SMM.US.x.H9                 -------C--------------T--------------------------------------------C--G--------A------------------------------------------C----------------G--C-----------A--G--G--A------ 4799
SMM.US.x.PBJA               -------C--------------T--------------------------------------------C--G--------A------------------------------------------C----------------G--C-----------A--G--G--A------ 5092
SMM.US.x.PGM53              -------C--------------T--------------------------------------------C--G--------A-----------------------------------------------------------G--C-----------A-----G--A------ 5243
SMM.US.x.SME543             -------C--------------T--------------------------------------------C--G-------GA------------------------------------------C----------------G--------------A--G--G--A------ 5329
SMM.US.x.pE660.CG7G         -------C--------------T--------------------------------------------C--G--------A------------------------------------------C----------------G--C-----------A--G--G--A------ 5315
STM.US.89.STM_37_16         ----------------------T--------C-----------------------G-----------------A------------------------------------------------G----------------------A--------A--G------------ 4973
```

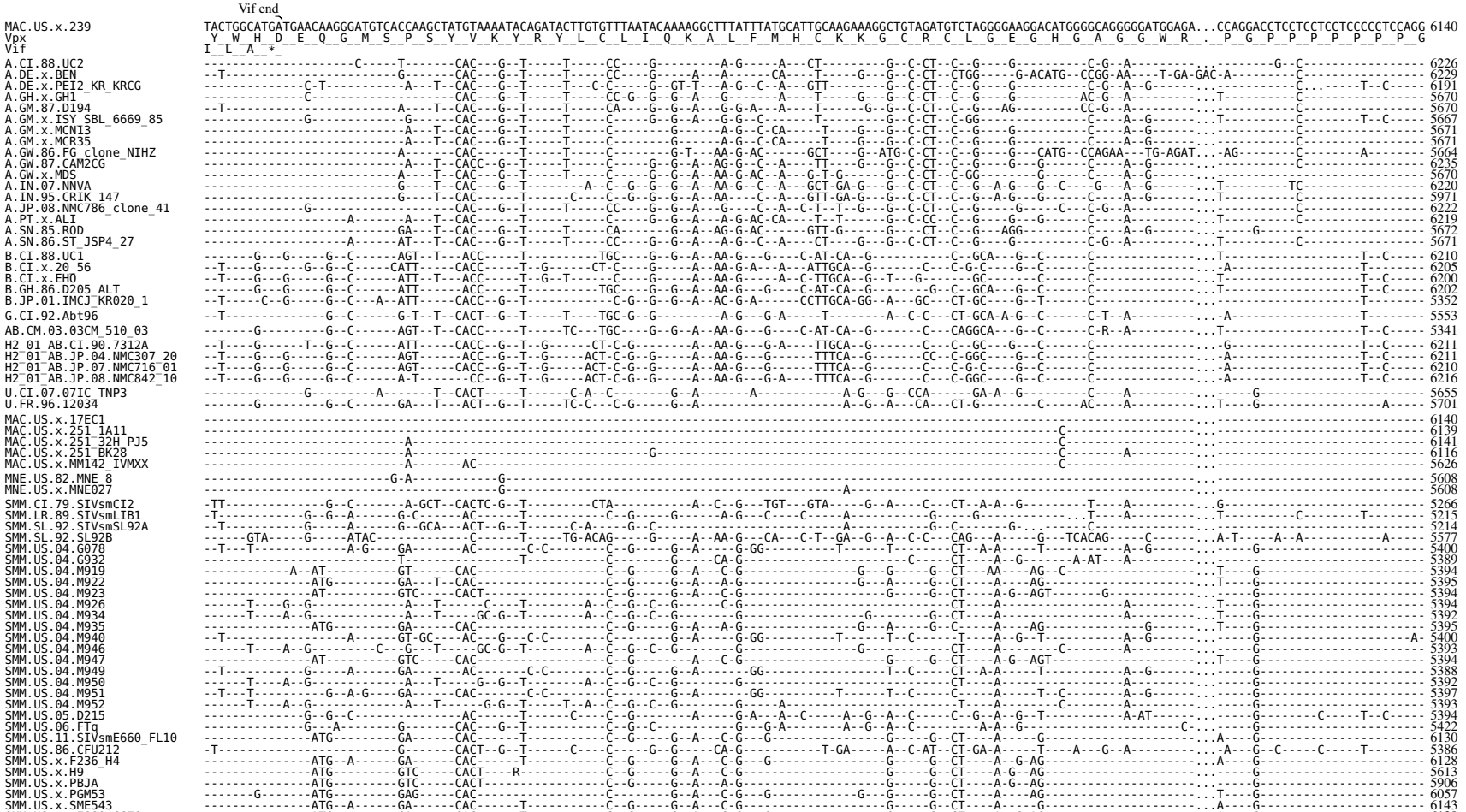
```
                              Vif start                                                                      Pol, Gag-Pol, and p31 Integrase end
MAC.US.x.239                 ATTATCAAAGATTATGGAGGAGGAAAAGAGGTGGATAGCAGTTCCCACATGGAGGATACC...GGAGAGGCTAGAGAGGTGGCA.......................TAGCCTCATAAAATATCTGAAATATAAAACTAAAGATCTACAAAAGGTTTGCTATGTGCCCC 5469
Pol                          _K__I__I__K__D__Y__G__G__G__K__E__V__D__S__S__S__H__M__E__D__.__T__G__E__A__R__E__V__.__.__.__.__.__.__.__.__A
Vif                                      _M__E__E__E__K__R__W__I__A__V__P__T__W__R__I__P__.__E__R__L__E__R__W__H__.__.__.__.__.__.__.__.__.__S__L__I__K__Y__L__K__Y__K__T__K__D__L__Q__K__V__C__Y__V__P__
A.CI.88.UC2                  -------GG--C---------A-GC-----C-------T---C-----C------G-G--...A-G----A-G------------.......................------TG-C-----C--A-----C-G---A-----C---G-GGG---GC-------T---- 5552
A.DE.x.BEN                   -------GG--C---------A--C-G--AC-------T---C-----C------G-G--...A-G----A-G----AA-----TGCCCTTGTCAAGTACCTGAAATA.......................C-G---A-----C---G-GG----GC-------T---- 5552
A.DE.x.PEI2_KR_KRCG          --C----G---C---------A-GCG------------T---------T------G----...A-G----A-G----A------.......................------TG-C--G-----A---C-C-G---A--------GG--GG---G--------T---- 5520
A.GH.x.GH1                   -------G---C---------A-GC----AC----...T---------C------G-G--AGG-AG----A-G------------.......................---T--TG-C--G--C--A-----C-G---A-G---C---G-GG----GC-------T---- 4996
A.GM.87.D194                 -------GG--C---------A-GC----AC-------T---------C------G-G--...A-G----A-G------------.......................------TG-C--GC-C-----G--C-G---A-----CT--G-GG----GC-------T---- 4996
A.GM.x.ISY_SBL_6669_85       --C----G---C------CC-A-GC-----A--------G--------C------G-G--...A-G----A-G----AA-----.......................------T--C--G-----A-----C-G---A---------G--C----GC-------T---- 4993
A.GM.x.MCN13                 --C--------C---------A-GC----AC-------TG--------T------G-G--...A-G----A-G----AA-----.......................------TG-C--------A-----C-G---A---------G--G-C--G--------T---- 4997
A.GM.x.MCR35                 --C--------C---------A-GC----AC-------TG--------T------G-G--...A-G----A-G----AA-----.......................------TG-C--------A-----C-G---A---------G--G-C--G--------T---- 4997
A.GW.86.FG_clone_NIHZ        --C--T-----C---------A-GC-----A-------TG-------TT------G-G--...A-G----A-G-----A-----.......................----T-GG-C--G--CT-------C-G---A-----CT--G-G-----A--------T---- 4990
A.GW.87.CAM2CG               --C----G---C---------A-GC-----C-----------------C------G-G--...A-G---AA-G-----A-----.......................------TG-C--G-----A-----C-----AGG-------G--C-----G--------T---- 5561
A.GW.x.MDS                   --C----G---C---------A-GC----CA-------TG--C-----T------G----...A-G----A-G-----A-----.......................------TG-C--G---T-A-----C-----A---------G--G----GC-------C---- 4996
A.IN.07.NNVA                 --C----GG--C---------A-GC-------------TG--C-----T------G----...A-G----A-G-----A-----.......................------TG-C--G-----A-----C-----A--------GG---G---G--------C---- 5546
A.IN.95.CRIK_147             -------GG--C---------A--C-----A-------TG--C-----T----A-G----...A-G----A-G-----A-----.......................------TG-C--G-----A-----C-----A---------G-------G--------C---- 5297
A.JP.08.NMC786_clone_41      -------G---C---------A--C----AC-------T---------C------G-G--...A-G----A-G----AA-----TAGCCTTGTCAAGTACCTGAAATA---.......................---A---------G-G-----GC-------T---- 5548
A.PT.x.ALI                   --C----G---C---------A-GC-----C-------TG--C-----T------G-G--...A-G----A-G-----A-----.......................------TG-C--G--C--A--G--C-G---A---------G-------G--------T---- 5545
A.SN.85.ROD                  --C----G---C---------A--C-----A-------TG--------C------G-G--...A-G----A-G----AA-----.......................------TG-C--G-----A-----C-----A--G------G-------G--------T---- 4998
A.SN.86.ST_JSP4_27           --C--------C---------A-GC-----A--------G-----A--T------G-G--...A-G----A-G------------.......................------T--C--G-----A-----C-G---AGG-------G-G-----G--------T---- 4997
B.CI.88.UC1                  --C--A-GGC-C--------G--------AT-----T-TG--A--G--------------...A-GC----------AA-----CAGTCTGGTCAAGTACCTGAAGCA---.......................---A----GT-G---C----C-CT-----C--T- 5533
B.CI.x.20_56                 --C--A-GGC-C-----------------AT-----T-TG--G--G--G-----------...A-GC-----------A-----CAGTCTGGTCAAGTACCTGAAGCA---.......................---A----GT-G---C----C-CT-----C---- 5528
B.CI.x.EHO                   --C--A-GGA-C-----------------AT-----T----G--G--G------------...ATGC-----------------CAGTCTAATTAAATACCTGAAATA---......................G--A----CT-G---C----C-CT-----T--T- 5523
B.GH.86.D205_ALT             -----A-GGC-C----------------GAT-----T-T---G--G--------------...A-GC-----------A-----CAGTCTGATTAAGTATCTTAAGTA---.......................---AGG---GT-G---C----C-CT-----C--T- 5525
B.JP.01.IMCJ_KR020_1         --------GA-------------------AT-----T----G--A---------------...--GC----------AA-----CAGCCTAGTCAAATATCTGAAACA---.......................---A----AT-GG--C----CACT-----T--T- 4675
G.CI.92.Abt96                --C--A----------------------------------A--A----------------...A-GC--A--G----A-----.......................---T--G-------TCT-A--G-T------A-----C-----G----CGGTT-----T--T- 4879
AB.CM.03.03CM_510_03         -----A-GGC-C----------------GAT-----T-T---G--G-------A-------...A-GC-------C---A-----CAGTCTGATCAAATACCTTAAGTA--A.......................---AGG---GT-G---C----C-CT-----C--T- 4664
H2_01_AB.CI.90.7312A         --C--A-GGC-C--------G--------AT-----T-T---A--G--G----A-------...A-GC----------AA-----CAGTCTAGTCAAGTATCTGAAGCA---.......................---AGG---AT-G---C----C-CT-----C---- 5534
H2_01_AB.JP.04.NMC307_20     --C--A-GGC-------------------AT-----T-TG--A--G--G------------...A-GC----------AA-----CAGTCTGGTCAAGTACCTGAAGCA---.......................---A-G---A------C----C-CT-----T---- 5534
H2_01_AB.JP.07.NMC716_01     --C--A-GGC-------------------AT-----T-TG--A--G--G------------...A--C----------AA-----CAGTCTAGTCAAGCACCTGAAGCA---.......................---A-G---A------C----C-CT-----T---- 5533
H2_01_AB.JP.08.NMC842_10     --C--A-GGC------------------GAT-----T-TG--A--G---------------...A-GC----------AA-----CAGTCTGGTCAAGTACCTGAAGCA---.......................---AGG---G------C----C-CT---A-C---- 5539
U.CI.07.07IC_TNP3            --C----------------------------------TG-----A--C------G-G--...A-GC---A-G-----A-----CAGCTTGGTGAAATACCTTAAGCA---.......................---A--G--AT-GAGT-----A--T-----A---- 4984
U.FR.96.12034                --C----G---------------------AT-------T------A--T------G-G--....-AGA---T-------A-----CTCCCTGATCAAACACCTG.AGTA--A......................C--A-----G-----G--A--G--T---------- 5030
MAC.US.x.17EC1               ------------------------------------------------------------..------------------------.......................-------------------------------------------------------- 5469
MAC.US.x.251_1A11            ------------------------------A-----------------------------..------------------------.......................-------------------------------------------------------- 5468
MAC.US.x.251_32H_PJ5         ----------------------G-------------------------------------..------------------------.......................-------------------------------------------------------- 5470
MAC.US.x.251_BK28            ------------------------------------------------------------..------------------------.......................-------------------------------------------------------- 5445
MAC.US.x.MM142_IVMXX         ------------------------------A-------T---------------------..------------------------.......................----------T-------------------------C------------ 4955
MNE.US.82.MNE_8              ------------------------------------------------------------..------------------------.......................-------------------------------------------------------- 4937
MNE.US.x.MNE027              ------------------------------------------------------------..------------------------.......................-------------------------------------------------------- 4937
SMM.CI.79.SIVsmCI2           --C-----G-------------------C---G--A----C--T--T-----A--C--...-AG----A-G--A--A-----CAGCCTGAT..............--A.........-C----T--G-TC--C--A-----AT---------G--T-----A---- 4589
SMM.LR.89.SIVsmLIB1          --C-----G----------------------------TG--------T----A------...A-G---AT-------A-----.......................------A------C----A--G-----C--------G--GA-C--A-CC--------A---- 4547
SMM.SL.92.SIVsmSL92A         --C--A-----------------------A-----------G--G-------AA-------...AC-C---T-------AC----CAGCTTAGTGAAATATCTAAAGTA--A.......................---AG-------GA-G--A--AGT---C--T---- 4549
SMM.SL.92.SL92B              --C----------------A-GC-----A---G-------G--AGT---------CAGTAGAAGA-TAG-G---C-C-----.......................----TG---T----TC-AT--G--------A-G---G---G-G--A-CC-----------A- 4903
SMM.US.04.G078               ---------------C----------------G---TG--------C--------C--..--G-----------A-----.......................-------------C--T---G-----C-------G-----G--A--A------------ 4729
SMM.US.04.G932               ----------------------------------------C-----T-----------...--G----T-------A-----.......................-------------C-C--------C-C--------G-----G------------------ 4718
SMM.US.04.M919               -----------C-----------------AT-------TG--------T-----------..--G---------A--------.......................---------C-----CT----------------G--C-----G---C---T---------- 4723
SMM.US.04.M922               -----------C-----------------AT-------TG--------T-----------..--G--AA--------------.......................---------C-----CT---------------GG--C-----G--A-C---T-----A---- 4724
SMM.US.04.M923               -----------C-----------------AT-------TG--------T-----------..--G-----------A------.......................---------C---C-C------------C--------C-----G--A-C---T-----A---- 4723
SMM.US.04.M926               -------G----------------------A-------TG--C-------------C--..--G------------------.......................----A--------C--T----------C--C-----C-----G-----C-----------T- 4723
SMM.US.04.M934               ---G--------------------------------TG--C-----G--------C--..--G-----------A------.......................----------------------G-----C--------C-----G-----C------------ 4721
SMM.US.04.M935               -----------C-----------------AT-------TG--------T-----------..--G------------------.......................---------T-----C------------------G--C-----G--A-C---T-----A---- 4724
SMM.US.04.M940               -----T----------C---------------G---TG--------C--------C--..--G------------A------.......................-------------C-C-----G-----C--A-GG--G-----G--A--A------------ 4729
SMM.US.04.M946               ---G--------------------------------TG--C-----G--------C--..--G-----------A------.......................-------------------G-----C--------C-------G----------------- 4722
SMM.US.04.M947               -----------C-----------------A--------TG--------T-----------..--G-----------A------.......................---------C-----C-------------C--------C-----G---C---T---------- 4723
SMM.US.04.M949               ----------------C---------------G--ATG--C-----C--------C--..--G------------A------.......................-------------C--------------T--C-GG--G-----G--A--A--------A---- 4717
SMM.US.04.M950               ------------------------------------TG--C-----G--------C--..--G------------------.......................-------------C--T---G-----C--------C----------CC-------------- 4722
SMM.US.04.M951               ----------------C---------------G---TG--------C--------C--..--G-----------AA------.......................-------------------T-A--G-----C--C-G---G-----G--A--G--------A---- 4726
SMM.US.04.M952               ------------------------------A-------TG--C-----G--------T--..--G------------------.......................-------------C--T---G-----C--------C-----------C-------------- 4722
SMM.US.05.D215               ---G-T-----C----C---------------------T--------C-----------...A-G--A--G---A------.......................---T---------C--T-A--G-TC--T-----G--C----------C---------A---- 4723
SMM.US.06.FTq                -----------C----C------------AT-------TG--A-----G-----------...A-G-----G----A------.......................------------GC--T-----------C-------G-----G--A-CC-------------- 4751
SMM.US.11.SIVsmE660_FL10     -----------C-----------------AT-------TG--------T-----------..------------------------.......................---------T---C-C-----------C--------C-----G-T--C---T---------- 5459
SMM.US.86.CFU212             --C--A------------------------------TG--------T-----------...A-G-----G---AA------.......................------------GC--T-A--------C--------C-----GC---CA--T-----C---- 4715
SMM.US.x.F236_H4             -----------C-----------------AT-------TG--------T-----------..------------------------.......................---------C---C-C-----------C--------C-----G-T--C---T---------- 5457
SMM.US.x.H9                  -----------C-----------------AT-------TG--------T-----------..--G-----------A------.......................---------C---C-C-----------C--------C-----G---C---T-----A---- 4942
SMM.US.x.PBJA                -----------C-----------------AT-------TG--------T-----------..--G-----------A------.......................---------C---C-C-----------C--------C-----G---C---T-----A---- 5235
SMM.US.x.PGM53               -----------------------------AT-------TG--C-----T----A------..--G-----------A------.......................---------C-----CT-----T--G-----G--C-----G---C---T-----A---- 5386
SMM.US.x.SME543              -----------C-----------------AT-------TG--------T-----------..------------------------.......................---------C---C-C-----------C-----G--C-----G-T--C-------------- 5472
SMM.US.x.pE660.CG7G          -----------C-----------------AT-------TG--------T-----------..------------------------.......................---------C---C-C-----------C--------C-----G-T--C---T---------- 5458
STM.US.89.STM_37_16          --C-------------------------------TG--------C-----------...A-G------G------------.......................-------------C--T-A--G-----C--------G---AGC----CA------------ 5116
```

```
MAC.US.x.239                ATTTTAAGGTCGGATGGGCATGGTGGACCTGCAGCAGAGTAATCTTCCCACTACAGGAAGGAAGCCATTTAGAAGTACAAGGGTATTGGCATTTGACACCAGAAAAAGGGTGGCTCAGTACTTATGCAGTGAGGATAACCTGGTACTCAAAGAACTTTTGGACAGATGTA 5639
Vif                         H__F__K__V__G__W__A__W__W__T__C__S__R__V__I__F__P__L__Q__E__G__S__H__L__E__V__Q__G__Y__W__H__L__T__P__E__K__G__W__L__S__T__Y__A__V__R__I__T__W__Y__S__K__N__F__W__T__D__V_
A.CI.88.UC2                 -CCAC-----G----------------T--------G-----A---------A-A-G--A---T---C----GA----G-CA------A-CC-A--G-----------A------TCCT--C--T----A--AC----------TA--G-A--G--C-----------T 5722
A.DE.x.BEN                  -CCAC-----G----------------T--------G-----A-----------A-G-AA---T---C-----A----G-CA------A-CC-A--------------A------TCCT--C-------A--AT----------TA--G-A--G--C-----------T 5722
A.DE.x.PEI2_KR_KRCG         -CCA------G----------------T--------G-----A------T----A-G-AAT--T--CC----GA----G-CA------A-CC-A--------------A------TCCT----------A--A-----------TA--G---GG--C-----------T 5690
A.GH.x.GH1                  -CCAC-----G--------G-------T--------G-----A-----T---A-A-G--A---T---C----GA----G-CA------A-CC-A--G-----------A------TCCT--C--T-------A-----------TA--G-A-GG--C-----------T 5166
A.GM.87.D194                --CAC-----A----------------T--------G-----A---------G-A-G--A---T---C----GA----G-CA------A-CC-A--------------A------TCCT--C--T----A---T----------TA--G-A--G--C-----------T 5166
A.GM.x.ISY_SBL_6669_85      -CCA------G--G-------------T--------G-----A------T--A-A-G-AAC--T---C----GA----G-CA------A-CC-A--------------A------TCCT-----T----A--A--G--T---------G-AGGG--C-----------T 5163
A.GM.x.MCN13                -CCA---A--A----------------T--------G-----A------T--A---G-AAC--T---C-----A----G-CA------A-CC-A--G-----------A------TCCT-----T----A--A--G--T-----TA-GG-A-GG--C-----------T 5167
A.GM.x.MCR35                -CCA---A--A----------------T--------G-----A------T--A---G-AAC--T---C-----A----G-CA------A-CC-A--G-----------A------TCCT-----T----A--A--G--T-----TA-GG-A-GG--C-----------T 5167
A.GW.86.FG_clone_NIHZ       --CA------G--G-------------T--------G-----A--------GA-A--GAAC--T---C-G--GA------CA------A-C--A--------------A------TCCT--C-CT-------A-----T-----TA--G-A--G--C-----------T 5160
A.GW.87.CAM2CG              -CCA------A----------------T--------G-----A------T--AGA-G--AC--T-G-C----GA----G-CA------A-C--A--------------A------TCCT----------A--A--G--T------A--G-A--G--C-----------C 5731
A.GW.x.MDS                  -CCA------G----------------T--------G-----A------T-GA-A-G-AAC--T---C----GA----G-CA------A-CC-A--------------A-----TTCCT-A--------A---C-G--T------A--G-A--A--C-----------T 5166
A.IN.07.NNVA                -CCA------G--G-------------T--------G-----A------T--A-CA-CAAT---A-CC----GA----G-CA------A--C-A--------------A-----TTCCT----------A--------T-----TA--G-A--A--C-----------T 5716
A.IN.95.CRIK_147            -CCA------G--G-------------T--------G-----A------T--A-A--CAAT--T--CA----GA----G-CA------A--C-A--------G----A-----TTCCT----------A---------------A--G----T--C-----------T 5467
A.JP.08.NMC786_clone_41     --CAC-----G----------------T--------G-----A-----T---G-A-G-AA---T---C----GA----G-CA------A-CC-A--------G-----A------TCCT----------A--A---------------A--G-A-GG--C-----------T 5718
A.PT.x.ALI                  -CCA------G----------------T--------G-----A------T----A-G-A----T---C----GA----G-CA------A-CC-A--------------A------TCCT----------A--A-----T-----TA--G-A--G--C-----------T 5715
A.SN.85.ROD                 -CCA------G----------------T--------G-----A------T--A-A-G-AAC--T---C----GA----G-CA------A-C--A--------------A------TCCT-----T----A--A-----T------A--G-A--G--C-----------T 5168
A.SN.86.ST_JSP4_27          -CCA------G--------G-------T--------G-----A------T--A-A-G--A---T---C-G--GA----G-CA--C---A-CC-A--------------A------TCCT-C---T----A--AC----T-----TA--G-A--A--C-----------T 5167
B.CI.88.UC1                 -CCAC-----G--------C-------T-----T--------A-----C---A-A---A-GCAT-C-----G--C------------A--C----C--G----GG--A-TCT-G--CT-C-----T--A--AC----A-----TAA--G--G-----AT--------G 5703
B.CI.x.20_56                -CCAC-----A--------C-------T-----T--------A-----C---A-A-----GCAT-CC----GA-C--G---------A--C----C--G----GG--ACTCT-A---T-C-----T----A----A------AG-G---G-----AC--------- 5698
B.CI.x.EHO                  -CCA------A--------T-------T-----T--------A--T--C--GA-A-----GCA---C-------C-----A--C--A-CC----C-------GG--A-TCT-G---T-C-----T--A--AC---A-----TGAG-G--G-----AT--------- 5693
B.GH.86.D205_ALT            -CCAC-----A--------T-------T-----T---A----A--T--C---A-CA----GCATGGC-------C-----A------A-CC-A--C-------GG--A-TCT-G--CT-C-----T--A--AC---A-----TGAG-G--------AT--------- 5695
B.JP.01.IMCJ_KR020_1        -CCA------A--------T-------------T--------A-----C---A-A-----GCAT--C----GA-C-----A------A-CC----C--G----GG--A-TC--G---G-C-----T--A--AC---A-----TAAG--A-GT--C-AT--------G 4845
G.CI.92.Abt96               --CAC--A--T--------T-------T--------G-----A------T--ACAA---A-GCA-----G---A--------------A-C-----C--T-----G--A---T-A--CCAG--------A--AT------------A---GA--A----AT--------G 5049
AB.CM.03.03CM_510_03        -CCAC-----R--------T-------T-----T--------A--T--C---G-CA---R-GCAY-CC----R--C-----A------A-CC-A--C--G----GG--ACTCT-G--C--C-----T--A--AC---A-----YGAG-G--------AT--------G 4834
H2_01_AB.CI.90.7312A        -CCAC-----G--------C-------T-----T--------A-----C---A-A-G---GCATGCC----G--C------------A--C----C--G----GG--A-TCT-G---T-C-----T----AT---A-----TGA-G---G-----AC--------- 5704
H2_01_AB.JP.04.NMC307_20    -CCAC-----G--------C-------T-----T--------A-----C---A-A--A--GCATGGC----G--C-----A------A--C----C--G----GG--A-TCT-G--CT-C-----T--A--AT---A-----TGA-G---G-----AC--------- 5704
H2_01_AB.JP.07.NMC716_01    -CCAC-----G--------C-------T-----T--------A-----C---A-TA----GCATGGC----GA-C-----A------A--C----C--G----GG--A-TCT-G---T-C-----T--A--AT---A-----TGA-G-------C-AC--------- 5703
H2_01_AB.JP.08.NMC842_10    -CCAC-----G--------C-------T-----T--------A-----C---AGAA----GCATGGC----GA-C-----A------A--C----C------G-G--A-TCT-G--CT-C-----T--A--AT---A-----TAA-GG--G---C-AC--------- 5709
U.CI.07.07IC_TNP3           --CAC--A--A--------C-------T--------G-----A--T--CT-GG-A-G--A-TCA---------A---G-CC------A-----------G--G--A----TTCA--A--------C--A------------TA---G-GGA--C---T-G--C--- 5154
U.FR.96.12034               --CAC-----A--------T-------T--------G-----A--T--T---A-AA----TCA-----G--G--G---CC------A-----------G--G--A----ATCCT-----------A---T----T------A---GAGGT--------------- 5200
MAC.US.x.17EC1              ------------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 5639
MAC.US.x.251_1A11           ----------------------------------------------------------------------------------------------------------------------------------G--------------------------------------- 5638
MAC.US.x.251_32H_PJ5        ---A---------------------------------------------------------------------------------------------------G-----------------T------------------G------------------ 5640
MAC.US.x.251_BK28           ------------------------------------------C--------------------------------------------------------------G--------------------------------G------------------ 5615
MAC.US.x.MM142_IVMXX        --CA-----------------------------------------------------------------------------A---------------G-----------------------------------------G------------------ 5125
MNE.US.82.MNE_8             --CA----------------------------------------------A--AA-----A-----------------------A-------------G---------------C---------A-------------G------------------ 5107
MNE.US.x.MNE027             --CA----------------------------------------------A---A-----G-----------------------A---A---------G-------------------------A-------------G------------------ 5107
SMM.CI.79.SIVsmCI2          --AAC-----A----------------T-------------A--T--T-GG-AA--A-TCA--C---------------C--A--C-A--------G----A-----TC-T----------A----C--T-----T---G------C-----------C 4759
SMM.LR.89.SIVsmLIB1         --CAC-----A--------G-------T-----T----G----T--C--G--A-G--A-GCAT--C---------------C--A---A--------G----A----T--C--C-----T--A--AC----T-----TAGT-GA--T------------G 4717
SMM.SL.92.SIVsmSL92A        --CAC--A--A----------------T--------------A----T--GG-AA--A-GC---------G--------A-----A--C---------------A----GTCCTG---------A---C-G--------TA-C-GA--T--C-----------C 4719
SMM.SL.92.SL92B             --CAC--A--A------------AC--AGC-TCT--G--G--T-----CT-GG----G-----T--C--G--G-------T---C--A--C----C-----------A---T-ATCA-GC--------T-----T--T---------G-A--A--------------- 5073
SMM.US.04.G078              --CA------T--------G-------T--------------------T----A-G--A--C----------A-------A-----A-CC----------------A-------------------A--A-----------A-C-G-GGA--C------------ 4899
SMM.US.04.G932              -C-AC--A--T--------------------------------------T--G--A---AC---T-C--------------A-----AG-----------G-G---------G----------A--A----------------G------------------ 4888
SMM.US.04.M919              --CA---A--T----------------T------------------------AG---T-AT-CT----G-----------A-----A---------------------A------------------A--A-------------C-GA--T--C------------ 4893
SMM.US.04.M922              --CA---A--T----------------T-------------------T--AGA--TAAC-CT--C--G-----------A-----A---------------------A----T-------------A--A-------------C-GA--T--C------------ 4894
SMM.US.04.M923              --CA---A--T----------------T--------------T---T--A----T-AG-CT----G--------G--A-----A-C--A-----------------------------------------A-------------C-GA-----C------------ 4893
SMM.US.04.M926              --CAC-----T-------------------------------A--T---T---A-G--AG------------A--------------A--C---------G-C---A------------------A--A-------------C-G------------------G 4893
SMM.US.04.M934              --CA---A--T-------------------------------A--T---T-G--A-G--A------------A--------------A--C---------G------------------------A--A-------------C-G------------------G 4891
SMM.US.04.M935              --CA---A--T----------------T-------------------T--AGA--T-AT-CT----G-----------A-----AG------------------A----------------A--A-------------C-GA--T--C------------ 4894
SMM.US.04.M940              --CA------T--------G-------T-------------------T----A-G--AG-C---CC----GA--------------A-CC---------G----A------------------A--A-----------TA-C-G---T--C------------ 4899
SMM.US.04.M946              --CA------T-------------------------------A--T---T----A-G--A------------A-------A-----A--C------------------A-----------------A--A-------------C-G------------------G 4892
SMM.US.04.M947              --CA---A--T----------------T--------G-----------T--A---GT-AGGCT-----G--------G--A-----A---------------G----------------------A--A-------------C-GA--T--C------------ 4893
SMM.US.04.M949              --CA------T----------------T-------------------T----A-G--AG-C---C-----GA--------------A--C-A--------G----A------------------A--A-----------TA-C-G---T--C------------ 4887
SMM.US.04.M950              --CA---A--T-------------------------------A--T---T-G--A----A---T----C-----A--------------A--C-A--------G----A------------------A--.-------------C-G------------------G 4891
SMM.US.04.M951              -CCA------T--------G-------T-------------------T----A-G--A-GC---C-----GA--------------A-CC-A----------------A------------------A--A-----------TG-C-G---T--C------------ 4896
SMM.US.04.M952              --CA---A--T-------------------------------A--T---T-G--A----A------------A--------------A--C---------G----A------------------A--A-------------C-G------------------G 4892
SMM.US.05.D215              --CA------A----------------T--T--------------T--C---G-AA---A-TCT--C------------------A-C--------------------A-----------C-----T--A--A----A-----TA-T-GA------------------ 4893
SMM.US.06.FTq               --CA------T--------T-------T-------------G---------T--G-A-GG-A-TCT--C-----G--------A-----A-C-----------G--G--A-------------------A--A--C-----------A-C-G--GG--C---T--------- 4921
SMM.US.11.SIVsmE660_FL10    --CA---A--T----------------T-------------T-----T--AGA--T-AG-CT----G----------A-----A----G------------A------------C--------A--A-------------C-GA--T--C------------ 5629
SMM.US.86.CFU212            --CA---A--T----------------T--------G---------T---T----A-G--A-TCT-----------------C--A-C--A--------G----A------------------T-A-------T-----AGC-GA--T--------------- 4885
SMM.US.x.F236_H4            --CA---A--T----------------T-------------T-----T--AG---T-AG-CT----G----------A-----A----G------------A------------------A--A-------------C-GA--T--C------------ 5627
SMM.US.x.H9                 --CA---A--T----------------T-------------------GT--AG---T-AGGCT----G----------A-----A------------GR-----------------------------A-------------C-GA--T--C-----------C- 5112
SMM.US.x.PBJA               --CA---A--T----------------T-------------------T--A----T-A-GCT----G--------G--A-----A------------G-------------------------------A-------------C-GA--T--C-----------G 5405
SMM.US.x.PGM53              --CA---A--T----------------T-----------------------AGA--T-A-TCT----G----------A-----A----G------------A------------------A--A-------------C-GA--A--C------------ 5556
SMM.US.x.SME543             --CA---A--T----------------T-------------T-----T--AG---TAAG-CT----G----------A-----A--------------------A-----------C-------A--A-------------C-GA--T--C------------ 5642
SMM.US.x.pE660.CG7G         --CA---A--T----------------T-------------T-----T--AGA--T-AG-CT----G----------A-----A----G------------A-----------C-------A--A-------------C-GA--T--C------------ 5628
STM.US.89.STM_37_16         -CCA------T--G-------------T-----T-----G--T--T--CT-G--A-G--A-GCA--CC-------T-----------A----A--------------A---------GAA-----T-----------T-----TA-C-GA--T--C---T--------- 5286
```

```
MAC.US.x.239               ACACCAAACTATGCAGACATTTTACTGCATAGCACTTATTTCCCTTGCTTTACAGCGGGAGAAGTGAGAAGGGCCATCAGGGGAGAACAACTGCTGTCTTGCTGCAGGTTCCCGAGAGCTCATAAGTACCAGGTACCAAGCCTACAGTACTTAGCACTGAAAGTAGTAAG 5809
Vif                        _T__P__N__Y__A__D__I__L__L__H__S__T__Y__F__P__C__F__T__A__G__E__V__R__R__A__I__R__G__E__Q__L__L__S__C__C__R__F__P__R__A__H__K__Y__Q__V__P__S__L__Q__Y__L__A__L__K__V__V__S
A.CI.88.UC2                --C---G---G---------CC--A-A--------C------TG-------G-G--A--T-----A-----A---------A--G---A-GT-AT----C-------AC-AT--CCA---------AGCA---------TCA--T--A---C----C--AGTG-----GCA 5892
A.DE.x.BEN                 --C---G---G---------CC--A-A---------------T----------G--A--T-----A-----A---------A--G---A-GT-AT----C-------AC-AT--CCA---------AGCA---------TCA--T--A---C----C--AGT-------CA 5892
A.DE.x.PEI2_KR_KRCG        --C---G---G-------TCCC--A-A---------------T----T-----G-----T-----A-----A---------A--G---A-GT-A-----C-------AT-A---CCA---C----GA-CTA-------GTTA--C--A-TTC-G--CT-AGT------GCA 5860
A.GH.x.GH1                 --C---G-------------CC--A-A---------------T----------G--A--T-----A-----A---------A--G---A-GT-AT----C-------AC-AT--CCA---------AGTA---------TCA--T--A---C----C--AGTG-----GCA 5336
A.GM.87.D194               --C---G---G-------TCCC--A-A--C------------T----------G--A--T-----A-----A---------A--G---A-GT-AT----C-------AC-A---CCA---------AGCA---------TCA--T--A---C----C--AGTG-----GCA 5336
A.GM.x.ISY_SBL_6669_85     --C---G---G--------CCC--A-A--C------------T----------G--A--T-----A-----A-----------------A-GTCAT----C-------AT-AT--CCA---C-------C-A------GTCA--C--A-TTC-G--CT-AGT-------CA 5333
A.GM.x.MCN13               --C---G---G-------TCCC--A-A---------------T--------------A--T-----A-----A---------A--G---A-GT-AT----C-------AT-AT--CCA---C----GAGC---------GTCA--C--A-TT--G--CT-AGTG-----GCA 5337
A.GM.x.MCR35               --C---G---G-------TCCC--A-A---------------T--------------A--T-----A-----A---------A--G---A-GT-AT----C-------AT-AT--CCA---C----GAGC---------GTCA--C--A-TT--G--CT-AGTG-----GCA 5337
A.GW.86.FG_clone_NIHZ      --C---G---G------T-CCC--A-A---------------T----------G--A--T-----A-----A---------A--G---A--T-AT----C-------AA-A---CC----G----GA-C---------GTCA--T--A-TTC-G--CT-AGTG-----GCA 5330
A.GW.87.CAM2CG             --C---G---G--------CCC--A-A-----T---------T--------------A--G-----A-----A---------A--G---A-GT-AT----C-------AA-AT--CC----C----GA-C----------TCA--T--A-TTC-G--CT-AGTG-----GCA 5901
A.GW.x.MDS                 --C---G---G-------TCCC--A-A---------------A--------------A--T----CA-----A---------A--G---A-GT-AT----C-------AT-AT--CCAG--C-----A-CT-------G-CA--T--A-TTC-G--CT-AGTG-----GCA 5336
A.IN.07.NNVA               --C---G-----------TCCC--A-A--G------------G--------------A--T-----A-----A---------A--G---A--G-AT----C-------AT-AT--CCAG--C------AC--------G-CA--T--A-TT--G--CT-AGT-------CA 5886
A.IN.95.CRIK_147           --C---G---G-A-----TCCC--A-A---------------G--------------A--T-----A-----A---------A--G---A-GA-AT----C-------AT-AT--CCAG--C----GAAC--------G-CA--T--A-TT--G--CT-AGT------GCA 5637
A.JP.08.NMC786_clone_41    --C---G-------------CC--A-A---------------T----------G--A--T-----A-----A---------A--G---A-GT-AT----C-------AC-AT--CCA---------AGTA---------TCA--C--A---C----CT-AGTG-----GCA 5888
A.PT.x.ALI                 --C---G---G-------TCCC--A-A--G------------T----------G--A--T-----A-----A---------A--G---A-GT-AT----C-------AT-AT--CCAG--C-------CA-------GTCA--C--A-TTC-G--CT-AGTG-----GCA 5885
A.SN.85.ROD                --C---G---G------TG-CC--A-A----------------------------A--T-----A-----A---------A--G---A-GT-AT----C-------AT-AT--CC---------GAGC---------GTCA--T--A-TTC-G--CT-AGTG-----GCA 5338
A.SN.86.ST_JSP4_27         --C---G---G---G---TCCC--A-A---------------T----------G--A--C-----A-----A---------A--G---A-G--AT-A--C-------AC-A---CCA---C-----------------GTCA--C----TTC-G--CT-AGTG-----GCA 5337
B.CI.88.UC1                -----TG--GTG------CAGC--------G-GT--------T-------------CAAT----A--G--A---------------A-GA-AT----C-A-----AC-AT--ATC------CG-AGGG-----------T------TTC----C--A-GG--CA--CA 5873
B.CI.x.20_56               -----TG-TGTG------CAGC--T-----G-GT--------T-------------CAAT----A--G--A---------------A-GA-AT----C-A-----AC-AT--ATC---C--CG-CAGG----------TT------TTC----CT-A-G---T---CA 5868
B.CI.x.EHO                 --T--TG-TGTA------CGA---------G-GT--------T---CG--------TAAT----A--G--A---------------A-GA-AT----CCA-----AC-A---ATC--------CAGG----------TT------TTC----C--A-G---T---CA 5863
B.GH.86.D205_ALT           -----TG-TGTG------CAGC--------G-GT--------T-------T-----CAAT----A--G--A---------------A-GA-AT----C-A-----AC-AT--ATC------CG-AGGG-----------T------TTC----C--A-GG--C---CA 5865
B.JP.01.IMCJ_KR020_1       -----TG-TGTA------CACC--T-----G-GT--------TT------------TAAT----A--G--A---------------A-GA-AT----C-A-----AC-A---ATC---C----GAGGA----------TT------TTC----C--A-GG--T--GCA 5015
G.CI.92.Abt96              ------G-GACA-----TCAA---T-----G-AT-------TGA------------T--T-----A-----A-------------G--GA-AT----C-------AC-AT--A-C---------AAGG-----G--T---T-G--A-TTC----CT-AC-------GCA 5219
AB.CM.03.03CM_510_03       -----TG-TGTG------CAGC--------G-GT--------T-------------CAAT----A--G--A---------------A-GA-AT----C-A-----AC-AT--ATC------CG-AGGG-----------T------TTC----C--A-GG--CA--CA 5004
H2_01_AB.CI.90.7312A       -----TG-TGTG------CAGC--T-----G-G---------T-------------CAAT----A--G--A---------------A-GA-AT----C-A-----AC-A---ATC------CG--GGG----------TT------TTC----CT-A-G---T---CA 5874
H2_01_AB.JP.04.NMC307_20   -----TG-TGTG------CAAC--T-----G-G---------T-------------CAAT-------G--A---------------A--A-AT----C-A-----AC-A---ATC------CG--AGG----------TT------TTC----CT-A-G---T---CA 5874
H2_01_AB.JP.07.NMC716_01   -----TG-TGTG------CAAC--T-----G-A---------T-------------CAAT-------G--A---------------A--A-AT----C-A-----AC-A---ATC------CG--AGG----------TT------TTC----CT-A-G---T---CA 5873
H2_01_AB.JP.08.NMC842_10   -----TG-TGTG------CAAC--T-----G-G---------T-------------CAAT-------G--A---------------A--A-AT----C-A-----AC-A---ATC------CG--AGG-----------T------TTC----CT-A-G---T---CA 5879
U.CI.07.07IC_TNP3          --G--CG-T--------TCAGC--T-A---G-G---------T-----AC-----A--G--G--A----A--------A--------GT-AT-A--C-------AC----ATCG---------AGGG--A----------T----A-T-C-G---T-A--GA-----CA 5324
U.FR.96.12034              -----TG-T-GC------CAA---------G-G-G-------T---------------T-------G--A--------A--------GT-AT----C-------AC--T--ATCG-------GAC-G---------TCTT----A-TTC--------C-G--C-T-GCA 5370
MAC.US.x.17EC1             ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 5809
MAC.US.x.251_1A11          ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------- 5808
MAC.US.x.251_32H_PJ5       ------G-------------------------------------------------------T--------------------C-------A------------------------G---------------T------------------G--------- 5810
MAC.US.x.251_BK28          ------G-------------------------------------------------------------------------------A------------------------G-------------------------------A---------- 5785
MAC.US.x.MM142_IVMXX       ------G-A-------T--------------------------------------------------------G------------------A------------C-----------T-----------------G--------- 5295
MNE.US.82.MNE_8            ------G---G-------------------------------------------------------------------------------------------------------AC-----------T-----------------G--------- 5277
MNE.US.x.MNE027            ------G--------------------------------------------------------------------------------------------------------A------------T-----------------G--------- 5277
SMM.CI.79.SIVsmCI2         ------G-TGTG-----TCAA---A-A--C--T-----C---T----T-A----CA---------A-----A------------G--GA-AT----C-------A--AT--ATCG--C------AGG-----------T-G--A-TTC-G--T--C-GGA-TC-CCA 4929
SMM.LR.89.SIVsmLIB1        -----TG----------TG-G---------G-G------------------T--A----------------A------------------GT-A--------------------A---------A--T----------TT-G---T-------T-A---------GA 4887
SMM.SL.92.SIVsmSL92A       --G---G-TCTG----------AC-GA-A---G-AG--------T----T-AC--GCA----------A-----A----------G--G----A-AT-A--C------AC-----ATCT------C-A--TA-----T----TT----A-T---G--TT-A-G--CCT-GCA 4889
SMM.SL.92.SL92B            --C---G-TGTA-----TCAG---GCA-----T------------------G-----CAC-CG------CAA--------A-----G--GG--T-A--C-A---TG-C-A-G--GT---C---C-C-CTAGT--C-A----------CTG--G--C--A--G------TT 5243
SMM.US.04.G078             ------G-T--------TG-GC--------G----------TA--------T---A----------C---A-----------------GT---A-------------------A------C--AA-------------T-G-----TC-----T-A--------G-- 5069
SMM.US.04.G932             ------G---G--------C-----------------------------------T--------------C-----------------------G-----------------T---A---C------A------------T---------C--------G--------- 5058
SMM.US.04.M919             ------G-T-T--------C----------G-------------------T---A---------AC------------------A--T--------C-----A-------A----------A--------------------TC-------A-C------G-- 5063
SMM.US.04.M922             ------G-T----------C----------G-------------------T---A---------AC------------------A--T-------------------------A----------A--------------------TC-------A-C------G-- 5064
SMM.US.04.M923             ------G-T----------C----------G-------------------T---A---------AC-----------------GA--T---------------A-------A----------A-T------T----T--------TC-------A-C------G-- 5063
SMM.US.04.M926             ------G---------T-CC--------C-----------------------G---------C-------------------T-------------------------A---------AC-------------------A-TT-----TT-A-GG--------- 5063
SMM.US.04.M934             ------G-T----------CC------------------------------G---------C-------------------T-------------------------A----------C-A------------------A-TT-----T--A-GG--------- 5061
SMM.US.04.M935             ------G-T----------C----------G-------------------T---A---------AC-----------------GA--T--------------------------A----------A------------T-------TC-------A-C------G-- 5064
SMM.US.04.M940             ------G-T--------TG-GC--------G----------TA--------T---A----------C---A-----------------GT---A-------------------A------C--AA-------------T-G-----TC-----T-AG-------G-- 5069
SMM.US.04.M946             ------G-T----------CC------------------------------G---------C-------------------T-------------------------A----------C-A------------------A-TT-----T--A-GG--------- 5062
SMM.US.04.M947             ------G-T----------C----------G-------------------T---A---------AC-----------------GAG-T---------------A-------A----------A-T-------------------TC-------A-C------G-- 5063
SMM.US.04.M949             ------G-T--------TG-GC--------G----------TA--------T---A----------C---A-----------------GT---A-------------------A------C-GAC-------------TT-G-----TC-----T-A-G------G-- 5057
SMM.US.04.M950             ------G-T----------CC------------------------------T---G---------C-------------------T-------------------------A----------C-A------------------A-TT-----T--A-GG--------- 5061
SMM.US.04.M951             ------G-T--------TG-GC-------------------TA--------T---A----------C---A-----------------GT---A-------------------A------C--AA-------------T-G-----TC-----T---G------G-- 5066
SMM.US.04.M952             ------G-T----------CC------------------------------T---G---------C-------------------T-------------------------A----------C-A------------------A-TT-----T--A-GG--------- 5062
SMM.US.05.D215             ------G-T-G------TG-AC--T----C-AT--A---------------T-G--A--C-----A-----A-----------------GT-AT----------------T---A--------C-A--------------TT-G-----T---------A--------GGA 5063
SMM.US.06.FTq              ------G------------CG---------G---------------T------------------C-------------------T---------C--------------A----------A---------------T-------TC-----T-A-G-A------- 5091
SMM.US.11.SIVsmE660_FL10   ------G-T----------C-----------------------------T---A---------AC-----------------GA--T---------------A-------A---------AA-T-------------------TC-------A-C--------- 5799
SMM.US.86.CFU212           ------G-T--------T-CA---------G-AT-------T-------------------A----A-----------------GT-A---------------A-----A---------GAA-T------------T--G--A--TC---------GT-------- 5055
SMM.US.x.F236_H4           ------G-T----------C-----------------------------T---A---------AC-G-A-------------GA--T---------------A-------A---------AA-T-------------------TC-------A-C--------- 5797
SMM.US.x.H9                ------G-T----------C----------G-------------------T---A---------AC------------------A--T---------------A-------A----------A-T-------------------TC-------A-C--------- 5282
SMM.US.x.PBJA              ------G-T----------C----------G-------------------T---A---------AC------------------A--T---------------A-------A----------A-T-------------------TC-------A-C------G-- 5575
SMM.US.x.PGM53             ------G-T----------C----------G-------------------T---A---------AC-----------------GA--T---------------A-------A----------A--------------------TC-------A-C------G-- 5726
SMM.US.x.SME543            ------G-T-G--------C-----------------------------T---A---------ACA--A-------------GA--T---------------A-------A---------AA-T-------------------TC-------A-C--------- 5812
SMM.US.x.pE660.CG7G        ------G-T----------C-----------------------------T---A---------AC-----------------GA--T---------------A-------A---------AA-T-------------------TC-------A-C--------- 5798
STM.US.89.STM_37_16        ------G---G------TCAGC--------G-------------------------------------A--------------A-GT-A-----------------A---A----------A------------T------------TT-A--------GGA 5456
```

Vpx start

```
MAC.US.x.239              CGATGTCAGA......TCCCAGGGAGAGAATCCCACCTGGAAACAGTGGAGAAGAGACAATAGGAGAGGCCTTCGAATGGCTAAACAGAACAGTAGAGGAGATAAACAGAGAGGCGGTAAACCACCTACCAAGGGAGCTAATTTTCCAGGTTTGGCAAAGGTCTTGGGAA 5973
Vpx                          M  S  D__.__.  P  R  E  R  I  P  P  G  N  S  G  E  E  T  I  G  E  A  F  E  W  L  N  R  T  V  E  E  I  N  R  E  A  V  N  H  L  P  R  E  L  I  F  Q  V  W  Q  R  S  W  E_
Vif                       __D__V__R__.__.__S__Q__G__E__N__P__T__W__K__Q__W__R__R__D__N__R__R__G__L__R__M__A__K__Q__N__S__R__G__D__K__Q__R__G__G__K__P__P__T__K__G__A__N__F__P__G__L__A__K__V__L__G__
A.CI.88.UC2               AC-AAATGAC...AGAC-----A--A--GG---G--A------------------C--T-------------G------G---G--TA---A-CCT-G---------------------G--CC-A-----C-----------G----------C---AG- 6059
A.DE.x.BEN                AC-AAATGAC...AGAC-----A--A--GG-A--G--A------------------C--T-------------G------G-G--G--CA---A-CCT------G-----A--G-----T--G--CC-A-----C-----------G----------C---AG- 6059
A.DE.x.PEI2_KR_KRCG       AC-AAATG-C...AGAC-----AA-A-C-G-A-----A------G----C----TA--T--T--------T----CT------G-A--G-TG-------CC--------------A--G--------G--TC------T-----------G----------C---AG- 6027
A.GH.x.GH1                AC-AAATGAC...AGAC-----A--A--GG-A--G--A------------------C--T-------------G------G---G--TA---A-CCT-------------A--G-----T--G--CC-A-----C-----------G----------C---AG- 5503
A.GM.87.D194              AC-AAATG-C...AGAC-----A--A--GG-G--G--A------------------C--T-------------G------G---G--TA---A-CCT------G-----A--G-----T--G--CC-A-----C-----------G----------C----C- 5503
A.GM.x.ISY_SBL_6669_85    GC-AAATGAC...AAAC-----A----C---A-----A----------C-----A--T--C-A------------T------G---G--G-----A-CC---------------A--G--------G--C-------T-----------G----------C---AG- 5500
A.GM.x.MCN13              GC-AAATGAC...AGAC-----A--A-CGG-A--C--A----------C-------T--C-A---------A-T-----G---G--G-----A-CC---------------A--G--T-----G--CC-A-----T-----------G-----G-----C---AG- 5504
A.GM.x.MCR35              GC-AAATGAC...AGAC-----A--A-CGG-A--C--A----------C-------T--C-A---------A-T-----G---G--GA----A-CC---------------A--G--T-----G--CC-A-----T-----------G-----G-----C---AG- 5504
A.GW.86.FG_clone_NIHZ     AC-AAATGAC...AGA---------A-C-G-G-----A-------C--C--G----T--C-A---------C------G---G----------CT---------------A--G--T----AC---C-A-----C-----------C-----G-----C---AG- 5497
A.GW.87.CAM2CG            AC-AAATGAC...AGAC-----A----CCG-A-----A----------C-------T--C-A---------CC-----G---G-------A-CT---------------A--G--T-----G--CC-------T-----------G----------C---AG- 6068
A.GW.x.MDS                AC-AAATG-C...AGAC-----A----C-G-A-----A-------C--C-------T--C-AGC--------CC-----G---G-------A-CC------------------G--T-----G--CC-A-----T-----------G----------C---AG- 5503
A.IN.07.NNVA              AC-AAATGAC...AGA------A----C-G-A-----A-------C--C-------T--C-------------CT-----G-G--G--G-----A-CC---------------A--G--T-----G--C--------C-----------G----------C----G- 6053
A.IN.95.CRIK_147          AC-AAATGAC...AGA------A----CGG-A-----A-------C--C-------T--C-------------CT-----G---G--G-----A-CC------------------G--T-----G--C-----A--C-----------C-----G-----C----G- 5804
A.JP.08.NMC786_clone_41   AC-AAATG-C...AGAC-----A--A--GG-A--T--A------------------C--T-------------G------G---G--TA---A-CCT-G-----------A--GG----T--G--CC-A-----C-----------G----------C---AG- 6055
A.PT.x.ALI                AC-AAATG-C...AAAC-----A--A-C-G-A-----A----------C-------T--C-------------G------G---G---------CCT------------------G--T-----G--CC-------C-----------G----------C---AG- 6052
A.SN.85.ROD               AC-AAATGAC...AGAC-----A----C-G-A-----A-------C--C-------T--C-------------CC---------G-------A-CC------------A--A--G--T--------CC-A--A--T-----------G-----G-----C---AG- 5505
A.SN.86.ST_JSP4_27        AC-AAATG-C...AGGC-----A----C---A-----A----------C-----A--T--C-------------G------G---G--G-----A-CC---------------A--G--------G--CC-A-----T-----------G----------C---AG- 5504
B.CI.88.UC1               G--A-GA-A-GATGGA------------G-G-----A-------C-AC-----A-----------A--A-----T------G-A------A--ACA---C-C-----G-TA--A--C------T-G--CC-A--A--C-----------C-----G----G-----C- 6043
B.CI.x.20_56              A--A-GA-A-AATGGA------------G-G-----A-------C-AC--G-----G-----------A----C-------G-A------C---CA--AC-C----AG--A--A--C--T--TT-G--CC-A--A--C-----------C------A--------C- 6038
B.CI.x.EHO                A--A-GA-A-GATGGA------------G-A-----A-------C--C-----A---G-----------A------------G-G-C----T----AC-TC-C-------TA--A--C--T---T-G--C-----A--T-----------C-----G-A--------C- 6033
B.GH.86.D205_ALT          G--A-GA-A-AATGGA------------G-G-----A-------C-AC-----A---G-------A--A---CT------G-A------A--ACA---C-C-----G-TA-----C-----TT-G--CC-A--A--T-----------C-----G----------C- 6035
B.JP.01.IMCJ_KR020_1      A--A-GAG--AATGGAC-----------G-G--G--A-------C--C-----A-----------A--A-------------G-A------A--ACA---C-C-------TA--T--C--T--T--G--C-----A--T-----------C-----G----------C- 5185
G.CI.92.Abt96             AA-A-GACAT...GGA---A---------G-M-----A-A------C-AC--G--------C--------A-------------G-A-------MACA-----C-----GATA--A--C--T--TT----C--------T---------------G------C------ 5386
AB.CM.03.03CM_510_03      G--A-GA-A-AATGGA---------A-----A-----A--G-----C-AC-----A-----------A--A-----K------G-A------A--ACA---C-C-----G-TA--A--C------T-G--CC-A--A--C-----------C----------G-----C- 5174
H2_01_AB.CI.90.7312A      A--A-GA-A-AATGGA------------G-G-----A-------C-AC---------G-----------A----C-------G-A------A---TA--AC-C-----G--A--A--C-----TT-G--CC-A--A--T-----------C----------------C- 6044
H2_01_AB.JP.04.NMC307_20  A--A-GA-A-AATGGA---A--------GG-G-----A-------C-AC---------------------A----T------G-A------A--ACA---C-C-----G--A--A--C-----TT-G--CC-----A--C-----------C-----G----------C- 6044
H2_01_AB.JP.07.NMC716_01  A--A-GA-A-AATGGA---A--------GG-G-----A-------C-AC---------------------A----T------G-A------A--ACA---C-C-----G--A-----C-----TT----CC-----A--C-----------C-----G----------C- 6043
H2_01_AB.JP.08.NMC842_10  A--A-GA-AGAATGGA---A---------G-G-----A-------C-AC---------------------A---CT------G-A------A--ACA---C-C-----G--A--A--C--T--TT-G--CC-----A--C-----------C----------------C- 6049
U.CI.07.07IC_TNP3         G-----------......C-----A--A-----A-----------TCA-----C--A---G-G-A---A--------------G-GTC------CACA-CC------CA------A-----T--T--G--C-----A--T-----------G-----------C----G- 5488
U.FR.96.12034             A----GG---......---A--A-------------A-----TTCA-A------A--TG----------T---AGT------G-A-----T--GAGA--T--C-----G-CA--A--GC-A--T--T--C-----A--T-----------C-----------C---C-- 5534
MAC.US.x.17EC1            ----------......-------------------------------------------------------------------------------------------------------------------------------------------- 5973
MAC.US.x.251_1A11         ----------......-------------------------------------------------------------------------------------------------------------------------------------------- 5972
MAC.US.x.251_32H_PJ5      TT--------......---------------------------------G------A--------------------------------------------------------------------------------------------------- 5974
MAC.US.x.251_BK28         ----------......-------------------------------------------------------------------------------------A-----------------G-------------------------------- 5949
MAC.US.x.MM142_IVMXX      TC--------......----------------------------------------A------G----------------------------------------------------G--------------------------------------- 5459
MNE.US.82.MNE_8           TT--------......------A---------------------------------A------------------------------------A------------------T-G------------------------------------- 5441
MNE.US.x.MNE027           TT--------......------A---------------------------------A------------------------------------A--------------------G------------------------------------- 5441
SMM.CI.79.SIVsmCI2        -A-ACAG-ATGGGAAG-G-A--A-------------A---C------G-----A-------------T-----G-----T---C-----G--A-CAG-------G-GA--A-----A--T--G--C--------G-----------C----G------C------ 5099
SMM.LR.89.SIVsmLIB1       GC----------......------A------------------------A--T-----------------G------G-A--------G--A--A---------C----A--G-----T-----------AT-G-----------C----G------C------ 5051
SMM.SL.92.SIVsmSL92A      A...-ATG--......---A--A--ACAG-----T--A--G--------G--C----------------------C------G-A-----------A--AG-C------ACA--A--G-----TT-G-----------C-----------C----G------C------ 5050
SMM.SL.92.SL92B           AC-AAATGAC...AGAC--A-----A-----------A-------------------A-----------------------T--C---AT-----G----CCC-G--TCA-ACA--A---C-A-----------A---T-------------C---AG-----G------- 5410
SMM.US.04.G078            TT--------......------A-------------------------------------------------G-----GC-------G----CA-----------G-----T--------T-----------A--G---------------G------------- 5233
SMM.US.04.G932            TC--------......------A-----G-------------------------------------A--T---------T-------G--G-----A--T---------------A--G--T--T---------------G-------------------------- 5222
SMM.US.04.M919            TC--------......------------G---------------------A-----------A-----------------G-----------------T--------G-CA--A--G--T---T-G--G--------G-----------C----G------C------ 5227
SMM.US.04.M922            TC--------......---------A--G------------------A---C--------------C-----G-------------A-----------G-C---A--G--T---T-G--G--------T-----------C----G------C------ 5228
SMM.US.04.M923            TC--------......------A--------T----------------A-------C------------------G-A---------------A--------G-CA--A--G--T--TT-G--G-----A--------------C----G------C------ 5227
SMM.US.04.M926            TC---G----......-------------------------T----------AA-------G-----CGA--G-----C--A--AG-------G--A--A-----T--T-----------AT-G-----------G----G---A--------- 5227
SMM.US.04.M934            TC--------......----------------------------------A-----------------C-A--------C-CA--A--------G--A--A--G--T---------------T-G-----------G----G------------- 5225
SMM.US.04.M935            TC--------......---------A--G-----------G-------------A--G----------------C------G---------------A--A--------G-CA--A--G--T---T-G--G--------------------C----G------C------ 5228
SMM.US.04.M940            -T--------......------A-----G--------C-------------G--A-------------T-----G-----C-A--------G--A--A--------G-----T-------T-----------A----------------G----G------------- 5233
SMM.US.04.M946            TC--------......------A-------------------------------A--T--------------C-A--------C-CA--A--------G--A--A--G--T---------------T-G-----------G----G------------- 5226
SMM.US.04.M947            TC----------....C--------A--G-----------------------G----C----------C-----G-A-----------A--C--------G-CA--A--G--T--TT-G--G--------T--------------------G------C------ 5227
SMM.US.04.M949            TT--------......---------A-------------------------------G----G--A--------C-----G-A------G-----A--A--G-----G-----C-----T--T-----------A--------------G----G------------- 5221
SMM.US.04.M950            TC--------......------A------------------------------------------G-----C-A--------C--A--C--------G--A--A--G--T--T-----------AT-G-----------G----G------------- 5225
SMM.US.04.M951            T---------......------A---------------------------A---G----G--A------------C--------G--------C--------G-----C--G-----T-----------A--G-----------G----G------------- 5230
SMM.US.04.M952            TC--------......------A-----------------------CAG---------A-T-C----G-----GC-----------C--A---G--------G--A-----G-----T-----------T-G-----------G----G------------- 5226
SMM.US.05.D215            AC------A-......------A------------A-------------------T--G--A--T-------------G-A--------G--A--C-----------TA--A-----T--T--------A----G-----------------G------C------ 5227
SMM.US.06.FTq             -C--------......------A----------G------------A---------------------------G-G--G-----C--A--T-------G-CA--A--------------------A--G-----------C----G------C------ 5255
SMM.US.11.SIVsmE660_FL10  TC--------......------------G-------------------G-----A------A----------------------------------A-GA--------G-CA--A-----T-----G--G------------------------G------C------ 5963
SMM.US.86.CFU212          TC---G----......------------G--------A--------------A---------A-------G-----G-------G--G--A--AC-------G-TA--A-----T-----G-----A-----------------------G-----G------- 5219
SMM.US.x.F236_H4          TC--------......------------G--------------------A-------A------------------------------------A--A--------G-CA--A-----T-----G--G------------------------G------C------ 5961
SMM.US.x.H9               TC-----RSR......------------R----------------------A--RR----C-----Y-----C------C--------S----A--A--------G-CA--A--G--T--TT-G--R-------------------C----G----K-C------ 5446
SMM.US.x.PBJA             -C--------......-----------------------------------A--------C-----------C------G--------------A--A--------G-CA--A--G--T--TT-G--G-----A--------------C----G------C------ 5739
SMM.US.x.PGM53            TC--------......------A-----G-------------------A--------------A---------------------------A--A--------G-CA--A--G--T---T-G--G----------------------------G------------- 5890
SMM.US.x.SME543           TC--------......------A-----G----------------G--A-------A-----------------------------------A-GA--------G-CA--A-----T-----G--G------------------------------C------ 5976
SMM.US.x.pE660.CG7G       TC--------......------------G----------------G--A-------A-----------------------------------A-GA--------G-CA--A-----T-----G--G------------------------G------C------ 5962
STM.US.89.STM_37_16       AC--------......------A--------A--G--C------------------G------AG--A--A-----G------C-------G---------C-------C-------A--------T-----G-----A--G-------------C----G------------- 5620
```



Draft of 30 September 2015, not for public release
hiv2dna_main.tex r8876–10546

```
                                          Vpx end             Vpr start
MAC.US.x.239              ACTAGCATAA.............ATGGAAGAAAGACCTCCAGAAAATGAAGGACCACAAAGGGAACCATGGGATGAATGGGTAGTGGAGGTTCTGGAAGAACTGAAAGAAGAAGCTTTAAAACATTTTGATCCTCGCTTGCTAACTGCACTTGGTAATCATATCTATAAT 6297
Vpx                       __L__A__*_
Vpr                                              _M__E__E__R__P__P__E__N__E__G__P__Q__R__E__P__W__D__E__W__V__V__E__V__L__E__E__L__K__E__E__A__L__K__H__F__D__P__R__L__L__T__A__L__G__N__H__I__Y__N_
A.CI.88.UC2               T----TC---.........................--...........................-TGG-A---G-T-G-A-----T--T--.................-A-A----------C-A--GG----------C---------T---T----AC--CT----------- 6336
A.DE.x.BEN                T----TC---..TGACTGAAGCACCAAC---GTTT--C------G---GGACC----GG-----CTT-G--AG---C------A-A--AAC----AGG---A-A--G--------C----G----------C---------T---T-----CT-CT-----C------ 6397
A.DE.x.PEI2_KR_KRCG       C----TC---..TGACTGAAGCACCA-C---GTTT--C--G--GG-----ACC----CG----GG---G-------G------A-A-GAA-C---AG----T-A-G---------------G--------C-------------A-T-----C--CT------G-GC- 6359
A.GH.x.GH1                T----TC---..TGACTGAAGCACCAAC---GTTT--C------G---GGACC----GG-------T-G---G---C------A-AAGAA------G----A-A--G--------C-----G----------C---------TC--T-----C--CT-----C---G- 5838
A.GM.87.D194              T----TC---..TGACTGAAGCACCAAC---GTTT--C------G---GGACC----GG--A--G-T-G--AG-ACC------A-A--AAC----A-G---A-C--G--------C-------------------CT---------T---T-----C--CT----------- 5838
A.GM.x.ISY_SBL_6669_85    TT---TC---..TGACTGAAGCACCA-C---GTTT--C--G--GG---GGACC----CG---------G-------G------A-A--AA-----AG----A-A--G-----------------G--------C-----------------T-----CT-CT---------C- 5835
A.GM.x.MCNI3              T----TC---..TGACTGAAGCACCAAC---GTTT--C--G--GG---GGACC----CG---------G-------G---A--A-A--AA-C---AG-A--A-A--GA-------------G--------C-----------------T-----C--CT-----C---C- 5839
A.GM.x.MCR35              T----TC---..TGACTGAAGCACCAAC---GTTT--C--G--GG---GGACC----CG---------G-------G---A--A-A--AA-C---AG-A--A-A--GA-------------G--------C-----------------T-----C--CT-----C---C- 5839
A.GW.86.FG_clone_NIHZ     TT---TC---..TGACTGAAGCACCAAC----CTC--C--G--GG--AGGACC----CG-----G---G------C-------A-A--AA-C---AG----A-AG--------------GG--------C-----------------T-----C-GAT---------C- 5832
A.GW.87.CAM2CG            GT---TC---..TGACTGAAGCACCAAC---GCTC---------GG---GGACC----CG-----G---G---------------A-A--AA-C---AG---CA-A------------CC----G--C-----C-----------------T-----CGGA---------GC- 6403
A.GW.x.MDS                TT---TC---..TGACTGAAGCACCAAC---GCTC--C--G--GG---GGACC----CG-----G---G-------G------A-A--AA-----AG---GA-A-C-----------C-T--G--------C-GCA--------------T-----C-GA----------C- 5838
A.IN.07.NNVA              TTGG-TT---.ATGACTGAAGCACCAAC---GCTT--C--G---G--AGGACC----CG-----G---G-------G------A-A--AA-C--AAG----A-A-----------------GG--------C--C--------G-----T-----C-GAT-------C-C- 6389
A.IN.95.CRIK_147          T----TC---..TGACTGAAGCACCAAC---GCTC--C--G---G-C-GGACC----CG-----G---G-------G------A-A--AA-C--AAG----A-A-----------------GG--------C-----------------T-----C-TAT-C------G-- 6139
A.JP.08.NMC786_clone_41   T----TC---..TGACTGAAGCACCAAC---GTTT--C------G---GGACC--G--G-------T-G---G---C---A--A-AAGAA----A--G---A-A-G---------C-----G-------GA--C---CC-----TC--T-----C--CT---------GC- 6390
A.PT.x.ALI                T----TC---TGACTGAAGCACC-ACAG--TTTCC--CGGCCGG----GGACC----C.-----G---.-------G-----A-A--A--C---AG----A-A-----------------GG--------C------A------T---T-----CGGCT---------C- 6387
A.SN.85.ROD               T--G-TC---..TGGCTGAAGCACCAAC---GCTC--C--G-TGG---GGACC----TG-----G---G-------G---A--A-A--AA-CT--AG----A-A------------------G--------C--------------T---T-----C--AT---------C- 5840
A.SN.86.ST_JSP4_27        T----TC---..TGACTGAAGCACCAAC---GTCT--C--G--GG--AGGACC----CG-----G---G-------G------A-A--AACC---AG---GA-A---T---------------G--C-----C--------------T-A-T-----C--CT---------GC- 5839
B.CI.88.UC1               C-----C---..TGGCAGAAGCAGCCCC---G-CT--------G------A-C-----G--A--G--------A--G-------A---C--CA----G---A-A--GC-------C----GG----------C-----------------G---------TT------C-G- 6378
B.CI.x.20_56              C-----C---..TGGCAGAAGCAGCCCC---G-CT--------G-----G-AC-----G--A--G--------A--G-------A---C--C-----G---A-A--GC-------------GG----------C-----------------G---------TT------C-G- 6373
B.CI.x.EHO                T-----C---..TGGCAGAAGCAG-CCC---G-TT-------GG--A--AAC--------A-----G-----AC-G------------C--C-----G---A-A---C-------C-----G--------C--C-----AT------------C---TT------C--- 6368
B.GH.86.D205_ALT          C--G--C---..TGGCAGAAGCAGCCCC---G-TC--------G--C--GAAC--------A-----G-----A--G---A---G----A-C-----G---A-A--GC-------C-----G---------------------------G---------CTT------C-G- 6370
B.JP.01.IMCJ_KR020_1      CT-G--C---..TGACAGAAGCAGCCCC---G-CT--------G-----G-AC--------A--G--------A--G-------A---C--C-----G---T-A--GC-------C----GG----------C-----------------G---------TT------CGG- 5520
G.CI.92.Abt96             T--------G..........ATGGCA-------TC--C------G----G-CC--C--G--A--G-------------------A--A------------A-C---------C--------------C-----------------G---------CT---------G-- 5713
AB.CM.03.03CM_510_03      C--G--C---..TGGCAGAAGCARCCYC---R-TC--------G--C--GA-C--------A----AG-----A--GA-AR---G-----A-C-----G---A-A--GC-------CY----G-------------------Y---G---------CTT------C-G- 5509
H2_01_AB.CI.90.7312A      C-----C---..TGGCAGAAGCAGCCTC---G-CT--------G-----G-AC-----G--A--G--------A--G-------A---C--C-----G---A-A--GC-------C----G----------C-----------------G---------CTT---T--C-G- 6379
H2_01_AB.JP.04.NMC307_20  C-----C---..TGGCAGAAGCAGCCCC---G-CT-------GG----GAAC-----G--A-GG--------A--G-------A---CA-C-----G------A--GC-------C---------------C-----------------G---------CTT------C-G- 6379
H2_01_AB.JP.07.NMC716_01  C-----C---..TGGCAGAAGCAGCCCC---G-CT--------G----GAAC-----G--A-GG--------A--G-------A---CA-C-----G--G--A--GC-------CC--------------C-----------------G---------CTT------C-G- 6378
H2_01_AB.JP.08.NMC842_10  C-----C---..TGGCAGAAGCAGCCCC---G-CT-------GG----GA-C-----G--A-GGT-------A--G-------A---CA-C-----G--G--A--GC-------C---------------C-----------------G---------CTT------C-G- 6384
U.CI.07.07IC_TNP3         CT-G------.............----C-----------G--GG-C-----C-----G--A--G--T--------G-----C--T-----CT----G---T-A--GC---------G------------C--CA---------------------CT---------G-- 5812
U.FR.96.12034             -T-GATC---.............---C-----T---------GG-C-GG-CC--------A--G--TG----C--G-----GAGA--AA-A-----G-----A------------G---GG-----------------------------G---------CT----------- 5857
MAC.US.x.17EC1            ----------.............------------------------------------------------------------------------------------------------------------------------------------------------------- 6297
MAC.US.x.251_1A11         ----------.............----------------------C---TG-------------G-------------------------G---G-------------------------------------------------------------------------------- 6296
MAC.US.x.251_32H_PJ5      ----------.............----------------------C-------------------------------------------------------------------------------------------------------------------------------- 6298
MAC.US.x.251_BK28         ----------.............----------------------C---------------------------------------T---------------------------------------------------------------------------------------- 6273
MAC.US.x.MM142_IVMXX      ----------.............----------------------C-------------G----------A-----A-------------------G-----------GC-T-----C------------------------------------------------------- 5783
MNE.US.82.MNE_8           ----------.............--------------G-------C-----------------------------------G-------------------------------------------------G--------------------------------------- 5765
MNE.US.x.MNE027           ----------.............--------------G-------C-----------------------------------G-------------------------------------------------G---------T--------G-- 5765
SMM.CI.79.SIVsmCI2        TT-G------.............----CT---CAC-------GG-------C----CC--A--G-------C-C---------A--T--A----G---A-C--------------G--G-------C-----T-------C--G---------T--------G-- 5423
SMM.LR.89.SIVsmLIB1       CT--------.............----C-----------G--GG-------C----CGC-A----------C--G-----------A---------G---G-A--------------C-------------------T---AC--G-----C---T--------G-- 5371
SMM.SL.92.SIVsmSL92A      TT-T--G---............TGTCTTT-TCCCC---A------G----G-C---C-----A--G--T--------------ACA--A--CT-A-----GA-T---AC------C--G----------C--C--------------T--------CT--------G-- 5373
SMM.SL.92.SL92B           TT-G--C---..................TG---CATG-------GG----GACCAAC-C---A-------------G---A--AGA--T--A---------T----------G--AC-G-----------C-----------T-----TT-A--A---T-CG-A---G-- 5730
SMM.US.04.G078            ----------.............----C---------------G-------C-----G--A--G-----------G------ACA--A--------G---A-A--GC--------C-G----------------------------G---------T----T---G-- 5557
SMM.US.04.G932            ----------.............---AC---G-----------G-C-----C-----G--A-----G-----C--G--------------C--------GA-A-------------------------A-------------G---------T--------G-- 5546
SMM.US.04.M919            ----------.............--AC----------------G------CC-----G-------G-------------------A--------G---A-A-----------CC-G---------------------------------------CT--------G-- 5550
SMM.US.04.M922            ----------...........A---AC----------------G------CC-----G----G--G----------------A--A--------G---G-A-----------C--G-----------C-----------------G--------CT--------G-- 5553
SMM.US.04.M923            ----------...........A---AC----------------G------CC-----G--A----G----------------A--A--------G---A-A-----------CC-G--C--------------------------G--------CT--------G-- 5552
SMM.US.04.M926            GT--------...........A----C---------------GG------CC-------------G----C--------C-A--A-------------G-A-------------------G-------C-----------------G---------T--------G-- 5552
SMM.US.04.M934            GT--------.............----C---G-----------G------CC-------------G---------------A-A--A-------------A-A---------T--C---G--C------------------------G---------T--------G-- 5549
SMM.US.04.M935            ----------...........A---AC----------------G------CC-----G-------G----------------A--A--------G---G-A-----------C--G-----------C-----------------G--------CT--------G-- 5553
SMM.US.04.M940            ----------.............----C--------------GG-------C-----G--A--GT---------G------AC---A--------G---A-A---C--------C-G--C--------------------------G---------T----T---G-- 5557
SMM.US.04.M946            GT--------.............----C---G-----------G------CC-------------G---------------A-A--A-------------A-A---------T--C---G--C------------------------G---------T--------G-- 5550
SMM.US.04.M947            ----------...........A----C----------------G------CC-----G-----G------------------A--A--------G---G-A-----------CC-G-----------C-----------------G--------CT--------G-- 5552
SMM.US.04.M949            ----------.............----C--------------GG-------C-----G--A--G-----------G------ACA--A--------G--GG-A---C--------C-G-CT-------------------------G---------T----T---G-- 5545
SMM.US.04.M950            GT--------.............---AC----------------G------CC-------------G---------------A-A--A-------------A-A---------T--C---G--C------------------------G---------T--------G-- 5549
SMM.US.04.M951            ----------.............----C--------------GG-------C-----G--A--G-----------G------ACA--A--------G--GG-A---C--------C-G----------------------------G---------T----T---G-- 5554
SMM.US.04.M952            GT--------.............----C---G-----------G------CC-------------G---------------A-A--A----------------A---------T--C---G--C-----------------------GA--------T--------G-- 5550
SMM.US.05.D215            C---------.............----C--G-GT---------G-------C-----G--A-------------------GA--A--------G--GA-A-----------C----G--------C-------A----A---G---------T--------G-- 5551
SMM.US.06.FTq             CT--------.............----C---------------G-------C-----G--A-------------G------------A-------------T-A-----------C-G--G--------A-----------------G---------T--------G-- 5579
SMM.US.11.SIVsmE660_FL10  ----------...........A----C----------------G------CC-----G----------------------------A--------G---T-A-----------CC-G-----------------------------G--------CT---T----G-- 6288
SMM.US.86.CFU212          C---------.........AAG----C----GCT--C-----GG-----G-CC-----G--A--G--G-------------A----A--C-----G--G--C-------------AA--GG----------T---------------G--------CT-C--T---G-- 5546
SMM.US.x.F236_H4          ----------...........A----C----------------G------CC-----G-------------------------A-----A--G---T-A-----------CC-G-------------------------------G--------CT---T----G-- 6286
SMM.US.x.H9               ----------...........A---AC----------------G------CC-----G-------G------------R--A--A--------G----KA-----------CC-G--C--K-------------------TY--G--------CT--------G-- 5771
SMM.US.x.PBJA             ----------...........A---AC----------------G------CC-----G-------G----------------A--A--------G---A-A-----------CC-G--C---------------------------G--------CT--------G-- 6064
SMM.US.x.PGM53            ----------..........AA---AC----------------G------CC-----G-------G----------------AA-A--------G---G-A------------C-G-----C------------------------G--------CT--------G-- 6216
SMM.US.x.SME543           ----------...........A----C----------------G------CC-----G------------------------------A--------G---A-A-----------CC-G------------------------------G--------CT---T----G-- 6301
SMM.US.x.pE660.CG7G       ----------...........A----C----------------G------CC-----G------------------------------A--------G---T-A-----------CC-G------------------------------G--------CT---T----G-- 6287
STM.US.89.STM_37_16       CT-------G...........A---AC-C-C-----------GG-------C-----G--A-------------G-----------A--------G---A-A---C-------CC-GCGG---------------------GC--G--------CT--------G-- 5945
```

```
                               Tat exon 1 start                                                                                                                       Vpr end
MAC.US.x.239              AGACATGGAGACACCCTTGAGGGAGCAGGAGAACTCATTAGAATCCTCCAACGAGCGCTCTTCATGCATTTCAGAGGCGGATGCATCCACTCCAGAATCGGCCAACCTGGGGGAGGAAATCCTCTCTCAGCTATACCGCCCTCTAGAAGCATGCTATAACACATGCTATT 6467
Vpr                       _R__H__G__D__T__L__E__G__A__G__E__L__I__R__I__L__Q__R__A__L__F__M__H__F__R__G__G__C__I__H__S__R__I__G__Q__P__G__G__G__N__P__L__S__A__I__P__P__S__R__S__M__L__*
Tat exon 1                    _M__E__T__P__L__R__E__Q__E__N__S__L__E__S__S__N__E__R__S__S__C__I__S__E__A__D__A__S__T__P__E__S__A__N__L__G__E__E__I__L__S__Q__L__Y__R__P__L__E__A__C__Y__N__T__C__Y__
A.CI.88.UC2               ---------A----------A--C--CA---G-------A------A--------C----C------CC------TG---G-T-G--G---A-----T---T--A-AA--A--A---C----TG-C------C----A--C-----G------AC----C--------- 6506
A.DE.x.BEN                --------------------A--C--CA---G-------A--C---A--------C------G---C-------CG-----T-A--G---A-----T---T--A-AA--A--A---C----TG-C-----GC----A--C-----GC-----AC-----A-------- 6567
A.DE.x.PEI2_KR_KRCG       -----------------C--AA-C--CA---G-------ATG----G--------C------G---C-------CA-----T-AAAT--A-----T------A-AA------AG-C----T-------------AA--C-----G------A-----C----T---- 6529
A.GH.x.GH1                --------------A-C---A--C--CA---G--------------A--------C------G---CC------CA--G--T-A--G---A-----TA-----A-AA--A--A---C----T--C-----GC----A--C-----G-----TAC----T--------- 6008
A.GM.87.D194              --------------------A-----CA---G---------G----A--------C------G---CA------CG-----TGA--G---A----AG------A-AA--A--A--GC----TG-C-----GC----A--C-----G------AC----T--------- 6008
A.GM.x.ISY_SBL_6669_85    --------------------A--C--CA---G--------GG----A--------C-------C---C-------CA-----TGG------A-----T--------AA----A-------------------------A--C------A-----A------A-----T-- 6005
A.GM.x.MCN13              --------------------A--C--CA---G-------ATG----A--------C-----------C-------CG-----T-GG-T--A-----T-----A-A-----A--C----T--C------C-T--A--C------C----A------A--------- 6009
A.GM.x.MCR35              --------------------A--C--CA---G-------ATG----A--------C-----------C-------CG-----T-GG-T--A-----T-----A-A-----A--C----T--C------C-T--A--C------C----A------A--------- 6009
A.GW.86.FG_clone_NIHZ     --G-----------------A--C--CA---G--------------G--------C-----GC---C-------CA-----TGG------A-----T----GA-AA---------------------------A--C-----G-----AC----T-C--------- 6002
A.GW.87.CAM2CG            --------------------AA----CA---G--------GG----G--------C-------C---C-----G-CA--------A-----A-----T-----A--A---------------------------AA--C-------A---TG-----T--------- 6573
A.GW.x.MDS                -----------T--------A--C--CA-------------T-A-----------C------CA--C-------CA-----TGG------A-----T----GA-AA---------------------------A--C------------A------AT-------- 6008
A.IN.07.NNVA              ---T-----------C----A--T--CA---G---------G--T-G--------C--T----CA--C-------CA-----TGG------A--------T-----AA-A-------------------T--AA--C------------G-----T-G--------- 6559
A.IN.95.CRIK_147          ---T----------------A--C--CA---G---------G--T-G--------C--T--T-CA--C-------CA-----TGG------A--G------------AA---------------A--------A--C------------GC--G---T-------- 6309
A.JP.08.NMC786_clone_41   --------------G-----A--C---A--A-G---G--GAT---T-A--G----TC------C----C----A--CA--------G--GAC-A-AGG-TAAT--GA-AA---------C-T--T-AC---T-GC----A--C------A-----A------A---T--C- 6560
A.PT.x.ALI                --------------------AA----CA-G--G-------ATGC--A--------C-----------C-------CA----TGG--GT--A---G-T-----A-AA-----A-A-------------------A--C------A-----A----------------- 6557
A.SN.85.ROD               --------------------A--C--CA---G--------A-G---G--------C--T----C---C-------CA----TGG------A-----T----GA-AA---------------------------A--C------A-----A-----T--------- 6010
A.SN.86.ST_JSP4_27        --------------------A--C--CA---GG-------G-----A--------C---C--T----C-------CA------GG--G---A--G--T--T--G--CA-----C--------T-A-----------AA--C-----G-----G------A---T--C- 6009
B.CI.88.UC1               --G--------T-------CA----------G-------A---------G-----C------C-C--C------C---T--TCAA----A--G----A-----A--------------------------------------T---------G----T-A------C- 6548
B.CI.x.20_56              --G--------T-----C-CA------------------AG--T-------------------C-C--C----G--A--G---CAG----A--G--T--A-----A--------------------------------------T--G------A-------------C- 6543
B.CI.x.EHO                --G------A-T--------A----------G-------A-C-------G--G--A-----C-C--C-------G--T---CAA----A--G--T--A-----G--------------------------------------TA-G------AG---T--------C- 6538
B.GH.86.D205_ALT          --G--------T-------CA----------G-------A---------------C-----C-C--C-------C---T--TCAA----A--G--T--A---T-A-----------------------A----------C--TA-G------G----T--------C- 6540
B.JP.01.IMCJ_KR020_1      --G-----G--T---A---CA--------------------------T----G----------C-C--C---------A--G---CAA----G--G--T--A-----A-------------------T-------------T---------AGA--T-AG-----C- 5690
G.CI.92.Abt96             --G-------------C----------------------A------G--GA----T-------C----C--TC----A-AC--T-GATG------G--T--AGGGAAC-CA--C-----------A---A-----------------G--G--T-C---T-A---T---- 5883
AB.CM.03.03CM_510_03      --GT-------T-------CA--G--------G---G--A---------G-----C------C-C--C-------C---T--TCGA-----A--G-----A--G--A-----------------------R----------C--TA-G------A----T-A------C- 5679
H2_01_AB.CI.90.7312A      --G--------T-------CA--G---------------AG--T----G-------------C-C--C-----------T--TCAA----A--G--T--A-----A--------------------------C---------GAGG------A----T--G------- 6549
H2_01_AB.JP.04.NMC307_20  --G----------------CA------A-G--------C-A---T-----A---A-----C-C--C-----------T--TCAA----A--G--T--A-----A--------------------------------C--G-GG-----TA----T-G-------- 6549
H2_01_AB.JP.07.NMC716_01  --G----------------CA--------G--------C-A---T-----A---A-----C-C--C-----------T--TCAA----A--G--T--A-----A--------------------------------C--G-GG-----TA----T-AG------- 6548
H2_01_AB.JP.08.NMC842_10  --GT---------------CA--------G--------C-A---T-----A---A-----C-C--C-----------T--TCAA----A--G--T--A-----A--------------------------------C--G-GG-----TA----T-A-------- 6554
U.CI.07.07IC_TNP3         --GT----------------A-----------G--T--C-A-----G--A---C----TG---C---C----A----TGCG--T-----------GGG--AAA-A-------------G-----G---------------A-G--T-----T-A------C- 5982
U.FR.96.12034             --G--------T--------A-----------G-------A-----G--A---AA----C-C--C--------G-----TCGG-----A--G------AGCAGCA----------------A--A--G-----------T--G--G----C---T-A-------- 6027
MAC.US.x.17EC1            ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 6467
MAC.US.x.251_1A11         ------------------------------------------G--G--A----------C----------C---------------T-----------A------A----------------------------------------------------------- 6466
MAC.US.x.251_32H_PJ5      -------------------------------------------------------------------------A------------------------------------A---------------G--G------------------------------ 6468
MAC.US.x.251_BK28         -------------------------------------------------------------T-----------A------------------------------------A------------T--G--G------------------------------ 6443
MAC.US.x.MM142_IVMXX      --------------------------------------------------------A--------A---C---G--T-----------------------------------A--------------------------------------------------- 5953
MNE.US.82.MNE_8           ---------------------------------------A----------G------------C---------C---C-----T-----------T---A-----------------A----------------------A--------------------- 5935
MNE.US.x.MNE027           ---------------------------------------A----------G------------C---------C---A-----T-----------T---A-----------------A------------------------------------G------- 5935
SMM.CI.79.SIVsmCI2        --G-----------------A-------G----T--C-AG----G--------A----T--C--C-------CA----TC-T----------T--G--CAG-------------------A--G-----------C--G-AG----C----CGC--T--C- 5593
SMM.LR.89.SIVsmLIB1       --GT-------------C--A-------G--G-------AG----G--------A-----------C---------A--C--T-GG---------T-----GT-A--A---A--------GG--A--------------G-AG----T---T-A--------- 5541
SMM.SL.92.SIVsmSL92A      --GT----------------A-------------T----A----T-G--GA----C--G--C-C--C---------A--G--T-CAA---A-----T--ATGCAGAA-A------------G--A------------------A-G----G---T-A-----T-- 5543
SMM.SL.92.SL92B           -CTT-C--G--T--AA-A--A--G-----G---A-A----AG--AT-A---AAG--AT-G---T-A--C--T--GCAT-------C---T--A--G--T--A--G--AA-----CC------C--GGG-T---------T--G-----GA-G--T-------GC--T-T-- 5900
SMM.US.04.G078            --------------------------------G-------A------G-------C-----------C-----AAT----TG--------C---T-----AGC--A-------------T-GG--A------------------GAAG----G---T---------- 5727
SMM.US.04.G932            ---T----------------------------G-------A---------G-------------CT--C-------------C--TG------------------------A----------------A------------------AG----C-------------- 5716
SMM.US.04.M919            --------------------A-------------------------------G----------C-A--C--------A-----T-CT-------------------AG--A----------------G--A------------------G--G----T---T-----T--C- 5720
SMM.US.04.M922            --------------------A-----------------------------G------------C-A--C--------A-----TG-T---------G-------AG--A---------------GG---------------------G--G----C---T--------C- 5723
SMM.US.04.M923            --------------------A-----------G-----------------GA---A-------C-----------T--------GA-----------------GT---A--------------------A------------------G--G----------AG-----C- 5722
SMM.US.04.M926            ---T----------------------------G-------AG---------------------A--C--------A--C--T-CT-----------------AG--A--G-------------G---A------------------G--G----------------T--C- 5722
SMM.US.04.M934            ---T----------------------------G-------AG-------------A-------A--C--------A--C--T-CT-----------------AG--A-----------------G--A--G---------------G--G----C---TG-T--T--C- 5719
SMM.US.04.M935            -----------T-----G--A--------------------------------------C-A--C--------A-----TG-T-------------------AG--A----------------GG---------------------G--G----C---T-----T--C- 5723
SMM.US.04.M940            --------------------------------G--T---A------G--GA---C-----------C------AA-------C----------G-------G-CA-A-------------T-GG--A------------------GAGG----G---T-----T---- 5727
SMM.US.04.M946            ---T----------------------------G-------AG-------------A---------A--C--------A--C--T-CT-----------------AG--A-----------------G--A-----------G-----G--G----C---TG-T--T--C- 5720
SMM.US.04.M947            --------------------A-----------G-------------T--GA---A--------C-----------T-------G-------------------T---A-----------------T---A------------------G--G----------AG-----C- 5722
SMM.US.04.M949            -----------T--------A-----------G-------A------G---A---T------T---C-------C-----TGC-------G---------G-CA---------------T-GG-------------------G-AG----G---T---------- 5715
SMM.US.04.M950            ---T----------------------------G-------AG--------------A-----------C--------A--C--T-CT-----------------AG--A-----------------G--A------------------G--G----T---T--T--T--C- 5719
SMM.US.04.M951            -----------T--------------------G-------A------G---A---A-----------C----A--C------C---------G-------G-C---------------T-GG--A---------------------AG----G---T---------- 5724
SMM.US.04.M952            ---T----------------------------G-------AG---------------A---------A--C--------A--C--T-CT-----------------AG--A-----------------G--A------T---------G--G----T---T--T--T--C- 5720
SMM.US.05.D215            -----------------C--A------------------AG----T-------T-----------C----G--------CAA-----------T----------A-------------GG--A---------------------G--G-------T-G--------- 5721
SMM.US.06.FTq             --G-----------------------------G-------A-----A--A---A-----C-C--C-------A--C---C--T--A--------AC-AA-A-A-------------T---A--G---AA---------A-G----G---T--------C- 5749
SMM.US.11.SIVsmE660_FL10  --------------------A-------------------------G--GA---A-------C--------A-T----TGC--------------------T--A-A-----------------A---A------------------GA------------AG-----C- 6458
SMM.US.86.CFU212          --G-----------G--C-------C-----G-------AG--A----G-------------C--C--------T----TGC------A-------T--C--AA-A-------------T-GG--A----------A-----GATG--T-G---TT-C------- 5716
SMM.US.x.F236_H4          --------------------A-------------------------G--GA---A-------C--------A-T----TGC--------------------T--A-A--C.......................................G---------AG-----C- 6420
SMM.US.x.H9               --------------------A-----------G-----------------GG---A-------C-----------T------GA-----------------GT---A--R--------------A-----------KG---G--G--T-------AG--Y--C- 5941
SMM.US.x.PBJA             --------------------A-----------G-----------------GA---A-------C-----------T------GA-----------------GT---A------------------A------------------G--G--T-------AG-----C- 6234
SMM.US.x.PGM53            --------------------A-------------------G---------------------A--C--------A----TGCT-----------------AG-A-A-------------G---A------------------G--G----C---T-----T--C- 6386
SMM.US.x.SME543           --------------------A---------------------------A--GA---A-------C--------A-T----TGC--------------------T--A-A------------------A-----------------GC-------------AG-----C- 6471
SMM.US.x.pE660.CG7G       --------------------A---------------------------G--GA---A-------C--------A-T----TGC--------------------T--A-A-------------A---A------------------GA-------------AG-----C- 6457
STM.US.89.STM_37_16       --------------------A-----------G-------AG--------G-----C-----------C-----------T---CG--------G--T--------A--A-----------T-GG--A----------A-----G-AG--T-G-----AG----T-- 6115
```

Rev exon 1 start
Tat Rev exon 1 end
Tat Rev intron start
Env start
MAC.US.x.239
Env
Tat exon 1
Rev exon 1
A.CI.88.UC2
A.DE.x.BEN
A.DE.x.PEI2_KR_KRCG
A.GH.x.GH1
A.GM.87.D194
A.GM.x.ISY_SBL_6669_85
A.GM.x.MCN13
A.GM.x.MCR35
A.GW.86.FG_clone_NIHZ
A.GW.87.CAM2CG
A.GW.x.MDS
A.IN.07.NNVA
A.IN.95.CRIK_147
A.JP.08.NMC786_clone_41
A.PT.x.ALI
A.SN.85.ROD
A.SN.86.ST_JSP4_27
B.CI.88.UC1
B.CI.x.20_56
B.CI.x.EHO
B.GH.86.D205_ALT
B.JP.01.IMCJ_KR020_1
G.CI.92.Abt96
AB.CM.03.03CM_510_03
H2_01_AB.CI.90.7312A
H2_01_AB.JP.04.NMC307_20
H2_01_AB.JP.07.NMC716_01
H2_01_AB.JP.08.NMC842_10
U.CI.07.07IC_TNP3
U.FR.96.12034
MAC.US.x.17EC1
MAC.US.x.251_1A11
MAC.US.x.251_32H_PJ5
MAC.US.x.251_BK28
MAC.US.x.MM142_IVMXX
MNE.US.82.MNE_8
MNE.US.x.MNE027
SMM.CI.79.SIVsmCI2
SMM.LR.89.SIVsmLIB1
SMM.SL.92.SIVsmSL92A
SMM.SL.92.SL92B
SMM.US.04.G078
SMM.US.04.G932
SMM.US.04.M919
SMM.US.04.M922
SMM.US.04.M923
SMM.US.04.M926
SMM.US.04.M934
SMM.US.04.M935
SMM.US.04.M940
SMM.US.04.M946
SMM.US.04.M947
SMM.US.04.M949
SMM.US.04.M950
SMM.US.04.M951
SMM.US.04.M952
SMM.US.05.D215
SMM.US.06.FTq
SMM.US.11.SIVsmE660_FL10
SMM.US.86.CFU212
SMM.US.x.F236_H4
SMM.US.x.H9
SMM.US.x.PBJA
SMM.US.x.PGM53
SMM.US.x.SME543
SMM.US.x.pE660.CG7G
STM.US.89.STM_37_16

Draft of 30 September 2015, not for public release
hiv2dna_main.tex r8876–10546

```
                                                                                                                                          _V1 loop start
MAC.US.x.239                GAAGTGGCCCTTAATGTTACAGAAAGCTTTGATGCCTGGAATAATACAGTCACAGAACAGGCAATAGAGGATGTATGGCAACTCTTTGAGACCTCAATAAAGCCTTGTGTAAAATTATCCCCATTATGCATTACTATGAGATGCAATAAAAGTGAGACAGATAGATGGGG 6968
Env                         _E__V__A__L__N__V__T__E__S__F__D__A__W__N__N__T__V__T__E__Q__A__I__E__D__V__W__Q__L__F__E__T__S__I__K__P__C__V__K__L__S__P__L__C__I__T__M__R__C__N__K__S__E__T__D__R__W__G
A.CI.88.UC2                 ---A-AC-TT-A-----G------GCT--------A---G----------A--------A--------A-----C---AGG--A--------A--G-----A--A-----C--G--GA-A--TC----TG-GG-A----AT--T---CC-GT-ACAGGGA---ACACCAA 7007
A.DE.x.BEN                  ---A-AATTT-A-----G-----GGCT--------A-------------G--------A---G----A-----C-----T--A--------A--------A--A-----C--GC--A-A--TC----TG-GG-A----AT--T-GC-GGGT-C-AGGGA---CCAC-AC 7068
A.DE.x.PEI2_KR_KRCG         ---A-AC-TT-A-----G-----GGCT-----C--A-----C--------A--------A---G----A-----C---A-T--A--------A---G----A--A-----C------A-A--C-----TG-GCAA---GA---T--C-GC-CAAGT-----G--CA-TAA 7030
A.GH.x.GH1                  ---A-AA--T-G-----G-----GGCT--C--C--A---G----------A--------A---G----A-----C---AGT--A--------A--------A--A--C--C---C--A-A--TC----TG-AG-A-----T--T--C-GC-CCACT-ACA-C-C-ACCAC 6509
A.GM.87.D194                ---A-AA--T-A-----G------GCT--------A---G----------A--------A--------A-----C---AG---G--------A--------A--A-----C--G--GA-G--CC----TG-GG-G----AT--T----T--C-TCAGGGAC--CCGC-AC 6509
A.GM.x.ISY_SBL_6669_85      --GA-AC-TT-G-----A-----GGCT-----C--A---G------T---A--------A---G----A-----C---A-T--A--------A--------A--A-----C------A-G--T-----TG-A--A----AC--T--CGC----ACAGAGAGCGC-GTT-C 6506
A.GM.x.MCN13                --GA-AA-TT-G-----G-----GGCT--C-----A---G----------A--------A--------A-----C---A-T--A--------A--------A--A-----C------A-G--T-----TG-AG-A--------T--C--C-CA--TG--AGG-ACACAAC 6510
A.GM.x.MCR35                --GA-AA-TT-G-----G-----GGCT--C-----A---G----------A--------A--------A-----C---A-T--A--------A--------A--A-----C------A-G--T-----TG-AG-A--------T--C--C-CA--TG--AGG-ACACAAC 6510
A.GW.86.FG_clone_NIHZ       ---A-AA-T--A-----G-----GGCT--C-----A--------------A--------A---G---------C---A-T--A--------A--------A--A-----C------A-A--C-----TG-AG-A----AC--T-CA-GG-ACAT---CACAT-GACA-- 6503
A.GW.87.CAM2CG              --GA-AC-TT-A-----G-----GGCT--C-----A---G-------A-A--------A--------------C---A-T--A--C-----A-------A--A-----C------A-A--T-----TG-AG-A----A---T--C-T----ACA-GT--C-C-ACTAT 7077
A.GW.x.MDS                  ---A-A--TT-G-----G-----GGCT--------A---G----------A--------A---G----A-----C---AGT--A--C-----A-------A--A-----C--G---A-A--C-----TG-AG-A----AC--T-GC-G-CTCAGTGACAGCGC-A-CAC 6515
A.IN.07.NNVA                --GA-A--TT-A-----G-----GGCT--C-----A---G-C--------G--------A---G----A-----C---A-T--A--C-----A----C--A--A-----C------A-A--C-----TG-AG-A---GAG-----G--C--A-------GG-ACGCAAC 7063
A.IN.95.CRIK_147            ---A-A-TTT-G-----GG----GGCT--C-----A---G-C--------A--------A--------A-----C---A-T--A--------A----C--A--A-----C------A-A--C-----TG-A--A-----------C-C----ACA--CACA-CCACAAC 6810
A.JP.08.NMC786_clone_41     ---A-A---T-G-----G-----GGCT--C--C--A---G----------A-----------------A-----C---AGC--A--------A--------A--A--C--C---C--A-A--TC----TG-AG-A-----T--T--CCGC-CCAGC-ACCCA-CCACACC 7067
A.PT.x.ALI                  ---A-A--TT-G-----G-----GGCT--C-----A---G----------A--------A---G----A-----C---AG---A--------A--------A--A-----C--G---A-A--T-----T--AG-A----AG--T-GC--C-TAAGC-----G--CACAAC 7058
A.SN.85.ROD                 ---A-AA-TT-G-----A-----GGCT--------A--------------A--------A--------A-----C-----T--A--C-----A--------A--A-----C---C--A-A--T-----TG-AG-A----A-----GC-GC-CA----GCAGC-C-G--AA 6502
A.SN.86.ST_JSP4_27          ---A-A--TT-A-----G-----GGC---C--C--A--------------A--------A---G----A-----C---AGT--A--------A--------A--A--C--C---C--A-A--C-----TG-AG-A---C-T--T--C-GC-CAACTG--A-A-ACACAAC 6510
B.CI.88.UC1                 --GA-CAGTG-A---A-A-----GGCT--------A--------------A-----------G---T----G--AGT--T-----A-----C-----A--A-----T--GC-GA-G--GC-G---G-GG-A--------T-----C-CA-GA--CA---CTACTAC 7049
B.CI.x.20_56                --GA-CAGTG-G---A-A-----GGCT--------A--------------G-----------G---T----G--A-T--T-----A-----CC----A--A-----T--GC-GA-G--GC----TG-GG-A----AT--TTC--C--A-A-C---AGA-CTAACAA 7038
B.CI.x.EHO                  ---A-CCAAT-A---A-A-----GGCT--------A---G----------G----T------C-A-------G--AGT---------------------A--A-----T-----A-----C-G--TG-G--A----AG--T-----G-CGTG--GCTCAGCTA-CAA 7033
B.GH.86.D205_ALT            --GA-CAGG--A--CA-A-----GGCT--------A---G----------G---C-------G---T----G--AG----------A-----C-----A--A-----C---C--A-----C-G--TG-GG-A----AC--T-G-----CC--A---A-CCC-G--AA 7035
B.JP.01.IMCJ_KR020_1        ---A-C--G--G---A-A-----GGCT--------A---G---C------G-----------G--C-----G--A-T--A--------T--------A--A-----T-----A-----C----TG-A--A----AT--T-G---CCT-A-C--CACAGCCAACAC 6185
G.CI.92.Abt96               ---C-A--AT-G-----A------GCT------------G---C------A-----------G---A-----T--A-T--G--------A--T-C--A--------T-----A-A--TC-G-----G-C-----------------C-------A-C-A------ 6381
AB.CM.03.03CM_510_03        ---A-AA-GT-A-----G------GCT--------A---R----------A--------A--------------C---AG---A--------A--------A--A-----C--G--GA-A--CC-R--YG-GG-A----AT--T-CA-CC-CG--AG-CACACCCCA-AA 6168
H2_01_AB.CI.90.7312A        ---A-A-----G-----G-----GGCT--C-----A--------------A-----------G---A-----C---AGT--A--------A--------A--A--C--C---C--A-A--TC----TG-AG-A-----T--T--C-GC-CCACCG-CACC-CTACACC 7044
H2_01_AB.JP.04.NMC307_20    ---A-A---T-G-----G-----GGCT--C-----A---G----------A--------A---G---------C---AGT--A--------A--------A--A-----C---C--A-A--TC----TG-AG-A----AC--T-CC--C-CCACT-GCA-C-C-AACAC 7047
H2_01_AB.JP.07.NMC716_01    --GA-A---T-G-----G-----GGCT--C-----A---G----------A--------A---G---------C---A-T--A--------A--------A--A-----C--GC--A-A--TC----TG-AG-A----AC--T-CC--C-CCACT-GCA-C-C-AACAC 7046
H2_01_AB.JP.08.NMC842_10    ---A-A-----G-----G-----GGCT--C-----A--------------G--------A---G---------C---A-T--A--------A--------A--A-----C---C--A-A--TC----TG-AG-G----AC--T-CC--C-TC-CT-GCA-C-C-AACAC 7052
U.CI.07.07IC_TNP3           ---T-A--TT-A---A-C-----GGCT--------T---G-C--------A-----G-------------------A-T--T-----A-----T--------A-----C--GC-GA-A--CC-C--T--AG-A-----G--T-------CG--------C-A------ 6477
U.FR.96.12034               --GT-A--AT-A-----C-----GGC--------------------C---A-----G--A---G-G--A-----C---AGT--T--------A--T-----A--A--------G--GA-A--CC-T--T--AG-A--------T-G-----C---A---A---------- 6522
MAC.US.x.17EC1              ---A------------------------------------------------------------------------------------------------------------------------------------------------------------------- 6968
MAC.US.x.251_1A11           ---T-------------------------------T---G-G-----------------------A---------------------------------------------A----A--------------------------------------------------- 6967
MAC.US.x.251_32H_PJ5        ---T----------C------------------------G-G----------------------------------------------------------------------------------------------------------------A------ 6969
MAC.US.x.251_BK28           ---T-------------------------------T---G-G------------------------C------------------------------------------------------------------------------------------------ 6944
MAC.US.x.MM142_IVMXX        ---T-------------------------------T---G-G------------------------C-----------------------------------------------------------------------------------------A------ 6454
MNE.US.82.MNE_8             ---T-----------A-------------------T---G-G------------------------------T----------------------------A----A-------------------A------C---------------A------ 6436
MNE.US.x.MNE027             ---T-----------A-------------------T---G-G------------------------------T----------------------------A----A-------------------A------C--------A------A------ 6436
SMM.CI.79.SIVsmCI2          ---A-AAGAA-G-----G-----GTC---------T---G-C--C-----A-----------G----------T---A-TT-A-----A-----T--C-----A-----T-----GA-G-----G-----G-A---C-T-------G-C---A-------A------ 6091
SMM.LR.89.SIVsmLIB1         ---A-AAGTA-C-----C-----GGCT------------G-C--C-----T-----G-A--------------T---A-C--------A--A--C--T-----G--------GC--A-T--------T---G-A---AG--T-------A---A---A----G----- 6033
SMM.SL.92.SIVsmSL92A        ---T----TA-------C-----GTCA--------A---G----------A-----G----------------T---A-T--T---------T--T--A--C-----------GA-------G--T---G-A---CAG--TG---G--A-------AGA-------- 6047
SMM.SL.92.SL92B             ---T-------------------GGC---C-----T---G----------A--T-----A--------A-----T---A-T--------A--A--T--T--A----------G---A-T--------T---G-----A---T-------A---------C-------- 6395
SMM.US.04.G078              ---T-----A-------C------GCA--------T---G-C---------------G----------------------A-T----------------T-----A--C--------G---A-T--C---------G-C----A---------G-CA--C-----C-A------ 6228
SMM.US.04.G932              ---T----T---------------CAA------------------C------------------------------A-T--T-----A--------------A--C--------------A-----C-------G-C--------------------T------------- 6217
SMM.US.04.M919              ---T----TA-A--------------A--------------------------------------------------A-T------------------T-----A--C--------------A----C----------C----A---------G-CA--A-----CCA------ 6221
SMM.US.04.M922              ---T-----A-A--------------A------------G-C-----------------------------------A-T------------------T----------------------A----C--G--------C----A---------G-CA--A-----C-A------ 6224
SMM.US.04.M923              ---T----AA-------C-----GGCT--------T--------------------A--------------------A-C--------A--A--C--T-----C----------C-CA-----C----T--AG-A-----G--T-------A---A-----C-------- 6217
SMM.US.04.M926              ---T-----A-A--------------A----------------C-------T-----G-------------------A-T------------------T-----A--C--------G--GA----C-----------C----A---------G-CA--A-------------- 6223
SMM.US.04.M934              ---T----TA-A--------------A----------------C---------------------------------A-T------------------T-----A--C--------------A----C---------G---------------G-CA--A-------A------ 6220
SMM.US.04.M935              ---T-----A-A--------------A----------------C---------------------------------A-T-------------T--T---------C--------------A----C-----------C----A---------G-CA--A-----C-------- 6224
SMM.US.04.M940              ---T-----A-------C------GCA--------T---G-C---------T-----G-------------------------------------G--C-----A--C--------G---A----------T---G-C----AG-------GG-CAC-A---A---A------ 6228
SMM.US.04.M946              ---T----TA-A--------------A------------G-C--------------------------G---A-C-----------------T-----A--C--------------A----C---------G---------------G-CA--A-------A------ 6221
SMM.US.04.M947              ---T----AA-------C-----GGCT--------T---G--------------A-------------G---A-C--------A--A--C--T-----C--C-----C-CA----------T--AG-A-----G--T----------C------------- 6217
SMM.US.04.M949              ---T-----A-------C------GCA--------T---G--------------G-----------------------C--------A--A--C-----A--C--------G---A-----C---------G-C-----G--T---GGG-CA--A---A---A------ 6216
SMM.US.04.M950              ---T----TA-A--------------A----------------C-------------------------G---A-T------------------T-----A--C--------------A----C---------G---------------G-CA--C-------A------ 6220
SMM.US.04.M951              ---T-----A-------C-----GGCA--------T---G-C---------------G--------------------------------A--A--T-----A--C--------G---A----------T-----C-----G--T----G-CA--A-------AG----- 6225
SMM.US.04.M952              ---T----TA-A--------------A------------G-C-----------------------------------A-T------------------T-----A--C--------C--A----C---------G---------------G-CA--A-------------A- 6221
SMM.US.05.D215              ---A--C-TA-------------GGCT--------T---G----------G------A--T-----------G---A-C---------T-A--C--T------------------A--------------AG-C-----G--T----G-----C-----C-------- 6222
SMM.US.06.FTq               ---T-A--T--C-----C-----GGC---------T---G-G--C-----------------G-------------A-C------A---T--T----A--C-----T---C--A-A---C----T--AG---------------GG-CA-----A--C------- 6247
SMM.US.11.SIVsmE660_FL10    ---T----AA-C-G-A-C-----GGCT--------T--------------------A-------------G---A-C--------A--A--C--T-----C---------C-CA-----C----T--AG-A-------T------C-----------G----- 6953
SMM.US.86.CFU212            ---C-A--AA-C-----C-----GGCT-----------------C-------------G-A-------------C---A-T-----------A--C-----A--A-------------A----------T--AG-A-------------GC---------C-AG----- 6211
SMM.US.x.F236_H4            ---T----AA-C-----C-----GGCT--------T---G--------------A-------------G---A-C--------A--A--C--T-----C---------C-CA-----C----T--AG-A-------T------C-----------G----- 6915
SMM.US.x.H9                 ---T----AA-------C-----GGCT--------T---G--------------A-----------------A-C--------A--A--C--T-----C---------C-CA-----C----T--A--A-----G--T-----------M---MC-------- 6436
SMM.US.x.PBJA               ---T----AA-------C-----GGCT--------T---G--------------A-----------------A-C--------A--A--C--T-----C---------C-CA-----C----T--A--A-----G--T-----------------C-------- 6729
SMM.US.x.PGM53              ---T-----A-A--------------A----------------C-----------------------------A-T------------------T-----A--C--------------A----C-----------C---------------G-CA--A-----C-------- 6887
SMM.US.x.SME543             ---T----AG-C---A-C-----GGCT--------T--------------------A-------------G---A-C--------A--A--C--T-----C---------C-TA-----C----T--AG-A-------T------C-----------G----- 6966
SMM.US.x.pE660.CG7G         ---T----AA-C---A-C-----GGCT--------T--------------------A-------------G---A-C--------A--A--C--T-----C---------C-CA-----C----T--AG-A-------T------C-----------G----- 6952
STM.US.89.STM_37_16         ---T----AA-A-----C-----GGCT-----------G-C--------------A-------------C---A-C--T-----A--A--C--T--A--------------GA-T---C----T---G-A-----G--T-----G-A---A-----C-A------ 6610
```

```
MAC.US.x.239                ATTGACAAAATCAATAACA..........................................ACAACAGCATCAACAACATCAACGACAGCATCAGCAAAAGTAGACATGGTCAATGAGACTAGTTCTTGTATAGCCCAGGATAATTGC...ACAGGCTTGGAACAAGAGC 7093
Env                         __L__T__K__S__I__T__.__.__.__.__.__.__.__.__.__.__.__.__.__.__T__T__A__S__T__T__S__T__T__A__S__A__K__V__D__M__V__N__E__T__S__S__C__I__A__Q__D__N__C__.__T__G__L__E__Q__E__
A.CI.88.UC2                 TGCC---GCCAA-CC---C...................................................G--G-C-G-C----C--CAACC-TT--T-CC-TAC---AA-A-----A-G-TC-A-C---G---G-GCA--C------...-----A----GGG-C--AG 7123
A.DE.x.BEN                  CCC--ATCCCAGG-CCT-G................................................AGT--C-----C--G-GAC--C-CA--T-CGC--CCTC---AA-A-----A---TC-AAC--C----AAA-CA-C-CA---...G----A--A-GGT-T---G 7187
A.DE.x.PEI2_KR_KRCG         CAGC----GTGAGGGG-GC...........................................................................A--GTCCC---G--AT-A--C--A----C---A--C---A--A-CA-C-GC---...T---AT--A-GGAGT--AG 7122
A.GH.x.GH1                  -ACAGGG-GCA---C-GGG....................................................................................ATGTCA-AG-TAA-TGAA-C----CCATC-TA-TCA--C--C---...-----A----GGA-G--AG 6592
A.GM.87.D194                CCC--GTCC-C---AC-TT....................................................................................ACA--AA-AG----A-A-TC-A-C-------G-G-CA-C--C---...-----A--A-GGA----AG 6592
A.GM.x.ISY_SBL_6669_85      -ACT----GCC--TCTGG-.................................................................................CCT--T---A-A-----T---GA-C-A--C---CAATT-A-C------...T----AC--AGGG-G--AG 6592
A.GM.x.MCN13                CACAC-C-C-A--GC-T-C...............................................................C--CGTA--AT-AA-CCC--GACAGA-A-A-G-----A-TCC--A--C---CG-GCAA-C--C---...T----A----G-G----AG 6614
A.GM.x.MCR35                CACAC-C-C-A--GC-T-C...............................................................C--CGTA--AT-AA-CCC--GACAGA-A-A-G-----A-TCCA-A--C---CG-GCAA-C--C---...T----A----G-A----AG 6614
A.GW.86.FG_clone_NIHZ       GAG----G-CA-CCA--AT....................................................................................AT--CAA-T-TAA-TGAC-C---CCA-GCGCGTGCA--C--C---...-----A--AA-GG-G--AG 6586
A.GW.87.CAM2CG              GA-A-GG-C-A-C-C-C-C......................................................................AGCA----T---AG-CACCAA-A-G---C-A-TC-C----C---CG-ACAA-C--C---...T----A--A--GG----AA 7175
A.GW.x.MDS                  -AG--AC-C-A-C-C--AC.................................GCCACCGCAG-C---A-CA---TC-G-G-C--A--CAT-G--C---GGAACACA-CCA-A-G-----A-TC-C-A--C---CG-GCA--C--C---...T----AC-A--GG-G--AG 6649
A.IN.07.NNVA                CACA--CGC-C---C---C...............................................................--AC---AGATTAATG-GAC-ACTGA-A-A-G-----AC-C---A--CG--CG-GGA--C--C---...T----A--AAG-G-G---G 7167
A.IN.95.CRIK_147            CGCAC-C-C-AGC-C--GT.........................................................G--GGG-GC-----CA--C-C---CCCATA-----A-------AC-C---A--C--GCA-GCAA-C--C---...T----A----G-G-G--AG 6920
A.JP.08.NMC786_clone_41     -GCAGGC-C-A---AGT-C...............................................................TTCC--A-CC---G-G-GAATA-A-CCA-A-----A-A-GA-A-A--C---CG-G-AA-C--C--T...-----A----GGG-G--AG 7171
A.PT.x.ALI                  CACAT-CCCGAGCCC-GGG........................................................................AGCA--CTCAA-CC-C--A-A-------GCGA-C-A--C---AAGGCA--C--C---CCC-GG--AC-A-GGG-T--AG 7156
A.SN.85.ROD                 CAAC----CC----AG-GC................................................A--AGC-----CA----C---C-CA---ACC-G-AGC-AGA-A-A-G----GA--C-C-A--CGC-CG-GCA--C--C---...T----A----G-G-G--AG 6621
A.SN.86.ST_JSP4_27          C-CC---CC-A---CC--C........................................................................A--A--GC-AACACA-CAA-AGGA----A-TC---A--C---CG-ACA--C--C---...-----G----G-G----AG 6605
B.CI.88.UC1                 TAAAC-T-TTA-T-C-C-T............................................................ATT--T--TA-TAA-C--TC--A-A--T-AC----C--C--A---C-A-----TAAAA-T--C-CC---...C-G--AA-A-G--T---AA 7156
B.CI.x.20_56                TACT--TGC---T-CT--T...............................................................--TGGTAAT--TA-T-C-CC-ATAG-A--------AG-A-T-C-A---G-TAAAGCTA-C--C---...T-G--AA-A-G--T---AG 7142
B.CI.x.EHO                  -GA---C-CTA-GTCCT-C............................................................G-GT-CCTCAG---TT-T-CTCAGAC-C-AC----C--AGA---CAAA----TCAAA-T--C-GC--T...G---GA-A-G--T---AG 7140
B.GH.86.D205_ALT            TGCC-GT-GTA-T-CC--C......................................................--T-AGC-T--T--CA-C--TCGTGGGC-GA-A-C-A-T--C--A--AGACC-A--C---AAAA-T--C-GC---...-----AC-A-G-G-----G 7148
B.JP.01.IMCJ_KR020_1        CACC---GCCAAC-AC-TC..................................................................-----C--TAAG--CA--AGT---C------GA-A-GAC--A-----CCGAG-T--C--C---...-----TA-A---TTT---A 6286
G.CI.92.Abt96               TC-A---GGGA-C-CC--T..........................................GT---CC-TG-----GTTC-C--A--TAA-ATG-TG-C--CT--AC-A--G---AGC-----CCAG--CT--ATGT-C--------T...-----TA-AC-GTC---AT 6506
AB.CM.03.03CM_510_03        TACC----C-CGG-AT--C....................................................................................ACA--AA-AG-G--T-A-GACC--------AGTAGAA-C--C---...-----A--A---A-T--AG 6251
H2_01_AB.CI.90.7312A        -CCA-GC-CCA-T-AC-AC......................................................-----CA----AGAGC-CA--A--GG--GGCCAGA-A-A-----A---TT-C-A--C--GCG-ACA--C--C---...-----A----G-G-G--AG 7157
H2_01_AB.JP.04.NMC307_20    -CCA--CCCT----AC-T-...........................................................................-A-GGGCCCACAC--A-A-----A---TC-C-A--C---CGAA-TA-C-GC---...-----A----G---G--AG 7139
H2_01_AB.JP.07.NMC716_01    -CCA--CTCTC---AC-T-...........................................................................-A-GGGCCCACAC--A-A-----A---TC-C-A--C---CGAA-TA-C-GC---...-----A----G---G--AG 7138
H2_01_AB.JP.08.NMC842_10    --CAG-CTCTC-CCC---C.....................................................................T-TC--A-C-C--A--GGC--A-A-----A---TC-C-A--C---CGAA-TA-C-GC---...-----A----G-G-G--AG 7150
U.CI.07.07IC_TNP3           G--A---GG-A--G--G--....................................ACAACA----GTC--A---AC---A--G-A-ACAT-A---G--G-TC---AGAAA-A----GA--AGA-C--------AG----A--------...-----A--A---TC----- 6608
U.FR.96.12034               ---A---GG-A--GC-G--...................................................A-------TAGCC-AG--ATGCA--A-TC-AGTA-AG-AA-A-----A--AGAGC--------CGAA-TA---GC---...T----A--A--------A- 6638
MAC.US.x.17EC1              -------------------..........................................-------------------------------------G-------------------------------------------------...G------------------ 7093
MAC.US.x.251_1A11           -------------TC----..........................................---A-CA--G--G--C---A-G--ATG---A---A--A-----A---------------------------TT--T-----------...------------------- 7089
MAC.US.x.251_32H_PJ5        -------------T-----..........................................---A--G--C-----G--C-A--G---G--T-----A-----------------------------------A-T--T-----------...------------------- 7091
MAC.US.x.251_BK28           -------------TC----....................................ACAATA---------G--C--------G-AC---T-----A---A---------------------------------T--A-----------...------------------- 7075
MAC.US.x.MM142_IVMXX        -------------TC----.................................ACAACAGCAT-----A--A--------A-G-A-A-T--GT--AG-C-AG-----A-----------------C----G--TT--T-----------...------------------- 6588
MNE.US.82.MNE_8             -------------TC----..........................................---------C-----G--AT-C-A---AA-G---AGGC-A--A-AG------------A----C-------AAT--T----------...------------------- 6561
MNE.US.x.MNE027             -------------TC----.......................................ACA---G-----C--------AA---A---A--A--AAGG--A----AG----------A-A----A-----G--AAT-GT---------...------------------- 6564
SMM.CI.79.SIVsmCI2          -------GG-AAC-CG--G..............GCAGCACCTACAACCACGACAACAACA---G---T-A--G----G--GTA--A--A-----GG--G-----TA-A-----T----C-GAA------AGGA-T---G--T..C.----AC-A----C---A- 6243
SMM.LR.89.SIVsmLIB1         G--A----G-A--TC----................................................A--A--C--T--A---A--C-AGA--AAT-T---------A--A-----A---TCAC-A-----GC-AA-CA----C--T...T----A------G---A- 6152
SMM.SL.92.SIVsmSL92A        T--A---GG-GGG-C-G-T.................................ACAACTACAG-TG--A-TA---AGG-------T--TTT-A---TCCCT--TA-GGAAT-A-----C--A----TC-----GA-AA-T----C---T...C----AA-A--G-----A- 6181
SMM.SL.92.SL92B             G-------G-G--GCT--T...........................ACTACTAGCTCACCA--T--TA-TAGCC-CTT-A-TG-TG-TAGCC--T--GG--AG--A--C--T--C--C----TG-------C-AAGA-CA-C-------...T-T---A-A--G--G--A- 6535
SMM.US.04.G078              ---A---GGGA---C----..................ACAACACAAAAAAACAACAACAACAGGCC---------A-G--AA-T-AG--A-GC---T-GC------AG-T--G-C---T-G----C-----G--AGGA-CA---G--T..----A---A--G--G--A- 6377
SMM.US.04.G932              -C-----GGGGT-TC----.................................ACAAAAAAACC-----AT-GT-G--C--A-TT-A-G-A--G-----CT-AA-CCAGAAT-A-----A--AGA-C--------AAAA-T----------..----A------G---A- 6351
SMM.US.04.M919              -C-A---GG-AT--C-C--....................................ACACTCT-----A--A-----CA---T--A---A--A--AAT-C-AC---AG-TA-A-----A-G----------C--AAAATC--------T..-------C-A-GG--G--A- 6352
SMM.US.04.M922              G--------GAATG-----............ACAGCAAGGAATGCAACAGCAACAACACCAG-----A--AA----C--G----AGT-AT-A---T--C-CC-A--G-T--A-------A-GACA-----G---AAA-CA----C---...C-------A--G--G--A- 6379
SMM.US.04.M923              T--A----G-A---C-G--........................TCAACAACAATACAAAAAT-----A-G---------C----A-T-CAGG---T-GC--A-A-GG-T--A-----C-G--C-C---------AAAA-T----G---T...-T---T--------G---- 6360
SMM.US.04.M926              ---A----GGCATGGC---...........................ACCGCAACAACCCCA---------G---A-T--A----A----A-AAG-T--C-AAT-TTGA---A-----T--ATCA------G--AAAA-TA-C-----T..------C-A--------A- 6363
SMM.US.04.M934              ---A---GG-A-CCC-T--...................................................C--------G--C-A--CATGCA-----C-----CTGA-A-A--------ATCA---------AAGA-T----G---T..-------A-A-----G--A- 6336
SMM.US.04.M935              ---A---GGGAAT-C----.......................................ACA-----A--A--T--T--G----AG--A--G---T--CGCC-A-A--T--A-------G-GACC-----G----AA-CA-------T..-------A--G--G--A- 6352
SMM.US.04.M940              GC-A----GGGAC-C-GA-ACAGGGAACACAACAGCAACAACACCAACTACGACACCATCAC-----A--AG--AC---A--GGA-ACA--A--T-G-GGAAT--GGAAA-A-GG-CA-A-GA-A----C---AGGA-T--C-----T..-------A-A--G-----A- 6395
SMM.US.04.M946              ---A----GGGGCCC-T--..............................ACAACAACCCCAT--T--A--GTG-A---GA--G-A-A-A--CG--T-CC------TGA-C-A--------ATCA------G--AGGA-T--C-----T..-------A-A----G--A- 6358
SMM.US.04.M947              C--A---GG-A--CG----..................ACAATGAAACCAACAACAACAGAAG--T--A--GGG---G-----C-A---A--C--AT--C--C-A-AG-TT-A-----A-G-GA-C-A-------AGTTTA---------..----A------------- 6366
SMM.US.04.M949              -C-A---GGGA--GCT-T-..............................ACAACACCTACAC-C---A--AGT------AA--AT-GGA--A--A--C-G-ATCGGGACC-AG-G--T-A----A----C---AGGA-T----CA--T...G-T---A-A--G-----A- 6353
SMM.US.04.M950              ---A---GGGA-CCC-T--.......................................ACA------A--G-C-GT---A----AG-----A--C--GC--C--TTGA-C-A-----A--AGAC---------AAGA-T--C-----T..-------A-A-----G--A- 6348
SMM.US.04.M951              -C-A---GC-A---CT--T..........................................C-GG--A--A---TTG-TA---GT-GTCT-A-CA---C-CCG--G-CAC-A-CA-GA-A-GAAA----C---AGGA-T--C-GC--T...-AG---C-A--------A- 6350
SMM.US.04.M952              ---A---GG-A-CCC-T--.......................................ACA---------AA---GC--A----A---C--A-----GT--C--TTGA-A-A--------ATCA------G--AGGA-T--C-----T..---------A-----G--A- 6349
SMM.US.05.D215              C------GG-GT-CAGG--.......................................ACA---C--A--G--T----------A---A--A--AAGGCTA----A-CA--A-C---T-AC---G-A---G--CAAA-CA---GC--T...C-T---A-A-----G--A- 6350
SMM.US.06.FTq               -C-A--GGGCAACTCCGAC..........................................------A--GT----C-TAGCG-C-A-A-TA--A--G--AA--GTGAAA-A-----C---GACC-C--CT-CAAAGGTA-C--C--T..-----GA-A-G-AGT--AG 6372
SMM.US.11.SIVsmE660_FL10    T-------GGAACGC-GGG..............................ACAACAACAACA-----A--A-------G--G-A---C--AGT-TGGC--A-A-TG-TA-A-----A-G--A-C----C---AAAA-TA---G--T...G------------G---- 7090
SMM.US.86.CFU212            -C-A---GG-A--GC-G--.................................CCACCAACA---AGA--CT--AG---A--T-A---A--AA-C--CC----CT-GAA--A-----T--CGAAC-C---G--AAA-TTA-C--C---..-----A-------C---A- 6345
SMM.US.x.F236_H4            T------GG-AACGC-GGG....................................ACAACA---------AT------A--G-A---C--AGT-T-GC--A-A-TG-TA-A-----A-G--A-C----C---AAAA-TA---G--T...G------------G---- 7046
SMM.US.x.H9                 T--A---GG-GS-CC-G--....................................CCAACA-----CA-A--T-----A--C-AC--T--C--AT--T--C-A-GG-T--A-----C-G-GA-C--------AGAAGTA-------T...--------------G---- 6567
SMM.US.x.PBJA               T--A---GG-A--CC-G--.....................CCAACAACAACACAAACAACA------CA-G--T-----A--C-A---T--C--AT--C--C-A-GG-T--A-----C-G-GA-C--------AAAATTA-------T...--------------G---- 6875
SMM.US.x.PGM53              G--A-----GAATGA--T-...........................CCAACAACAACAACA------A--A---A--GCC---AAG---A-G--AT--C--C-A-AG-TA-A-----A-G-GA-C--------AG-A-CA-------T...---------A--G--G--A- 7027
SMM.US.x.SME543             T------GG-AG-GC-GAG..............................ACAACAACAACAG-G-A-T--A-----T--A----A---A--A---T--C-CC-A-GG-TA-A-----AGG-GA------C---AAAA-TA---G--T...G------------G---- 7103
SMM.US.x.pE660.CG7G         T-------G-AACGC-GGG.................................ACAACAACA---T--A--A--------G--G-A---C--AGT-T-GC--A-A-TG-TA-A-----A-G--A-C----C---AAAA-TA---G--T...G------------G---- 7086
STM.US.89.STM_37_16         -------GGGAA--C-GT-..........................................-------TGA--C-----G--G-AG-----G--A----GCC---GT-A--A-----A--A--C--C---G--AGTA-CA-C--C---...------------G----AT 6735
```

```
                                 V1 loop end┘   └V2 loop start                                                                                                      V2 loop end┘
MAC.US.x.239              AAATGATAAGCTGTAAATTCAACATGACAGGGTTAAAAAGAGACAAGAAAAAAGAGTACAATGAAACTTGGTACTCTGCAGATTTGGTATGTGAA.....................CAAGGGAATAACACTGGTAATGAAAGTAGATGTTACATGAACCACTGTAACACT 7242
Env                       Q__M__I__S__C__K__F__N__M__T__G__L__K__R__D__K__K__K__E__Y__N__E__T__W__Y__S__A__D__L__V__C__E__.__.__.__.__.__.__.__Q__G__N__N__T__G__N__E__S__R__C__Y__M__N__H__C__N__T_
A.CI.88.UC2               GG---G-T-AT-----G-----T-----G--AC-GG--CA---T----T---G-G---T-C---C--A--------A-AT---G----T-----C.....................AGCACA--C--G--A--C-CAA-C-CG-C-----------GA--T--------A 7272
A.DE.x.BEN                -G-----GCAA---G-G-----T----AG------G--CA---T-------GGAG---T--G--C--A-----T-TA-A----G----T-----C.............................-ACACA-CA-CTG-C-C-----------GA--T--C-----A 7324
A.DE.x.PEI2_KR_KRCG       -GG-AG-CGAT---CGG-----T--------AC--C--CT---T---CC-C-GC-A--T-G------A--------AAAG---G-A--T-----C.........................-CA-CT-A---G-CCAGCC-C-A---------------T--C-----A 7265
A.GH.x.GH1                -G--AG-C-AT---C-G---T-T--------A---G-G-----T---------C----T--------A--------AAA----G-A--T------.....................TC-AAT----C--AA-A-GGGA---AC-----------------T--C-----A 6741
A.GM.87.D194              -GG--G-TGAG---G-G-----T-----G---C--G--CA---T-------GGA----T-----CG-A--------AAG----G----T-----C..........................--G--AAACGGAAC-G-C-C-----------GA--T--C-----A 6732
A.GM.x.ISY_SBL_6669_85    -C---G-CGAG---C-G-----T--------AC--G-GTT---T---------C----T-G------C--------AAA----G----T------.......................TCAG------GCACAG-CCG--AA-----------------T--C-----A 6738
A.GM.x.MCN13              -GG--G-C-AT---C-------T--------A---G-G-----T--------GC-A--T-G---G--A--------GAAG---G-A--T------..........................GGA-A---C-CCAC-GA--C----------------T--C-----A 6754
A.GM.x.MCR35              -GG--G-C-AT---C-------T--------A---G-G-----T--------GC----T----GG-A--------GAAG---G-A--T------..........................GGA-A---C-CCAC----C-C----------------T--C-----A 6754
A.GW.86.FG_clone_NIHZ     -------CGA----C-G--T-GT--------A---G-G----------G----C----T-C----G-A--------AAA----G----T-----C...........................------CTCA-G-C-G--C-AG--------------T--C-----A 6729
A.GW.87.CAM2CG            -G--AG-C-AG---C-C-----T--------A---G-------T---------C-------------A--------AAGC---G----T-----T..................AATAGCACAG--C-G--CACA-----G-CC-CG--------------T--C-----A 7327
A.GW.x.MDS                -G---G-C-AG---C-------T--------A---C-G-T---T---------C----T--C-----A--------AAGC---G----T----G-.....................G--AT---------CACA--CC----C--G--C-------G---T--C-----A 6798
A.IN.07.NNVA              -G--AG-C-TA---C-G-----T--------A----T------T--CC----AG---T--------A--------AAGC---G----T------..................CCAG-TA-C-CC-CT-GCAC-CGCA-C-A---C-CC-CAGGCC-G---ATGCG---G 7319
A.IN.95.CRIK_147          -GG-AG-C-AT---G-G-----T--------A---GTG-----T--ACC---GACA--T--------A--------ACGT---G---AT------......................CCAG---G---CACA--CTCG--A-AG--------------T--C-----A 7066
A.JP.08.NMC786_clone_41   -G---G-C-A----C-G-----T--------A---G-G--G-------C----CG---T--------A--------AAG----G-A--T------..................CCAA-TAATC-G-C-G-A-A-GGGA--GAA-A---------------T--C-----A 7323
A.PT.x.ALI                -G---G-C-AT---CGG--------------A---C-G-----T---CC----C----T--------A--------AAA----G----T------.....................-C-TTT--C-C---CACA--CC-G-CC--G--------------T--C-----A 7305
A.SN.85.ROD               ---C---C-AT--CC-G-----T--------A---G-------T---------C----T--------A--------AAA----G----T-----G......................ACA-----T-GCACA---C-G-CCCAG--------------T--C-----A 6767
A.SN.86.ST_JSP4_27        -G---G-CGA----C-G-----T--------A---G-G--G-T--------CTA--T--------A--------AAA----G-A--C------......................TCA---G----CAAG--A--G-AA-C-------------------C-----A 6751
B.CI.88.UC1               -C-CAG-TGA----T-C--T-----------AC---G------TG--------C-A--T--G--C--C------GAAAA------A-AG--CA-C..........................GG--ACA-C-CCAGC-CC-T-----------GAACT-----T--C 7296
B.CI.x.20_56              -TG--G-T-A-----CC--T-----------AC---G-CA---TG---G----C-A--T-----C--C-----AGGAG------A-AG--C--G......................................-GT-CC--------------GAACT-----T--A 7270
B.CI.x.EHO                --------GA----C-------A--------A-----------TG--TC----C-A--T--G--C--C-----TAAACA------A--G------......................AA-GGG-CA-GGA----------C-A------T--A--AAC------T--A 7286
B.GH.86.D205_ALT          ----A--GCAA-----T--T-GT-----G--AC---G------TG--CT----C-A--T--A--C--C--------A-A------A-AG---A-T.........................----C--GGAAGT--ACC--C-----C--T--A-GAAC---C-----A 7291
B.JP.01.IMCJ_KR020_1      -C---G---CA---CG------T--------A------GT---TG-ACCG---AC---T-----C--C-----TG-A-A----G-ACAG---A-T...................................-GG--C-----CC-T--A--AAC---------A 6414
G.CI.92.Abt96             CC---G-TG----C-----T-----------C------G--TC---G---G--A--------G--------T--ACA-------CA---Y---.....................--GTCC-----G-G--AG-----G--------C--T----GG--------T--A 6655
AB.CM.03.03CM_510_03      -G---G-T-CA-----C----GT-----------G--CA---T--------GAG---T-----G--A-----T--ACAG---G----T-----C..................AAAA-CAATTCA-G-GAAACATTAA-C-AA-C-----------GA--T--------A 6403
H2_01_AB.CI.90.7312A      -G---G-CGAT---C-G--------------A---G-G-----T----C----C-A--T-G------A--------AAA----G-A--T-----G.....................TC-AAT--CGC--G--A-GGGAG-GAC-----C--------T--T--C-----A 7306
H2_01_AB.JP.04.NMC307_20  -G--AG-CCAG---C-G--G-----------A---G----G--T--------GCG---T--------A-----T--AAA----G-A--T-----G.....................ACCAAT---GGGC-AAA-G-GAC---C-----C--T--------T--C-----A 7288
H2_01_AB.JP.07.NMC716_01  -G---G-CCAG---C-G--TG------T---A---G-G--G--T--------CG---T--------A--------AAA----G----G----G-..........................G-GGC--A---C--CAT---C-----C--------T--T--C-----A 7281
H2_01_AB.JP.08.NMC842_10  -GG-AG-CCAG---C-G--TG----------A---C-G-----T--------CGA--T-----G--A-----T--AAA----G-A--T-----G.........ACCAATAATAAGAC-AAT-----G--AAA-G-GAC--AC-----C------------T--C-----A 7311
U.CI.07.07IC_TNP3         C----G--G-----C-G-----T---------C---G------TC-AG-G--GC----T-G---G--C-----T---CGT--CA-A--C------..................CAGAGTAAT----CTGA-AA---GAGC--C-----C--T-----T--T-----T--C 6760
U.FR.96.12034             C-T--G-G--T----G---T--T--------C--------G--T------C-------T--A-----A-----T---CAT--CA-A--T------......................CA-----CT-G--AA-G--------A-------------------C-----C 6784
MAC.US.x.17EC1            -------------------------------------T---------------------------C--------------------------------.....................------------------------------------------------------ 7242
MAC.US.x.251_1A11         --------G-------------------------------------------------------------------------------------.....................----------G---------------------------T----------- 7238
MAC.US.x.251_32H_PJ5      --------G------------------------------C--G-------------------A---------T-----.....................----------G----A-------------C-------T----------- 7240
MAC.US.x.251_BK28         -------------------C-------------------C--G-------------------A---------T-----.....................----------G----A---------C----C-------T----------- 7224
MAC.US.x.MM142_IVMXX      C----------------------------------------G--------------------------T-----.....................----------G--------------------------T------T--- 6737
MNE.US.82.MNE_8           C---------------------------------------G-G-------------------------T-----.....................-----T---G----AA---------------------T----------- 6710
MNE.US.x.MNE027           C------------------------------------------G------------------------T-----.....................-----T---G----AAG--------------------T----------- 6713
SMM.CI.79.SIVsmCI2        CCC-A--TG----C-----T--T--------AC-T-----G-----A--G-----A--T-----G--A-----T--ACAG-----A--C------......................C-A---GGT-A---C-G------C--------T--------T--------A 6389
SMM.LR.89.SIVsmLIB1       C-------G-A----GG--------------AC----G--G-----------G---------G--A-----T---AG---------G------......................CAA-----T-A--AG---------------C--T----G---T--C--T--A 6298
SMM.SL.92.SIVsmSL92A      C-------G-A-----G--T-GT-----T---C---GG--------------GC-A--T-----G--C-----TG-CAGG--A-T--G------......................-AAG-C---CAGAC-GGCA-C---C-C-----T----GG--T--------C 6327
SMM.SL.92.SL92B           C-------G-T---C----T-----------AC-----------C------GGC-------------C--------CAG---C-A--C-----G.........................C-AGGAGGAAA-G-AAGC-----------------T--T--C--T--A 6678
SMM.US.04.G078            C------------------T--T-------------------T--A--G-----A--T--------A-----T--AAG---------------......................CAA--G-G--A---A-G--------AC---------C-T--T--------C 6523
SMM.US.04.G932            C----G--------------------------A----GG-----T-G--GG--G-----T-------C---------A----C--A--T------....................GC-AAT-----T----CC---------------C--------T--T--------C 6500
SMM.US.04.M919            -----G-------------T---------------C-GTG-----A--G-----A--T--C-----A-----T--CA----C---A--------CAAACTACTGAAACTAATGAAACTAAT--A--TGGAA-----A------AT-----T-----T--T--C--T--C 6522
SMM.US.04.M922            -----G---A---------T--T---T-------G-----G-----A--G-----A--T--------A-----T--AAG------AA--------......................CAA---CTGA---A----------AC-----------T--T--------C 6525
SMM.US.04.M923            CC---G-------CC----T--------------------------A--G-----A--T--------A-----T--AAG------A--T------....................---AAC--C-CTGACA----AACC-----------T--------T--------C 6509
SMM.US.04.M926            CC-------------G---T-----------TC----------T-----------A--------G--C-----T--------C-----C-----G............CAAGGTAATGCTAAT--A-CA-G-AA-...-------AT-----T-----T--T--------- 6518
SMM.US.04.M934            C----G-------------T------------C----------T---G---G--A--------G--------T---A----------T-----G......................CAA---GT-----G-----------------T-----T--T--------- 6482
SMM.US.04.M935            C----G-------------T--------------G-----------A--G--G--A--T--------A-----T--AAG------A---------......................CCAG--GCT-AA--G----------AC-----------T--T--------C 6498
SMM.US.04.M940            -----G--CAA---C----T--T-------------C-------------G--G-----------------------CT----C-----T------......................C-A---GCT----AGGG-A----C-----C--T-----T--T--------C 6541
SMM.US.04.M946            C----G-------------T------------C----------T---GG--G--A--------G--------T---A----------T-----G..............CGAAATGGT-A----GGT----AAG-----------------T-----T--T-----T--- 6513
SMM.US.04.M947            CC--------T--------T--------------------G-----A--G--G--A--T--------A-----T--AAGG--A-A--T------......................CAA-------A--AATC---------A------T--------T--C------ 6512
SMM.US.04.M949            C----G--CAA--------T--T-----------------G-------G--G---------------C------G-C-----C-----T------......................CAA---GTT---AA---------C--------T-----T--T--C------ 6499
SMM.US.04.M950            C----G-------------T------------C----------T-------G--A--------G--------T--CA-----------T-----G............CAAAATAATGGT-AA---GT--A-CA-----------------T-----T--T--------- 6506
SMM.US.04.M951            C----G--CAA--------T--T---------------G---G-------GCGG-------------A--------C-----C-----T--C---......................-----T---GCT----A---------C--------T-----T--T--C-----C 6499
SMM.US.04.M952            C----G-------------T------------C----------T---G---G--A--------G--------T---A----------T-----G.....................---AAT---GGT--A-CA------------------T-----T--T--------- 6498
SMM.US.05.D215            C----G--TCA--C-G---T------T-C---C----G----------C------A--------G--C---------AGG---A-A----------......................------GC-G-A-A--A--CA--G---------C--T--------T-----T--A 6499
SMM.US.06.FTq             --T--G--TCT--C-----T--------------G--------A--G--G-----------G--C--------AAGG--------C------..................CGAG-C--A---G-A-TA-CA---------------C--------T--T--C--T--- 6524
SMM.US.11.SIVsmE660_FL10  CC------G-T--------T--------------------G-----A-GG-T---A--T--------A-----T--AAG------AA-C-----G......................CA-TCAGCG-A--AA-G---G----A----------C-T--T--------C 7236
SMM.US.86.CFU212          C-------G-A-----------T--------A--G-------G--G------G--------------A---------AGT--------C-----G......................-----TGG-----ACACA---------C--------T-----T--T-----T--A 6494
SMM.US.x.F236_H4          CC------G-T--------T----------------T--G-----A--G-----A--T--------A-----T--AAG------AA-C-----G......................CA-TCAGCG-A--AA-G---G----A----------C-T--T--------C 7192
SMM.US.x.H9               CC---G----T--------T-------------------------A--G-G---A--T--------A-----T--AAG------A--G------..................CAAA-TA-C---GGA-A--AA-C---C----A---C--T---------T--------C 6719
SMM.US.x.PBJA             CC---G----T--------T-------------------------A--G-G---A--T--------A-----T--AAG------A--T------......................CAA----G--A--AA-C---C----A---C--T---------T--------C 7021
SMM.US.x.PGM53            C----G---------G---T-------------------AG-----A--G-----A--T--------A-----T--AAG----A-A---------......................---A-T-C--CT-G-A-A-----------A-----------T--T--------C 7176
SMM.US.x.SME543           CC------G-T--------T-----------------G--G-----A--G-T---A--T--------A-----T--AAG------AA-C-----G......................CA-CCAGCA-A---A-G---G----A----------C-G--T--------C 7249
SMM.US.x.pE660.CG7G       CC------G-T--------T--------------------G-----A-GG-T---A--T--------A-----T--AAG------AA-C-----G......................CA-TCAGCG-A--AA-G---G----A----------C-T--T--------C 7232
STM.US.89.STM_37_16       C-C-AG--G-T-----------T--------AC----------T--A--G-G--------------G---------AGC---C-AA-C------......................CAA---GT------CG-A--G--C-----C-------GA--T--C--T--C 6881
```

```
MAC.US.x.239                TCTGTTATCCAAGAGTCTTGTGACAAACATTATTGGGATGCTATTAGATTTAGGTATTGTGCACCTCCAGGTTATGCTTTGCTTAGATGTAATGACACAAATTATTCAGGCTTTATGCCTAAATGTTCTAAGGTGGTGGTCTCTTCATGCACAAGGATGATGGAGACACA 7412
Env                         _S__V__I__Q__E__S__C__D__K__H__Y__W__D__A__I__R__F__R__Y__C__A__P__P__G__Y__A__L__L__R__C__N__D__T__N__Y__S__G__F__M__P__K__C__S__K__V__V__V__S__S__C__T__R__M__M__E__T__Q
A.CI.88.UC2                 --A--C---A-G--A--A-----T---------------T----G-AG-----A--C--CA----A--G---------C--A--A-----C-----T--C--C--C--G------GCA-----T--CC----A--A--A-CTG----------T-----------A--G-- 7442
A.DE.x.BEN                  --AA-C---A-------A-----T--G--C--------------G--G-----A--C--------A--G--C-T---CC-AT-A-----------T--C--C------------GAA-----G--CA----A--A--A-CTG-----------------------A--G-- 7494
A.DE.x.PEI2_KR_KRCG         --A--C---AC------A-----T--G--C--------------G--G-----A--C--------A--G----TATGC-----A-----C-----T--C---------------GA---C--G---C----A--A--A-CTG--A----------A---------A--G-- 7435
A.GH.x.GH1                  --A--C---AC------A-----T--G--C--------------A-AG-----A--C--------A--G---------CC-A--A-----C-----T--C--------------CGA---A--G--------A--A--A-CT---A--------------------A--G-- 6911
A.GM.87.D194                --A--C---A-------A--------G--C--------------G-AG-----A--C--------A--G----T---CC-A--A-----C-----T--C--C------------GAA-----G--C-----A--A--A-CTG-----------------------A--G-- 6902
A.GM.x.ISY_SBL_6669_85      --A--C---AC------A--------G--C--------------G--------A--C--------A--G----T--TC-----A--G--C-----T--C-----C---------GA---C--T--C-----A--A--A-CT---A----T-----A---------A--G-- 6908
A.GM.x.MCN13                --G--C---AC------A--------G--C--------------G--G-----A--C--------A--------T---CC-A--A-----C-----T--C---------------GC---C--T--C--------A--A-CTG--A-------C--A---------A--G-- 6924
A.GM.x.MCR35                --G--C--TAC------A--------G--C--------------G--G-----A--C--------A--G----T---CC-A--A-----C-----T--C---------------GC---C--T--C--------A--A-CTG--A-------C--A---------A--G-- 6924
A.GW.86.FG_clone_NIHZ       --A--C---AC------A-----T--G--C--------------G--G-----A--C--------A--G----T---C--A--A-----C-----T--C---------------GCA-----C--C-----A--A--A-CTG--A-------C--A---------A--G-- 6899
A.GW.87.CAM2CG              --A--C---AC------A-----T--G--C--------------G--G-----A--C--------A--G----T---CA-A--A-----C-----T--C--G--C---------GCA--C--C-----------A--A-CT---A-------C--A---------A--G-- 7497
A.GW.x.MDS                  --A--C---A-------A-----T--G--C---------A----G--G-----A--C--------A--G----T---C-----A-----C--------C----------G---CA---C--T--------A--A--A-CT---A----T--C--------------A----- 6968
A.IN.07.NNVA                -GGTAC--GAG---T--A-----T--G-----------A---G--G-----A-----C-----A--G---------C-----A-----------T--C-----C---------GA---C--T---A----A--A--A-CTG----------C-------------A--G-- 7489
A.IN.95.CRIK_147            --A--C---AC-----A-----T--G--C---------A---A--G-----A-----------A--G---------C-----A-----C-----T--C-----C---------GCA--C--T--------A--A--A-CT--CA-T-----C-------------A--G-- 7236
A.JP.08.NMC786_clone_41     --A--C---A---A--A-----T--G--C---------A---A--------A-----------A--G---------CC-A--A-----C-----T--C---------------CGC---A--T--C--------A--A--T-----------------------------G-- 7493
A.PT.x.ALI                  --A--C---AC-----A-----T--G--C--------------A--G-----A--C--------A--T-----C--CC-A--A-----CG----T-TC---------------GCA--C--T--C-----A--A--A-CTG--A-----------------------G-- 7475
A.SN.85.ROD                 --A--C---AC---A--A--------G--C--------------A--G-----A--C--------A--G---------CC-AT-A-----------T--C---------------GCA--C--C--------A--A--A-CT---A-------C-----------A--G-- 6937
A.SN.86.ST_JSP4_27          --A--C---AC------A--------G--C---------A---G--G-----A-----------A--G----T---CC----A-----C-----T--C---------------GA---C--T-----------A--A-CTG--A----T--------------A--G-- 6921
B.CI.88.UC1                 --A--G-----------A-----T--------C-----CAGCT-A-----------C-----T--C-----A------C----A------------------C--C-------------------A------AG---A--A--A--G--C--T--------A-------------- 7466
B.CI.x.20_56                --A--G-----------A-----T--------C-----CAGCT-A-----------C-----T--C-----A-----A-A--G-----------T--C--C-----------C----AC--C---AG---A--A--A--G--C--T--------A-------------- 7440
B.CI.x.EHO                  --AA---------A--A--------------------AG-T-A----------------T--C-----A-T---------A-----C-----T--T--A-----------C-----C--C---AG-----A--A--A---CTG-A-...--A--------A----- 7453
B.GH.86.D205_ALT            A-AA-------------A-------------------CAGCT-A--G-------------T--C--G--G-T-TT-C-A--A-----------T--C--C-----------C-----C--C--CAG-----A--A-CG--C--C--------A--------A----- 7461
B.JP.01.IMCJ_KR020_1        --AA-------------A-------------------CCAAT-A----------------C--C-----C-T------TA-G-----C-----TT-T--C----------------C--C---AGC-----A--A--A--C--C--T-----A--------A----- 6584
G.CI.92.Abt96               --A--C--T--G--A--A-----T-----------------------A--------A--------T--G-----G-T---------A-----------T--T--C--------A---GCA-----T--C-----A--A----------A----T-----A--------A----- 6825
AB.CM.03.03CM_510_03        --A--C---A----A--A-----T--G--C---------A----G--------A--C---A----A--G---------C--A--A-----C-----T--C--C-------------CAA-----T--C-----A--A--A-CTG---------------A-----A--G-- 6573
H2_01_AB.CI.90.7312A        --A--C---AC---A--A--C-----G--C--C-----------A--G-----A--C--------A--G----T----C----A---------------C--C-----------------C--C---AG-----A--A--G--C--T--------A--------A----- 7476
H2_01_AB.JP.04.NMC307_20    --A--C---AC------A--C-----G--C--------------A--G-----A--C--------A--G--C-T----------A---------------CG-C-----------C-C---C--C---AGC--A--A--A--A--C--T--T-----A-------------- 7458
H2_01_AB.JP.07.NMC716_01    --A--C---AC------A--C-----G--C--------------A--G-----A--C--------A--G--C-T----------A---------------C-GC-----------C-C---C--C---AGC--A--A--A--A--C--T--T-----A-------------- 7451
H2_01_AB.JP.08.NMC842_10    --A--C---AC------A--C-----G--C--------------G--G-----A--C--------A--G--C-T----------A---------------C-GC-----------C-----C--C---AGC--A--A--A--A--C--T--T-----A-------------- 7481
U.CI.07.07IC_TNP3           --------T-------C--------------------A--A-----------C-----C--G-----G-T---------A--G---------T--TG-------------C-GC-----G--C---------A--A------A----T--------------A----- 6930
U.FR.96.12034               ------------------------------------C--A--------------------G-----C------C--T-A-----------T--T-----C-----A-----------T--C-----A--A--A------A-----------A----G--GA----- 6954
MAC.US.x.17EC1              ---------A------------------------------------------------------------------------------------------------------------------------------------------------------------ 7412
MAC.US.x.251_1A11           ---------------------------------------------------------------------------------------------------------------------------------------------------------------------- 7408
MAC.US.x.251_32H_PJ5        ---A--------------------G-----------A--------------------------------------------------------------------------------------------------------------------------------- 7410
MAC.US.x.251_BK28           ---------------A-----------------------A------------------------------------------------------------------------------------------------------------------------------ 7394
MAC.US.x.MM142_IVMXX        ----------------G---------G--------------------G---A--------------------------------------------------------------C---------A------------------------------------------ 6907
MNE.US.82.MNE_8             --------T----A----------------------------------------------------------------------------------------------C--------------------------A------------- 6880
MNE.US.x.MNE027             --------T----A------------------------------------------------------------------------G-------------C--------------------------A------------- 6883
SMM.CI.79.SIVsmCI2          --A-----A--G--A--A--------------------C--------------------A---------------A----------T-----------------GAA-----T--CA-------A---------A----T--T--A--------A----- 6559
SMM.LR.89.SIVsmLIB1         AG---C--T-----A--C-------------------------A-------A--C--------A----A----------------------T-----------------GCA-----T--C--C-----------G----------------------------- 6468
SMM.SL.92.SIVsmSL92A        AG-A-C--T--G--A-----------------------------G------C--A-----------------------------------T-----------------GCA-----T--C--------A--A--T----------T--A--------A----- 6497
SMM.SL.92.SL92B             AG------A--------A--------G--C----------C--C----A-----C--------A----C----------------------T-----------------GCA-----T---AG-----A--A--A--A-----T-------------A----- 6848
SMM.US.04.G078              AG------T--G--A--C--------G---------------------------A-----------G-----------------------TT----------------GCT-----C-----------A-----T--------------A----------G-- 6693
SMM.US.04.G932              --------T--------------T-------------------------G-----A----------------C-----C---------C-----TT------------------------C-----------A----------------------------G-- 6670
SMM.US.04.M919              AG------T-----A--C--------G---------------------------A-----------G-----------------G--------T-----------------GCT-----T-----------A-----T--------------A--------A----- 6692
SMM.US.04.M922              AG------------A--C--------G--C------------------------A-----------G-----------------G--------TT----------------TCT-----C---------A--A-----T--------------A--------A----- 6695
SMM.US.04.M923              AG------T-----A-----------G---------------------------A-----------G-----------------G--------TT----------------GCT-----C---AG------A--A--T--------------A--------A----- 6679
SMM.US.04.M926              --------T-----A--C----------------------------------A--C--------------C-------------G--------------GC--------T------------T--C--------A--A-----A------------------------------ 6688
SMM.US.04.M934              --------T-----A--C----------------------------------A-----------------------------------T--------------------T----T--C-----------A-----A------------------------------ 6652
SMM.US.04.M935              AG------T-----A--C--------G--C------------------------A-----------G-----------------G--------TT--G-------------TCT-----C---------A--A-----T--------------A--------A--G-- 6668
SMM.US.04.M940              -----G--------A--C---------------------------------------G--------C--------------------T-----C---A--------GAA-----G--C---------A--A-----C--------T--A-----------G-- 6711
SMM.US.04.M946              --------T-----A--C----------------------------------A-----------------------------------T-----------------------A----C--C-----------A-----A------------------------------ 6683
SMM.US.04.M947              AG------T--G--A--C----------------------------------A-------T--G-----------------G--------T-------------T---GCT-----C-----------A--A--T--------------A--------A----- 6682
SMM.US.04.M949              -----G--T-----A--C------------------------C-A-----------------------C----------------------T-----C---A--------GAA---G---C---------A--A---------------T-------------G-- 6669
SMM.US.04.M950              --------T-----A--C----------------------------------A-----------------------------------------C----------------------T--C--------A--A-----A------------------------------ 6676
SMM.US.04.M951              -----G-----G--A--C----------------------------------A--------------C----------------------T-----A--A--------GAA--------C---------A--A-----C--------T-------------G-- 6669
SMM.US.04.M952              --------T-----A--C---------------------------G-----A-------------------------------------------C----------------------T--C--------A--A-----A------------------------------ 6668
SMM.US.05.D215              AG---C--T-G--------------G--C------------G---------------------A-----A-----------C----------T--T-----------A---GCA-----T---AG---A--A--A--T-----T--T-----A--------A----- 6669
SMM.US.06.FTq               -----A-----G--A-----------G----------------------A-------A--------A--T--C-----------------------T-----------------GCA-----T--CAG-----A-----------------------------A--T-- 6694
SMM.US.11.SIVsmE660_FL10    AG------T--G--A--C--------G---------------------------A-----------G-----------------G--------TT----------------GCT-----C-----------A-----T--------------A--------A--G-- 7406
SMM.US.86.CFU212            AG---A-----------C----------------------------------A--C--------A-----C-----------------------TT-----------T------------T-----------A--A-----------T--------------A----- 6664
SMM.US.x.F236_H4            AG------T-----A--C--------G---------------------------A--C-------G-----------------G--------TT--------T------GCT-----C-----------A-----T--------------A----------G-- 7362
SMM.US.x.H9                 AG------T-----A--C--------G--------------------Y--A-----------G-----------------G--------T--C---------------GCT-----C---A-------A--A--T--------------A--------A----- 6889
SMM.US.x.PBJA               AG------T-----A--C----------------------------------A-----------G-----------------G--------TT----------------GCT-----C---A-------A--A--T--------------A--------A----- 7191
SMM.US.x.PGM53              AG------T-----A--C--------G---------------------------A--C-------G-----------------G--------T-----------------GCT-----C-----------A-----T--------------A----------G-- 7346
SMM.US.x.SME543             AG------T--G--A--C--------G---------------------------A--C-------G-----------------G--------TT----------------GCT-----------------A-----T--------------A----------G-- 7419
SMM.US.x.pE660.CG7G         AG------T--G--A--C--------G---------------------------A-----------G-----------------G--------TT----------------GCT-----C--C--------A-----T--------------A-------------- 7402
STM.US.89.STM_37_16         AG---A--T-----A--C---------------------CT-A--G----------------A--T--------------C-----C-----T---------A--------GCA-----T--C---------A--A-----C--G--T--T--A-------------- 7051
```

V3 loop start
MAC.US.x.239 GACTTCTACTTGGTTTGGCTTTAATGGAACTAGAGCAGAAAATAGAACTTATATTTACTGGCATGGTAGG...GATAATAGGACTATAATTAGTTTAAATAAGTATTATAATCTAACAATGAAATGTAGAAGACCAGGAAATAAGACAGTTTTACCAGTCACCATTATGT 7579
Env T S T W F G F N G T R A E N R T Y I Y W H G R D N R T I I S L N K Y Y N L T M K C R R P G N K T V L P V T I M
A.CI.88.UC2 7609
A.DE.x.BEN 7661
A.DE.x.PEI2_KR_KRCG 7602
A.GH.x.GH1 7078
A.GM.87.D194 7069
A.GM.x.ISY_SBL_6669_85 7075
A.GM.x.MCN13 7091
A.GM.x.MCR35 7091
A.GW.86.FG_clone_NIHZ 7066
A.GW.87.CAM2CG 7664
A.GW.x.MDS 7135
A.IN.07.NNVA 7656
A.IN.95.CRIK_147 7403
A.JP.08.NMC786_clone_41 7660
A.PT.x.ALI 7642
A.SN.85.ROD 7104
A.SN.86.ST_JSP4_27 7088
B.CI.88.UC1 7633
B.CI.x.20_56 7607
B.CI.x.EHO 7620
B.GH.86.D205_ALT 7628
B.JP.01.IMCJ_KR020_1 6751
G.CI.92.Abt96 6992
AB.CM.03.03CM_510_03 6740
H2_01_AB.CI.90.7312A 7643
H2_01_AB.JP.04.NMC307_20 7625
H2_01_AB.JP.07.NMC716_01 7618
H2_01_AB.JP.08.NMC842_10 7648
U.CI.07.07IC_TNP3 7097
U.FR.96.12034 7121
MAC.US.x.17EC1 7579
MAC.US.x.251_1A11 7575
MAC.US.x.251_32H_PJ5 7577
MAC.US.x.251_BK28 7561
MAC.US.x.MM142_IVMXX 7074
MNE.US.82.MNE_8 7047
MNE.US.x.MNE027 7050
SMM.CI.79.SIVsmCI2 6726
SMM.LR.89.SIVsmLIB1 6635
SMM.SL.92.SIVsmSL92A 6664
SMM.SL.92.SL92B 7015
SMM.US.04.G078 6860
SMM.US.04.G932 6837
SMM.US.04.M919 6859
SMM.US.04.M922 6862
SMM.US.04.M923 6846
SMM.US.04.M926 6855
SMM.US.04.M934 6822
SMM.US.04.M935 6835
SMM.US.04.M940 6878
SMM.US.04.M946 6853
SMM.US.04.M947 6849
SMM.US.04.M949 6836
SMM.US.04.M950 6843
SMM.US.04.M951 6836
SMM.US.04.M952 6835
SMM.US.05.D215 6836
SMM.US.06.FTq 6861
SMM.US.11.SIVsmE660_FL10 7573
SMM.US.86.CFU212 6831
SMM.US.x.F236_H4 7529
SMM.US.x.H9 7056
SMM.US.x.PBJA 7358
SMM.US.x.PGM53 7513
SMM.US.x.SME543 7586
SMM.US.x.pE660.CG7G 7569
STM.US.89.STM_37_16 7218

Draft of 30 September 2015, not for public release
hiv2dna_main.tex r8876–10546


```
                                                                V3 loop end
MAC.US.x.239              CTGGATTGGTTTTCCACTCACAACCA......ATCAATGATAGGCCAAAGCAGGCATGGTGTTGGTTTGGAGGAAAATGGAAGGATGCAATAAAAGAGGTGAAGCAGACCATTGTCAAACATCCCAGGTAT.........ACTGGA...ACTAAC...AATACTGATAAA 7728
Env                       S__G__L__V__F__H__S__Q__P__.__.__I__N__D__R__P__K__Q__A__W__C__W__F__G__G__K__W__K__D__A__I__K__E__V__K__Q__T__I__V__K__H__P__R__Y__.__.__.__T__G__.__T__N__.__N__T__D__K_
A.CI.88.UC2               -A---CATAGG--T--T--T--GG--GTC...-----CA-A-AA--T-GA-----------C------AA---C--C-----A-GA--C--GC-G-----------A---C---CAGG-----------C.........-AA---...--C--T...G----GAG---G 7761
A.DE.x.BEN                -A-----A--G--T-----T--G--.......-----CACA-----T-G-----------CC----------C-G-----G---A--C--GC-G-----------A---C----AC----------A--C.........-AA---...-TC--T...G----A-GG--- 7810
A.DE.x.PEI2_KR_KRCG       -A--GAGA--G--T-----C-G---GATC...--------A--A--C--------------C-----C-----TG-T------A-A--C--GC-G---------A--A---C----G------------------.........-GA---...--C---...G-C--AC-G--- 7754
A.GH.x.GH1                -A--G--A--G--T-----C--G---.......------ACA--A--T-G---A---------C-----CAA---C--G----G---A--C--GC-G-----A-----A---C--A-A-----------------.........-AA---...--C---...G-C--AA-G--T 7227
A.GM.87.D194              -A--GCGAAGG--T-----T-GG---GTC...TA---CA-A-AA--TGG-------------------CA---C--C----TA--A--C--GCGG-----------A---C---CA-------------C........GGA---...--A--T...G----A-GA--- 7221
A.GM.x.ISY_SBL_6669_85    -A--CCGCAGA--T-----C--GAAGATC...------A-A-AA--C-G---A--------CC----CAA---CG-G----G---A--C--GC-G---------A--A---C----A----------------.........-AA---...--C--T...G-C--AA----- 7227
A.GM.x.MCN13              -A--G--AAGG--T-----C--G--GGTC...------A-A--A--C-GA--A---------------CAA---TG---------GA--C--GC-G---------G-A---C---CA---------------.........-AA---...--C--T...G-A--AA-G--T 7243
A.GM.x.MCR35              -A--G--AAGG--T-----C--G--GGTC...------A-A--A--C-GA--A---------------CAA---TG---------GA--C--GC-G---------G-A---C---CA---------------.........-AA---...--C--T...G-A--AA-G--T 7243
A.GW.86.FG_clone_NIHZ     -A--G--TAAG--T-----C--G---GTC...------A-A-AA--C-G---A---------------C-A---CC-----------A--C--GC-G---------G-----C---CG-----------A---.........-AA--G...-AC-GGAGCCGC--A--G--T 7221
A.GW.87.CAM2CG            -A--GCAAAGA--T-----C-GG---ATC...--T---A-A-----C-G---A--------C-----CAA---C--C----CA--A--C--GC-G-----A-----A---C---CGG-G------------.........-AA---...--C--A...--C-TAACGG-T 7816
A.GW.x.MDS                -A-----A--G--T--T--C--G--G.......------AGA--A--T-G---A--------C-----CAAG--C--C----G-AGA--C--GC-G-----A---G-A-----A-A-----C---------.........-AG---...--C-GT...--C-TAACA--T 7284
A.IN.07.NNVA              -A--C--A--G--T--T--C--G---.......------ACC--A--T-GA--A--------C-----CAA---C-G-----------A--C--GC-G-------------A-------CAG----C---------.........-GG---...--C--T...-T---A--A--- 7805
A.IN.95.CRIK_147          -A----ATAAG--T--T--C-GG---GTA...------ACA--A--T--A--A--------C-----CAA---CCGG--------C--C--GC-G-----A---A-A---G---CAG-------------..........G-G-T-...--C--G...G--C-AA-G--C 7552
A.JP.08.NMC786_clone_41   -A--GGCAAGG--T-----C--G---GTA...--T---ACA-AA--T-----A--------C-----CAA---CG-G-----AA-G--CT-GG-G---------AG-A---C--AAAG-------------.........-GA---...C-C--T...G-C--AA-GC-- 7812
A.PT.x.ALI                -A--G--AA-A--T-----C--G---.......------A-A--A--C-GA--A--------C-----CAA---CG------G-A-A--C--GC-G---------G-A---C----A----------------.........-AA---...--C--T...G-C--AA-CC-- 7791
A.SN.85.ROD               -A---CAT--G--T-----C--CTACCAGCCG------A-A--A--C-GA--A--------C-----CAA---C--------A--C--C--GC-G---------G-A---C---CA---------------.........-GA---...--C--T...G-C--AAGG--T 7259
A.SN.86.ST_JSP4_27        -A--G--A--G--T-----C--G---.......------AGA--A--C-G---A--------C-----CAA---CG-G--------A--C--G--G----------TA---C---CA---------------.........-AA---...--C---...G-C--A--A--- 7237
B.CI.88.UC1               -A---C--AA------T--G--G--T......C-----ACC--A--C-GA-----T-----C------AAG-----C----TA--A--C----GG-----A---G-------CA-A----------------.........-AA---...--A--T...--C--C--G-GG 7782
B.CI.x.20_56              -A--GCGAAGA-----T--G-GG--TGTT...---------G--A--C-GA-----T-----C-------AG-----C----CA--G--C--G-GG---------G-----G-CA-G---------------.........--A---...-TA--G...--C-TCACG--- 7759
B.CI.x.EHO                -A--TA-TC-C-----T-----G--T.......------A-A--A--T--A--A--T-----C------AA------C----CA--A--C---C-G-------AG-------AAA--T--------A---.........T-A---...--A-CA...----TCTCAC-G 7769
B.GH.86.D205_ALT          -A---C-AC-------T-----G--T.......------A-G--A--C-GA--A--T-----C------AAG-----C----CA--A--C-----G-------AAG------CA-A----------------.........-AA---GGTG-A--A...----TCACA-GC 7780
B.JP.01.IMCJ_KR020_1      -A--TC-AA-C-----T-----G--T.......--T---C-A--A--C-G---A--C-----C---------------C----CAC-G--T-----G----------G---------CA--G--------A---.........--A---...--A---...G----C-CA--- 6900
G.CI.92.Abt96             -A--GC-A--C--------T------.......--------A--A--C-GA-----T---------------C-----------A--C--GC-G-----A---G-A---G-G-----G--C-----C---.........-AA---...--A--T...G-C--AA-G--M 7141
AB.CM.03.03CM_510_03      -A---C-A--G--T-----T--G---.......-----CCGC-----T-GA-----------C------A---C--C----CA--A--C--G--G-----------A---C----A--G--------A--C.........-AG---...--C--T...G----GACA--- 6889
H2_01_AB.CI.90.7312A      -A--G--A--G-----T--C--G--T.......------A-A--A--T-G---A--------C-----CAA---CG-G----G---A--C--GC-G---------A--A---C--A-A----------------.........-AA---...--C--T...G-C--AAGG--T 7792
H2_01_AB.JP.04.NMC307_20  -A--G--A--G--------C--G---.......------ACA--A--T-G---A--------C-----CAA---C--G--------A--C--GC-G-------------A---C--A-A----------------.........-AC---...-AC-GA...--C-TAAG---- 7774
H2_01_AB.JP.07.NMC716_01  -A--G--A--G--T-----C--G--C.......------A-C--A--T-G---A--------C-----CAA---C--G----G---A--C--GC-G-------------A---C-CA-A------T--------.........--A---...--C--A...--C-TAAG-TT- 7767
H2_01_AB.JP.08.NMC842_10  -A--G--A--G--------C--G---.......------A-C-----T-GA--A--------C-----CAA---C--G----G---A--C--GC-G-------------A---C--A-A----------------.........-AC---...-AC--A...--C-TAAG---- 7797
U.CI.07.07IC_TNP3         -A--G--A-----------T--G---.......--------A--A--C-----A--C-------------T--G-GT-----A--G-----CC-G-----A---G-A---C-G--A-----------A---.........----------...-TC--T...G----AA-A--G 7246
U.FR.96.12034             -A---C-A--G-----T--T---G-T.......------A-C-AT-----A--A--T-------------T-----T-----A-GG-----CC-G---------G---------C---C--------------.........CA---T...--C--G...----T-AG-C-G 7270
MAC.US.x.17EC1            -----------------------G-.......----------------G--------------------------------------------------------------------------.........------...------...------------ 7728
MAC.US.x.251_1A11         --CA----------------------.......------------------------------A---------------------------------------------------------------.........------...------...------------ 7724
MAC.US.x.251_32H_PJ5      --------------------------.......----------------------------------T---------------------A-----------------------------------.........------...------...------------ 7726
MAC.US.x.251_BK28         --------------------------.......---------------------------------------------------------------A-----------------------------.........------...------...------------ 7710
MAC.US.x.MM142_IVMXX      ---C----------------------.......G-------G------------------A-----------T-------G--------------------------------------------.........------...------...------------ 7223
MNE.US.82.MNE_8           --------------------------.......----------------A-----------A------A------T-------G---------------A--------------------------.........------...------...------------ 7196
MNE.US.x.MNE027           --------------------------.......--------A-------A-----------A------A------T-------G------------------------------------------.........------...------...------------ 7199
SMM.CI.79.SIVsmCI2        -A--G--A-----------C-----T.......--------A--A-----A--A--T-------------G-----G-----A--G---C-----G---------G-A---C----------C---------.........-AG---GGA--C--T...G-----AGG--G 6878
SMM.LR.89.SIVsmLIB1       -A--------A--T--T-----G---.......--------G-------A--A----------------T--GTC-----GAA-A------C------C---G-A---T-A------------------.........----------C--T...G----AA-C--G 6784
SMM.SL.92.SIVsmSL92A      -A-----A-----T-----C------.......--A-----G--A-----A--A--T-------------AA---GG-G------A-G--T----G-----A--AG-A--------C---------------.........----------A--T...-GA--A--A--G 6813
SMM.SL.92.SL92B           -A--T-----C--T--T--G------.......--A-----G-------GA--A--------C-------------------GA--A--T--GC-G-----T---A-A-----A--------C---------.........----------C--T...G-----AGG--- 7164
SMM.US.04.G078            -A--G-----C-----T--G-----C.......--A-----G--A-----A-----C-----C-------------GT-----A-G--C--CC-G-----A--AG-------------G------------.........--A---...-----T...G-------G--- 7009
SMM.US.04.G932            -A--G-----C--T--T--------T.......--A-----G-----------A--T-----CC-----------G-C----GA-GA--C--G--G-----C---G-------AA-T-----------A---.........--A---...-----T...G-A--AA-G--T 6986
SMM.US.04.M919            -A--G-----C-----T--G-----C.......--A-----G--A-----A--A--C-----C------AA------T----CA--G--C--C--G---------G-A----------------------A---.........----------...-----T...G-----A-GC-- 7008
SMM.US.04.M922            -A--G-----C-----T--G-----C.......--A-----G--A-----A-----C-----C-------------GT-----AC-G--C--CC-G---------G-A----------------------A--C.........-AC---...-----T...G---T-A-A--G 7011
SMM.US.04.M923            -A--G-----C-----T--G-----C.......--A-----G--A-----A-----C-----CC-----------G-C-----A-GG--C--CCGG--A------G-C---T-G-------------------.........--G---...-----T...G-G---A-A--G 6995
SMM.US.04.M926            -A----------T-------------.......--------G--A---GA--A---------A------------GC------A-A------C-G--A------G-A----------------------A---.........--A---...-----T...------A-G--- 7004
SMM.US.04.M934            -A--------C--T--------G---.......--------G------GA--A----------------------GT------GA--------G--A------G-A---C-------T--------A---...AATGAA-G---...-----G...G---T-AG-C-- 6977
SMM.US.04.M935            -A--G--A--C--T--T--G-----C.......--A-----G--A---GA-----C--------------------G-----AC-G--C--CC-G-----A---G-A----------------------A---.........-AG---GGA-----T...G-----A-A--G 6987
SMM.US.04.M940            -A--G-----C--T--------G---.......--A-----G--A---GA--A--------C---------G---G-G------GA-----C--G--A--C---G-A---T-GA---------------AATAATAAA----------...--C--T...--A--A--AG-G 7036
SMM.US.04.M946            -A--------C--T--------G---.......--------G------GA--A----------------------GT------GA--------G--A--A---G-A---C-------T--------A---.........CAG---...-----A...----T-AG-C-- 7002
SMM.US.04.M947            -A--G-----C-----T--G-----C.......--A-----G--A---GA-----C-----CC------A----------CGA--G--C--CC-G--A------G-A---T-G-------------------.........--G---...-----T...------------ 6998
SMM.US.04.M949            -A--G-----C--T------------.......--A-----G--A-----A--A--G-----C-T---CA-G---G-C-----A-A-----CC-G------C---G-----T-G-----G-----------.........--G---...--C--T...G----AA-ACGG 6985
SMM.US.04.M950            -A--------C--T--------G---.......--------G------GA--A------------C------------------A--------G--A------G-A-----------G--------A---.........-AA---...----GA...G-A-T-A-A--T 6992
SMM.US.04.M951            -A--G-----C--T------------.......--A-----G--------G---A--------C------A-G---G-T------GA-----CC-G------C---G-----T-G--CGC-------------.........GA---...--C--T...G-A--AA-A--C 6985
SMM.US.04.M952            -A--------C--T--------G--G.......--------G------GA--A----------------------GC-----GAA-A------G-G--A------G-A---C-------G--------A---.........CA---...-----T...G-----AGG--- 6984
SMM.US.05.D215            -A--GC-A--C--T--TG----G---.......--A-----G--A--T-G---A--------C-----------C--G-----AC-G--------G------C---G-A---T-G-------------------.........--A---...--C--T...G----AAGC--G 6985
SMM.US.06.FTq             -A--C-----C-----T--C------.......--A-----C-------GA--A--C-------------T-------------ATCA--T--GC-G---------A--A---C----A-----------A---.........----------...--C--T...G-----A-A--C 7010
SMM.US.11.SIVsmE660_FL10  -A--G-----C-----T--G--G--C.......--A-----G--A-----A-----C-----C-------------GC-----A--G--C--CC-G--A------G-A---T-G-------------------.........--G---...-----T...G-----A-G--- 7722
SMM.US.86.CFU212          -A--G--A--------T--------C.......--A-----G------GA-----------C-------------------TCA-----------------G----T-G-------------------.........--A---...--C--T...G----AAG---- 6980
SMM.US.x.F236_H4          -A--G-----C-----T--G--G--C.......--A-----G--A-----A-----C-----C-------A----GC-----AA-G--C--CC-G--A------G-A---T-G-------------------.........--G---...-----T...G-----AGG--- 7678
SMM.US.x.H9               -A--G-----C-----T--G-----C.......--A-----G--A-----A-----C-----C------------G-------A--G--C--CCGG--A-----GMAM--T-G-------------------.........-AG---...-----T...R-K-Y-R-AC-- 7205
SMM.US.x.PBJA             -A--G-----C-----T--G-----C.......--A-----G--A-----A-----C-----C------------G-------AA-G--C--CC-G--A------G-A---T-G-------------------.........--G---...-----T...--G-----AC-- 7507
SMM.US.x.PGM53            -A--G---A-C-----T--G-----C.......--A-----G--A-----A--A--C-----C-------------GT----GA--G--C--CC-G---------G-A----------------------A---.........--G---...-----T...G-G---A-G--- 7662
SMM.US.x.SME543           -A--G-----C-----T--G-----C.......--A-----G--A-----A-----C-----CC-------------C----GC--G--C--CC-G--A------G-A---T-G-------------------.........--G---...-----T...G-----AGG--- 7735
SMM.US.x.pE660.CG7G       -A--G-----C-----T--G-----C.......--A-----G--A-----A-----C-----C-------------GC-----A--G--C--CC-G--A------G-A---T-G-------------------.........--G---...-----T...G-----A-G--- 7718
STM.US.89.STM_37_16       -A--C-----C-----T--G-----C.......--T-----G--A-----A--A--T-----C------------G------GA-GA-------------T---G-A---T-G-------------------.........--A---...--C--T...G-C--G-CA--- 7367
```

```
                                                                                      V4 loop start
MAC.US.x.239              ATCAATTTG...ACGGCTCCTGGAGGA...GGAGATCCGGAAGTTACCTTCATGTGGACAAATTGCAGAGGAGAGTTCCTCTACTGTAAAATGAATTGGTTTCTAAATTGGGTAGAAGATAGGAATACAGCTAACCAGAAG..............CCAAAGGAACAGCA 7877
Env                       _I__N__L__.__T__A__P__G__G__.__G__D__P__E__V__T__F__M__W__T__N__C__R__G__E__F__L__Y__C__K__M__N__W__F__L__N__W__V__E__D__R__N__T__A__N__Q__K__.__.__.__.__.__P__K__E__Q__H
A.CI.88.UC2               --T--C--T...GT-AAA--A----T-GGCTC---C--A--G--G--A-A---------T--C-----------A--TT----------C----C------C--C--------G--GA-C----CG......-G---A--A.......................... 7893
A.DE.x.BEN                --T--C--T...---AAA--G----C-GGCTC---C--------GG-A--T--------T--C-----------A--T----------C----C------C--C---------------C-A---C..........--A-CA.......................... 7939
A.DE.x.PEI2_KR_KRCG       --T--C--T...-ACAA--A---AA-GGTTC----G-A-----GGTA-A---------T--C-----------A--T--A-----C--C----C-C----C--C------A-----A-C---GCA.......................................--C-C 7882
A.GH.x.GH1                --T--C--T...-AAAA--A---A--GGCTC---C--A-----GG-A-A---------T--C-----------A--T--------C--C----C------C--C-------------A-C--ACCG......--T----CA.......................... 7359
A.GM.87.D194              --T--C--T...---AAG--A---AT-GGTTC---C--A-----G--A-A---------T--C-----------A--T----------T----C------C--C-------------A---A--CG..........-----A-CA.......................... 7353
A.GM.x.ISY_SBL_6669_85    --T--C--T...-A--A--A-A-AA-GACTC---C--A-----AG-A-AT--------T--C-----------A--------T--C--C----C------C--T-------------A-C-A--CG......GGT--A.......................... 7356
A.GM.x.MCN13              --T--C--T...-A--A--A---AAGGGCTC---C--A--G--GG-A-A---------T--C-----------A--T--------C--C----C------C--C------A-----A---A--CA.......................... 7366
A.GM.x.MCR35              --T--C--T...-A--A--A---AAGGGCTC---C--A--G--GG-A-A---------T--C-----------A--T--------C--C----C------C--C------A-----A---A--CA.......................... 7366
A.GW.86.FG_clone_NIHZ     --T--A--T...AA--A--A--GA--GGCTC---C--A-----A--A-A---------T--C-----------A-CT--------C--C----C------C--C----------GA-C----CG.....GGT-----A.......................... 7353
A.GW.87.CAM2CG            --T-CC--C...AA--A--A-A-A--GGTTC---C--A-----G--A-A-------T-T--C-----------A--TT-------C--C----CC-----C--C-------------A---AACC-......--TAC--CA.......................... 7948
A.GW.x.MDS                --T-CC--C...-----A--A---AA-GGCTC---C--A-----AG-A-A---------T--C-----------A--T--------C--C----C------C--C-------------A-C--A-C-.....G-TA--CCA.......................... 7416
A.IN.07.NNVA              --T--A--T...-A--A--G---AA-GGTTC---C--------GG-A-A---------T--C-----G-----A--T--------C--T----C------C--C------A----GA-C--A-CA.......--TG-A-CA.......................... 7937
A.IN.95.CRIK_147          --T-CC--T...G-A--A--G---AA-GGCTC---C--------AGAA-A---------T--C-----G-----A--T--------CG-C----C------C--TG-----A----GA-C--ACC-.......-CGAG----GCA.......................... 7687
A.JP.08.NMC786_clone_41   --TCGC--T...-TAAAA--A----A-GGCTC---CG-A-----G-AA-A---------T--C-----------A--T--------C--C---G-------C--C-----------G---A--T--A......GGTA-A.......................... 7941
A.PT.x.ALI                --T--C--T...-AAAA--A---A--GGCTC----G-A-----GGTA-AT--------T--C-----------A--T---C-T--C--C----C------C--C---------G---A-C-AA-CG.....GGT---G-A.......................... 7923
A.SN.85.ROD               --T-GC--T...G-A--G--A---AA-GGCTC---C--A-----AG-A-A---------T--C-------------T--------C--C----C------C--C------A----GA---A--CA.......................... 7382
A.SN.86.ST_JSP4_27        --TCG---T...-TA--G--A----A-CGCTC---C--A-----GG-A-A---------T--C-----------A--T--------C--T----C------C--C-------------A-C--A-CG......--T----CA.......................... 7369
B.CI.88.UC1               --A-GA---...GTA-GG--CTCG-C-GGTTC---C--A--G--G-GACAT--------T--C--T--G-----A---T----T-----T----CC------T----C-----G--GA-C--A-CA.....GGTACA-CT.......................... 7914
B.CI.x.20_56              --A-------...GTA-GG--CTCG-C-GGTTC---C--A--G-CA-GA-AT--------T--C--T--G-----A---T----T-----C----CA------T----C-----------GC-AA---......GGAACC.......................... 7888
B.CI.x.EHO                --A-GG--A...G-A-AG-AC-CGA--AGCTC------A-----A-GA-AT--------T--C--T--G--------T-----------T----C--TT--CT----C-----------A----A-C-......GGG-TC.......................... 7898
B.GH.86.D205_ALT          G-A--G--AGTAT-A-AA-A----AA-GGTTC------A---ACA--A-AT--------C-----T---------A---T-A---------T----CC-----CT----C---------GA-C-AA-CA.......--TACA-CT.......................... 7915
B.JP.01.IMCJ_KR020_1      --A--G--A...G-A-AG-ACTCGCA-GGCTC------A-----A-GA-AT--------T--C--T-----------TT-G---------T----C------C-----C---------A------CA.....GGAATT.......................... 7029
G.CI.92.Abt96             --A-CC--T...-AA-A--A----A-GGTTC------A-----A-AG-----------T-------------A---T-G--T----------C-----------------A----G--C--A----TGA--CT--TA---.................--CC------GA-- 7293
AB.CM.03.03CM_510_03      --T--C--T...---AAA--A----A-GGCTC---C--A-----GG-A-A---------T--C-----------A--T--------C--C----C------C--C------------GA------CG.....GG-ACA.......................... 7018
H2_01_AB.CI.90.7312A      --T-CC--T...-AAAA--A---AC-GGCTC---C--A-----GG-A-A---------T--C-----------A--T--------C--C----C------C--C-------------A-C--A-CG.....GGT----CA.......................... 7924
H2_01_AB.JP.04.NMC307_20  --T--CC-T...AAAAA--A----A-GGCTC---C--A-----GG-A-A---------T--C-----------A--T--------C--C----C------C--C-------------A-C--A-CG......--T----CA.......................... 7906
H2_01_AB.JP.07.NMC716_01  G-T--C--T...-AAAA--A----A-GGCTC---C--A-----GG-A-A---------T--C-----------A--T--------C--C----C------C--C-------------A-C-AA-CG......--T----CA.......................... 7899
H2_01_AB.JP.08.NMC842_10  --T--C--T...-AAAA--A---ACTGGCTC---C--A-----GG-A-A---------T--C-----------A----------C--C----C------C--C-------------A-C--A-CG.....GGT---.......................... 7926
U.CI.07.07IC_TNP3         --T------...-ACAG--A------GGA...--C--A---CC---------------T--------------A--TT-G--T-----G--------------G-------------C--A...GGT-AA---G-C-C-AAAGAC...CAAAACATT-GA--G-GA-- 7404
U.FR.96.12034             --A---C--...G-A-AA--A-C-----...--T---T-A--G--A-AA--T--------T--C--------------A---------------G---CA-----C--C-----------A-C----GC-TG-AGGGA--A-C-...............T---G---GAGA-- 7419
MAC.US.x.17EC1            ---------...---------A-----...------------------------------------------------------------------------------------------------------------------...............-------------- 7877
MAC.US.x.251_1A11         ---------...---------------...------------------------------T---------------------T--------------------------------------------------...............-----------A-- 7873
MAC.US.x.251_32H_PJ5      ---------...---------A-----...-------------------------------------------------------------------------------G--GT-A---------G-...............-----------G--- 7875
MAC.US.x.251_BK28         --------A...---------------...-------------------------------------------------------------------G------G--GT-A---C-----G-...............-----------G--- 7859
MAC.US.x.MM142_IVMXX      ---------...---------A-----...------------------------------T--------------------------------------------------G-CT-A---C------...............-----------G--- 7372
MNE.US.82.MNE_8           ---------...---------------...------------------------------T------------------------------------------A----CTGA--GGAACT-CC.........CAGAAG---C------G--- 7351
MNE.US.x.MNE027           ---------...---------C-----...------------------------------T------------------------------------------A----CTGA--GGAACT-CC.........CAGAAG---C---------- 7354
SMM.CI.79.SIVsmCI2        --------C...-A-AG--A---A--AAGTCT-----A-----AG-A--------------C----------A----------C--G---CA-----C--C----------G--A-----C-TGA-AG-A...............CAAAAC--GG-----A-AA- 7030
SMM.LR.89.SIVsmLIB1       --T------...-A----A------...--------A-----A--A--T--------------------------T-----G----------C--G----------G----AG-C---AA-C--GCT-GATGGACATCACAAACTAGC--A----GAA- 6948
SMM.SL.92.SIVsmSL92A      --TT-CC-A...-----C--A------...--G--C--A--G-------T----------C----------A-------T-----------------G-----------C--A---TTGA-ACT--CA-GATGGAAAAACCAAAAG---G-----A-AGC 6977
SMM.SL.92.SL92B           --T---C-A...-----C--G------...--G-----------C--A--------------------------A----T-----------------C--C----------G-----------GAGC-G--CA-G-TGGACAACTCAAACCAA------------ 7328
SMM.US.04.G078            ---------...-A----AA-----...--G-----A--------T--T----------C--TC-------A---T-A--T--C---------------C--T----------G-----A--CGA-AAAGGTTCT-TATGGGAACAACAAAACAAG--A--G--AA- 7173
SMM.US.04.G932            --T-C-C--...-AC---A-----...--------------C-----T----------C----------A--T----T--C--------------GG----------T--GA--TTTGG--ACA--G-T-CT-G-TGGAAACAGCAAAACT---GA---A---- 7150
SMM.US.04.M919            --T---C--...-A----A-----...--------A--------T--T-----------------T--------A--T--T--T--C---------------C--T----------T-----A--CGA-A-AGGGTCC-GATGGATGCAACAAACA-AG--A--G-GAA- 7172
SMM.US.04.M922            --T---C--...-A-----------...--------A------------------------------------------T--C---------------C--T-------------------AG-C-T-AAAGGTTCT-GATGGAAGCAACAAAACGTG--A--GAGAA- 7175
SMM.US.04.M923            --T---C-A...-A----AA-----...--------A--------T---------------TC----------T-A--T--C-----------------T--------------C-TAG-C-T-A----AG-T-GATGGACAAGACAAAAC--G--A--G-GA-- 7159
SMM.US.04.M926            --T------...----------------...--------T--------T---------------T-----------T-----T--C---------------C--G--------------C--A----T-T-AC-T-CT-GATGGACACAGCAAAAGAA-CGA--GAGAG- 7168
SMM.US.04.M934            G-T------...----------------...--------A----C---------------T--------A--T-----T--C---------------C-------------G-----A--CGA-A-CGGT----GATGGACGGATCAAAAGT-T--A--G-GAA- 7141
SMM.US.04.M935            --T---C--...-A--------------...--------A--G--C--------------T------------------T--C---------------C--T-------------------A--C---AG-GCT-CT-GATGGAAGCAGCAAAAGAAG--A--GAGAG- 7151
SMM.US.04.M940            --TT--C--...-----C--A-----T...--G--C--A--G-------T----------C----------A------------G----------C----------T--G--C--A--C---T-GC---CA-GATGGACGCAACAAAAGAA---------A-- 7200
SMM.US.04.M946            G-T------...----------------...--------A----C---------------T--------A--T-----T--C---------------CT-------------G-----A--CGA-A-C-GT-TT-GATGGCTGAAACAAAAT--T--A--G-GAA- 7166
SMM.US.04.M947            --TT--C-A...-AT---A-----G...--------A--------T---------------TC-------A--TT-A--T--C---------------C--T-----------G--C......---AAAGGTT-C-GATGGAGCAGCCAGAACAAG--A--TA---- 7156
SMM.US.04.M949            --T---C--...-A--C--A-----T...--G-----A----------T-----------------------A--T---------------------C-----------T-----C--A------T-G-G---A---...............AA--GA-----A-T 7134
SMM.US.04.M950            --T---C-A...----------------...--------A-----C------------------T--------A--T---------C--------C--------G-----------G-----A--CCA-A-CC-TTCT-GATGGAAGCAACAAAAG--T--A--G-GA-- 7156
SMM.US.04.M951            --T---C--...-A--C--A-----T...--G-----A----------T-----------T-----------A--T---------------------C-----------T-----C--A--CTT-A-GC---TC--AAAGCATGGAGCAAAA-T---------T--- 7149
SMM.US.04.M952            --TT-----...----------------...--------A-----C--------------------T-----------A--T-----T--C---------------C-----------------G-----AT---A-AAC-G-TCT-GATGGAAGCAACAAAAGG-T-GA--GA-AG- 7148
SMM.US.05.D215            --T--GC--...-AAAG--A-----...------C--A-----C------------------------T-----T------------------CT------------------A--A---G-TA-AG-GAT---TGCGAGAAACAAAAACAACCG----A--A- 7149
SMM.US.06.FTq             --A--C--T...-T-AG--A---AA-ACCTCT--C--A-----AG-T--T--------------------A-----------T--------C-----C--C----------T-----C--A--C-TGT--TCA--A--C...............AAGGCA-----AA- 7162
SMM.US.11.SIVsmE660_FL10  --T---C-A...-A----A-C----...--------A-----C--T--T------------T-----------A---T-A-----C--------------------T-----------G--C--AG-CCA-AAG-GTAGC-GATGGAGACAACAAAATA-G-GA--G-GA-- 7886
SMM.US.86.CFU212          --T---C--...-A--A--A------...------C--T--G--C-------------T-----------A---------T------------------C---------------------A--C-G-A-GC-GAT-G--GCGAGAAACAACAACGAT--A--GAG-AG 7144
SMM.US.x.F236_H4          --T---C-A...-A----A-C----...--------A-----C--T--T-------------TC-------A---T-A--T--C-----------------T-----------G--C--AG-CCA-AAGGGTGGC-GATGGAAACAACAAAATAGG--A--G--A-- 7842
SMM.US.x.H9               --T--GC-A...-A----A------...--------A---------T-------------TC-------A---T-A--T--C-----------------T-------------Y-...-TA-A-AA-GGTTCT-GATGGACAAGTCAAAAC-AG--A----GAT- 7366
SMM.US.x.PBJA             --T--GC-A...-A----A------...--------A---------T-------------TC-------A---T-A--T--C-----------------T------------A--...-TACA-AA-GGTTCT-GATGGACAAGTCAAAAC-AG--A--G-GA-- 7668
SMM.US.x.PGM53            --T---C--...-A----A------...--------A-----C--T--T------------T-----------T-----T--C---------------C--T-------------------A--C---T-AGGTTCTCTATGGACACAACAAACAAAGGCA--CG-AA- 7826
SMM.US.x.SME543           --T---C-A...-A----A-C----...--------A-----C--T--T-------------TC-------A---T-A-----C-----------------C-----------G--C--AG-CCA-AA--GTA-C-GATGGAAACAACAAAAGAAGCCA--G--A-- 7899
SMM.US.x.pE660.CG7G       --T---C-A...-A----A-C----...--------A-----C--T--T------------T-----------A---T-A-----C-----------------T-----------G--C--AG-CCA-AAG-GTAGC-GATGGAGACAACAAAATA-G-GA--G-GA-- 7882
STM.US.89.STM_37_16       --A-GGA-A...GT---------G---...--------A--G--C--------------------C-T-----A--T--T--T--C--------------CT--------A-----A-C--A-G----T-AG-AAT--GAGATTGGAACAAAAACAA---------A-- 7531
```

```
                                        V4 loop end                                                                                                                               V5 start          V5 end
MAC.US.x.239                TAAAAGGAATTACGTGCCATGTCATATTAGACAAATAATCAACACTTGGCATAAAGTAGGCAAAAATGTTTATTTGCCTCCAAGAGAGGGAGACCTCACGTGTAACTCCACAGTGACCAGTCTCATAGCAAACATAGATTGGATTGAT.........GGAAACCAAACTA 8038
Env                         __K__R__N__Y__V__P__C__H__I__R__Q__I__I__N__T__W__H__K__V__G__K__N__V__Y__L__P__P__R__E__G__D__L__T__C__N__S__T__V__T__S__L__I__A__N__I__D__W__I__D__.__.__.__G__N__Q__T__
A.CI.88.UC2                 .C-GC-C--C--T-C------C-----A--G--G-----T--T--C--------G-----GC--T----A----------T--G--A--G--GT-G--C--C-----A-----A-----CA-A--T--T-----T--CACAGA--GA.............--T--G--C- 8050
A.DE.x.BEN                  .CGGC-C--C--T......--C-----A-AG--G-----T--T--C--------------G--------A----------T--G--A--G--GT-GG-C---G-A--A-----A-----CA-A--T--T-----T--CATAGA-A-A......AATC-G-CT--T--C- 8096
A.DE.x.PEI2_KR_KRCG         AC-GC-C-----T-CA--G--C-----A--G--------T--T--C-------G-------C-----A-A----------T--G--A--G--AT-GGTC--C-----A-----A-----CA-A--T--T-----T--CAT-T-----.................G--A- 8040
A.GH.x.GH1                  .C-GCAC-----T-C---G--C-----A--G--------T--T--C-----C--G-----G--------A----------T-----A---C-GT-G--C--C-----A-----------CA-A--T--T-----T--CGTA-A-AG-.................G--A- 7516
A.GM.87.D194                .C-CG-C--C--T-C------C-----A--G--G-----T-----C--------G-----G-C------A----------T--G--A--G--GT-G--C--C--T--A-----A-----CA-A--T--T-----T--C-CAGA--GA.............--T--G--C- 7510
A.GM.x.ISY_SBL_6669_85      .C-GCAT--C--T-----G--C-----AGAG--------T--T--C--------G-----G--------A----------T--G--A-----GT-GT-C--CG-A--A------------A---T--T-----T---GTTGA--GA..........AT----GG--A- 7516
A.GM.x.MCN13                .C-CC-C-----T--A--G--C-----A-----------T-----C--------G-----G--------A----------C--G--A--G--GT-G--C--C-----A-----A--T--CA-A--T--T-----T---GCA-A--GA.............TA-T--A- 7523
A.GM.x.MCR35                .C-CC-C-----T--A--G--C-----A-A---------T-----C--------G-----G--------A----------C--G--A--G--GT-G--C--C-----A-----A--T--CA-A--T--T-----T---GCA-A---A.............TA-T--AG 7523
A.GW.86.FG_clone_NIHZ       .C-GC-C-----T-CA--G--C-G---A--G--------T--T--C-----C-GG-----G------C-A----------C--G--A--G--GT-G--C--C-----A--G--------CA-A--T--C-----T---GC-GGA---.................A- 7507
A.GW.87.CAM2CG              .--GC-C-----T-CA-----C-----A--G--------T-----C--------G-----G--G-----A----------C--G--A--G--GT-G--C--C-----A-----------CA-A--T--T-----T--CGA--GG---..........AATC-GAC---A- 8108
A.GW.x.MDS                  .TGGCAC-----T--A-----C-----A-AG--------T-----C--------------A-G------A----------C--G--A-----GT-GGTC--C-----A-----------CA-A--T--T-----T--CGAC-AG---.............T--G--A- 7573
A.IN.07.NNVA                .--GCAC-----T--------C-----A-AG--G-----T----------------C---G-G-C----A----------C--G--A--C--GT-G--C---G----A--T--------CA-A--T--T-----T--A........G..............T-----A- 8088
A.IN.95.CRIK_147            .TGGC-C-----T--------C-----A--G--------T-------------G-----G---C----G----------C--G--A--G--GT-G--C--C-----A--T--------CA-A--T--T-----T---GTA-AC--A......ATA-AT--A-GG--A- 7850
A.JP.08.NMC786_clone_41     .GTGC-C-----T-----G--C-----A--T--------T-GA--C-----C--G-----G-T------A----------T--G--A-----GT-GGTC--C-----A--G--------CA-A--T--C-----T--CAT-GG-A-A..............TC---A- 8098
A.PT.x.ALI                  .C-GCAC-----T-CA--G--C-----A-AG--------T--T-TC-----C----C---G--------A----------T--G--A-----GT-G--C--C-----A-----A-----CT-G--T--T-----T--CAC-GA--GC.................G--A- 8080
A.SN.85.ROD                 .C-CC-C-----T-CA--G--C-----A-AG--------T-----A-------G-----G-G------A----------C--G--A--G--G--GT-C--C-----A-----A-----CA-A--T--T-----T--C---CAAA-C.........AAT--T--G--A- 7542
A.SN.86.ST_JSP4_27          .C-GCAC-----T--------C-----A-AG--------T--T--C-----C--G-----G--------A----------T--G--A---C-GT-A--C--C-----T-----------CA-A--T--T-----T--CG-AGGA--G.................G--A- 7526
B.CI.88.UC1                 .C-G-A------T---A-G--C--C--C-A---G---G------G-----C--G--T--A---T-C--A--------------G-----TACG---T-C-----T----GT--C--------T-----C-----T--CGT-TA-T--GATGGCAAT-AT-C-A-G--C- 8083
B.CI.x.20_56                .-----A-----T-----G-----C--C-----G---G------G-----C--G--T--A---T-C--A--------------G-----CCTG---T-C-----------T--C------A-T-----C-----T--A--------------.........A-C---G-G--C- 8048
B.CI.x.EHO                  .--G--A-----T-C-T---C--C--C-----G---G------G-----C--GA-T--A-G-----G-------------------T--A---T-C-----T-----T--T-----C--------C-----T--C-----A---.........AAG----TT---- 8058
B.GH.86.D205_ALT            .CGG--A-----T-C-------------C-----G-----------G-----C--G--T--G-----A-A------------------A--T--G---T-A-----T-----T-----T--C--------C-----CA-C-CAGAC............AAC-GTACC--C- 8072
B.JP.01.IMCJ_KR020_1        .--G-AA-----T-----G--C--C--C-----GG----------G--------G--T--A--------A------------------A--TC----T--C-----T-----T--T------A-------C--T--T-------CA.........AATAAT--TAGG---- 7189
G.CI.92.Abt96               A-G-C-C-----T--A--C--C--------G--G-----T-----A-----C--------A--------G--C-----C--C-----A--T--T--G--------T--A-----C-----C--A-----G--T-----AGA--A-AGG......TCTAAC-----T--A- 7457
AB.CM.03.03CM_510_03        .C-GCAC-----T-C---G--C-----A-AG--G-----T--T--C--------------G--------A----------T--G--A-----AT-G--T--C--T--A-----A-----CA-A--T--T-----T--CACA-AA..................T--G--C- 7175
H2_01_AB.CI.90.7312A        .C-GCAC-----T-C---G--C-----A-A---------T--T--C-----C--G-----A--------G----------T--G--A---C-AT-G--C--C-----A-----------CT-G--T--T-----T--CGTAGAC-TA......GGTAAT----GG--A- 8087
H2_01_AB.JP.04.NMC307_20    .C-GC-C-----T-CA--G--C-----A-A---------T--T--C-----C--G-----G--------A----------T--G--A-----GT-G--C--C-----A-----------CA-A--T--T-----T---G--CCG---......AATAT--G------A- 8069
H2_01_AB.JP.07.NMC716_01    .C-GC-C-----T-CA--G--C-----A-AG--------T--T--C-----C--G-----G--------A----------T--G--A-----AT-G--C--C-----A-----------CA-A--T--C-----T-------GA---......AATA-T-G----G-A- 8062
H2_01_AB.JP.08.NMC842_10    .C-GCAC-----T-CA--G--C-----A-AG--------T--T--C-----C--G-----G--------A----------T--G--A-----GT-G--C--C-----A-----------CA-A--T--T-----T------C-G---..........AC---------A- 8086
U.CI.07.07IC_TNP3           C-----A--C--T--A--T-----C--------GG----T--T--A---A---G-----A--G-----A--------------G--A-----T-----T-----T-----------A-----G-----C-------------C---..........AC---T-G----- 7565
U.FR.96.12034               ---------C--T-----T--C--------G--GG----T-----A-------------G-G-------G------------------A-----T-----C--C--T--T-----A--T--CA-G--T-----T--T---C-----A-C.........AA---T-----A- 7580
MAC.US.x.17EC1              --------------------------C-----------------------------------------------------------------------------------------------------------------------------.......------------- 8038
MAC.US.x.251_1A11           ----------------------------------------------------------------------------------------------T-------------------------------------------C---.......------------- 8034
MAC.US.x.251_32H_PJ5        --G-------------------------------------------------------------------------------------------------------------------------------------------C---.......------------- 8036
MAC.US.x.251_BK28           --G------------G------------------------------------------------------------------------------------------------------------------------------C---.......------------- 8020
MAC.US.x.MM142_IVMXX        -------------A--------------------------------------------------------------------------------------------------------------------A-----C---.......------------- 7533
MNE.US.82.MNE_8             --------------------------------------------------------G------------------------------------T--------------------------------------------------.......------------- 7512
MNE.US.x.MNE027             -----------------------------------------------------------------------------------------------------------A------------------------------------.......------------- 7515
SMM.CI.79.SIVsmCI2          G-G-CAT-----T-CA--C------------------A--T---------------T-----C--A----------C-----A--T--T--G--A--C--T--A-----A-------T-----C--T--CA-C----A---C.............TGGC--A- 7191
SMM.LR.89.SIVsmLIB1         G----A---C--T-----T--C--C-----G---G----T--T--C--------G-----A--G-----G--------------G--A-----T-----------T-----T--A--A-----A--T-----T-----------AA--.........AAT--TG-G---- 7109
SMM.SL.92.SIVsmSL92A        ------------T--A--T-----C--C------G-G--T--T---------G-------------------------------A-------GCAA----------G--A--A---T-A------G-T------------A--.........-A---TG-G--C- 7138
SMM.SL.92.SL92B             C-----A-----T--A--T--C--C--C--G--------T--T--G-----C-GG-----A-------C--------------G--A------T-A--T-----------------C--A--------T--T-----------C.........AAC--TG-G--C- 7489
SMM.US.04.G078              G-----A-----T-------------------G---------T--G-----C--G-----T--------G----------T--G--A-------G--A-----T-----T--A--T--C-----------T-----C-----A---.........AAT-----G---- 7334
SMM.US.04.G932              C--------C--T-----T--C-----A--G--GG----A-----A-----C-G------A--------G----------T-----A--A-T--TGA-----T-----T------------G-----G-A-----A-----A--.........AA---T--G--C- 7311
SMM.US.04.M919              G--G--A-------------------------G-----------G-----C-----------------G----------T--G--A-------G--A-----T--T--T--A--T--C-----------G-A---A-A---A-AG-.........AAC-----G---- 7333
SMM.US.04.M922              G-----A-------------------------GG----------A-----C--G--------C---G--------C--T--G--A-------G--A-----T-----T-----T--C-----------G-G-----G-C--GG--C.........-------------- 7333
SMM.US.04.M923              ---G--A-----T-------------------G-----------G-----C-----------------G----------T--G--A-------G--A-----T-----T--A--T--C-----------CG-A-----------CA--.........-AC--TG-G--C- 7320
SMM.US.04.M926              G-----A--C--T--A----------------GG----------G-----C-G---------------G----------T--G--A-------A--A-----T-----T--A--T--CA----------G-T------------------.........AA---TG-G--C- 7329
SMM.US.04.M934              G-----A-----T-------------------------------G-----C-----------------G----------T--G--A-------G--A-----T-----T-----T--C--T--------T---A---ATGA-A-G......GAGAAG---ACT--C- 7305
SMM.US.04.M935              G--G--A-------------------------GG----------G-----C-GG--------------G----------T--G--A-------G--A-----T-----T--A--T--C-----------G-A-----C------C--.........-AC--TG-G---- 7312
SMM.US.04.M940              C--------C--T-----T--------A--------T-----T--A-----C-------T--------G----------T---TCA---A----TGAA--C--T-----T--A--------A--------T-----C-----AT-C.........AA---TG-G--C- 7361
SMM.US.04.M946              G--G--A-----T-------------------------------G-----C-----------------G----------T--G--A-------G--A-----T-----T-----T--C--T--C--T--------T-----G-ATGA---G...CATAAAAAT--AACT--C- 7333
SMM.US.04.M947              ------------T-------------------GG----------G-----C-G---------------G----------T--G--A-------G--A--C-CT---AT--A--T--C-----------G-A----------G-C---.........AAC--TGTG--C- 7317
SMM.US.04.M949              -----AA--C--T-----T--------A-----GG-T--T--T--A-----C--G-----T-------G----------C--GTCA---A----TGAA--C--T-----T--------CA-A--------T-----C-----AA--.........-A---T--G---- 7295
SMM.US.04.M950              G--G-AA-----T-------------------------------A-----C-----------------G----------T--G--A-------G--A-----T-----T-----T--C-----------G-T------ATG-CAG-......AGCAA---TG-G--C- 7320
SMM.US.04.M951              C--------C--T-----T--------A-----G--T-----T--A-----C--G-----T-------G----------T---TCA---A----TGA----C-----AT--------CA-A--------T-----AAT---CA--.........AA-G-T--G--C- 7310
SMM.US.04.M952              G-----A-----T-------------------------------G-----C-----------------G----------T--G--A-------G--A----CT----AT-----T--C--T--C--T--------T-------ATG-C---.........AA-G-TG-G--C- 7309
SMM.US.05.D215              A---CAC-----T-----G--C--C--------------T--T--C-----C-G-----A--G----G----------T-----A-------G--A---------T--T--------A-----------G-A---------C--A--.........AAC--TG-G---- 7310
SMM.US.06.FTq               A-G-C-T--C--T-----T--C--C---C-G--G-----T------------C--G-----G--------A----------T--G--A----A--G--A-----T--A-----A--------T--T--T------T------------.........------AG---A- 7323
SMM.US.11.SIVsmE660_FL10    G--G-AA-----T-------------------------------G-----C-----------------A----------T--G--A-------G--A-----T-----T--A--T-----------G-G---------CCA--.........AAC--TG-G--C- 8047
SMM.US.86.CFU212            A--G-----C--T----------C--------GG-G-----T--A-----C-G-----T--------A-------G--T-----A-------G--A-----------T-----------A-----G-A-----C---TA---.........AA------------- 7305
SMM.US.x.F236_H4            G--G-AA-----T-------------------------------G-----C-----------------A----------T--G--A-------G--A--C--T-----T--A--T---------G-G-----------CA--.........A-C--TG-G--C- 8003
SMM.US.x.H9                 CC-G-AA-----T--------------R----G-----------G-----C-----------------G-R--------T--G--A-------G--A-----------T--A--T--C-----------G-A-----------CA--.........--C--TA-G--C- 7527
SMM.US.x.PBJA               A-GG--A-----T-------------------G-----------G-----C-----------------G----------T--G--A-------G--A-----T-----T--A--T--C-----------G-A-----------CA--.........--C--TG-G--C- 7829
SMM.US.x.PGM53              G--G--A--C--T-------------------G-----------G-----C-G-----T---------G----------T--G--A-------G--A-----T-----T--A--T--C-----------G-T-----C----------.........--C----G---- 7987
SMM.US.x.SME543             G--G--A-----T-------------------------------G-----C-----------------A----------T--G--A-------G--A-----T-----T--A--T-----------G-G-----------CA--.........AAC--TG-G--C- 8060
SMM.US.x.pE660.CG7G         G--G-AA-----T-------------------------------G-----C-----------------A----------T--G--A-------G--A-----T-----T--A--T-----------G-G----------CCA--.........AAC--TG-G--C- 8043
STM.US.89.STM_37_16         A--G--------T--A-----C--C--C------G----T--T--A-------------G-----------------------C-A--------T--T-----T-----T---------A-A-----------T--C----C-A-C.........AAT--TG-G---- 7692
```

```
                                                                                              Rev Responsive Element (RRE) start                Env gp120 end Env gp41 start
MAC.US.x.239                 ATATCACCATGAGTGCAGAGGTGGCAGAACTGTATCGATTGGAATTGGGAGATTATAAATTAGTAGAGATCACTCCAATTGGCTTGGCCCCCACAGATGTGAAGAGGTACACTACT...GGTGGCACCTCAAGAAATAAAAGAGGGGTCTTTGTGCTAGGGTTCTTGGGT 8205
Env                          N__I__T__M__S__A__E__V__A__E__L__Y__R__L__E__L__G__D__Y__K__L__V__E__I__T__P__I__G__L__A__P__T__D__V__K__R__Y__T__T__.__G__G__T__S__R__N__K__R__G__V__F__V__L__G__F__L__G_
A.CI.88.UC2                  -C--T---T-T-------------------------C-----A-----------C--C-----GA----A--A--A-----------C--A--T---AG--AA--A--A--TT-C......TCG-C-C-AG-G--G--C--------T--G--C---------C-------- 8214
A.DE.x.BEN                   -C--T---T-T--------A--------------C-----A---C-------C--C------A----A--A--A-----------C--A--T------CA--GA------T-C......TCAACTC-AGTG--G--C--------T--G--C------------------ 8260
A.DE.x.PEI2_KR_KRCG          GC--T---T-T--------------------A--C--------------------C-----------A--A--A-----------C--A--T---TCA-AA--A-----TT-C......TC--CTC-ACAG--G-----------T--G-----------AG-------- 8204
A.GH.x.GH1                   ----T---T-T--------------------A--C-----A---------------------A----AG-A--A-----------C--A--T--GAGA-A---A--A--TT-C......TCG-CTC-AGTG--G--C--------T--G--------------------- 7680
A.GM.87.D194                 -C--T---T-T-------------------------C-----A--------G--C--C-----GA----AG-A--A------CCG--C--A--T---A-A-A---A--A--TT-C......TCG-CTC-AGTG--G--C--------T--G--C------------------ 7674
A.GM.x.ISY_SBL_6669_85       ----T---T-T--------------------A--C--------------G-----------------AG-A--A-----------C-----T----CA-AA--A--A---T-C......TC--CTC-AGGG---C----G-----T--GC--------------C-A--- 7680
A.GM.x.MCN13                 ----T---T-T--------------------A--C-----A--G-----------------G-----A--A--A-----------C--A--T----CA-AA--A--A---T-C......TC-ACTC-AATG--G--C--G-----T--G--C------------------ 7687
A.GM.x.MCR35                 ----T---T-T--------------------A--C-----A--G-----------------G-----A--A--A-----------C--A--T----CA-AA--A--A---T-C......TC-ACTC-AATG--G--C--G-----T--G--C------------------ 7687
A.GW.86.FG_clone_NIHZ        ----T---T-T----------C---------A--C--------------G-----C-----------A-----A-----------C--A--T---TCA--A-----A---T-C......TC--CTCA-CAG---C---C------T--G---------------C-A--- 7671
A.GW.87.CAM2CG               ----T---T-T--------T---------------C--------------G-----------G-----A--A--A-----------C--A--T---TCACAA-----A---T-C......CC--CTCA-GGG---CC---------T-CG--------------------- 8272
A.GW.x.MDS                   ----T---T-T--------T-----------A--C-----A--------------------------A--A--A-----------C--A--T---CCA-AA--AC-A---T-C......TC--CTCA-GGG---C----------T--G--------------------- 7737
A.IN.07.NNVA                 --G-T---T-T-------G---------C--A--C-----A--------G-----C-----------A--A--A-----------C--A--T---CCA-AA--AC-A---T-C......TT--C-T--AGG---C-A-C------T--G--C------------------ 8252
A.IN.95.CRIK_147             ----T---T-T--------------------A--C---G-A--------G-----------G-----AG-A--A-----------C--A--T---TCA-AA--AC-----T-CT--.....G-C---GGG----C---C------T--GC-C-------AA--------- 8015
A.JP.08.NMC786_clone_41      ----T---T-T--------------------A--C-----A--------------------------AG-G--A--------T--T--A--T---TCACAA--A--A---T-C......TC-ACTC-GGGG--------------T--AC-C------------------ 8262
A.PT.x.ALI                   ----T---T-T--------------------A--C-----A--------G--------------------A--A-----------C--A--T---TCA-AA-G---A---T-C......TC-ACTC-AAGG--G-----------T--G--C--------------A--- 8244
A.SN.85.ROD                  -C--T---T-T--------------------A--CA-------G-----------------G-----A--A--A-----------C--A--T---A-A-AA--A--A---T-C......TC--CTCA-GGG---C---C------T--G--C------------------ 7706
A.SN.86.ST_JSP4_27           ----T---T-T--------------------A--C-----A--------G-----------GA----AG-A--A-----------T--A--T---CCA--A--A--A---T-C......TC--CTC-AGTG--G-----------T--A--C--------------A--- 7690
B.CI.88.UC1                  ----T--------------A----G---------CA---------C---G--C--C-----------A--A--A--G--------T--A--T-----GA-A--A--A--TT-C......TCAACG--AC-G--G-----------T--AA-G-----------------A 8247
B.CI.x.20_56                 ----T--------------A----G----------A---------C---G--C--C-----------A--A--A--G--------T--A--T---A--A-A--A--A--TT-C......TCA-CA--AC-G--G----GG-----T--AA-G------------C----A 8212
B.CI.x.EHO                   ----T--TG----------A---T-----------AA--------C---G--C--C--------------A--A-----------T--A--T---AG-A-A--A--A--TT-C......TCA-TG--AC-G--G-----------T--AC-------------T-----A 8222
B.GH.86.D205_ALT             ----TT-TG----------A---T-----------A---------C---G--C--------------A--A--A-----------C--A--T--------A-GA--A---T-C......TCA-TA-AAC-G--G-----------T--AA-G--------A--------A 8236
B.JP.01.IMCJ_KR020_1         ----T--------------A---T-----------A---------C---G--C--C---C-------A--A--A-----------T--A--T---AG---A--A-----TT-C......TCA-CG--AC-G--G-----G-----T--ACCG-----------T-----A 7353
G.CI.92.Abt96                ------TTT-C-----------------------------A--------G-----------------A--A--------------T--A------A---A------A--TT-C......TC--TG--AC-G-A------------T--A--------------T-----A 7621
AB.CM.03.03CM_510_03         -C--T---T-C-------------------------C-----A--G--------C--C-----GA----A--A--A-----------C--A--T---AC--AAGG------TT-C......TCAACTC-AGGG--G-----------T--A-----------A--------A 7339
H2_01_AB.CI.90.7312A         ----T---T-T-------------------------C-----A---------------------A----AG-G--A--------T--C--A--T---TCA-AA--A--A---T-C......TC-ACTC-GGGG---C----------T--A--C------------------ 8251
H2_01_AB.JP.04.NMC307_20     -C--T---T-C-------------------------C-----A--------G------------A----AG-G--A-----------C--A--T----GA-AA--A--A---T-C......TC--CTC-AGGG---C----G-----T--G--C--------------A--- 8233
H2_01_AB.JP.07.NMC716_01     -C------T-C-------------------------C--------------G-----------------AG-G--A-----------C--A--T----GA-AA--A--A---T-C.........TCTC-AGGG---C-C-------ATA-GC-C------------------ 8223
H2_01_AB.JP.08.NMC842_10     -C--T---T-C-------------------------C--------------G------------A----AG-G--A-----------C--A--T-----A-AA--A--A---T-C......TC--CTC-AGAG---C----G-----T--G-----------------A--- 8250
U.CI.07.07IC_TNP3            ------------------------------------------------C----T-----C---C----------A--A-----------T--T------AG------A--A--TT-C......TCA-CA--AC-G--G--------G--A--G--------T------------ 7729
U.FR.96.12034                -----T-------C-----A-------------------A----A--G-----G-----------G-----A--A-GA-----------T--T------CCA-----A--A--TT-C......TCA-TG--AC-G--G--------G--A--A--------T------------ 7744
MAC.US.x.17EC1               ----------------------------------------------------------------------------------------------------------------------...--------------------------------------------------- 8205
MAC.US.x.251_1A11            -----------------------------------------------------------------------------------------------A------G-------------...--------------------------------------------------- 8201
MAC.US.x.251_32H_PJ5         ----------------------------------------------------------------------------------------------------------------------...--------------------------------------------------- 8203
MAC.US.x.251_BK28            G------------------------------------------G--------------------------------G-----------------------------------------...--------------------------------------------------- 8187
MAC.US.x.MM142_IVMXX         G------------------------------------------------------------------A---------------------------A----------------------...--------------------------------------------------- 7700
MNE.US.82.MNE_8              -----------------------------------------------------------------------------------------------A----------------------...-----------------------------------------------A--- 7679
MNE.US.x.MNE027              -----------------------------------------------------------------------------------------------A----------------------...---------C-------------------------------------A--- 7682
SMM.CI.79.SIVsmCI2           --------T-T-----C-----------GT----CA-------G-----G--C--C-----G-----A--T--A-----------T--A------A-------A--A---T-CT-G...-CAACGC-TG-C--G-----------T--A--------------------A 7358
SMM.LR.89.SIVsmLIB1          --------------------A-T---------------------------------C--------T-----T-----------A--------A---A-------A------T--............---ATG---CCA--------------------------------A 7267
SMM.SL.92.SIVsmSL92A         -------------------A----------------------------------C-----G-----A--------------T--------T---A-------A--------C......ACCA-T--------------G-----------------G------------ 7302
SMM.SL.92.SL92B              ----T--------------A---------T-------------------G-----------G---------------------A----T------C-------A--------C......ACAA-T-------AG-----------A------------------------ 7653
SMM.US.04.G078               -C--T------------------------------A-------------G--C--C-----------A--T------------A-------A---AG------A---------......AC---T--------------------------------------------- 7498
SMM.US.04.G932               ----T--A-----------T-------CC--------T-----------G--C--C-----G-----A--A--A---------A-------A---AG--A------------C......AC--CT-----------C--G-----------------------------A 7475
SMM.US.04.M919               ----------------------------------CA-------G-----------C--------------T------------------------A------G---------C--A...AC-A-TG-------------G------------------------------ 7500
SMM.US.04.M922               -C--T-------------------------------A-------G-----------------------A--T------------------------A----A-GA--------C-AC...-C---T--------------G------------------------------ 7500
SMM.US.04.M923               -------------------A------C----------------G-----------C--------------T--------------C-----------G--A-GA--------C--A...AC---TG-------------------------------------------- 7487
SMM.US.04.M926               -------------------------------------------G-----------C--------------T------------A-----------A-------A--------C--A...ACCA-TG-------------G------------------------------ 7496
SMM.US.04.M934               -------------------T-----------------------G-----------C--------------T------------C-----------AG------------------A...-C-A-TG-------------G------------------------------ 7472
SMM.US.04.M935               -C--T-------------------------------A-------G-----------------------A--T------------------------A---A-G----------C-AC...AC---TG--C----------G------------------------------ 7479
SMM.US.04.M940               -C----------------------GG--GT-A--------A--------G-----C-----G-----A--A-----T------A-------A---AG------A---------.........A-T--------G-----------------------G------------ 7525
SMM.US.04.M946               -------------------------------------------G-----------C--------------T------------------------AG------------------A...-C---TG-------------G-----------------G------------ 7500
SMM.US.04.M947               -------------------A--------G--------------G-----------C--------------T-----G------------------AG---A--A--------C--A...ACC--TG-------------------------C------------------ 7484
SMM.US.04.M949               -C--T-------------------GG--GT-A-----------------G--C--C-----G-----A--A-----T------A-------A---AG---A--A---------......A----T-----------------------------------------A--- 7459
SMM.US.04.M950               -------------------------------------------G---------------------------------------------------AG------------------A...AC-A-TG-------------G------------------------------ 7487
SMM.US.04.M951               -C--T---C---------------GG--GT-A-----------------G-----C-----G--------A-----T------A-------A---AG------A-------A--G-GGTACCTCA--------------------------------------------- 7480
SMM.US.04.M952               -------------------------------------------G-----------C---------------------------------------AG------------------A...AC-A-TG-------------G------------------------------ 7476
SMM.US.05.D215               ----------------G--C-------------C--------------G-----C-----------A--------------------T-----CAG------------T--C......ACACAG---------C-A-G------A-----------------------A 7474
SMM.US.06.FTq                --------T-C-------------------CA-------G-----G-----C-----------A--T--------------T--A--T--C-----A--A--A--TT-CT--...-TAACG--AC-G--G-----------T--G--C-----T-----------A 7490
SMM.US.11.SIVsmE660_FL10     -------------------------------------------G-----------C--------------T------------------------AG---A-GA--------C--A...AC---TG-------------G------------------------------ 8214
SMM.US.86.CFU212             -------------------------------------------------G-----C-----------A---------------A-A-----A---A----A-----------C---...AC-ACT---------GC---------------------------------A 7472
SMM.US.x.F236_H4             -------------------------------------------G-----------C------A-------T------------------------AG---A-GA--------C--A...AC---TG-------------G------------------------------ 8170
SMM.US.x.H9                  -------------------------------------------G-----------C--------------T--------------C---------AG---A--A--------C--A...AC---TG-------------------------------------------- 7694
SMM.US.x.PBJA                -------------------------------------------G-----------C--------------T----------C---C---------AG---A--A--------C--A...AC---TG-------------------------------------------- 7996
SMM.US.x.PGM53               -C-------------------------------CA-------G--------------------------T------------------------AG---A-G---------A--.......--TG-------------G------------------------------ 8151
SMM.US.x.SME543              -------------------------------------------G-----------C--------------T-----------------------------A-GA--------C--A...AC---TG-------------G------------------------------ 8227
SMM.US.x.pE660.CG7G          -------------------------------------------G-----------C--------------T------------------------AG---A-GA--------C--A...AC---TG-------------G------------------------------ 8210
STM.US.89.STM_37_16          -C-----TGCA------------------------------------------------C-----------A--A--------------------A--GA----------A-----A......AC-A-T----------C---G-----------C------------------ 7856
```

Draft of 30 September 2015, not for public release
hiv2dna_main.tex r8876–10546


```
MAC.US.x.239          TTTCTCGCAACGGCAGGTTCTGCAATGGGCGCGGCGTCGTTGACGCTGACCGCTCAGTCCCGAACTTTATTGGCTGGGATAGTGCAGCAACAGCAACAGCTGTTGGACGTGGTCAAGAGACAACAAGAATTGTTGCGACTGACCGTCTGGGGAACAAAGAACCTCCAGAC 8375
Env                   _F__L__A__T__A__G__S__A__M__G__A__A__S__L__T__L__T__A__Q__S__R__T__L__L__A__G__I__V__Q__Q__Q__Q__Q__L__L__D__V__V__K__R__Q__Q__E__L__L__R__L__T__V__W__G__T__K__N__L__Q__T
A.CI.88.UC2           --------G--A-----------------------CC-------T-A----------G------C---C-------------------------------A-A-----------------C----------------------G--A--T------G- 8384
A.DE.x.BEN            --------G--A---------------------CG---CC-------T-A--C--------G------C---C-------------------------------A-----------------A----------------------G--A---------G- 8430
A.DE.x.PEI2_KR_KRCG   ----------A-----------------------C--------T-G-----TC--G--CT--AC-G-------------------------------------------------------A----------------------A--T--------- 8374
A.GH.x.GH1            --------G--A-----------------------CC-------T-G----------G------C---C-------------------------------A-----------------A----------------------G--A--T------G- 7850
A.GM.87.D194          --------G--A-----------------GC-----C--------T-G----------G------C---C-------------------------------------------------A-------T----------------G--A--T------G- 7844
A.GM.x.ISY_SBL_6669_85 ------A-G--A------G---------G--------TC-------T-G--------T--G------CCG-------------------------------------------------A----------------------T--A--------G- 7850
A.GM.x.MCN13          ----------A-----C-----------------C--A------T-G--------T--G------C---C-------------------------------G---A----------------------A--T------G- 7857
A.GM.x.MCR35          ----------A-----C-----------------C--A------T-G--------T--G------C---C-------------------------------G---A----------------------A--T------G- 7857
A.GW.86.FG_clone_NIHZ ----------A----------C-----------C--------T-A--------T--G----------C-------------------T-------------A----------------------A--T------G- 7841
A.GW.87.CAM2CG        ------A----A-----AGT---------A------C--------T-A--C----T--G----------C-------------------A-------------------------------------A-TT-----G- 8442
A.GW.x.MDS            ------A----A-----G-----------------C--------T--------T--G------C---C-------------------------------------------------A----------------------A--T------G- 7907
A.IN.07.NNVA          --------C--A----A-----------------C--------T--------C----G----C---C-------------------------G--------A----------------------G--A--T------G- 8422
A.IN.95.CRIK_147      ----------A-----------------A-----C--------T--------C----G----C---C-------------------------------A----------------------G--A--T------G- 8185
A.JP.08.NMC786_clone_41 ------A----A-----AG-----------------C--------T-G--------T--G-----------------A-----------A----------------A----------G------A--T------G- 8432
A.PT.x.ALI            --------G--A-----------------A----AG-T--A------T-T--------T--G----------C-------------------------------G---A----------------------G--A--T------G- 8414
A.SN.85.ROD           ----------A-----------------------CC----CG--T-G----------G------C---C-------------------------------------C----------------------G--A---------G- 7876
A.SN.86.ST_JSP4_27    ------A-G--A----AG-----------------C--------T-G--------T--G----------C-------------------------------------A----------------------A--T------G- 7860
B.CI.88.UC1           C----T----T-------------------AA---C--------T-A----------G------C-----------------G----------C----------A-------G---C------G----------G--A----------- 8417
B.CI.x.20_56          -----G-----------------------A----C--------T-A----------G----------------------G----------C-----C----A-------G---C------G----------G--G--A--------T- 8382
B.CI.x.EHO            --C--T--G-----------------------C------T--T-G----------G------C-----------------G----------CG----------A----------C------G----------G--G--A--------G- 8392
B.GH.86.D205_ALT      --C--T----T-------------------AA---T--------T-A----------G------C---C----------G----------CTG---------A-------G---C------G----------G--G--A--------G- 8406
B.JP.01.IMCJ_KR020_1  -----T-----------------------------C------T--T-G--C-------G-----------------------G----------C----------A-------G---C------G----------G--G--A--------G- 7523
G.CI.92.Abt96         -----T----T-----------------A-A-----------------T-G----------G----G-------------------------T----------------G---A----------------------G--G--A---------- 7791
AB.CM.03.03CM_510_03  --C--T--G-----------------------A----C--------T-A----------G------C-----------------G----------C----------A-------G---C------G----------G--G--A--------G- 7509
H2_01_AB.CI.90.7312A  ------A-G--A----AG-----------------C--------T-G--------T--G------C---C-------------------------------A-----------------A----------------------A--T------G- 8421
H2_01_AB.JP.04.NMC307_20 ------A-G--A-----GGT---------------C--------T-A----------G------C-----------------------------------A----------------A----A--T----------------A--T------G- 8403
H2_01_AB.JP.07.NMC716_01 ------A-G--A----AGT---------------C--------T-G----------G------C-----------A-----------------------A----------------A----A--T----------------A--T------G- 8393
H2_01_AB.JP.08.NMC842_10 ------A-G--A-----AGT---------A------C--------T-A----------G------C-----------------------------------A----------------A----A--T----------------A--T------G- 8420
U.CI.07.07IC_TNP3     --------T---------A---------A------T--------T-T--------T--G----------------------A-----T----------A----------GC------G----------G--T--A---------- 7899
U.FR.96.12034         --------C-------------------T----------------T-A----------G------C-----------------------------------A----------A------G----------T--A---------- 7914
MAC.US.x.17EC1        ---------------------------------------------------------------------------------------------------------------------------------------------------------- 8375
MAC.US.x.251_1A11     ----------------------------------------------------------G-----------------------------T-------------------------------------------------------------- 8371
MAC.US.x.251_32H_PJ5  ----------------------------------------------------------A----G-------------------------------------------------------------------------------------- 8373
MAC.US.x.251_BK28     ---------------------------------------------------------------G-------------------------------------------------------------------------------------- 8357
MAC.US.x.MM142_IVMXX  ------------------------------------------CG--------------G-------------------------------------------------------------------------------------- 7870
MNE.US.82.MNE_8       ---------------------------------------------------------------G-------------------------------------------------------------------------------------- 7849
MNE.US.x.MNE027       ---------------------------------------------------------------G-------------------------------------------------------------------------------------- 7852
SMM.CI.79.SIVsmCI2    -----T--G---------------------A----C--------T-T--------T--G-------------------------------------------------C----------C----------G----------T--A---------- 7528
SMM.LR.89.SIVsmLIB1   ----------A-----------------A----C---------G----------G----G-------------------------------------------------------------------T--A---------G- 7437
SMM.SL.92.SIVsmSL92A  -----T--G---------------------A----------T----------G-------------------------G----------C----------------------------------T--A---------- 7472
SMM.SL.92.SL92B       --------G---------------------A-----C-A------T--------G-------------------A-----------------------------------------------T--A--T--T----- 7823
SMM.US.04.G078        -----T--G---------------------A----C--------T-G--------G----G-------------------------T----------------T----------T--A---------- 7668
SMM.US.04.G932        -----T--G-------------------------------------------------G----G-------------------------------------------------------T--A---------- 7645
SMM.US.04.M919        --------G-----------------------C-------T-----------G----G-------------------------T----------------------------T--A---------- 7670
SMM.US.04.M922        --------G--A--------------------C-------T-----------G----G-------------------------T----------T----------C--A---------- 7670
SMM.US.04.M923        --------G--A--------------------C-------T-G---------G----G-------------------------T----------------------------T---------- 7657
SMM.US.04.M926        ------------------------------------C-------------T--G------C-------------------------------------------------------A---------- 7666
SMM.US.04.M934        ------------------------------------------------T--G------C-------------------T-------------------------------------A---------- 7642
SMM.US.04.M935        --------G--A--------------------C-------T-A---------G----G-------------------------T----------G----------T----------C--A---------- 7649
SMM.US.04.M940        --------G-----------------A----C--------T--C-------G------C-------------------------T----------------------------T--A---------- 7695
SMM.US.04.M946        ------------------------------------------------T--G-------------------------------------------------------------A---------- 7670
SMM.US.04.M947        --------G--A--------------------------------T-G---------G----G-------------------------T----------------------------T---------- 7654
SMM.US.04.M949        -----T-------------------T--A---T------T---T--C---------G------C-------------------------T----------T----------G--T--A---------- 7629
SMM.US.04.M950        ------------------------------------------------T--G------C-------------------------------------------------------A---------- 7657
SMM.US.04.M951        --------G-----------------A----C----------T--C------G------C-------------------------T----------------------------T--A---------- 7650
SMM.US.04.M952        ------------------------T--------A----------------T--G------C-------------------G----------------------------------A---------- 7646
SMM.US.05.D215        ----------A-----------------------C--------T-G----------G-------------------------------------------------------T--A---------- 7644
SMM.US.06.FTq         ------------------------A----------T----------G----G-------------------------A----------C----G----------G--T--A-----T----- 7660
SMM.US.11.SIVsmE660_FL10 --------G--A-----------------C--------T-G----------G----G-------------------------T----------------------------T---------- 8384
SMM.US.86.CFU212      ------------------------------------T-------T--G------C-------------------------------------------------------T--A---------- 7642
SMM.US.x.F236_H4      --------G--A-----------------CG-------T-G----------G----G-------------------------T----------------------------T---------- 8340
SMM.US.x.H9           --------G--A-----------------------C--------T-G----------G----G-------------------------T----------------------------T---------- 7864
SMM.US.x.PBJA         --------G--A-------------A---------CG------T-G----------G----G-------------------------T-------------------------G-T---------- 8166
SMM.US.x.PGM53        --------G--A-----------------------CG---T-G-T------------G----G-------------------------T-------A----------T----------T---------- 8321
SMM.US.x.SME543       --------G--A-----------------------C--------T-G----------G----G-------------------------T----------T----------T---------- 8397
SMM.US.x.pE660.CG7G   ---T----G--A-----------------------C--------T-G----------G----G-------------------------T----------------------------T---------- 8380
STM.US.89.STM_37_16   --------G--A-----------------A-----C---------G----------G--------A-------------------------------------------------C---------- 8026
```

```
                                                                                                                        RRE end
MAC.US.x.239              TAGGGTCACTGCCATCGAGAAGTACTTAAAGGACCAGGCGCAGCTGAATGCTTGGGGATGTGCGTTTAGACAAGTCTGCCACACTACTGTACCATGG............CCAAATGCAAGTCTAACACCAAAGTGGAACAATGAGACTTGGCAAGAGTGGGAGCGAAAGG 8533
Env                       __R__V__T__A__I__E__K__Y__L__K__D__Q__A__Q__L__N__A__W__G__C__A__F__R__Q__V__C__H__T__T__V__P__W__.__.__.__.__P__N__A__S__L__T__P__K__W__N__N__E__T__W__Q__E__W__E__R__K__
A.CI.88.UC2               A--A--------T-----------------------A-----A---T-A---------A-------G--G-------------------T---............GT-----ACTC-T-------T-GA---------AT---A-----G--A-----AAA-C-A- 8542
A.DE.x.BEN                A--A--------T-----------C-----C-T-----A-----A---T-A-------------------G-------------------G---............GT-----ACTC-T--T-G--TG-C-----A---AT---A-----G---------AA-C-A- 8588
A.DE.x.PEI2_KR_KRCG       A--A--------T--------A---C---G-----------G---A---T-A-----------A---------------T----C-------T---............GA----AAC--CA--GT---TG-T--------CAT---G-----G--A-----A-A-C-AA 8532
A.GH.x.GH1                A--A--------T-----------------------A---T-A---T-A-----G-----------G--G--------------------------............GT-----ATTC-T--T-G--TG-C---------AT---A-----G--A-----AAA-C-A- 8008
A.GM.87.D194              A--A--------T--------A--------------A-----A---T-A-----------------G--G--------------------------............GT-----ACTCCT-------TG-C---------AT---A-----G--A-----AAA-CGA- 8002
A.GM.x.ISY_SBL_6669_85    A--A--------T--T---------C--GCA----------GA--A---T-A--------------------------------------------............GT-----AC-CCT-------TG----------CAT---A--------A-----A-AC--AA 8008
A.GM.x.MCN13              A--A--------T-------------------------------A--A---T-A-----------A------------------------------............GT-----AT-CCT----G--TG----------AT---G--------A-----AG-C--AA 8015
A.GM.x.MCR35              A--A--------T-------------------------------A--A---T-A-----------A------------------------------............GT-----AT-CCT----G--TG----------AT---G--------A-----AG-C--AA 8015
A.GW.86.FG_clone_NIHZ     A--A--------T--A---------C------------------A---T-A---------------------------T-----------------............GT-----AT-CCT-G-----TG-T---------AT---G-----G--A-----A-A---A- 7999
A.GW.87.CAM2CG            A--A--------T--A---------C-------T-----A-----A---T-A--------------------------------------------............G------AGTC-T-------TG-T---------AT---A-----G--A-----A-A----- 8600
A.GW.x.MDS                A--A--------T--A---------C----------------A--A---T-A-----G--------------------------------------............GT----AGTTCCT--GA---TG-T---G-A---AT---G-----G--A-----A-A----- 8065
A.IN.07.NNVA              A--A--------T--A-----A---C------C-------------AG--T-A--------------------------------------------............GA-----AT-CCT--GT---T--T-----A---AT---G--CA----A-----A-A----- 8580
A.IN.95.CRIK_147          A--A--------T--A-----A---C-------T-------------A---T-A--------------------------------------------............GT-----ATTCCT-G-----TG-T-----A---AT---A---T----A-----A-A---A- 8343
A.JP.08.NMC786_clone_41   A--A--------T--T-----A--------------------A--A---T-A-----G--C--------G---------------------------............GT-----AC--CT-GT----TG-T---G-T--CAT---A------C-A-----A-ACC-AA 8590
A.PT.x.ALI                A--A--------T--------A-------------------G---A---T-A-----------A---------------------------------............GT----AACTCCT---A---TG-T---G----CAT---G-----------------A-A-C-A- 8572
A.SN.85.ROD               A--A--------T--A----------C--C------------G---A---T-A--------------------------------------------............GTT----ATTCCT--G----TG-C---G-----AT---G-----G--A-----AAA-C-A- 8034
A.SN.86.ST_JSP4_27        A--A--------T--------A------------------------A--A---T-A--------------------------------------------............GT-----AC-CCT----G--TG-T---------CAT---A-----G--A--------A-CGAA 8018
B.CI.88.UC1               A--A-----------------A---C------------A-TA--A---T-G-----------T-----G--G--T--T-----G-------------........................A--C---C------G-C---G-A--CAT---A------C-------AAAGCGA- 8575
B.CI.x.20_56              A--A-----------------A---C------------A-T---A---T-G-----------T-----------G----T-----G-------------........................A-TC---C-----C--C----CAG-CAT---A-----GC-A------GAG---- 8540
B.CI.x.EHO                A--A-----------------A--TC-C--A-------A--A--A---T-A-----------T--C-----G----T-----G-------------............GT-----A-TCC--T-AG---G-C-----T--CAT---A------C-A------A-GC-A- 8550
B.GH.86.D205_ALT          A--A-----------------A---C-------T----A--A--A---T-A-----------T--C-----G----T-----G-------------........................A--CC--C--------T-----T--CAT---A------C-------AAAGC-A- 8564
B.JP.01.IMCJ_KR020_1      A--A-----------------A---C----A-------A-TA--A---T-G---------C--T--C-----G----T-----G-------------..................CAA--A-T-C--T-----C---G---CAT---A-----GC---------AACC--- 7681
G.CI.92.Abt96             ---A-----------------A--TC------------A-G---A---T-A-----------------G-----------------T---GATGCCCTTGGTG-T---AA--CAT-GGAG---C-A-----T---AT---A--------A-----AAAGC--A 7961
AB.CM.03.03CM_510_03      A--A--------------------TC-------T----A--R--A---T-A-----------------G--G----------------R------............GT-----A-TCCT----G--T--C---G-----AT---A-----GC-A------GA----- 7667
H2_01_AB.CI.90.7312A      A--A--------T--T-----A-------------------A--A---T-A-----------------G---------------------------............GT-----AC--CT-G-----TG-T---G----CAT---G------C-A-----AAA-C-AA 8579
H2_01_AB.JP.04.NMC307_20  A--A--------T--T-----A---C---------------A--A---T-A-----------A---------------------------------............GT-----AC--CT-GG----TG-C---G----CAT---G------C-A------AA-C-AA 8561
H2_01_AB.JP.07.NMC716_01  A--A--------T--T-----A---C---------------A--A---T-A-----------A---------------------------------............GT-----AC--CT-GG----TG-C---G----CAT---G------C-A------AA-C-AA 8551
H2_01_AB.JP.08.NMC842_10  A--A--------T--T-----A---C---------------A--A---T-A-------C--------------------------------------............GG-----AC---T-G-----T--A---G----CAT---G------C-A-----AAC-C-AA 8578
U.CI.07.07IC_TNP3         ---A-----------------A---C-------T-----AAGT--A-----A-----------------G--G--------T--------------------............GT----AATTC-T--CA---TG-C---------CAT---G-------------AGT---A- 8057
U.FR.96.12034             ---A-----------------A---C-------T-----AAGT--A-----A-----------C-------G--------T--------G------............AT---C-AC-CAT-------T--T---G-T---AT---G--------A------GA----- 8072
MAC.US.x.17EC1            ---------------------------------------------------A----------------------------------------------............-----------------T------------------------------------------- 8533
MAC.US.x.251_1A11         ---------------------A----------------------------------------------------------------------------............---------------------------------------------------------------- 8529
MAC.US.x.251_32H_PJ5      --------------------------------------------------------------------------------------------------............----------------------G-C-----------T-------------------------- 8531
MAC.US.x.251_BK28         ---------------------------------------------A----------------------------------------------------............----------------------G-C-----------T-------------------------- 8515
MAC.US.x.MM142_IVMXX      -------T-------------------------------------A----------------------------T-----------------------............----------------------G-T-------------------------------G---- 8028
MNE.US.82.MNE_8           ---A-----------------------------------------A-------------A------------T-------------------------............-------------A-----------T-------------------------------------- 8007
MNE.US.x.MNE027           ---A-----------------------------------------A-------------A------------T-------------------------............------------------------T-------------------------------------- 8010
SMM.CI.79.SIVsmCI2        ---A-----------------A---C-------T-----A-G---A-------------A---------------------G-----T---............GA----AAC-C-T-GCA---TG-T-----T---AT------------A-----AAT----A 7686
SMM.LR.89.SIVsmLIB1       A--A--------T-----------------------A-----A---T-A-----------T----G-------T-----------GAG---........................AA--CAT--CA---T--T---G-G---AT------------A-----AA-G---- 7595
SMM.SL.92.SIVsmSL92A      ---A--------T-----------C-------T-----A-----A---T---------C--A----G--G-------T--C------------............-------ATTC-T-G------G-T-----T---AT-------------------AA-G---- 7630
SMM.SL.92.SL92B           ---A-----C-----------A---C-------T-----A-----A---T-A---------A----G--------------------T---............-------AC-----TGTC---G-C---------AT---A--------A------AA----- 7981
SMM.US.04.G078            ---A--------T-----------C-------T-----A-----A---T------------A-----------------C------------............-------AT--CT-G--G--T--T-----T---AT---C-----------------A------- 7826
SMM.US.04.G932            ---A--------T-----------C-------T-----A-----A---T-A----------A-------G-----------------------............-------AT-C-T---T----G-C-----T--CAT-----------------------AAA----- 7803
SMM.US.04.M919            ---A--------T-----------C-------T-------------A---T-A---------C--T-----G--G------------------------............--------CTCGT-GGTG--T-GT---------AT----------A-A------A--C--- 7828
SMM.US.04.M922            ---A--------T-----------C-G-----T-------------A---T-A-----------T-------G-----------------A---............----------CAT-GCTG--T--T---G-A---AT-------------------AA------ 7828
SMM.US.04.M923            ---A--------T-----------C-------T-----A-----A---T-A-----------T----G--G--T--------------------............-------AC-C-T-G-----TG-C---------AT-------------------AAA-C--- 7815
SMM.US.04.M926            ---A-----------------A-----------T-----A-----A---T-------------A--C----G-----------------A---............-------AT-----G-T----G-C---------AT-------------------A------- 7824
SMM.US.04.M934            ---A-----------------A-----------T-----A-----A---T-------------A----G--G-----------------A---............-------GT-----GGTG---G-C---------AT-------------------A------- 7800
SMM.US.04.M935            ---A--------T-----------C-------T-------------A---T-A-----------A-------G--------------G------............-------A--CAT-GGTG--T--T----G---AT-------------------AA-GC--- 7807
SMM.US.04.M940            ---A--------T-----------C-------T-----A-----A---T-------------A------------------C------------............-------AT-CCT---T---TG-T-----T---AT---C---------------A------- 7853
SMM.US.04.M946            ---A-----------------A-----------T-----A-----A---T-------------A----G--G-----------------A---............-------AT-----GGT----G-C---------AT------------A-----A------- 7828
SMM.US.04.M947            C--A--------T-----------C-------T-----A-G---A---T-A-----------T-------G--------------------------............-------A--CAT-G--G--T--T---------AT-------------------AA--C--- 7812
SMM.US.04.M949            ---A--------T-----------C-------------A-----A---T-------------A------------------C------------............-------AC---T-G-T---TG-T-----T---AT---C-----GC-------A---C--- 7787
SMM.US.04.M950            ---A-----------------A-----------T-----A-G---A---T-------------A-------G-----------------A---............------CAGT-----G-TG---G-A---------AT-------------------A---C--- 7815
SMM.US.04.M951            ---A--------T-----------C-------T-----A-----A---T------------------------------C------------............-------AT---T-G-----TG-T-----T---AT---C---------A-----A------- 7808
SMM.US.04.M952            ---A-----------------A-----------T-----A-----A---T-------------A----G--G-----------------A---............-------AT-----GGT----G-C---------AT-------------------------C--- 7804
SMM.US.05.D215            ---A--------T-----------C-------T-----A-----A---T-A-----------T----G-----------------C--G-----............-------AT-CAT-G--------T---------AT---A----------------AA--C--- 7802
SMM.US.06.FTq             C--A-----------T-----A---C--C----T-----AA----A---T-A-----------A--C--G----------------------G---............GTC----ATTCA--CCA----G-C---------AT------------A------AA-C--- 7818
SMM.US.11.SIVsmE660_FL10  ---A--------T-----------C-G-----T-------------A---T-A-----------T----G--------T------------------............-------A--CAT-GGTG--T--T----G---AT-------------------AA--C--- 8542
SMM.US.86.CFU212          A--A-----------------------------------A-----A---T-A-----------------G--G---------------CGA---...............GCC----A-G-G------G-T-----T--CAT---C----------------A------ 7800
SMM.US.x.F236_H4          ---A--------T-----------C-G-----T-------------A---T-A-----------T----G--------T------------------............-------A--CAT-GGTG--T--T---------AT-------------------AA--C--- 8498
SMM.US.x.H9               ---A--------T-----------C-------T-------CRN-A---T-A-----------T----G--G--------------K----------............-------AC-CAT-G-----T-RC---------AT--Y----------------AAA-C--- 8022
SMM.US.x.PBJA             ---A--------T-----------C-------T-------------A---T-A-----------T----G--G---------------------A--............-------AC-CAT-G-----T--C---------AT-------------------AAA-C--- 8324
SMM.US.x.PGM53            ---A--------T-----------CG------T-------------A---T-A-----------T----G--G--------------------------............--------CTCAT-GGTG--T--T---------AT------------A-----AA--C--- 8479
SMM.US.x.SME543           ---A--------T-----------C-G-----T-------------A---T-A-----------T----G--------T------------------............-------ACTCAT-GGTG--T--T---G-----AT-------------------AG------ 8555
SMM.US.x.pE660.CG7G       ---A--------T-----------C-G-----T--------C--A---T-A-----------T----G--------T------------------............-------A--CAT-GGTG--T--T----G---AT-------------------AA--C--- 8538
STM.US.89.STM_37_16       ---A--------T--------A---C-------T-----A-----A---T---------------------G-----------C------------............-------ATTC-T-GGT---GG-T---------AT---A-----G--A------A------ 8184
```

```
MAC.US.x.239              TTGACTTCTTGGAAGAAAATATAACAGCCCTCCTAGAGGAGGCACAAATTCAACAAGAGAAGAACATGTATGAATTACAAAAGTTGAATAGCTGGGATGTGTTTGGCAATTGGTTTGACCTTGCTTCTTGGATAAAGTATATACAATATGGAGTTTATATAGTTGTAGGA 8703
Env                       V__D__F__L__E__E__N__I__T__A__L__L__E__E__A__Q__I__Q__Q__E__K__N__M__Y__E__L__Q__K__L__N__S__W__D__V__F__G__N__W__F__D__L__A__S__W__I__K__Y__I__Q__Y__G__V__Y__I__V__V__G_
A.CI.88.UC2               -CCG--A-C-A--G-C------C-GTCAAAG-T----A-------------------A-----T----------------A--A--------------C--------C--------TT-AA-C--C-----C--A-----T--------G-----------A------ 8712
A.DE.x.BEN                -CCG--A-C-A--G-C------C-GTCAAAGTT----A--A--C--------------A-----T----------------A--A-----------A-TC-------C---------T-AA-C--C---G-C--------T-----------GC-------A--G--- 8758
A.DE.x.PEI2_KR_KRCG       CCCG-GA-C-A--G-C------C-GTAGATCGT-----C----------C-----------A--T--------GC-------A--A---------------T--------C--------TT-AA-C--C-----T--------T--G-----------G--A-AA----- 8702
A.GH.x.GH1                -CCG--A-C-A--G-C------C-GTCAAAGTT----AC-------------------A-----T-----------------C-A---------------C--------C--------TT-AA-C--C-----C--------T-------------------A------ 8178
A.GM.87.D194              -CC---A-C-A--G-C------C-GTCAAAGTT----AC-------------------A-----T---------C-------AC-A---------------C--------C--------TT-GA-C--C-----C--A-----T-------------------A------ 8172
A.GM.x.ISY_SBL_6669_85    -CCG----C-A--G-C------C-GT-AGAGTT----AC---------C--G-----A-----T--------GC-G------C-A---------------T-------------------T-AA-C--C-----C--------T-----------CATG-----A------ 8178
A.GM.x.MCN13              -CCG-GA-C----G-C------C-GTCAA-AAT----AC-A----------G---------T---------C---------A---------------T-----T--C---------T-AA-C--C-----C--------T-------------------A-AA----- 8185
A.GM.x.MCR35              -CCG-GA-C----G-C------C-GTCAA-AAT----AC-A----------G---------T---------C---------A---------------T---A-T--C---------T-AA-C--C-----C--------C-------------------A-AA----- 8185
A.GW.86.FG_clone_NIHZ     -CCG--A-C----G-C------C-GTCAAAGT-----AC-A----------G-----A-----T--------GC-------A--A---------------T--AC------C-----T-CA-C--C---G-C-G------T---------------G----A------ 8169
A.GW.87.CAM2CG            -CCG--A-C----G-C------C-GTCAAAGTT----A--------C----G-----A-----T---------C-------A--A--C-A---------T--AC---C---------T-AA-C--C-----C-GC-----T-------------------A------ 8770
A.GW.x.MDS                -CCG--A-C----G-C------C-GTCAAAAGT----A----------C--G--G--AC-------------GC-------A--A--C---------A-T---------------------T-AA-C--C-----C--------T---------------C-----A------ 8235
A.IN.07.NNVA              -CCG-CA-C----G-C------C-GTCAAAGTT----AC---------C--G--G--A-----T---------C-------A--A--------------A-T--------C---------T-AA-C--C-----C--------T---------------C-----A------ 8750
A.IN.95.CRIK_147          -CCG--A-C----G-C------C-GTCAAAGTT----AC---------C-----G--A----G---------GC-T-----A--A---------------T--------C---------T-AA-C--C-----C--------T--G----------GC-----A------ 8513
A.JP.08.NMC786_clone_41   -CAGT---C-A--G-C------C-GT-AAAAAT----AC-A--------C-T------A-------------------------A--A---------------T--------C---------T-AA-C--C-----C-GA-----T--G----------T------AA---C- 8760
A.PT.x.ALI                -CCGT-A-C-A--G-C------C-GT-AA-AGT----ACG------------G-----A-----T-C-------C-------A--A---------------T---AC---C---C-----T-AA-CG-C---G-C--------T-------------------A-A------ 8742
A.SN.85.ROD               -CCG--A-C----G-C------C-GTAAAAGTT----AC------------G--------A--T---------C-------A--A---------------A-T--------------------T-AA-C--C---G-C--------T-----------GCT----A-A----C- 8204
A.SN.86.ST_JSP4_27        -CCG-AA-C-A--G-C------C-GT-AAAGTT----AC---------C--G-----A----------------C-------A--A---------------T--------C--------TT-AA-C--C-----C--A-----T--G-------------------A------ 8188
B.CI.88.UC1               -CA-T--------T-C----------------GT----A-----T--G--A-------A-G------------------G--A--A---------------T-----T--------------T-CA-C--C-----GGCA--C--CAGG-TA---C-A---G----A-C---- 8745
B.CI.x.20_56              -CC---A------T-C-----------CAAT-AT----A--------G--A--G--G--A-----T--------GC---G--A------CAT--------T---A-T---------------T-CA-C--C-----GGCA--C--CAGG-TA---C-G---G----A------ 8710
B.CI.x.EHO                -CCG---------T-C-----------AAAT-A-----A--------G--A--------A-----------------G-----G--A--A---CAA-----A-T--CA-T---------------T-CA-C--C-----GGCA--C--CAGG-TA---T-A--------AA----- 8720
B.GH.86.D205_ALT          -CC-T--------T-C---------G--T--GT----A-----T--G--A-----G--A-----------------------G--AA-A---------------T-----T----------------CA-C--C--------A-----C--T-TA---C-A--------A-C---G 8734
B.JP.01.IMCJ_KR020_1      -CCG---------T-------------AAAT-AT----A-T------G--A--------AG---------CA-G-----G--AC-A--CCAA---------T---A-T--------------T-CA-C--C-----GC------C---ATA---T-A---G--A-A-----G 7851
G.CI.92.Abt96             -CA----T-----G--C--C--C---CGG--AT----A--A-------------G-----A--A--T--------G-----G--A--A---------------A---------------------AA-C--C---G-T--A--CG-TT-C-TA--CC-----G----G-C---- 8131
AB.CM.03.03CM_510_03      -CCG-GA-C-A--G-C------C-GTCAAA-TT----AC----G-----------------A-----T-----------------AC-A---------------C---A----C--------TT-AA-C--C-----C-GA-----T-------------------A-C---- 7837
H2_01_AB.CI.90.7312A      -CCG-GA-C----G-C------C-GT-AAAGT-----AC----------C--G-----A-------------------------A--A---------------T--------C--------TT-A--C--C---G-C--A-----T--G-----------------A------ 8749
H2_01_AB.JP.04.NMC307_20  -CCG--A-C-A--G-C------C-GT-AAAGTT----AC----------C--G-----A-------------------------A--A---------------T--------C--------TT-GA-C--C-----C--A-----T--G---------CTC-----A------ 8731
H2_01_AB.JP.07.NMC716_01  -CCG--A-C-A--G-C------C-GTATAAGTT----AC----------C--G-----A-------------------------AC-A---------------T--------C--------TT-GA-C--C-----C--A-----T--G---------CTC-----A------ 8721
H2_01_AB.JP.08.NMC842_10  -CCG--A-C-A--G-C------C-GT-AAAGT-----AC---G-----C--G-----A-------------------------AC-A---------------T--------C--------TT-AA-G--C-----C--A-----T--G---------CTC-----A------ 8748
U.CI.07.07IC_TNP3         -AA-AGA------G--------T----T-T-AT-G--A--A--T-------------A-----T------C-G--------A------------------T-----T--------A--T---A-C--C---G----A---G--T-C-TA---T-----G----A-C---- 8227
U.FR.96.12034             -GA---A---------------C---CAG--AT----A-CA--T-----A-----------A------------------G--A--A---A------A-C-----T--C---------T--CA-C--A---G-T--A---G--T-C-TA---C-G---G----A-C---- 8242
MAC.US.x.17EC1            ------------------------------------------------------------------------------------------------A-------------------------------------------------------------------------- 8703
MAC.US.x.251_1A11         -------------G------------------------------------------------------------------------A------------------------------------------------------------------------------------ 8699
MAC.US.x.251_32H_PJ5      -------------G-----------------------A----------------------------------------------------------------------------------------------------------------------------------- 8701
MAC.US.x.251_BK28         -------------G-----------------------A---------------------------------------------------------------------------------------------------------------A-----G----------- 8685
MAC.US.x.MM142_IVMXX      -------------G-C---------G-----------A-------------------------------------------------------------------------------------A-----------------------------A--------A-------- 8198
MNE.US.82.MNE_8           -------------G-----------G----TT-----A--------------------A------------------------------------------------------------------------------G------------C---------------------- 8177
MNE.US.x.MNE027           -------------G-----------G----T------A--------------------A------------------------------------------------------------------------------G------------C---------------------- 8180
SMM.CI.79.SIVsmCI2        -CAGAGATC----G-C------CT---AATCAT-G---------------A--------A-----T-------G--G-----A--A-------------------------C--------TT--A-C------G-C--A-----CT-TATA---T-----G----A------ 7856
SMM.LR.89.SIVsmLIB1       ---------C-T--G-C-----------CAA--AT----A--A-----------------------A------------------AC-----------------A--------C-----C---T-GA-------G---A--C-----------G---ATGG-----T----- 7765
SMM.SL.92.SIVsmSL92A      -G------C-T--G-C-----------CTGA-GT----A--A--C---------------A-----T----------------G--A--A---------------C-------------------A--C------G-------C-----G-TG-----G--C------A----T 7800
SMM.SL.92.SL92B           ----A---C-A--G-C----------TCAAA-GT-G------A----G-C-A--G--------A------------------G--A-------------------------A-------------------A-C--C---G---GA--C-----G--------C-T-C----CA----- 8151
SMM.US.04.G078            -C----ATC-A--G-C----------T-AGA-GT-G--A--A-----G--C------------A--T--------G--G-----A--A----------AG---A--------------------------C---------A--C-----G--------C---C----------- 7996
SMM.US.04.G932            --A----TC-A--G-C--------T--CTAA-GT----A--A--------C--------A--------------------------A----------A---------T-----------------C--C-------GA-------G-----------------CA----- 7973
SMM.US.04.M919            ---------C-A--G-C----------TCAAT-AT-G--A--A------------G-----A-----T----------G-----A--A----------A--A-T--------------------------A-------------A--------G-----T--GCT-------C----- 7998
SMM.US.04.M922            -CA------C-A--G-C----------TCAAT-AT-G--A--A------------G-----A-----T--------G--G-----A--A-------------A-C-------------------------A-------------A--------G-----T--GCT-------C----- 7998
SMM.US.04.M923            -GA------C-A--G-C----------TCAAT-AT-G--A--A------------G-----A-----T--------G--G-----A--A---------------T----------------------AA-----------GA----------------T--ACTA-------T----- 7985
SMM.US.04.M926            -------TC-A--G-C-------------TA-G--------A-----------------A------------------G-----A----C----------A-------C-------------------------------GA----------------T--A---G----A------ 7994
SMM.US.04.M934            -------TC-T--G-C-------------TA-G--------A-----------------------------------G--G-----AC-A---------------A-------C-------------A-------------A----------------TA-A---G----------- 7970
SMM.US.04.M935            --------C-A--G-C----------TCAAT-AT-G--A--A------------G-----A-----T----------G-----A--A-------------A-C-------------------CA---G------GA--------G--------ACT-------C----- 7977
SMM.US.04.M940            -------TC-A--G-C----------CTTGA-GT-G--A--A--G--GG-C--------A-----T--------G--G-----A--A----------AG-------------------------A----C--------A--C----------------C-T-C----------T 8023
SMM.US.04.M946            -------TC-T--G-C-------------T--G--------A-----------------A-----------------G--G-----AC-A---------------A-------C--------------------------A---------------T--A---G----------- 7998
SMM.US.04.M947            --------C-A--G-C----------TCAAT-GT-G--A--A------------G-----------T--------G--G-----A--A-------------A-C-------------------A------------GA----------------T--ACT-------T----- 7982
SMM.US.04.M949            -------TC-A--G-C----------C--G--AT-G--A--A--G--G--C--------A-----TC--------------G-----AC-A-----------------------------------A----C--------A--C-----G--------C---C----------T 7957
SMM.US.04.M950            --------C-T--G-C-------------TA-GT--------A-----------------A-----------------G--G-----AC-A---------------A-------C-----------A------------GA--C-----------T--A---------A----- 7985
SMM.US.04.M951            -------TC-A--G-C----------CTTGT-AT----A-----G--GG-------------A--T--------G--G-----A--A-------------------------------------A----C--------GA--C-----G--------C---C----------- 7978
SMM.US.04.M952            -------TC-T--G-C-------------T--G--------A-----------------A-----------------G--G-----AC---------------A-------C---------------------------A---------------T--A---G----------- 7974
SMM.US.05.D215            ----T--TC-A--G-C------T---CAAA-G-----A--A-----------------------------------G--G-----AC-A---------------C------------------A-C----------GA--C-----G--------AATC-----GT----T 7972
SMM.US.06.FTq             -A-CT-A-C-A--G-C------T---CAG-AAT-G--A--A--------A-T-----A-----T--------T--------A--A----T--------T------C-----------CA-G--A---G----A---G-CT-T-TC---T-C---G----A-C---- 7988
SMM.US.11.SIVsmE660_FL10  --------C-A--G-C----------TCAAT-AT----A--A------------G-----A-----T----------G-----AC-A-------------A-C-------------------A------------GA---------------T--ACTA------T----- 8712
SMM.US.86.CFU212          ------ATC-A--G-C----------GCTT--AT----A--A-----------------A-----------------G--G-----AC-A---------------A---------------T--CA-C------G----A--------G------A-A---------A------ 7970
SMM.US.x.F236_H4          --------C-A--G-C----------TCAAT-AT----A--A------------G-----A-----T----------G-----A--A-------------A-C-------------------A------------GA---------------T--ACT-------C----- 8668
SMM.US.x.H9               -GA------C-A--G-C----------TSAATYAT-G--A--A------------G-----AM----T----------G-----A--A----SK-----N-T--Y---------K--------A---Y---------A---------------T--ACTG------C----- 8192
SMM.US.x.PBJA             -GA------C-A--G-C----------TCAATCAT-G--A--A------------G-----A-----T-C--------G-----A--A-------------A-T-------------------A-------------A---------------T--ACTG------C----- 8494
SMM.US.x.PGM53            -----GA-C-A--G-C----------TCAGGCAT-G--A--A------------G-----A-----T----------G-----A--A-------------A-T-------------------A-------------A--------G-----T--ACT-------C----- 8649
SMM.US.x.SME543           --------C-A--G-C----------TCAAT-AT----A--A------------G-----A-----T----------G-----AC-A-------------A-C-------------------A------------GA---------------T--ACTA------T----- 8725
SMM.US.x.pE660.CG7G       --------C-A--G-C----------TCAAT-AT----A--A------------G-----A-----T----------G-----AC-A-------------A-C-------------------A------------GA---------------T--ACTA------T----- 8708
STM.US.89.STM_37_16       -------TC-T--G-C----------CAA--GT----A--A--T---G-------G--A-----T--------G--G-----AC-A---------------A---------------------A-C------G---GA--C--------------A---T-----A----- 8354
```

Tat Rev intron end_Tat Rev exon 2 start
MAC.US.x.239
Env
Tat exon 2
Rev exon 2
A.CI.88.UC2
A.DE.x.BEN
A.DE.x.PEI2_KR_KRCG
A.GH.x.GH1
A.GM.87.D194
A.GM.x.ISY_SBL_6669_85
A.GM.x.MCN13
A.GM.x.MCR35
A.GW.86.FG_clone_NIHZ
A.GW.87.CAM2CG
A.GW.x.MDS
A.IN.07.NNVA
A.IN.95.CRIK_147
A.JP.08.NMC786_clone_41
A.PT.x.ALI
A.SN.85.ROD
A.SN.86.ST_JSP4_27
B.CI.88.UC1
B.CI.x.20_56
B.CI.x.EHO
B.GH.86.D205_ALT
B.JP.01.IMCJ_KR020_1
G.CI.92.Abt96
AB.CM.03.03CM_510_03
H2_01_AB.CI.90.7312A
H2_01_AB.JP.04.NMC307_20
H2_01_AB.JP.07.NMC716_01
H2_01_AB.JP.08.NMC842_10
U.CI.07.07IC_TNP3
U.FR.96.12034
MAC.US.x.17EC1
MAC.US.x.251_1A11
MAC.US.x.251_32H_PJ5
MAC.US.x.251_BK28
MAC.US.x.MM142_IVMXX
MNE.US.82.MNE_8
MNE.US.x.MNE027
SMM.CI.79.SIVsmCI2
SMM.LR.89.SIVsmLIB1
SMM.SL.92.SIVsmSL92A
SMM.SL.92.SL92B
SMM.US.04.G078
SMM.US.04.G932
SMM.US.04.M919
SMM.US.04.M922
SMM.US.04.M923
SMM.US.04.M926
SMM.US.04.M934
SMM.US.04.M935
SMM.US.04.M940
SMM.US.04.M946
SMM.US.04.M947
SMM.US.04.M949
SMM.US.04.M950
SMM.US.04.M951
SMM.US.04.M952
SMM.US.05.D215
SMM.US.06.FTq
SMM.US.11.SIVsmE660_FL10
SMM.US.86.CFU212
SMM.US.x.F236_H4
SMM.US.x.H9
SMM.US.x.PBJA
SMM.US.x.PGM53
SMM.US.x.SME543
SMM.US.x.pE660.CG7G
STM.US.89.STM_37_16

```
                                          Tat end
MAC.US.x.239              TGGCAACAGCTCCTGGCCTTGGCAGATAGAATATATTCATTTCCTGATCCGCCAACTGATACGCCTCTTGACTTGGCTATTCAGCAACTGCAGAACCTTGCTATCGAGAGTATACCAGATCCTCCAACCAATACTCCAGAGGCTCTCTGCGACCCTACAGAGGATTCGAG 9043
Env                       __G__N__S__S__W__P__W__Q__I__E__Y__I__H__F__L__I__R__Q__L__I__R__L__L__T__W__L__F__S__N__C__R__T__L__L__S__R__V__Y__Q__I__L__Q__P__I__L__Q__R__L__S__A__T__L__Q__R__I__R__
Tat exon 2                V__A__T__A__P__G__L__G__R__*
Rev exon 2                _W__Q__Q__L__L__A__L__A__D__R__I__Y__S__F__P__D__P__P__T__D__T__P__L__D__L__A__I__Q__Q__L__Q__N__L__A__I__E__S__I__P__D__P__P__T__N__T__P__E__A__L__C__D__P__T__E__D__S__R
A.CI.88.UC2               ----T--GA--T----------C----A-C--C--A--GC---------A--TG----CT--T---------G------A-----T------GGA---A-----CGCGAAC-C--C--C--G--G--TG--CTC-----AT.............--GACAGCA--CA-G- 9040
A.DE.x.BEN                ---TT--GA--TG---------CA---A-C---G-G--G----------A--T-----CT----------TCG------A--A--T-------GA---A-----C-AGAAC-C--C--C--G--G--TG--CTC------T..............---ACAGCA--CA-G- 9086
A.DE.x.PEI2_KR_KRCG       ---AG-------G---------CA----C------A----------------G-----T-----------CG------A----GT-----AGGA---A-----C--GAGC-T--C--C--------T----CT-------T--TCAGAGAG-A---ACA-CA--CA-G- 9042
A.GH.x.GH1                ---TT--GA--TG-A-------C----A-C-----A------------A--T-----CT-----------CG------A--AG-T-------GA---A-----C-CGAAC-C--C--C--A--G--TG--CTC-----AT.............---ACAGCA--CA-G- 8506
A.GM.87.D194              ---TTT-G---TG---------CAC--A-C--C--A--A---------A--T-----CT-----------CG------A--A--G------GGG---A-----C-AGAAC-C--C--C--G--G--TG--CTC------T..............---ACAGCA--CA-G- 8500
A.GM.x.ISY_SBL_6669_85    ---A-G---A------------C------------A---------------TG----C------------CA-A----A--A--G------GGA---A-----C---C-C----T---------.............................C-CA--- 8485
A.GM.x.MCN13              ---AG-------G---------C----AG------A------------A---G-----T---------G-CG-A----A--A--T------G-A---A-----C--GA-C-C--T--C-----G----G-TT-------T--TCAGAG-G-A--GACAGCA--CA--- 8525
A.GM.x.MCR35              ---AG------TG-A-------C----AG------A------------A---G-----T-----------CG-A----A--A--T------G-A---A-----C--GA-C-C--T--C-----G----GCTT-------T--TCAGAG-G-A--GACAGCA--CA--- 8525
A.GW.86.FG_clone_NIHZ     ---AG----A--T---------C-----C------A---------------TG----T-----------CG------A--A--T------GGA---A-----C--GA-C-C--C---------------CT-------T---CAGAGAG-A---ACAGCA--CA--- 8509
A.GW.87.CAM2CG            ---AG----A-------------------------A----------------G-----C-----------CG-A----A--A--T------G-A---A-----C--GA-C-T-----C---------C--C------AT....................C-CA--- 9089
A.GW.x.MDS                ---AG----A--T-----C----C-----C------A--------------TG----T-----------CA-A--G-A--A--G------GGA---A-----C--GA-C-T--T--C---------C--CT-----AT....................C-CA--- 8554
A.IN.07.NNVA              ---AT----A--T-----C----C-----C------A-----------T--TG----T-----G-----CA-A----A--A--G-----AGGA---A-----C--GAGC-T--T--C---------C--C------AT....................C-CA--- 9069
A.IN.95.CRIK_147          ---TT----A--T--A--C----C-----C------A-----------A--TG--A--T-----G-----CA-A----A--A--G-----AGGA---A-----C--GAGC-T--T--C---------C--C------AT....................C-CA--- 8832
A.JP.08.NMC786_clone_41   ---AG------TG---------C----AG-------------------G-----T-----------CA-----A-GAG-T------G-G---A-----C--GA-C-T-----C--------TG--C................................A--- 9067
A.PT.x.ALI                ---GG----A-TG-----C----C---C-C----T-A----------A--TG--AGCT----------TCG----G-A-----T------GGA---A-----C--GA-C-C--C-------------G--CT---G---T--TCAGAGAG-G--GACA-CA--CA-G- 9082
A.SN.85.ROD               ---AG----A-A------C----C-----C------A--------------G-----T-----------CA-A----A-----T------GGA---A-----C--GAGC-T--T--C--------TC--CTA-----AT....................C-CA--- 8523
A.SN.86.ST_JSP4_27        ---AG---AT-GG-----C----C----AG------A--------------G-----T----------ACA-A----A--A--T------GGA---A-----C--GAGC-T-----C--------T---CTC------T--TCGGAGAG-AT-GACAGCAG-CA--- 8528
B.CI.88.UC1               G-CTT----A--T-----C-------------CGC---C--T-----T---------GGAA-------T------G-A--A-GG------G-A---A--G-T--AGACC-C---A-----------GC------A................................ 9054
B.CI.x.20_56              G-ATT----A--T-----C-------------CGC------T-----T---------GGAA-------T------G-A--A-GG------G-A---A---T---GACC-C---A-----------GC------A................................ 9019
B.CI.x.EHO                G---T----A--T-----C--------C-----C--C--C---C--T---------GGGA-------T---------A----GG-----------AT---A-AGACC-T---A-C-----------G-------A................................ 9029
B.GH.86.D205_ALT          G-A-T----A--------C--------------C-------T--AC-G-----G---GGAA-----A-T---------A--A-GG-----------A--G-T--AGACC-T---A--------C-A---CTCAACC-AT---CAA..................... 9052
B.JP.01.IMCJ_KR020_1      G-A-T----A----------------CAG---C--C--C---------T---------GGAA-----A-T---------A--A--G--------G---A--G-T--AGA-C-T---ACC---------T-G--T----A................................ 8160
G.CI.92.Abt96             ----T----A--T-----------------------------M-------G--AGGGAA------------------A-------------G----A-----C--GA-CYT-----CT--------GC-GT----A--AA----CAG------C--AGCA---A--- 8471
AB.CM.03.03CM_510_03      G-A-T----A--T-----C-------------CG-------T---C-G-----G----GGAA-----A-T---------A--A-GG-------------A--G-T--AGACC-C---A---------------------A................................ 8146
H2_01_AB.CI.90.7312A      G---T----A--T-----C--------C-----C--C--C-----A--T-----G----GGAA-------T---------A-GA-GG-------------A--G-T--AGACC-T---A-C----------GCT-----A................................ 9058
H2_01_AB.JP.04.NMC307_20  G---T----A--T--------------C-----C--C--C-----A--T-----G----GGAA-------T------G-A-GAAGG--------------------AGACC-T---A----------T-GCT----GA................................ 9040
H2_01_AB.JP.07.NMC716_01  G---T----A--T--------------C--G--C--C--C-----A--T-----G----GGAA-------T---------A-GAAGG--------------------AGACC-T---A----------A-GCT----GA................................ 9030
H2_01_AB.JP.08.NMC842_10  G---T----A-----------------C--G--C--C--C-----A--T-----G----GGAA-------T---------A-GAAGG--------------------AGACC-T---A-C--------T--CT-----A................................ 9057
U.CI.07.07IC_TNP3         ----TT--A---------------A---------C------------------G-----T------------------A------T-T----G-T-G-------CAGAGC-T-----G-------G-GGCCTG----.........AGAGTA--GACAGCA--CA--- 8558
U.FR.96.12034             ----T----A--T--------------AG------A-------------A--------T----G------------A-----G------GGA---------C--GA-T-G-----G------G--GT-GT-------C.......................--CA--- 8561
MAC.US.x.17EC1            ------------------------------------------------------------------------------------------------------------------------------------------------------------------------ 9043
MAC.US.x.251_1A11         ------------------------------------------------------------------------------------------------------------------------------------------------------------GA---------- 9039
MAC.US.x.251_32H_PJ5      ----------------------------------------------------------------------------------------------C-------------------------------------------G---------A------- 9041
MAC.US.x.251_BK28         ----------------------------------------------------------------------------------------------C-------------------------------------------------GA---G------ 9025
MAC.US.x.MM142_IVMXX      ----------------------------------------------------------------------------------------------C-------------------T--------------------------G---A------ 8538
MNE.US.82.MNE_8           ----------------------------------------------------A-------------------T--------------------------C-------------------T-------AT----CA------------A--C---- 8517
MNE.US.x.MNE027           ---------T------------------G---------------------------------------------------------------C-------------------T-------AT----CA------------AG-C---- 8520
SMM.CI.79.SIVsmCI2        ----T----A--T-----------A---------------------------G--A--T--T------------G-A--A---------GGA---AG---T---GA-C-T-----G-------GG---CT----ACTA.........GT-TCT--AGCAGCCA-GA 8187
SMM.LR.89.SIVsmLIB1       ----T----A--T-------------------------------------A---G---------------G-----G-A--A---T------G-T-G-----TC--GA-C-G-----A-----GG-A-C-GT-------A-----A-AAG-AA----ACCAG-CA--- 8105
SMM.SL.92.SIVsmSL92A      ---T-C---A-------------A-----------------------------G-------AT--T------------A--A--G-------G----A----T---GA-C-T-----G-----------C-TA------A----TAAGCCTGA--AGC-----C--T- 8140
SMM.SL.92.SL92B           ---AT----A-TG----------------------------------------G--------GA-A------------A--A---T-TG-CTCG---AGC---C----CC-------A--------A-C-GTG------A-----A-A--T-AGT---CC----A--- 8491
SMM.US.04.G078            ----T----A--T----------------------------------------GT-AG-T----------------G----A----------G-T-G---------GA-C-------G-----------CGT-------A-----A-A---AGC------A--C---- 8336
SMM.US.04.G932            ---------A--T----------------------------------------G----T--G--T------------A-GA--G-----------------T---GA-C-G-----C-----------G-GT-------------ACAAAG-G--G-----A--C---- 8313
SMM.US.04.M919            ---------A--T----------------------------------------G----T----A------------------A---------G---AT--C-T----A-------A-------------GCGT-------------TCA----AT-------AG-C-A-- 8338
SMM.US.04.M922            -----G---A--T----------------------------------------GT---T------------------------A--G-----G---GT--C----CA------------------G---G----------A-----CTT--TAGC------AG---A-- 8338
SMM.US.04.M923            ---------A--T----------------------------------------GT-A------------------------A--GT----GGAT-G-----T---GAGC-G---A---------G---G----------C-----AAG---G--G-----AG-C--T- 8325
SMM.US.04.M926            ---------A--T----------------------------------------G-----------------G----A--G------G-T-GT--C-T----A-C--------------G---G-GT-------A----TCA----AT-------AG-C-A-- 8334
SMM.US.04.M934            ---------A--T-----------------------------------------T---T-----------------G--------------G--T-GT--C-T----A-------------------GG-GT-------A----TC-----G--------AG-C-A-- 8310
SMM.US.04.M935            -----G---A--T-------------------C-----------------------GT---T--------------------------G------G---GT--C-T---CA--------------G---G----------A----TCT---TAG-------AG---A-- 8317
SMM.US.04.M940            ----T----A--------------------------A---------C-------------T-----------------G---A---------G-T-G----T---GA-C-T-----G------G--GG-GT-------A-----A-A---AG-G----CA--CA--- 8363
SMM.US.04.M946            ---------A--T-------------------------------------------------T-----------------G---A---------G--T-GT--C-T----A----------------------G-G----------A----TC---G-A--T-----AG-C-A-- 8338
SMM.US.04.M947            ---------A--T-------------------------A--------G---------------T-----G---A---T-----GAT-G-----T---GAGC-----------------GCG--------C----TAAA---A--------AG-C--T- 8322
SMM.US.04.M949            ----T----A-------G--------------------AC-------G-----T-----------------G---A--------A-G-T-G-----------GA-C-G-----G-----G---G-GT-------A-----A-A-G-AG------TAC------ 8297
SMM.US.04.M950            -----G---A--T-------------------C-----------A-------G-----T---------------------A--------G--T-GT--C-T----A-----------------G---G-GT-------A----TC-----TA--------AG-C-A-- 8325
SMM.US.04.M951            ----T----A--T-----G-----------G---C-----------C---------AG-T-------------------A--G----------T-G-----T---GA-C-G-----G-----------G-GT-------A-----A-A---AG-------A--C---- 8318
SMM.US.04.M952            ---------A--T-----C---------------------------------A----T----T------------------------A--------G--T-GT--C-T----A------------------------G-GT-------A----T---AG-A--------AG-C-A-- 8314
SMM.US.05.D215            ----T----A--T---------------G---C----C--T--AC------GT--C------------------A--A--G-------G-T-G-----ACAGA-C-G-----C-----G---G-GT--------A----A-AAGTT--C----CAG-C---- 8312
SMM.US.06.FTq             ----T----A--T----------------------------------------G-----T-----A-----------G-A-GA------------T-G-----T--ACAGC-------G-----G---G-G----G-G--A----CAGAG-GT-G-----AG-C---- 8328
SMM.US.11.SIVsmE660_FL10  ---------A--T-----------------------------------------------------------------G-----GGAT-G-----T---GACC-------------------G---------C-----AA----GTCG--A--AG-C--T- 9052
SMM.US.86.CFU212          -----T---A--T----------------------------------------GT-C--T---------------G-A-----GT---------------T---GA-C-------A---------G-G--------A-----CAGTGTG--G--AGCA--AA-G- 8310
SMM.US.x.F236_H4          -----G---A--T----------------------------------------GT-------------------------------G-----GGAT-G-----T---GAGC----------------G-G--------C-----AA----GT-G--A--AG-C--T- 9008
SMM.US.x.H9               ---------A--T-------R--------------------------------G--------------------------------GT----GGAT-G-----T---GANC-G---A-------G---G-G--------C-----AAG---G--G-----AGCC--T- 8532
SMM.US.x.PBJA             ---------A--T----------------------------------------GT-------------------------------GT-----GAT-G-----T---GAAC-G---A-C-----G---G-G--------C-----AAG---G--G-----AGCC--T- 8834
SMM.US.x.PGM53            ---------A--T----------------------------------------G-----T-----------------G---A---------G---AT--C-T----A------------------G-GT---------.........CT-T-------A--C---- 8980
SMM.US.x.SME543           -----G---A--T---------------------------A-------------------------------------------G-----GGAT-G-----T---GA-C-----------------G-G--------A-----AAG---G--G--A--AG----T- 9065
SMM.US.x.pE660.CG7G       -----G---A--T---------------------------A--------G-----------------------------------G-----GGAT-G-----T---GACC-----------------G-G--------C-----AA----GT-G--A--AG-C--T- 9048
STM.US.89.STM_37_16       ---A-T--A--------------A----------C-----------------G--AG------------------G-A--A-----T---GG-T-GT-----C--GA-C-------C------C----CCT----A---A----CAG--TA--G-----A---A--- 8694
```

```
                                  Rev end                 Nef start
MAC.US.x.239                      AAGTCCTCAGGACTGAACTGACCTACCTACAATATGGGTGGAGCTATTTCCATGAGGCGGTCCAGGCCGTCTGGAGATCTGCGACAGAGACTCTTGCGGGCGCGTGGGGAGACTTATGGGAGACTCTTAGGAGAGGTGGAAGATGGATACTCGCAATCCCCAGGAGGATT 9213
Env                               E  V  L  R  T  E  L  T  Y  L  Q  Y  G  W  S  Y  F  H  E  A  V  Q  A  V  W  R  S  A  T  E  T  L  A  G  A  W  G  D  L  W  E  T  L  R  R  G  G  R  W  I  L  A  I  P  R  R  I
Nef                                                                  M  G  G  A  I  S  M  R  R  S  R  P  S  G  D  L  R  Q  R  L  L  R  A  R  G  E  T  Y  G  R  L  L  G  E  V  E  D  G  Y  S  Q  S  P  G  G  L
Rev exon 2                        S  P  Q  D  *
A.CI.88.UC2                       -CTGG--G--ACT-A-GGC-G------G----------CGAG-GGA----A--A--TT----A--AA--GCA--GA------G---------------------A-G-GG-----TA--G-AG-GCA-C-CATC--G--GG-A--C------G----A-------C 9210
A.DE.x.BEN                        -CTGG--G--ACT-A-GGC-G--C-A--G----------CGAG-GGA----A--A--TT----A-AT--GC---GA--A---G---------------------TGG--------AG-AGCGC-AC-CATC--G--GG-A--------G-T--A--A--A--C 9256
A.DE.x.PEI2_KR_KRCG               -CTGG--G--ACT-AC-A-AG------G----------CGAG-GGA----A--A-T-C------T-C-TGCA--GA--A---G---------------A-------A-------G---G-G-AA-GG-AC-GATC--C--GG-A--------T--G-T--A--A-----C 9212
A.GH.x.GH1                        -CTGG--G--ACT-A--GC-G------G---------G-CGAG-GGA----A--A--AT----A--AT--GCA-AGA--A---G---------------A----A-----G-GG-----T-CAG-AG-GCAAC-C-TC--G--GG-A--------G----A--A--A----C 9676
A.GM.87.D194                      -CTGG--G--ACT-A-GGC-G------G----------CGAG-GGA----A--A---T--GA--AT--GCA--GA------G------A---------------A-G-GG-----T--AG-AGCGCAAC-CATC--G--GG-A--C------G----A--A-----C 9670
A.GM.x.ISY_SBL_6669_85            -CTGG--G--ACTCA-GGCAG------G--G---------CGAG-GGA----A--A---T--------C--GC---GGT-A-A-G---------A--A-----G--A-GAG---G----G-G----GG-AC--ATC-----GG----------G-T--AC-A-----C 9655
A.GM.x.MCN13                      -CTGG--A--A-----CGCAG-----T-G--G---------CGAG-GGA----A-GA---T-------T--GCA--GG--A---G-------------A-----A-------G----G-G-A--GCA-C-GATC--G--GG-A-----T--G----A--A--A--C 9695
A.GM.x.MCR35                      -CTGG--A--A-----CGCAG-----T-G--G---------CGAG-GGA----A-GA---T-------T--GCA--GG--A---G-------------A-----A-------G----G-G-A--GCA-C-GATC--G--GG-A-----T--G----A--A--A--C 9695
A.GW.86.FG_clone_NIHZ             -CTGG--G---CT-A--GCAG-------G--G---------CGAG-GGA----A--A---T-----A---C-TGCA--GA--A-A-G---------A------G--A-------G---AG-G-G--GCAAC--ATC--G--GG-------T--G----AC-A-----C 9679
A.GW.87.CAM2CG                    -CTGG--G---C--A--G-AG---T--G---------CGAG-GGA----A--A---T--------CTGCA--GG----A-G---------------CA-G---G-G----GA-TGT-ACAAC-CAT------GG-A--C--T--GG----A--A-----C 9259
A.GW.x.MDS                        -CCAG--GC-ACT-AG-A-AG---T--------------CGAG-GGA----A--A---T--------T--GC---GGT-A-A-G---------A--A-----CA-G-G---G----G---AT-GG-AC--AT---G--GG-A--------GG----A--A-----C 8724
A.IN.07.NNVA                      -CTGG--G---CT-AG-ACAG---T--G---------CGAG-GGA----A--A--AT--------C-GC---GA--A--GG---------------A-------CA-GAG---G----G---GT-GC-AC-CAT---G--GG-A--------GG-T--A--A--A--C 9239
A.IN.95.CRIK_147                  -CTGG--G--ACT-A--ACAG---T--G---------CGAG-GGA----A--A--AT----A---T--GC---GA-----GG-A----------A----AG-CA--A----G--T-G---AG-GGAA---AT---G--GG-A--------------A--A-----C 9002
A.JP.08.NMC786_clone_41           -CTGG--G--AGT-A--ACAG-AC----G----------CGAG-GGA----A--A---T----A--TT--GCA--GA-----G---------------------G--A--C---ACTG-CT-GC--C-CTTC--G--GG-A-----T--G----A--A-----C 9237
A.PT.x.ALI                        -CTGG--G--ACT-A--GCAG-------G--G---------CGAG-GGA----A--A---T--G----T-TGCA--GAT-----G---------A-AAA-A-C---A-------G----G-G-AG-GCA-T-G-TC--G--GA---------G----A--A-----C 9252
A.SN.85.ROD                       -CTGG--G--ACT-AG-ACAG---T-T-G----------CGAG-GGA----A--A--AT--------C-GC---GG--A-A-G---------------CA-G-G---G---AG-GTAT-GGAAC--ATC--G--GG-A-------GG-T--A--A-----C 9693
A.SN.86.ST_JSP4_27                -CTGG--G--ATT-A-CACAG-------G--------G-CGAG-GGA----A--A---T--GA---T--GC---GG--A--GG----------A-AAA---C---A---G---C----G---A--GG-ACA-AT---G--GG-A-----T--G----A--A-----C 9698
B.CI.88.UC1                       ..CCG------CT-TC---TG-A-----C--G------ATC----GG-----A--A--AA----A--A-CAAC---GG------G------------AA-A-AG-AA---C-C-----A-AG----C--A--GACA-CGGA-GCA--CA----G-------------C 9222
B.CI.x.20_56                      ..CCA------CT-TCG--TG-A-----C--G------ATC----GG-----A--A--AA----A--A-CAAC---GG------GG-----------AA----G-AA---C-C-------AG----C--A---ACG-CGG-GGCA--CA----GG----------A--C 9187
B.CI.x.EHO                        ..CCG------CT-CC--CTG-A-----C-G------ATC----GG-----A--A--AA----A--A-CAGCC--GG-----GG-----------A-----GC-A-GAC--CG----GAGTC--C--------CG-C-G--GA---CA-T------A----------C 9197
B.GH.86.D205_ALT                  ..CCA------CTCCCGG-TG-G-----C--G------ATC----GG-----A--A---C--GA--G-CAGCA--GG--A--GG---A---------A-T--CG-A-AGAC----------G----C-------CG-CG---GCA--CA----G----------A--C 9220
B.JP.01.IMCJ_KR020_1              ..CAG------ACT-CC---TG-A-----C----------ATCG---GG-----A--A--AA----A--A-CGAC---GG-----GG----------A-AAA-A-AG-AA-G-C-C--------G----C--A---ACA-CG--GGCA--CG----G----------C 8328
G.CI.92.Abt96                     --CAT---C--CTA---GCAG-T--TT-CAGC-------TCC---GGC----A--A---TG-ACC--A-CAACT--GG-----CA----------A-AA--A-----A---C-------A-A-----GG-A-G-T-------G----------A--G--A--A--C 8641
AB.CM.03.03CM_510_03              ..CCA------CTCCCGT-TG-A-----C--G------ATCG---GG-----A--A--A---C---GA--A-CAGC---GG--A---G----------A-T---G-A-AGAC---------AG----C---G-G-CG-CG---GC---C-----G--------A--A--C 8314
H2_01_AB.CI.90.7312A              ..CCA------CTCCTGT-TG-G-----C----------ATCG---GG-----A--A--A-----A--A-CAGC-G-GG--A--GG----------A--A-AG--A-GAC----------G----C-----GACG-CG--GG-A--CA----G-------A--A--C 9226
H2_01_AB.JP.04.NMC307_20          ..CCA------CTCCCG--CG-AC----C--G------ATC----GGC----A--A--A-----A----C-GC---GG----GG----------A--A----CA-CA-GAC-A--------G----C-----GACG-CG--GG-A--CA----------A--A--C 9208
H2_01_AB.JP.07.NMC716_01          ..CCA------CTCCCG--CG-AC----C--G------ATC----GGC----A--A--A-----A----CTGC---GG----GG----------A--A----CA-CA-GAC----------G----C--A--GACG-CG--GG-A--CG----------A--C 9198
H2_01_AB.JP.08.NMC842_10          ..CCA------CTCCCG--CG-AC----C----------ATC----GGC----A--A--A-----A----CTGC----G----GG----------A--A----CATCA-GAC-A--------G----C--A--GACG-CG--GG-A--CA----------A--A--C 9225
U.CI.07.07IC_TNP3                 --CAT--G---CT---GGGAG------R--G--------CCAG-GG-----A--A---C----A--A-G----AGG--R---G----------AA----------A----------GG-AC-G-T-------G----------G--------C 8728
U.FR.96.12034                     --CGT--GCACCT----A-AG-T-----------------CAG-----A-A--A--AT----A--GT--G---A-G----CG----------AGCC--A--G--A----AC----------GG--C-G-T---------C-T-GA--T--A--A--------C 8731
MAC.US.x.17EC1                    ---------------------------------------------------------------------------------------------------------------------- 9213
MAC.US.x.251_1A11                 ----------------------------------------A---------------------T--A-------------------------------------A---G-------------------------C----------T-----A--- 9209
MAC.US.x.251_32H_PJ5              ----------------------------------------------------G-------A-T--G-------------------------------------------------------------T--------- 9211
MAC.US.x.251_BK28                 ----------------------------------------------------A---G---------------A-------A-------------------------------------C----------T--------- 9195
MAC.US.x.MM142_IVMXX              ----------CT----------------------------------------A--A-------A----C-CA-------------------------A-------G----CAA--G-------------C----------T--------- 8708
MNE.US.82.MNE_8                   -------------------A-------------------------C-----A----------A-T--C-------------------------------------GG-A----T-------------------------T--------C 8687
MNE.US.x.MNE027                   -------------------A-------------------------C-----A----------A-T--C-------------------------------------GG-A--G-T-------------------------T--------C 8690
SMM.CI.79.SIVsmCI2                -CAGGA-----C--AGGA-CG----T--T-----------CAA--------A--A---TG-----AA-C-GCAC-GG-C---CG-----------AA-----G--A-----------------GG-A--GAT-------G--T---T---GC---G-----A--- 8357
SMM.LR.89.SIVsmLIB1               ----AA-T--A-GA---T-C-----T--T-------------GG---G-A--T---C----A--G-CA-----GG------G-----------A-------A-------C------GG-A--GAT-------G--T--TGG-------T---C-C--C 8275
SMM.SL.92.SIVsmSL92A              -TA----G--A-GA-C-G-AG-T---A-CAG----------CAA--C---ATC--AAGCC-G--A-AA-C-GCA-TCG------G-C-A-------TCA--A--G--A--ACAA-----C------GG-ACACAT-------G----CGCA--G-----A--------C 8310
SMM.SL.92.SL92B                   -GC-TG----A-GA---GCAGG---TA-C-GT---------AT--C---ATC--A--CTG-----AA-CA----------CA--AG--A--TC----------CT-A----------GG-AC-G-T-------G----CGCA--------A--------C 8661
SMM.US.04.G078                    -GT--A-----CT-AG-G-A--------G--G---------GG---A-A--A--A-----A--G-CG----A-------G----------AA-------A-----A----------GG-A--G-T-------G-T----------G---C-C--C 8506
SMM.US.04.G932                    -GA--A----A--------AG-----T-G----------C---GG---A-A--A--A-----A--A-CA---C-G------G----------A-----------A-----GGAAG-G-TC--G---GA---------G----G---C-C--C 8483
SMM.US.04.M919                    -----A----A-----GA-CT-------G----------CCT--GGC----A--A--A-----A--A-CG-CAC-GG------G----------A-AA--A----A-----------------GG-A--G-T----------G----------G-----A--- 8508
SMM.US.04.M922                    -----G----------GA-AT-------G----------CC---GG-----A--A--------A-C-------G-----G----------AA----------A----------GG-A--G-TC-----G-------T--------A--A------- 8508
SMM.US.04.M923                    -----A----AGT---GA-A----------G--------C----C-----A--A--A-CG--A--ATGG----A--T---CG----------AA----------A----------GG-A--G-T------GG-A----------G--AC-C--C 8495
SMM.US.04.M926                    -----A----------GA-AG----T--G--G-------CC---GG-----A--A--------A--A-CAACA--GG------G----------AA----------A--------C----------GG-AC-G-T-------G----------G-----A--- 8504
SMM.US.04.M934                    -----A--------AGCA-AG-------G----------CCAG-GGC----A--A----C---C--A-CG-AC--GG------G----------AA----------A----------GG-A--GAT-------G----------G--------- 8480
SMM.US.04.M935                    -----G----------A-AT-------G----------CC---GG-----A--A--A--A--A--A-CA-----GG------G----------A-AA----------A----------GG-AC-G-TC--------G---T-------G--A------ 8487
SMM.US.04.M940                    -TC--A-----CT----G-A-------G--G---------------GG---A-A--A--A-----A----CA----------GG----------------G---ACT----------G---C-T----TC-----GG-C--TGGG-------A---C-C--C 8533
SMM.US.04.M946                    -----A-------AGGA-AG-----T-G----------CCAG-GGC----A--A--------A--A-CA--C--GG------G----------AA----------A----------GG-A--G-T----------G----------G-----A--- 8508
SMM.US.04.M947                    -----A----A-T----A-A-------G--G---------C-T--C-----A--A--A-CG-----GTGG----A--T----CG----------AA----------A----------GG-AC-G-T------GG-AG-G--------T--G---C-A--C 8492
SMM.US.04.M949                    -----A-----CT----G--------G--G----------GG---A-A--A--A-----A--A-CA----------G----------AA-----G---ACT----------G------A-G-TC--G--GG-C--TGGG-------A--AC-C--C 8467
SMM.US.04.M950                    -----A----------GA-AG-------G----------CC-G-GG-----A--A--A-----A--A-CG-AC--GG------G----------A-AA----------A----------T--GG-A--G-T----------G----------G--------- 8495
SMM.US.04.M951                    -----A----ACT----G-A--------G--G----------GG---A-A--A--A-----A----CG----------GG----------------G---ACT--C-------A-A----C-A---CTC-----GG-C--TGGG-------A---C-C--C 8488
SMM.US.04.M952                    -----A----------GA-AG-------G----------CC-G-GGC----A--A--------A----CA--C--GG------G----------A-AA----------A----------GG-A--G-T----G----------G----G--------- 8484
SMM.US.05.D215                    -GA--A----------GACA-----T-G----------------T-GG-----A--A---A----A--G-T---TC-A--T---A-G----------AA----------A----------GG-A--G-T-------G----CGCA-------G--------C 8482
SMM.US.06.FTq                     -----A----ACT----------------C----------------GG-----A--A--A-CA--A--G-CA-TT-A-G-----G----------AA----------A----C-------T--AC-AC-G-T----G-------G--T-------G---C-C--C 8498
SMM.US.11.SIVsmE660_FL10          -----A----A-T--G-A-AG----------------C-T--------A--A--A--G-----GTGG----A--T----CG----------AA----------A----A----------GG-A--G-T-------G----------T---C-CG-C 9222
SMM.US.86.CFU212                  -GAC-A----A-----G--C-----TA-TGG---------CCTT-GG---A--A--A---C-G-----A-CAG-A--GG-----GG-C----------A----------A----------GG-A--G-T-------G--------TAAC-----G-----A--C 8480
SMM.US.x.F236_H4                  -----A----A-T----A-AG----------------C----------A--A--A--A--A--GTGG----A-CT----CG----------AA----------A----------GG-A--G-T-------G----------T---C-C--C 9178
SMM.US.x.H9                       -----A----AGT---GA-A----------G--------C-T--CC----A--A--A-CG--RRG-TGG----A--T--YRCG----------AA----------A----------GG-A--G-T------GG-A----------MR-AC-C--C 8702
SMM.US.x.PBJA                     -----A----AGT-C-GA-AG---------G--------C-T--CC----A--A--A-CG-----GTGG----A--T--TACG----------AA----------A----------GG-A--G-T------GG-A----------A--AC-C--C 9004
SMM.US.x.PGM53                    -----G----------GA-AT-------G----------CC---GG-----A--A--A--------A-CG----GG--CG-GG---A----------AA----------A----C------A----GG-A--G-T----------G----------A--- 9150
SMM.US.x.SME543                   -----A----A-T----A-A-------------------------C-T--------A--A--A-CA-----GTGG----A--T----CG----------A----------A----A----------GG-A--G-T-------G----------T---C-CG-C 9235
SMM.US.x.pE660.CG7G               -----A----A-T--G-A-AG----------------C-T--------A--A--A--G-----GTGG----A--T----CG----------AA----------A----A----------GG-A--G-T-------G----------G---C-C--C 9218
STM.US.89.STM_37_16               -----G----ACTA-G-GCAG----TT----G--------C-T--GGA----A--A--A-CG--A--A-CA------G-----GG----------A-----G--A----------A------G-A--G-T---G---A----CGGG--C-----A---C-C--C 8864
```

Draft of 30 September 2015, not for public release
hiv2dna_main.tex r8876–10546

3' LTR U3 start

```
MAC.US.x.239              TATCAGTG...AGGCCAAAAGTTCCCCTAAGAACAATGAGTTACAAATTGGCAATAGACATGTCTCATTTTATAAAAGAAAAGGGGGGACTGGAAGGGATTTATTACAGTGCAAGAAGACATAGAATCTTAGACATATACTTAGAAAAGGAAGAAGGCATCATACCAGAT 9541
Nef                       V__S__V__.__R__P__K__V__P__L__R__T__M__S__Y__K__L__A__I__D__M__S__H__F__I__K__E__K__G__G__L__E__G__I__Y__Y__S__A__R__R__H__R__I__L__D__I__Y__L__E__K__E__E__G__I__I__P__D_
A.CI.88.UC2               -T--T--T...-CAT---G---A--A------C------CC-----G------G----T-----A-----------------A-----------------G-T-------AGGGAT------------C-----T-G---C----------G-----GG-A--------- 9538
A.DE.x.BEN                -TC-T--T...-CA----G---A--A-GG---GA-----CC--T-------------T-----A-----------------A--------C-------G-T-------AGG--G-----------C----------C-------A--G-----G--A--------- 9584
A.DE.x.PEI2_KR_KRCG       -CC-T--C...-CA----G---A--AT----GG------CA--T--G------G----T--A--A------C----T-----------------T-----G---------AG--------------------------A-G---------------G--A--T------ 9540
A.GH.x.GH1                -TC-T--T...-CA----G---A--A------G------CC--T---------G----T-----A-----------...-G--AA---------T-----G-T-------AGGGAT------------C-----T-G---C----------G-----G--A--------- 9001
A.GM.87.D194              -TC-T--T...-T-----G---A--G--G---GA-----CC--T---C----------T-----A-----------------A--A-------------A-T-------AGGGAG------------C-----T-G-T-C----------G-----G--A--------- 8998
A.GM.x.ISY_SBL_6669_85    -C--T-AC...-CAT---G---A--AT-G---G------CA--T-G-A-----G----------AG----A--------T-----------------G----------AG--------------C-----------------------------G--A--T------ 8983
A.GM.x.MCN13              -C--T--C...-CA----G---A--A------GA-T---CAC-T-G---A--------TG----A-----------------A-----------------G----------AG--------------------------C-T--------------G--A--TG----- 9023
A.GM.x.MCR35              -C--T--C...-CA----G---A--A------GA-----CAC-T-G---A--------T-----A-----------------A-----------------G----------AG--------------------------C-TA-------------G--A--T------ 9023
A.GW.86.FG_clone_NIHZ     -CC-T--C...-CA----G---A--AT-----C------CA-T----------G----T-----------------------A-------------C-G-T---T---CAG--------------------T---------C........................ 8983
A.GW.87.CAM2CG            -TC-T-CC...-CA----G---A--A-------------CA--T---------G--------A-------------------A-------------C-G-T---------A----------------C-G--T-------------A--------G--A--TG----- 9587
A.GW.x.MDS                AC--T--C...-CA----G-ACA--A------G------CA--T---------G----T-----A-----------------A-----------------G-T-------A----------------C---T-----------------------G--A--T------ 9052
A.IN.07.NNVA              -CC-T--C...-CAT---GGACA-AA--G--GGA-----CC--T-----T------------------A--------T-----------------G----------AG-------GAGA----A----TT-------------------G--G--A--T-----C 9567
A.IN.95.CRIK_147          -T--T--C...-CA----GG--A--A-------------CA--T-G-A----------------A-----A-----------------------G-T-------A--....................................................... 9274
A.JP.08.NMC786_clone_41   -TC-T--T...-CAAGT-G---A--AA-G----------CC--T---------G--------A-----------------A-----------------A-T-------AGGGAG------------C-----C-G---C-------A--G-----A--A--------- 9565
A.PT.x.ALI                -C--T--C...-CA----G---A-AAT------------CA------------G----T-----A-----A---------G------------------G-T-------AG--------------------------------------------G--A--T------ 9583
A.SN.85.ROD               -T--T--C...-CA--------A--A------C------CAC-T-G------------T-----A-----A-----AC--G-----------------G-T-------A-----------A------A-T-----------------------G--A--TG----- 9021
A.SN.86.ST_JSP4_27        -CC-T--C...-CA----G---A--AT-----GA-----CA--T-GG-------G---T-----A-----G--------------------------C-G----------AT--G----G----G--C---------------------G--A--A--TGG---C 9023
B.CI.88.UC1               -C-AT--A...--A--C--TAGG--A--G---T------CA-----GA---------------------------------------------------C--T----AG--------------AC------C---TC----G--T--G--------TG--T-T-GA 9496
B.CI.x.20_56              CC-AT--A...--A--C-GTAAG--A--G---G------CA------------G----------------T--------------------------C---TC------AG--------------AC----T-C-C-TC----G--T--G-------G-TG--T-T-GA 9461
B.CI.x.EHO                -CCGT--A...--A--CGGG--C--A--G--GC------CA-T----C-A--G----------------T-------------A-------------TC--T----AG-----G----A---AC------C-----------T-----------TG-GT-T-GA 9471
B.GH.86.D205_ALT          CCAT-TGTAAG-CC----T---A--A--G---C------CA------C-A--G-----------------------C-----------------G-----T----AG--------C-----A------C---T--T--G--T-----------TG-GT-T-GC 9497
B.JP.01.IMCJ_KR020_1      -C-AT--A...--A--C--TAGG--A------C------CA--T-----------------------GG--------------------------------C-----AG--G----------T-AC------C---TC--------T--G--------TG--T-T-GA 8602
G.CI.92.Abt96             --C----C...TAC--C-----C--G------GT------------------------------A-------------------------------------------A----------------------C----------A-G--------GC----T--A--T------ 8969
AB.CM.03.03CM_510_03      CC-AT--A...--A--C-GTAGG--A--G---C------CA------ACA--------------------C-------------------T----------------AG----------------A-------C---T--T--G--T-----------TG-GT-C-GC 8585
H2_01_AB.CI.90.7312A      ---AT--C...-----CC-G--G--AT-G---C------CA----------------------------------------------------T---T----A------------G--A-----T-C---T--T--G--T--GAG---T--AG-TGGT-GA 9500
H2_01_AB.JP.04.NMC307_20  C--AT--C...-----C--TAAG--T--G---G------CA----------------------------------------------------T---T----AG---------------A-----T-C------T--G--T--G-----T--AG-CT-T-GA 9482
H2_01_AB.JP.07.NMC716_01  C--AT--C...-----T--TAAG--A--G---G------CAC-T-G---------------------------------------------A-T---T--C-AG------------G--A-------C------T--G--T--G-----T--AG-CT-T-GA 9472
H2_01_AB.JP.08.NMC842_10  C--AT--C...-----C--TAAG--A--G---GA-----CA----------------------------------------------------T---T----AG------------G--A-----T-C------T--G--T--G-----T--AG-CT-T-GA 9499
U.CI.07.07IC_TNP3         -------A...-------GG-----AT-G---C------C--------G-C--G----------A------T--------------------------------A-CA---------------A-----T--G-------------T--G-----G--AG-TG----- 9056
U.FR.96.12034             --G----A...--A----G---AG-A--G----------C-------------G----T-----A-----------------A---------T-----A-T---TTC-CAG-------------------T--C-------------T--G--G--A--A--C------ 9065
MAC.US.x.17EC1            --------...--------------------------------------------------------------------------------------------------------------------------------------------------------------- 9541
MAC.US.x.251_1A11         --------...---------------------------------------------------------------------------------------T----------------------------------------------------------------------- 9537
MAC.US.x.251_32H_PJ5      --C-----...-T----CG------------------------------------G---T--------------------------------------------------------------------G-------------G--------------------------- 9539
MAC.US.x.251_BK28         --------...-----------------------G------C--------------T--------------------------------------------------------------------G------------------------------------------ 9523
MAC.US.x.MM142_IVMXX      --C----T...GA-G-C-G----------------------------------------T---------------------------------------------------------------------------------------------------------------- 9036
MNE.US.82.MNE_8           --C-----...------CG-------T------T---A-------------G---T-------------------------------------------A----------A-----------G-----------------------------G------ 9015
MNE.US.x.MNE027           --C-----...------CG-------T-----G-------C-----------G---T-------------------------------------------A----------A----------------------------------------G-G------ 9018
SMM.CI.79.SIVsmCI2        --G-T---...--A------A-A--A--G--GGA-----C-----------------------------------.-------------A---------AG--G-------A----C-G--T--G-----G-----T--G-----A--A--------- 8684
SMM.LR.89.SIVsmLIB1       -GG----C...-T-AAT-G-A-C--T------G-----TC---T-----------------A---------------------------C--------CCAT--G----------TC-G--T-------------------G-----G--AG-------- 8597
SMM.SL.92.SIVsmSL92A      --G----A...-----T--GA-A--AT----GG------CA--------------------A------------------------------T---A-----------A---A--G--TTC-------T-T--T--G------------------ 8612
SMM.SL.92.SL92B           --GTT--C...T----------A---------G-----TCA--------------------A------------------T---------------AG--G-------------C----T------C-------A--G-----T------C------ 8992
SMM.US.04.G078            -C-----A...TAC---C-T-CC---T----GG-C---TCA---------TG---------------------------------------------CAG-------------------------------------------------G-C------ 8834
SMM.US.04.G932            -------A...T------G-AC---T--G---C-C----CA---------------------------------------------------------AG--G-----------A-----------C---------------A--A----T--- 8811
SMM.US.04.M919            --A-G--A...TAC--------G--AT-----G-C---TCA-----------G----------------------------------A--AC---------------A-------------------------------------C------ 8836
SMM.US.04.M922            -------A...TAC-----G--G--AT-----G-C---TCA-----------G-------------------A------------------AAC---------------C-----------------------------------G-T------ 8815
SMM.US.04.M923            GTC-----...--A----G-------T-------C----CA----------------------------------------------------AC--G-------A---A----TC-G-------------------A--A--------- 8823
SMM.US.04.M926            -------A...TAC--------A--AT-----C-C---TCA-----------G-----------------------------------------AC---------------------------------------G--------T--C--G--- 8832
SMM.US.04.M934            -------A...T--G-------G---T-----G-C---TCA-----------G----------------------------------------CAG-----------------------------------------------TG-T------ 8808
SMM.US.04.M935            --C----A...CAC----GG--G--AT-----G-C----CA-----------G-----------------------------------------AC-----------------------------C--------------------G-T------ 8815
SMM.US.04.M940            -------C...CAT----GG-----AT-G---G------C---T-----A----------------------------------------------AG-----G--------A-----TC-G---C------------------G-----T--T--- 8867
SMM.US.04.M946            -------A...T--G-------G---T-----G-C---TCA-----------G-----------------------------------------CA-------------------------------------------------TG-T------ 8836
SMM.US.04.M947            -TC-----...-CC----G--CC---T----GG-T----CA------A------------------------------------------------AC--G-----------A--G--TC--------------------A--A--------- 8820
SMM.US.04.M949            -------C...CAT----G------AT-----G------C---T-----A--------------------------------------A-----------A-----G--------A-----TT----C----G--A--------G---G-T--T--- 8801
SMM.US.04.M950            -------A...CAC--------A--TT-G---G-C---TCA-----------G-----------------------------------G-------GA--AG--G---------------------T-------------------T--T------ 8823
SMM.US.04.M951            -------C...CAC--T-G-----AAT-G---G------C---T-----A--------------------------------------A-----------AG-----G--C-----A-----T--C---C-----------G-----A--------T--- 8822
SMM.US.04.M952            -------A...CAC--------G--TT-G---G-----TCA------------------------------------------------------AG--------------T-----T-----------------------T--T------ 8812
SMM.US.05.D215            -GG-T--A...C-C-----G--C--T--G---G------CA-----------G-----------------------------------T--C-C---------------A--------C---C-------T----G--G--TG----T--- 8807
SMM.US.06.FTq             --C----A...CACT-------C---T-----TGC----C---------A--G---------A-----------------------------A-CAG-----------------AC-------T-----------------G-----A-----C--T--C 8826
SMM.US.11.SIVsmE660_FL10  -C------...CAC----G---C---T----GG-C----CA------------------------------------------------------A--AG-A-----------A-----T--G--A-G--------------A--A--------- 9550
SMM.US.86.CFU212          --G----A...CAT-----------AT-----G-----TCA--------------------A---------------------T--------------AG-------------G--TC-----G--C----------------A--A--------- 8808
SMM.US.x.F236_H4          -C------...CAC----G---C---T----GG-C----CA------------------------------------------------------A--AG-------------A-----T--G-----G--------------A--A--------- 9506
SMM.US.x.H9               GTC-----...--C----G---C---G------TC----CA-------------------------------------------G-----------ATT----------A---A-----TC-R-----------R--------A--A--------- 9024
SMM.US.x.PBJA             GTC-----...--C----G---C---G------TC----CA-------------------------------------------------------AT----------A---A-----TC-G------------------A--A--------- 9326
SMM.US.x.PGM53            -------A...CAC-----G--G--AT-----G-C---TCA-----------G---------------------------------------------AG-------------------------------------------------CC------ 9478
SMM.US.x.SME543           -C------...CAC--------C---T----GG-C----CA------------------------------------------------------A--AG-A-----------A-----T--G--A-G--------------A--A--------- 9563
SMM.US.x.pE660.CG7G       -C------...CAC----G---C---T----GG-C----CA------------------------------------------------------A--AG-A-----------A-----T--G--A-G--------------A--A--------- 9546
STM.US.89.STM_37_16       --G----A...CAT----G---C--A--T--GGA-----C------G------------C----A-----------AGC---------------------------AG--------------A--G-----G-----------A--------G--AG-T------ 9192
```

Draft of 30 September 2015, not for public release
hiv2dna_main.tex r8876–10546

```
                                                                        TCF-1 alpha binding                                                                      Nef end
MAC.US.x.239                  AGAGGTTCTAGCATGGAAGTTTGATCCAACTCTGGCCTACACTTATGAGGCATATGTTAGATACCCAGAAGAGTTTGGAAGCAAGTCAGGCCTGTCAGAGGAAGAGGTTAGAAGAAGGCTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAAGGAAACTCGCTGA. 9868
Nef                           __E__V__L__A__W__K__F__D__P__T__L__A__Y__T__Y__E__A__Y__V__R__Y__P__E__E__F__G__S__K__S__G__L__S__E__E__E__V__R__R__R__L__T__A__R__G__L__L__N__M__A__D__K__K__E__T__R__*_
A.CI.88.UC2                   G---ACA----TG---CG----A-CT-C-TG--------TGAG--CA----C-TCAC-CT-----------A-----GCA---------AT--C-----A-----TGG-AGGC--AA--G-AA-------GA-A-CAT-T-GT..................GAA-A-. 9847
A.DE.x.BEN                    G---ACAT---TG---C-------CT-C-TG--------T-AC--CA----C-TCAC-CTG-----------------GCA---------AT--C-----A---ATGG-AGGC--AA--G-AA-------GA-A-CAT-T-GT..................GAA-A-. 9893
A.DE.x.PEI2_KR_KRCG           G--AACAT--ATG----G------C--T-GG--------TGAG---AC---T-T-AA-C-------------A-----GTAT------------C----A------TGG-AGGC--AA--G-AA-------GA-A-CATTT....................A-T-A-. 9837
A.GH.x.GH1                    ---AACA---CTT---CG------C--C--G--------TGA---CA----T-TCA--CTGC-T---------------GCA---------AT--C-----A-----TGG-AGGC--AA--G-AA-------GA-A-CAT-T....................A-T-A-. 9307
A.GM.87.D194                  G---ACA----TT---CG------C--C-TG--------T-G---CA----C-TCA--CTGC-----------A-----GCA---------AT--C-----A-----TGG-AGGC--AA--G-AA-------GA-A-CAT-T.................-G-GAA-A-. 9307
A.GM.x.ISY_SBL_6669_85        G---ACA----TC-----------C--C-TG-----TC-TGAG--CAC-A-C-T-A--CT---------G--A-----GCA---------AA--GA---A--T--CTGG-AGGC--AA--G-AA-------GA-A-CATTT....................A-T-A-. 9289
A.GM.x.MCN13                  G--AACA--GAT------------C--C--G-----TC-TGA---CA----T-T-A-CCTGC-------G--A-----GCAT------------C----A-----CTGG-AGGC---A--G-AA-------GA-A-CATTT....................A-T-AG. 9329
A.GM.x.MCR35                  ........................C--C--G-----TC-TGAG--CA----T-T-A-CCTG--------G--A-----GCAT------------C----A-----CTGG-AGGC--AA--G-AA-------GA-A-CATTT....................A-T-AG. 9204
A.GW.86.FG_clone_NIHZ         .............................................A----T-T-AC-CTG--------G--A-----GCAT--T-----A---C-----A-----TGG-AGGCG--A--G-AA-----G--AA-A-CATTT....................A-T-A-. 9085
A.GW.87.CAM2CG                G---ACA--G-TC-----------CT-C-TG------CTT-AG---------C-T-AC-C-------------A-----GCA-------------AC----A--T---TGG-AGGCG-AA--G-AA-------GA-A-CATTT....................A-T-A-. 9905
A.GW.x.MDS                    G--AACAT---TT----G------CT-CCTG---------GAG---ACA--C-T-AAACTG--------G--A-----GCA-C-----------AC----A--G--ATGG-AGGCG--A--G-AA-----G--AA-A-CCTTT....................AA--A-. 9355
A.IN.07.NNVA                  G---ACA----T-----GA--------C--AT-A-----TGA-----TA--C-T-AGGCT-C--------G--A-----GCA-------------AC-----------TGG-AGGCG--AT---AG-----G--A--A-CATTT....................AAT-A-A 9874
A.IN.95.CRIK_147              ..........................................GA---CAT---C-TCAGATT-----T---G--A------CA-----------T-AC-----------TGG-AGGCG--A--G-AA-----G--AA-A-CCTTT....................A---A-. 9382
A.JP.08.NMC786_clone_41       G--AACA--GATG-----------CT-C-TG--A-----T-AC--CATA--C-TCAC-CT---T-------------GTA-C-------AT--C----A--G---TGG--GGC--AA--G-AA-------GA-A-CAC-T....................AA--A-. 9871
A.PT.x.ALI                    G---ACAT---TT----GA-----C--T-AG-----TC-TGA---CA-A--C-T-A--CT-C-------G--A-----GTA-------------C----A--T---TGG-AGGC---A--G-AA-------GA-A-CATTT-GTAAGA---G---C......A-----. 9904
A.SN.85.ROD                   G---ACA----TC---G----------CTTG-----T--T-G---C-----T-T-A--C-G-------G--A-----GCA-------------C-----------TGG-AGGCG--A--G-AA-------AA-A-CATTT....................A-T-A-. 9327
A.SN.86.ST_JSP4_27            ---AACAT---TT----G------C--C--G--A--T-TT-GC--C-----C-T-A--C---------G-------GTA-------------C-------T--ATGG-AGGC---A--G-AA-------GA-A-CGTTT....................A---A-. 9329
B.CI.88.UC1                   G---ACC--T--C---CG------C--CCTC-----AC-TGAC----TA--C-TCAAC--G-----------------GTATC-------AT-AC-----A-G---TGA-AGGCT--A----AA-------GA-A-C--CA....................GA--AG. 9802
B.CI.x.20_56                  G--AACC--T-TC---C-------C--CCTC-----AC-TGAC---ATA--C-TCAAC-AG-T----------------GTATC-------G--AC-----A-----TGG-AGGCT--A--G-AA-------AA-A-CC-CA....................GAT-AG. 9767
B.CI.x.EHO                    ----ACC--T-TC---C-------CT-CCTC--A--A--TGAC----T---C-TCAGC--G-T------G--------GTATC-------AA--C-----A-----TGG-AGGCT-AA--G-GA-------AA-A-C--CA....................GAG-AG. 9777
B.GH.86.D205_ALT              G---AC---TATC---C--------T-CCTC-----A--TGA-----T---T-TCAA---G-TT-----------C--GTA-C-------G--AC-------G---TGG-AGGCT--A----AA-------GA-A-C--CA....................GAT-AG. 9806
B.JP.01.IMCJ_KR020_1          ----ACC--TATC---C-------CT-CCTC-----A--TGAC---ATA--C-TCAA---G--------------------TA-C-------G--AC-----A-----TGG-AGGTT-------AA-------GA-A-C--CA....................GAT-AG. 8908
G.CI.92.Abt96                 C--A--GT-G--G-----------C--T--AT-A-----TGAC--CAG-----T-A-ATTGC-T--T--G-----CA-GT-G---------G--AC-----A-C--TATGG-A-GAG-AA---AACAG-----A--G-C--TA....................GAA-A-. 9275
AB.CM.03.03CM_510_03          G---AC---T-TC--------------CCTC-----GC-TGA-----T---C-TCAA---G-T-----------C--GTA-C-A-----G---C-----A-----TGG-AGGCT-AA---AA-------GA-A-C--CA....................GAT-AG. 8894
H2_01_AB.CI.90.7312A          G---AC---T-TC---C--------T-CCTC-----A--TT-AG-------C-TCAAC--G-T----------------GTATC-------AT-AC-----A-G--ATGG-AGGCT--A----AA---------A-A-C--CA....................GAG-AG. 9809
H2_01_AB.JP.04.NMC307_20      G---AC---T-TC---C----C----CTTC-----A--TT-A--------C-TCAAC-AG-T----------------GTATC-------AT-AC-----A-G--ATGG-AGGCG--A--G-AA---------A-A-C--CA....................GAG-AG. 9791
H2_01_AB.JP.07.NMC716_01      G---AC---T-TC---C----C----CCTC-----A--TT-A--------C-TCAAC--G-T---------G---GTATC-------AT-AC-----A-G--ATGG-AGGCG--A--G-AA---------A-A-C--CA....................GAG-AG. 9781
H2_01_AB.JP.08.NMC842_10      G---AC---T-TC---C--------T-CCTC-----G--TT-A--------C-TCAAC--GC-----------------GTATC-------AA--C----AA-G--ATGG-AGGCT--A--G-AA---------A-A-C--CA....................GAG-AG. 9808
U.CI.07.07IC_TNP3             ---A--AT-G---------------G----G-----T--TGAC---A-A--G-TCA-ACTG--T---------------T--T--------A---C-----A---------AG--------G-------------T-G------------------------A-----. 9383
U.FR.96.12034                 -----CCT---T-------------T-CT-------A--T-GC---C------T-A---AG-G---------------G--T---------T-------------------AG--G-------------------C-CAGTA-A-AA-...............T-----. 9377
MAC.US.x.17EC1                ---------------------------------------------------------------------------------------------------------------------------------------------------------------------------. 9868
MAC.US.x.251_1A11             --------C---------------C------------------------------------------------------------------------------------------------------------------------------------------A-A---. 9864
MAC.US.x.251_32H_PJ5          ---------------------------------A-----------------------------------------------------------C---------------------------------------------------------------------------. 9866
MAC.US.x.251_BK28             ------------G--------------------A----------------------C---------------.----------------------------------------------------------------------------------G-------------. 9849
MAC.US.x.MM142_IVMXX          ---------------------------------A--------------------A--------------------------------------------A---------A----------G------------------G-A----------G----------A-----. 9360
MNE.US.82.MNE_8               -----------T---------------------A--------------------A--------------------------------------------------------------------------------CT-A--A-------------G-------A-----. 9342
MNE.US.x.MNE027               -----------T---------------------A--------------------A--------------------------------------------------------------------------------CT-A--A---------------------A-----. 9345
SMM.CI.79.SIVsmCI2            G-----CT-G--------------C---CAAT----T--T-AC---------TCA-A-------T----------TTC---------G----------T------AG-A-------------------T-G--A-------------------A-A-----. 9011
SMM.LR.89.SIVsmLIB1           ------G--G------GC---------T-AA--A--T----C---------------------------------A----T-----------T--------T------AG--------------------T-A--A---------------------A-------. 8924
SMM.SL.92.SIVsmSL92A          ----ACC-----C---------A----T-TGT-A--T----C----------T-----------------T-G--------A--A-------------AGC-T----GG-TAGC--GAAGAAG-C----CACAAAA-GC-AC-T-GCG-AGAAGA-G 8939
SMM.SL.92.SL92B               -A----C-----C---------A----TCAC--A--------A-----A----T---A--GC---------------GT-G---------T--A-------------GAG------T--G-TAAC-AGCCAAAA-CGC-G-A-AAGATGGC-G-TA-G-AGGAAAC-A 9320
SMM.US.04.G078                G-----------G---------------CAA--A--------A--------T-T-A-----------------------T-TGC---------T----------T------CAG----------------------T-A--A------------GGA-G--ACA-..... 9157
SMM.US.04.G932                ---AAG------C--------------C--AT-..............-------T-A--A-----------A-----T-------------T-----------------AG----A------------------T-A--A---------------------A-..... 9121
SMM.US.04.M919                -------T-G--------------------AA--A--------G--------T-T-A-----C----------------T--TG-T------T-----A-----------A-AG-A----------------------T-A--A---------------------AA-..... 9159
SMM.US.04.M922                -------T-G---------------T--CAG--A-------AC--------T-T-A-----C----------------T-AG---------T-----------------A-AG-------------A-----------T-A--A---------------------AA-..... 9138
SMM.US.04.M923                ------AT-G------------------GAGT-A--T--T-AC---A------T-A--GAGC-----------------T--T-T-------T----------------G-AG-------------------------T-A--A---------------------AA-..... 9146
SMM.US.04.M926                ----AC-T-----------------T--CAG--A-------GA---------C-T-A----C-----------------T--T------------------------------AG-----------------------T-A--A---------------------AA-..... 9155
SMM.US.04.M934                ---A-----G------------------CAG-----T------A-----A----T-------C-----------------T--TG--------T--C---------------AG-----------------------T-A--A------------G------AA-..... 9131
SMM.US.04.M935                ---------G---------------T--CAA--A-------AA--------T-T-----A-------------------C--T---------T-----------------A-AG-----------------------T-A--A---------------------AA-..... 9138
SMM.US.04.M940                T--AACC--G--------A-----------A-----TC-T-GC-----A----TCA-------------G---------T-G-GC-----G-----------------------AG-A---------------------T-A--G---------------A-G--ACA-----. 9194
SMM.US.04.M946                ---A-----G------------------CAA-----T-----A-----------T-A-----C-----------------T--T--------T--C---------------AG-----------------------T-A--A---A------------G--AA-..... 9159
SMM.US.04.M947                ------A--G------------------GAAT-A--T--T-AC----T-----T-A---A--G-----------------T--T--------TT----------------------AG-----------------------T-A--A---------------------AA-..... 9143
SMM.US.04.M949                T--AACC--G-T------A-----------A-----TC-T-GA-----------TCA-------------G---------T-G-GC-----G-----------------------AG-----------------------T-A--G------------GGA-G--ACA-----. 9128
SMM.US.04.M950                ---A-----G---------------T--CGC-----T-----G-----------T-A---A-------------------C--T---------T------------------------G-----------------------T-A--A---------------------AA-..... 9146
SMM.US.04.M951                T--AACC-----G-----------------A-----TC-T-AC-----A----TCA-------------G---------T-G-GC-----G-----------------------AG-----------------------T-A--G-------------GA-G--ACA---A-. 9149
SMM.US.04.M952                ---A-----G------------------CAG-----T-----G-----------T-A---A--G------G---------T--T--------T-------------------------C-----------------------T-A--A---------------------AA-..... 9135
SMM.US.05.D215                ------A-----------A-----G-----C-----T--T-AC-----------T-A---A-------------A-----T---------------------------------------GC-C--------------------T-A--A------------GG-A--CAG....... 9128
SMM.US.06.FTq                 -A----CT-G--------------C--C--------T--T--C-----------A--AG-----------T-----GGA----------T-----------------------AG-----------------------T-A--A---------------------A-..... 9149
SMM.US.11.SIVsmE660_FL10      ------A--G------------------GAAT-A--T--T-GC---A------T-A---AG-----------------T--T---------T-----------------A-AG-----------------------A---A---------------------AA-----. 9877
SMM.US.86.CFU212              ---A--C-----------------------AT-A--T--T--C---C-----T----A-------------------TAT-----------------------------------AG-----------------------T-A--ATG-CTGACA-GA-G-A--C-AG...... 9130
SMM.US.x.F236_H4              ------A--G------------------GAAT-A--T--T-GC---A------T-A---AG-----------------T--T---------T-----------------A-AG------------------------.----A---------------------AA-----. 9832
SMM.US.x.H9                   ------A--G------------------RAGT-A--T--TRAC---A------T----GAGC-----------------T--TC--------T-----A---------GCAG-----------------------T-A--A---------------A----AA--AG. 9351
SMM.US.x.PBJA                 ------A--G-------------------AGT-A--T--T-AC---A------T----GAGC-----------------T--TC--------T-----A---------GCAG-----------------------T-A--A---------------A----AA--A-. 9653
SMM.US.x.PGM53                ---------G---------------T--CAA--A-------GA--------T-TCA-----C-----------------C--T---------T-----------------A-AG-----------------------T-A--A---------------------AA-----. 9805
SMM.US.x.SME543               ------A--G------------------GAAT----T--T-GC---A------T-A---AG-----------------T--T---------T-----------------A-AG--------------------A--TA---A------------G-----AA-----. 9890
SMM.US.x.pE660.CG7G           ------A--G------------------GAAT-A--T--T-GC---A------T-A---AG-----------------T--T-------A-T-----------------A-AG-----------------------A---A---------------------AA-----. 9873
STM.US.89.STM_37_16           ---A--A--G-TT-----A---------CTAT-A---C-T-----------C-T-------C-----------------TTCT---------T--C--A-----------GAG--------------------------A--G-------------------AA-----. 9522
```

```
                                                                                                   NF-κ-B-II      NF-κ-B-I                                                 TATA box
MAC.US.x.239               ...........................................................AACA..GCAGGGACTTTCCACAAGGGG.ATGTTACG..GGGAGGTACTGGGGAGGAGCCGGTCGGGAACGCCC.ACTTTCTTGATGTATAAATATC.ACTGCATTTCG 9969

A.CI.88.UC2                .....CAGGAACAACCATACTTGGTCAAAGCAGGAAGTAGCTACT......AAGAAATAGCTG-G-CT--------------G------.C---A--CAA------G--AT---------T---G----------.T-A-A---AC---------G-ACC-GCT.-CC--10003
A.DE.x.BEN                 .....CAGGAACAACCATACTTGGTCAAGGCAGGAAGTAGCTACT......AAGAAACAGCTG-GGCT--------------G------.C---A--CAA------G--AT---------T---G----------.T-A-A---AC---------G-ACC-GCT.-C-T-10049
A.DE.x.PEI2_KR_KRCG        ..AGACAGGAACAGCTATATTTGGTCAGAACAGGAAGTAGATGAT.......GAA..........-CT--------------G------.C---A--CAG------G--GT-------A-----G----------.T-A-A---.C---------G-ACC-GCT.GC-T- 9984
A.GH.x.GH1                 ..CAACAGGAACAGCCATACTTGGTCAAGGCAGGAAGTAGCTACT.GAGAACAGCTGAGGCTGCAGCT--------------G------.C---A--CAA------G--AT---------T---G----------.T-A-A---AC---------G-ACT-GCT.-C-TA 9471
A.GM.87.D194               .....CAGGAACAACCATACTTGGTCAGGGCAGGAAATAGCTACT......AAGAA.CAGCTG-G-CT--------------G------.C---A--CAA------G--AT---------T---G----------.T-A-AT-CTC---------G-ACC-GCT.-C-T- 9462
A.GM.x.ISY_SBL_6669_85     ..AAACAGGAACAACCATACTTGGTCAGGACAGGAAGTAGCTACT.......GAAAACAGCTG-G-CT--------------G------.C---A--CAG------G--AT---------T---G----------.T-A-A---TC---------G-ACC-GCT.GC--- 9447
A.GM.x.MCN13               ..AGACAGGAACAGCTATATTTGGCCAGGGCAGGAAATAACTACT.......GAAAACAGCTG-G-CT--------------.G------.C---A--CAG------G--AT-------------G----------.T-A-A---TC--------G--ACC-GCT.GC-T- 9486
A.GM.x.MCR35               ..AGACAGGAACAGCTATATTTGGCCAGGGCAGGAAATAACTACT.......GAAAACGGCTG-G-CT--------------.G---A--.C---A--CAG-----GG-AT------G------G----------.T-A-A---TC--------G--ACC-GCT.GC-T- 9361
A.GW.86.FG_clone_NIHZ      ..AAACAGGAACAGCTATACTTGGTCAGGGCAGGAAGTAGCTACT.......GAGAACAGCTG-G-CT--------------G------.C---A--CAG-A----G--AT---------T---G----------TT-A-A---TC---------G-ACC-GCT.GC-T- 9244
A.GW.87.CAM2CG             ..GGGCAGGAACAACCATACATGGCCAGGGCAGGAAGTAGCTACT.......GAAAACAGCTG-G-CT--------------G------.C---.--CAG------G-TAT---------T---G----------.T-A-A-C-TC---------TCACC-GCT.GC-T-10062
A.GW.x.MDS                 .GAGACAGGAACAGCTATACTTGGTCAGGGCAGGAAGTAACTACT.......GAAAACAGCTG-G-CT--------------G------.C---A--CAA------G--AT---------T---G----------.T-A-A-C-TC---------G-ACC-GCT.GC-T- 9514
A.IN.07.NNVA               GGAGACAGGAACAACTATACTTGGTCAGGGCAGGAAGTAGCTATT.......GAAAACAGCTG-G-CT--------------G------.C---A--CAG------G--AT-------------G----------.T-A-ATC-TC---------G-ACC-GCT.GC-T-10034
A.IN.95.CRIK_147           .GAGACAGGAACAACTATACTTGGTCAGAACAGGAAGTAGCTACT.......GAAAACAGTTG-G-CT--------------G-T----.C---A--CAG------G--AT---------T---G----------.T-A-ATC-TC---------G-ACC-GCT..--T- 9540
A.JP.08.NMC786_clone_41    ..CAACAGGAACAACTATACTTGGTCAGGGCAGGAAGTAGCTACT.......GCAAACAGCTG-G-CT--------------G--T---.C---A--CAA------A--AT---------T---G----------.T-A-A---AC---------G-ACC-GCT.-C-T-10029
A.PT.x.ALI                 .....................TTTGGTCAGGGCAGGAAGTAACTACT.......GAAAACAGCTG-G-CT--------------G------.C---A--CAG----A-G--AT---------T---G----------.T-A-A--C.C---------G-ACC-GCT.GC-T-10044
A.SN.85.ROD                ..AGACAGGAACAGCTATACTTGGTCAGGGCAGGAAGTAACTAAC.......AGAAACAGCTG-G-CT--------------G------.C---A--CAA------G--AT---------T---G----------.T-A-AT-CTC-----------ACC-GCT-GC-T- 9486
A.SN.86.ST_JSP4_27         ..AAACAGGAACAGCTATACTTGGTCAGGGCAGGAAGTAACTAAC......AGAAAACAGCTG-G-CT--------------G------.C-----CAG------G--AT-------------G----------.T-A-A---TC---------G-ACC-GCT.AC--- 9488

B.CI.88.UC1                ...GCAAGAGACAGCAGCATAAACAGGAACTAGCTGACACTGCACAAGAGGGA.AACTAGCTG---CC--------------A---A--.C---A--A.T------G--AT--------TT---T----------.-T-AAA-CTC---------G-ACC-GCT.-C-T- 9964
B.CI.x.20_56               ...GCAGGAAACAGCAGCTTAAAGAGGAACTAGCTGACAGTGCATAAACAGG..AACTAGCTG---CT--------------AT-----.CA--A--A.A------A--AT---------T---G----------.-T-AAAACTC---------G-ACC-GCT.--CT- 9928
B.CI.x.EHO                 ...ACAGGAAATAACAGCACAAAGAGGAACTAGCTAACACTGCATAGAGAAGGAAACTAGCTG-T-CT---.----------G------.C---A--AGT------G--AT---------A--.GA---------.-TCAAGACTC---------G-ACC-GCT.--GT- 9939
B.GH.86.D205_ALT           ...GCAAGAGACAGCAGCATAAACAGGAACTAGCTGACACTGCACAAGAAGGA.AACTAGCAG---CT--------------A---A--.C---A--A.T------AG-AA------G-TT---T----------.-T-AAA-CTC---------G-ACC-GCT.-C--- 9968
B.JP.01.IMCJ_KR020_1       ...GCAGGAAATAACAGCTTGAAGAGGAACTAGCTGACAGTGCATAAGAAAG..AACTAGCTG---CT-------------GAG-----.CA--A--A.A------A--AT---------T---G----------.-T-AAACCTC---------G-ACC-GCT.--CTA 9069

G.CI.92.Abt96              ............GCAGAAAGAGGACAAGCTGA............................C---.--------------GT-T---.C---A--AGG--AG-TACTAT------GA-T---T----------..-AAAA-CTC-----------AGCTGCA.----- 9397

AB.CM.03.03CM_510_03       ...ACAAGAGACGACAGCATAAACAGGAACTAGCTGACACTGCACAAGAGGGA.AACTAGCTG---CT--------------AG--A--.C---A--A.T------G--AT--A---G-TT---T----------.-T-AAA-CTC---------G-ACC-GCT.-C-T- 9056

H2_01_AB.CI.90.7312A       ..ACAGGAAAGCAGCAGCATAAAGAGGAACTAGCTGACGCTGCATAAGAAAGGAAACTGGCTG---CT--------------G------.C---A--AGT------G--AT--------TT---G----------.-T-AAGCCTC---------G-ACC-GCT.-C-T- 9974
H2_01_AB.JP.04.NMC307_20   ...ACAGGAAAACAACAGCATAAAGGGAACTAGCTGACGCTGCATAAGAAAGGAAACTGGCTG---CT--------------G------.C---A--AGT------G--AT--------TT---G----------.-T-AAT-CTC---------G-ACC-GCT.-C-T- 9955
H2_01_AB.JP.07.NMC716_01   ..ACAGGAAAACAGCAGCATAAAGGGGAACTAGCTGACACTGCATAAGAAAGGAAACTGGCTG---CT--------------G------.C---A--AGT------G--AT--------TT---G----------.-T-AAT-CTC---------G-ACC-GCT.-C--- 9946
H2_01_AB.JP.08.NMC842_10   ..ACAGGAAGACAGCAGCATAAAGGGGAACTAGCTGACACTGCATAAGAAAGGAAACTGGCTG---CT--------------G------.C---A--AGT------G--GT--------TT---G----------.-T-AAT-CTC---------G-ACC-GCT.-C-T- 9973

U.CI.07.07IC_TNP3          ........................................GCTGAC...AGAAGGAAACTAGCTG-GCTGA-------------G--T---.C---A--AAG---------AT---T-----T--.GA----------.T--AAAACTC-----------CAAC-GCT.C--T- 9508
U.FR.96.12034              ..........................CAGAAAGGAAACAAGCTGA...............G---.--------------C--A--.GA-G--AC.CA--GA-G-GCAA--GA-GAG-CA-G----------.T-AAATGATC---------G-ACC-GCT.C--T- 9498

MAC.US.x.17EC1             ...........................................................----.---------------------.--------.-------------------------------------.---------------------ACTGCA.----- 9969
MAC.US.x.251_1A11          ..........................................................G---.---------------------.--------.---------------------------A---.---------------------ACTGCA.---T- 9964
MAC.US.x.251_32H_PJ5       ..........................................................G---.---------------------.--------.---------------------------A---.---------------------ACTGCA.----- 9967
MAC.US.x.251_BK28          ..........................................................G-T-.---------------------.---------.........-GGG-A----CC-GTC...------A---.---------------------ACTGCA.----- 9939
MAC.US.x.MM142_IVMXX       ..........................................................G---.---------------------.----C-T-.--------------------T------A---.---------------------.----------- 9461

MNE.US.82.MNE_8            ..........................................................G---.--------------T------.----C-T-..---------------------------A---.---------------------.----------- 9443
MNE.US.x.MNE027            ..........................................................G---.--------------T------.----C-T-..---G--------------T---------A---.---------------------.----------- 9446

SMM.CI.79.SIVsmCI2         ..........................................................G---.-------------G-CG----.-G--G--AGT------AT-AA------G--T---G.---------.-TA-AA-CTC-----------AAGC----T--G-A 9114
SMM.LR.89.SIVsmLIB1        ..........................................................G---.---------------------.C---A--A.T----------T--------T---G---------T.-T--CT--TC-----------AACC-GCT.-ACT- 9026
SMM.SL.92.SIVsmSL92A       GAAGCTAGCTAA..............................................G---.--------------A------.C---A--A.T------A--AT------G--T---G----------.-TAAAA-CTC-----------A.C-A--T---GC 9053
SMM.SL.92.SL92B            GCTGA.....................................................G---.---------------------AG---A.-AA.------AGTCT---.C-G-GG-AA-----------CTAAACTC--TA----------ACC-GCTT-GC-- 9428
SMM.US.04.G078             ............................................................................................................................................................... 9157
SMM.US.04.G932             ............................................................................................................................................................... 9121
SMM.US.04.M919             ............................................................................................................................................................... 9159
SMM.US.04.M922             ............................................................................................................................................................... 9138
SMM.US.04.M923             ............................................................................................................................................................... 9146
SMM.US.04.M926             ............................................................................................................................................................... 9155
SMM.US.04.M934             ............................................................................................................................................................... 9131
SMM.US.04.M935             ............................................................................................................................................................... 9138
SMM.US.04.M940             ..............................GAAACAAG........................................................................................................................ 9202
SMM.US.04.M946             ............................................................................................................................................................... 9159
SMM.US.04.M947             ............................................................................................................................................................... 9143
SMM.US.04.M949             ...........................................................G-A-..CAG.......................................................................................... 9135
SMM.US.04.M950             ............................................................................................................................................................... 9146
SMM.US.04.M951             ..........................................................G----.-.............................................................................................. 9155
SMM.US.04.M952             ............................................................................................................................................................... 9135
SMM.US.05.D215             ............................................................................................................................................................... 9128
SMM.US.06.FTq              ............................................................................................................................................................... 9149
SMM.US.11.SIVsmE660_FL10   ..........................................................G---.---------------------.C---C-T-..-------------------T--CT.---------.----AT-CTC-----------A.----------- 9977
SMM.US.86.CFU212           ............................................................................................................................................................... 9130
SMM.US.x.F236_H4           ..........................................................G---.---------------------.C---C-T...-------------------T--CT.---------.----AT-CTC----------CA.----------- 9931
SMM.US.x.H9                ..........................................................G---.----------------A---G---------.-------T-----------T--CT.---------.----AT-CTC----------CA.----------- 9452
SMM.US.x.PBJA              ..............................GACAGC.AGGGACTTTCCACAAA.........G---.--------G--------A---.GATGCTACG.-------------------T--CT.---------.--G-AT-CTC----------CA.----------- 9776
SMM.US.x.PGM53             ..........................................................G---.----------------A---G---------.-------------------T--CT.---------.----AT--TC----------CA.----------- 9906
SMM.US.x.SME543            ..............................GACAGC.AGGGACTTTCCACA........................--------------------.C---C-T-..-------------------T--CT.---------.----AT-CTC----------CA.-----------10004
SMM.US.x.pE660.CG7G        ..........................................................G---.---------------------.C---C-T-..-------------------T--CT.---------.----AT-CTC-----------A.----------- 9973

STM.US.89.STM_37_16        ..........................................................G---.---------------------.C---A--A.G-------------AG-A-CTG-TG.-----------.TT--A---TC----------CA.--C--T-A-A- 9623
```

```
                                                  TAR element start
                                      3' LTR U3 end 3' LTR repeat start                                                                                            TAR element end                 Poly-A signal
MAC.US.x.239                CTCTGTA.TTCAGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCTCCAGCACTAGCAGGTAGAGCCTGGGTGTTCCCTGCTAGACTCTCA.CCAGCACTTGGCCGGTGCTGGGCAGAGTGACTCCACGCTTGCTTGCTTA..AAGCCCTCTTCAATA10135
A.CI.88.UC2                 -A-----.--T--C------------------------C--------A---------------------C---------C-----------------------.----TG--C------CA---------.C-G-------------------..---A------.----10167
A.DE.x.BEN                  -AT----.------------------------------C--------A----------------------------------------------G--------.----T-----------A---------.C-G-------------------..---A-----------10214
A.DE.x.PEI2_KR_KRCG         -AT----.------------------------------C----------------------------------------------------------------.----T-----------A---------.C-G-------------------..---A------.----10148
A.GH.x.GH1                  -AT----.--................................................................................................................................................................ 9480
A.GM.87.D194                -AT----.---............................................................................................................................................................... 9472
A.GM.x.ISY_SBL_6669_85      -AT----.--------------------------------------------------------------------------------------G--------.----TG---------C----------.C-G-------------------.A---A------.---- 9612
A.GM.x.MCN13                -AT----C-----------------------------------------------------------------------------------------------.--G-TG---------CA---------.C-G-------------------.A---A------.---- 9652
A.GM.x.MCR35                -AT-A--C-----------------------------------------------------------------------------------------------.--G-TG---------CA---------.C-G-------------------.A---A------.---- 9527
A.GW.86.FG_clone_NIHZ       -AT----.-------------.-C-------------------------------------------------------------------------------.----TG---------C----------.C-G-------------------..---A----.------ 9407
A.GW.87.CAM2CG              -AT----C-----------------------------------------------------------------------------------------------.----TG---------CA---------.C-G-------------------..---A------.----10227
A.GW.x.MDS                  -AT----C---............................................................................................................................................................... 9525
A.IN.07.NNVA                -AT----C------------------------------C---------AG.....--CT----.T----------------------...A------------.----TG---------C----------.C-G-------------------..---A-----------10191
A.IN.95.CRIK_147            -AT----C--------------------------------------------------------------------------------------G--------.----TG---------CA---------.C---------------------..---A------.---- 9705
A.JP.08.NMC786_clone_41     -AT----.------------------------------C--------A----------------------------------------------G--------.----T-----------A---------.C-G-------------------..---A------.----10193
A.PT.x.ALI                  -AT----.-----------------------------------------------------------------------------------------------.----TG---------CA---------.C-G-------------------..---A------.----10208
A.SN.85.ROD                 -AT----C---G----------------------------------------------------G--------------------------------------.--------------------------.C-G-C-----------------..--AA----C-T---- 9652
A.SN.86.ST_JSP4_27          -AT----.-----------------------------------------------------------------------------------------------.----TG---------CA---------CG-CTC-ACGCT-GCT-GCT-A-..---A------.---- 9653
B.CI.88.UC1                 -AT----.------------------------------C----------------------------------------------------------------.--------------------------.C-G-------------------..--A-------.----10128
B.CI.x.20_56                -AT----.AC----------------------------C----------------------------------------------------------------.-------------A------------.C-G-------------------..--A-------.----10092
B.CI.x.EHO                  -AT----.------------------------------C----------------------------------------------------------------.--------------------------.C-G-------------------..--A-------.----10103
B.GH.86.D205_ALT            -AT----.-----------------------------------------------------------------------------------------------.--------A----A---T--------.C-G-------------------..--A-------.----10132
B.JP.01.IMCJ_KR020_1        -AT----.AC----------------------------C----------------------------------------------------------------.--------------------------.C-G-------------------..--A-------.---- 9233
G.CI.92.Abt96               -------.-------------------------------------C---------------------------------------------------------.---------------------------C-G-------------------CA-C-A------.---- 9564
AB.CM.03.03CM_510_03        -AT----.AC------------------------........................................................................................................................................ 9089
H2_01_AB.CI.90.7312A        -AT----.-----------------------------------------------------------------------------------------------.--------------------------.C-G-------------------..--A-------.----10138
H2_01_AB.JP.04.NMC307_20    -AT----.-----------------------------------------------------------------------------------------------.--------------------------.C-G-------------------..--A-------.----10119
H2_01_AB.JP.07.NMC716_01    -AT----.--------------------------------------------------------------------------------------G--------.-------------A------------.C-G-------------------..--A-------.----10110
H2_01_AB.JP.08.NMC842_10    -AT----.--------------------------------------------------------------------------------------G--------.--------------------------.C-G-------------------..--A-------.----10137
U.CI.07.07IC_TNP3           -AT----.------------------------.......................................................................................................................................... 9540
U.FR.96.12034               -AT----.-----------------------------------------------------------------------------------------------.---AA-G........................................................... 9607
MAC.US.x.17EC1              -------.-----------------------------------------------------------------------------------------------.-------------------------------------------------..---------------10135
MAC.US.x.251_1A11           --A----.G----------------------------------------------------------------------------------------------.-------------------------------------------------..---------------10130
MAC.US.x.251_32H_PJ5        -------.-----------------------------------------------------------------------------------------------.-----------------------------G-------------------..---A-----------10133
MAC.US.x.251_BK28           -------.-----------------------------------------------------------------------------------------------.-------------A---------------G-------------------..---A-----------10105
MAC.US.x.MM142_IVMXX        -------.-----------------------------------------------------------------------------------------------.-----------------------------G-------------------..---A----------- 9627
MNE.US.82.MNE_8             -------.-----------------------------------------------------------------------------------------------.-----------------------------G-------------------..---A----------- 9609
MNE.US.x.MNE027             -------.-----------------------------------------------------------------------------------------------.--------------------T--------G-------------------..---A----------- 9612
SMM.CI.79.SIVsmCI2          --G-A-T.C-------------------------........................................................................................................................................ 9147
SMM.LR.89.SIVsmLIB1         -------.---------------------------------------A-T-------------G---------------------------------------.-------------------------G---G-------------------..---A------T---- 9192
SMM.SL.92.SIVsmSL92A        ACTGTAT.-------------------------......................................................................................................................................... 9086
SMM.SL.92.SL92B             -------.------------A-----------C-----A------C-A----------G-------TAGTA-GT--------------T--A---A-------A------ACA----AT-------T---.C-G------------------G......A----C.---- 9589
SMM.US.04.G078              .......................................................................................................................................................................... 9157
SMM.US.04.G932              .......................................................................................................................................................................... 9121
SMM.US.04.M919              .......................................................................................................................................................................... 9159
SMM.US.04.M922              .......................................................................................................................................................................... 9138
SMM.US.04.M923              .......................................................................................................................................................................... 9146
SMM.US.04.M926              .......................................................................................................................................................................... 9155
SMM.US.04.M934              .......................................................................................................................................................................... 9131
SMM.US.04.M935              .......................................................................................................................................................................... 9138
SMM.US.04.M940              .......................................................................................................................................................................... 9202
SMM.US.04.M946              .......................................................................................................................................................................... 9159
SMM.US.04.M947              .......................................................................................................................................................................... 9143
SMM.US.04.M949              .......................................................................................................................................................................... 9135
SMM.US.04.M950              .......................................................................................................................................................................... 9146
SMM.US.04.M951              .......................................................................................................................................................................... 9155
SMM.US.04.M952              .......................................................................................................................................................................... 9135
SMM.US.05.D215              .......................................................................................................................................................................... 9128
SMM.US.06.FTq              .......................................................................................................................................................................... 9149
SMM.US.11.SIVsmE660_FL10    -------.-----------------------------------------------------------------------------------------------.-----------------------------G-------------------..---A-----------10143
SMM.US.86.CFU212            .......................................................................................................................................................................... 9130
SMM.US.x.F236_H4            -------.-----------------------------------------------------------------------------------------------.-----------------------------G-------------------..---A-----------10097
SMM.US.x.H9                 -------.-------A------------------------------------------------------------------------------G--------.----------------------------RG-------------------..---A----------- 9618
SMM.US.x.PBJA               -------.-------A---------------------------------------------------------------------------------------.-----------------------------G-------------------..---A----------- 9942
SMM.US.x.PGM53              -------.-----------------------------------------------------------------------------------------------.-----------------------------G-------------------..---A-----------10072
SMM.US.x.SME543             -------.G----------------------------------------------------------------------------------------------.-----------------------------G-------------------..---A-----------10170
SMM.US.x.pE660.CG7G         -------.-----------------------------------------------------------------------------------------------.-----------------------------G-------------------..---A-----------10139
STM.US.89.STM_37_16         -------.-------A---------------------------------------------------------------------------------------.-------------A--------T------G-------------------..T-TAT---------- 9789
```

```
                               Poly-A signal 3' LTR R repeat end 3' LTR U5 start                                                          3' LTR U5 end
MAC.US.x.239                   AA.GCTGCCATTTTAGAAGTAAGCT.AGTGTGTGTTCCCATCTCTCCTAGCCGCCGCCTGGTCAACTCGGTACTCAATAA..TAAGAAGACCCTGGTCTGTTAGGACCCTTTCTGCTTTGGGAAA.CCGAAGCAGGAAAATCCCTAGCA 10279
A.CI.88.UC2                    --.-------A--.-----C---T-A------------------------T-------------.T-----GT---TCTGAG---C-------------------------C-C----------.T--A----G--------------- 10312
A.DE.x.BEN                     --.-------G--.-----C---T-A------------------------T-------------.T-----GT---TCTGAG---C-------------------------CTC----------.T--A-G------------------ 10359
A.DE.x.PEI2_KR_KRCG            --.-------G--.-----C---T-A........................................................................................................................... 10172
A.GH.x.GH1                     ..................................................................................................................................................... 9480
A.GM.87.D194                   ..................................................................................................................................................... 9472
A.GM.x.ISY_SBL_6669_85         --.-------G--.-----C---T-A........................................................................................................................... 9636
A.GM.x.MCNI3                   --.-------G--.-----C---T-A------------------------T------------...................................................................................... 9713
A.GM.x.MCR35                   --.-------G--.-----C---T-A------------------------T------------...................................................................................... 9588
A.GW.86.FG_clone_NIHZ          --.-------A--.-----C---T-A........................................................................................................................... 9431
A.GW.87.CAM2CG                 --.-------G--.-----C---T-A------------------------T-------------.T-----GT---CCTGAA---C-------------------------CT------------.--A-G------------------ 10372
A.GW.x.MDS                     ..................................................................................................................................................... 9525
A.IN.07.NNVA                   --.-------G--.-----C---TCA------------------------T-------------.G-----GT---CCTGAG---C-------------------------CT---------.--T--A-G--T--------------- 10336
A.IN.95.CRIK_147               --.-------G--.-----C---T-A-------------------T----T-------------.T-----GT---TCTGAA---C-------------------------CT------------C----G---------------T-- 9851
A.JP.08.NMC786_clone_41        --.-------A--.-----C---T-A------------------------T-------------.------GT---TCTGAG---C-------------------------...................................... 10301
A.PT.x.ALI                     --.-------A--.-----C-G-T-A-AG---------------------T-------------.T-----GT---CCTGAG---C---------------------------------------.-G--G------------------ 10353
A.SN.85.ROD                    --.-------G--.-----C-................................................................................................................................ 9671
A.SN.86.ST_JSP4_27             --.-------G--.-----C-................................................................................................................................ 9672
B.CI.88.UC1                    --..------A--.-----C---T-G------------------------T-------------.T--AA-G---T-CT-AAA.TA---------------------------------------.--A-G------------------ 10271
B.CI.x.20_56                   --..------C--.-----C---T-G------------------------T-------------.T--A--G---C-CTGAG-.TC--------------------------------C----G-.-T--G------------------ 10235
B.CI.x.EHO                     --.A-----TA--.-----C---ACA------------------------T-------------.T-----G---C-CT-G.....-A----------------------C--------------.--A-G------------------ 10242
B.GH.86.D205_ALT               --.-------A--.-----C---.-A------------------------T--.----------.------G---C-CTGA.....T------------------------CT------------.-------G--------------- 10269
B.JP.01.IMCJ_KR020_1           --.------TA--.-----C---T-G------------------------T-------------.------G-----CTGAGA.TA-------------------------...................................... 9339
G.CI.92.Abt96                  --.-------AA------------AA.--------GT................................................................................................................ 9599
AB.CM.03.03CM_510_03           ..................................................................................................................................................... 9089
H2_01_AB.CI.90.7312A           --.-------A--.-----C---TCA------------------------T-------------.------GT---GCTGAA-.-A-G-----------------------CT------------.--T-------------------- 10282
H2_01_AB.JP.04.NMC307_20       --.------TA--.-----C---TCA------------------------T-------------.T-----GTC--GCTTT.....-G-----------------------...................................... 10222
H2_01_AB.JP.07.NMC716_01       --.------TA--.-----C---TCA------------------------T-------------.T-----GTC--GCTTT.....-G-----------------------...................................... 10213
H2_01_AB.JP.08.NMC842_10       --.------TA--.-----C---TCA------------------------T-------------.T-----GTC--GCT-T.....-G-----------------------...................................... 10240
U.CI.07.07IC_TNP3              ..................................................................................................................................................... 9540
U.FR.96.12034                  ..................................................................................................................................................... 9607
MAC.US.x.17EC1                 --.----------------------.------------------------------------------------------..-------------------------------------------.----------------------- 10279
MAC.US.x.251_1A11              --.----------------------.------------------------T------------------------GG---..-------------------------------------------.----------------------- 10274
MAC.US.x.251_32H_PJ5           --.----------------------.------------------------T------------------------GG---..-------------------------------------------.----------------------- 10277
MAC.US.x.251_BK28              --.---------------------C.------------------------T------------------------GG---..---------------------------------------A---.----------------------- 10249
MAC.US.x.MM142_IVMXX           --.----------.-------................................................................................................................................ 9646
MNE.US.82.MNE_8                --.------T---.-------................................................................................................................................ 9628
MNE.US.x.MNE027                --.----------.----------C.--------C---------------T----.............................................................................................. 9664
SMM.CI.79.SIVsmCI2             ..................................................................................................................................................... 9147
SMM.LR.89.SIVsmLIB1            --.-------GA-.-----C----A.---------.................................................................................................................. 9224
SMM.SL.92.SIVsmSL92A           ..................................................................................................................................................... 9086
SMM.SL.92.SL92B                --G-----A-C--.-----C---TC............................................................................................................................ 9613
SMM.US.04.G078                 ..................................................................................................................................................... 9157
SMM.US.04.G932                 ..................................................................................................................................................... 9121
SMM.US.04.M919                 ..................................................................................................................................................... 9159
SMM.US.04.M922                 ..................................................................................................................................................... 9138
SMM.US.04.M923                 ..................................................................................................................................................... 9146
SMM.US.04.M926                 ..................................................................................................................................................... 9155
SMM.US.04.M934                 ..................................................................................................................................................... 9131
SMM.US.04.M935                 ..................................................................................................................................................... 9138
SMM.US.04.M940                 ..................................................................................................................................................... 9202
SMM.US.04.M946                 ..................................................................................................................................................... 9159
SMM.US.04.M947                 ..................................................................................................................................................... 9143
SMM.US.04.M949                 ..................................................................................................................................................... 9135
SMM.US.04.M950                 ..................................................................................................................................................... 9146
SMM.US.04.M951                 ..................................................................................................................................................... 9155
SMM.US.04.M952                 ..................................................................................................................................................... 9135
SMM.US.05.D215                 ..................................................................................................................................................... 9128
SMM.US.06.FTq                  ..................................................................................................................................................... 9149
SMM.US.11.SIVsmE660_FL10       --.----------.----------.A--------C---------------T-------------T----------G-C-CA.-------------------------------------------.----G------------------ 10287
SMM.US.86.CFU212               ..................................................................................................................................................... 9130
SMM.US.x.F236_H4               --.-------A--.----------.A------------------------T-------------T----------G-C-CA.---------------------------------------A---.----G------------------ 10241
SMM.US.x.H9                    --G-------G--.-------................................................................................................................................ 9638
SMM.US.x.PBJA                  --A----------.---------CA.------------------------T----.............................................................................................. 9996
SMM.US.x.PGM53                 --.------GC--.----------.A------------------------T-------------T----------G-ATCA.----------------------------------................................. 10184
SMM.US.x.SME543                --.----------.----------.A------------------------T-------------T----------G-C-CATA------------------------------------------.----G------------------ 10315
SMM.US.x.pE660.CG7G            --.----------.----------.A------------------------T-------------T----------G-C-CA.-------------------------------------------.----G------------------ 10283
STM.US.89.STM_37_16            --.-------A-------------A.----G-------------------T------------------------...........------------------------------................................. 9892
```

IV

PLV Complete Genomes

Contents

IV-1	Introduction	**223**
IV-2	Sequences	**224**
IV-3	Alignments	**226**

IV-1 Introduction

Complete genomes of HIV-1, HIV-2, and SIV sequences are represented in the alignment.

For HIV-1, we include a single sequence from groups M, N, O, and P. For HIV-2, we include a single representative from groups A, B, and G, plus both available subtype U genomes. For the SIVs, we include representatives of all species for which full genome sequences are available. The principle in deciding which sequences to include in this alignment was to provide a representative sampling of all primate lentiviruses which have had complete or nearly complete genomes sequenced. Many HIV-1, HIV-2, SIVsmm and SIVmac genomes were not included, but they can be found in the HIV-1/SIVcpz and HIV-2/SIVsmm alignments.

The alignment was based on the previous version published last year, produced by a hidden Markov model using HMMER and manual editing using the programs BioEdit and Se-Al. It is codon-aligned, meaning that the correct translation reading frame has been maintained as much as possible; in the case of overlapping frames, obviously one will not translate correctly. We have tried to "reset" the alignment so that each new gene starts in the first frame; this means some empty columns have been inserted. The HIV-2/SIVsmm *vpx* gene is postulated to be a duplication of the *vpr* gene and thus there may be two alternative alignments of this region of the genome, as there are for the duplicated stem-loops of the TAR element.

The annotation is based on known protein coding regions in HIV-1 and on annotations found in SIV sequence database entries. The protein cleavage sites that create Gag p17, Gag p24 and other mature peptides from the Gag and Gag-Pol precursor polyproteins have been experimentally determined for HIV-1 and at least one strain of HIV-2, but the study of analogous cleavages in SIV polyproteins have not been published. Two representative genomes have been translated: HIV-1 subtype B strain HXB2, and SIV Mac239. The translations are provided as a visual aid for finding landmarks in the genomes.

IV-2 Sequences

Primate lentivirus (PLV) Sequences in the complete genome alignments.

Name	Accession	Country	Author	Reference
H1B.FR.83.HXB2	K03455	France	Wong-Staal, F.	*Nature* **313**(6000):277-284 (1985)
H1O.CM.91.MVP5180	L20571	Cameroon	Gurtler, L.G.	*J Virol* **68**(3):1581-1585 (1994)
H1N.CM.95.YBF30	AJ006022	Cameroon	Simon, F.	*Nat Med* **4**(9):1032-1037 (1998)
H1P.CM.06.U14788	HQ179987	Cameroon	Vallari, A.	*J Virol* **85**(3); 1403-7 (2011)
CPZ.CD.06.BF1167	JQ866001	D.R.C.	Li, Y.	*J Virol* **86**(19):10776-10791 (2012)
CPZ.TZ.00.TAN1	AF447763	Tanzania	Santiago, M.L.	*J Virol* **77**(3):2233-2242 (2003)
GOR.CM.07.SIVgorCP2135con	FJ424863	Cameroon	Takehisa, J.	*J Virol* **83**(4):1635-1648 (2009)
MAC.US.x.239	M33262	United States	Kestler, H.	*Science* **248**(4959):1109-1112 (1990)
H2A.SN.86.ST_JSP4_27	M31113	Senegal	Kumar, P.	*J Virol* **64**(2):890-901 (1990)
H2B.CI.x.EHO	U27200	Cote D'Ivoire	Rey-Cuille, M.A.	*Virology* **202**(1):471-476 (1994)
H2G.CI.92.Abt96	AF208027	Cote D'Ivoire	Brennan, C.A.	*ARHR* **13**(5):401-404 (1997)
H2U.CI.07.07IC_TNP3	KC693505	Cote D'Ivoire	Ayouba, A.	*AIDS* **27**(15):2488-2491 (2013)
H2U.FR.96.12034	AY530889	France	Damond, F.	*ARHR* **20**(6):666-672 (2004)
ASC.UG.10.RT03	KJ461716	Uganda	Lauck, M.	*Retrovirology* **11**(1):55 (2014)
ASC.UG.10.RT08	KJ461715	Uganda	Lauck, M.	*Retrovirology* **11**(1):55 (2014)
ASC.UG.10.RT11	KJ461714	Uganda	Lauck, M.	*Retrovirology* **11**(1):55 (2014)
COL.CM.x.CGU1	AF301156	Cameroon	Courgnaud, V.	*J Virol* **75**(2):857-866 (2001)
COL.UG.10.BWC01	KF214240	Uganda	Lauck, M.	*Retrovirology* **10**(1):107 (2013)
COL.UG.10.BWC07	KF214241	Uganda	Lauck, M.	*Retrovirology* **10**(1):107 (2013)
DEB.CM.04.SIVdeb04CMPF3061	FJ919724	Cameroon	Aghokeng, A.F.	*Infect Genet Evol* **10**(3); 386-96 (2010)
DEB.CM.99.CM40	AY523865	Cameroon	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
DEB.CM.99.CM5	AY523866	Cameroon	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
DEN.CD.x.CD1_CMO580407	AJ580407	D.R.C.	Dazza, M.C.	*J Virol* **79**(13):8560-8571 (2005)
DRL.DE.11.D3	KM378563	Germany	Dietrich, U.	*ARHR* **31**(3); 353-7 (2015)
DRL.x.x.FAO	AY159321		Hu, J.	*J Virol* **77**(8):4867-4880 (2003)
GRV.ET.x.GRI_677_gri_1	M66437	Ethiopia	Fomsgaard, A.	*Virology* **182**(1):397-402 (1991)
GSN.CM.99.CN166	AF468659	Cameroon	Courgnaud, V.	*J Virol* **76**(16):8298-8309 (2002)
GSN.CM.99.CN71	AF468658	Cameroon	Courgnaud, V.	*J Virol* **76**(16):8298-8309 (2002)
LST.CD.88.SIVlhoest447	AF188114	D.R.C.	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.CD.88.SIVlhoest485	AF188115	D.R.C.	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.CD.88.SIVlhoest524	AF188116	D.R.C.	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.KE.x.lho7	AF075269	Kenya	Hirsch, V.M.	*J Virol* **73**(2):1036-1045 (1999)
MAC.US.x.251_1A11	M76764	United States	Marthas, M.L.	*J Med Primatol* **18**(3-4):311-9 (1989)
MND-1.GA.x.MNDGB1	M27470	Gabon	Tsujimoto, H.	*Nature* **341**(6242):539-541 (1989)
MND-2.CM.98.CM16	AF367411	Cameroon	Takehisa, J.	*ARHR* **17**(12):1143-1154 (2001)
MND-2.GA.x.M14	AF328295	Gabon	Souquiere, S.	*J Virol* **75**(15):7086-7096 (2001)
MND-2.x.x.5440	AY159322		Hu, J.	*J Virol* **77**(8):4867-4880 (2003)
MNE.US.x.MNE027	U79412	United States	Kimata, J.T.	*J Virol* **72**(1):245-256 (1998)
MON.CM.99.L1_99CML1	AY340701	Cameroon	Courgnaud, V.	*J Virol* **77**(23):12523-12534 (2003)
MON.NG.x.NG1	AJ549283	Nigeria	Barlow, K.L.	*J Virol* **77**(12):6879-6888 (2003)
MUS-1.CM.01.CM1239	EF070330	Cameroon	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
MUS-1.CM.01.SIVmus_01CM1085	AY340700	Cameroon	Courgnaud, V.	*J Virol* **77**(23):12523-12534 (2003)

Name	Accession	Country	Author	Reference
MUS-2.CM.01.CM1246	EF070329	Cameroon	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
MUS-2.CM.01.CM2500	EF070331	Cameroon	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
MUS-3.GA.09.09GabOI81	KF304707	Gabon	Liegeois, F.	*Viruses* **6**(7):2880-2898 (2014)
MUS-3.GA.11.11GabPts02	KF304708	Gabon	Liegeois, F.	*Viruses* **6**(7):2880-2898 (2014)
OLC.CI.97.97CI12	FM165200	Cote D'Ivoire	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
RCM.CM.00.SIVagi_00CM312	HM803690	Cameroon	Ahuka-Mundeke, S.	*J Gen Virol* **2010** Dec;91(Pt 12):2959-64 doi: 101099/vir0025767-0 Epub 2010 Aug 25
RCM.CM.02.SIVrcm_02CM8081[redacted]	HM803689	Cameroon	Ahuka-Mundeke, S.	*J Gen Virol* **2010** Dec;91(Pt 12):2959-64 doi: 101099/vir0025767-0 Epub 2010 Aug 25
RCM.GA.x.SIVRCMGAB1	AF382829	Gabon	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
RCM.NG.x.NG411	AF349680	Nigeria	Beer, B.E.	*J Virol* **75**(24):12014-12027 (2001)
SAB.SN.x.SAB1	U04005	Senegal	Jin, M.J.	*EMBO J* **13**(12):2935-2947 (1994)
SMM.CI.79.SIVsmCI2	JX860430	Cote D'Ivoire	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.LR.89.SIVsmLIB1	JX860431	Liberia	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.SL.92.SIVsmSL92A	JX860432	Sierra Leone	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.SL.92.SIVsmSL92B	JX860433	Sierra Leone	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.SL.92.SL92B	AF334679	Sierra Leone	Chen, Z.	*J Virol* **70**(6):3617-3627 (1996)
SMM.US.04.G078	JX860415	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.05.D215	JX860413	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.06.FTq	JX860414	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.86.CFU212	JX860407	United States	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.x.H9	M80194	United States	Courgnaud, V.	*J Virol* **66**(1):414-419 (1992)
SMM.US.x.PGM53	AF077017	United States	Novembre, F.J.	*J Virol* **72**(11):8841-8851 (1998)
SMM.US.x.pE660.CG7G	JX648292	United States	Lopker, M.J.	Unpublished
STM.US.89.STM_37_16	M83293	United States	Novembre, F.J.	*Virology* **186**(2):783-787 (1992)
SUN.GA.98.L14	AF131870	Gabon	Beer, B.E.	*J Virol* **73**(9):7734-7744 (1999)
SYK.KE.x.KE51	AY523867	Kenya	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
SYK.KE.x.SYK173_COMGNM	L06042	Kenya	Hirsch, V.M.	*J Virol* **67**(3):1517-1528 (1993)
TAL.CM.00.266	AY655744	Cameroon	Liegeois, F.	*Virology* **349**(1):55-65 (2006)
TAL.CM.01.8023	AM182197	Cameroon	Liegeois, F.	*Virology* **349**(1):55-65 (2006)
TAN.UG.x.SIVagmTAN1	U58991	Uganda	Soares, M.A.	*Virology* **228**(2):394-399 (1997)
VER.DE.x.AGM3	M30931	Germany	Baier, M.	*Virology* **176**(1):216-221 (1990)
VER.KE.x.9063	L40990	Kenya	Hirsch, V.M.	*J Virol* **69**(2):955-967 (1995)
VER.KE.x.AGM155	M29975	Kenya	Johnson, P.R.	*J Virol* **64**(3):1086-1092 (1990)
VER.KE.x.TYO1_patent	DJ048201	Kenya	Omori, T.	Patent: WO 2007049749-A 13 03-MAY-2007; DNAVEC CORPORATION
WRC.CI.97.97CI14	AM745105	Cote D'Ivoire	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
WRC.CI.98.98CI04	AM713177	Cote D'Ivoire	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
WRC.GM.05.Pbt_05GM_X02	AM937062	Gambia	Locatelli, S.	*Virology* **376**(1):90-100 (2008)

5' LTR U3 start

```
H1B.FR.83.HXB2                .....TGGAAGGGCTAATTCACTCCCAACGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAG.....  158
H1O.CM.91.MVP5180             .....-----T---T-----T-------TAAG---GC---A-----G-----C-----A--T-----T--G--A-T----------G----TGT--------G--A-----ACCT----TC--------A---------TG-TT--A--G--------GTCA...GCA  161
H1N.CM.95.YBF30               ..................................................................................................................................  0
H1P.CM.06.U14788              ..................................................................................................................................  0
CPZ.CD.06.BF1167              ACGCGT----------TA-TGG--T-C-AAG---GC--CA--A--A---ACC-----G-TTA----T-----AATA---------G------T---------------AG-T-------T---TGTAGA--T----TA-TT---T--------CA--CCCACCA  170
CPZ.TZ.00.TAN1                ..................................................................................................................................  0
GOR.CM.07.SIVgorCP2135con     ..................................................................................................................................  0
```

5' LTR U3 start

```
MAC.US.x.239                  .....--------A-TTA-T--AGTGC-A-------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-A-A--A----G----G-T-----CT----A-----A--T-----C---AA---A-----C---CTA-GG--AT----C--T--AA-TGTATCA  164
H2A.SN.86.ST_JSP4_27          ..................................................................................................................................  0
H2B.CI.x.EHO                  .....----------A-TT-CT-TAGTG-GA----G--TA-A--A--A--CACA-ACT-AG-AA-TGA-G-----ATTG-GT---GA-G--------------AT--------AG-A--------CAA-TT------C---CTA-GG--------G---A-AA-TATGATA  164
H2G.CI.92.Abt96               ..................................................................................................................................  0
H2U.CI.07.07IC_TNP3           ..................................................................................................................................  0
H2U.FR.96.12034               ..................................................................................................................................  0
ASC.UG.10.RT03                ..................................................................................................................................  0
ASC.UG.10.RT08                ..................................................................................................................................  0
ASC.UG.10.RT11                ..................................................................................................................................  0
COL.CM.x.CGU1                 ..................................................................................................................................  0
COL.UG.10.BWC01               ..................................................................................................................................  0
COL.UG.10.BWC07               ..................................................................................................................................  0
DEB.CM.04.SIVdeb04CMPF3061    ..................................................................................................................................  0
DEB.CM.99.CM40                ..................................................................................................................................  0
DEB.CM.99.CM5                 ..................................................................................................................................  0
DEN.CD.x.CD1_CMO580407        ..................................................................................................................................  0
DRL.DE.11.D3                  ..................................................................................................................................  0
DRL.x.x.FAO                   ..................................................................................................................................  0
GRV.ET.x.GRI_677_gri_1        ..........T---A--TA-T----TG--A-----G--A-G-----GA--T---ATGC--TGA--GAGTGG--AATAA-AGA-----G---AGCT---T-------C--G-----A--G--C--GAGAGT------C-TC----TT--------C-----G--CCTGCAT  164
GSN.CM.99.CN166               ..................................................................................................................................  0
GSN.CM.99.CN71                ..................................................................................................................................  0
LST.CD.88.SIVlhoest447        ..................................................................................................................................  0
LST.CD.88.SIVlhoest485        ..................................................................................................................................  0
LST.CD.88.SIVlhoest524        ..................................................................................................................................  0
LST.KE.x.lho7                 ---------------TA-TGG-----GA------G--C---T-AAT---C-ACTGTG-GA-TGAGTGG--TCT-A-TGGA--C-TCAT----------GAT--A-----A-CA-----C--T-----A---------TG-GGC--T---A------G-CT------  158
MAC.US.x.251_1A11             ..............A-TTA-T-TAGTGC-A-------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-A-A--A----G----G-T-----CT----A-----A--T-----C---AA---A-----C---CTA-GG--AT----C--T--AA-TGTATCA  164
MND-1.GA.x.MNDGB1             ..................................................................................................................................  0
MND-2.CM.98.CM16              ..................................................................................................................................  0
MND-2.GA.x.M14                ..................................................................................................................................  0
MND-2.x.x.5440                ..................................................................................................................................  0
MNE.US.x.MNE027               ..................................................................................................................................  0
MON.CM.99.L1_99CML1           ..................................................................................................................................  0
MON.NG.x.NG1                  ..................................................................................................................................  0
MUS-1.CM.01.CM1239            ..................................................................................................................................  0
MUS-1.CM.01.SIVmus_01CM1085  ..................................................................................................................................  0
MUS-2.CM.01.CM1246            ..................................................................................................................................  0
MUS-2.CM.01.CM2500            ..................................................................................................................................  0
MUS-3.GA.09.09GabOI81         ..................................................................................................................................  0
MUS-3.GA.11.11GabPts02        ..................................................................................................................................  0
OLC.CI.97.97CI12              ..................................................................-AT---TA------AG..GCAACAG-GG-GGT-GA--A-GGCAC-T-A--GAGA--TAA---GCCA--G----TA-GGG-T--TAGGAG---G--ACCATTG  100
RCM.CM.00.SIVagi_00CM312      ..................................................................................................................................  0
RCM.CM.02.SIVrcm_02CM8081     ..................................................................................................................................  0
RCM.GA.x.SIVRCMGAB1           ..................................................................................................................................  0
RCM.NG.x.NG411                ..................................................................................................................................  0
SAB.SN.x.SAB1                 ..............A-TTA-T--AGTG--A-G---A-GA-G---T-G-----C-ATGCTCTTA-TGAGTGG---ATAG-AGA--GA-G------T--T---GAT--------T-C---G--C---AA-TG-------------TT--A--G-----T--G--CCTGTCA  164
SMM.CI.79.SIVsmCI2            ..................................................................................................................................  0
SMM.LR.89.SIVsmLIB1           ..................................................................................................................................  0
SMM.SL.92.SIVsmSL92A          ..................................................................................................................................  0
SMM.SL.92.SIVsmSL92B          ..................................................................................................................................  0
SMM.SL.92.SL92B               ..................................................................................................................................  0
SMM.US.04.G078                ..................................................................................................................................  0
SMM.US.05.D215                ..................................................................................................................................  0
SMM.US.06.FTq                 ..................................................................................................................................  0
SMM.US.86.CFU212              ..................................................................................................................................  0
SMM.US.x.H9                   ..................................................................................................................................  0
SMM.US.x.PGM53                ..................................................................-A----G---A--T------T----------A--------C--TA--TT------C---CTA-GG--AT-G--C-----AAGTGTCTCA  90
SMM.US.x.pE660.CG7G           ..............A-TTA-T--AATG-GAA------TAGA--AT-A---A---AC--GG-AA-GGA-G----AATAA-A--A----G---A--T------T----------A-CT-----C--TA--TA--A---G---CT--GG--AT----C-----A--TGTCTCA  164
STM.US.89.STM_37_16           ..............A-TTA-T--AGTG-..................................................................................  22
SUN.GA.98.L14                 ...............CTGGT-----AG-ACC---G----A---T-G---T---ATGCAG-AA-TGAGTGG----T-A-AA---GC-G---AG-T-----CAA------T--AG-G---------AA-G-------C---CTA-GG------C------CACCATTGAT  164
SYK.KE.x.KE51                 ..................................................................................................................................  0
SYK.KE.x.SYK173_COMGNM        ..................................................................................................-A-C-C-G--C--GT-A-T----------TA-GGG-A-----T---AA----ATAGAG  61
TAL.CM.00.266                 ..................................................................................................................................  0
TAL.CM.01.8023                ..................................................................................................................................  0
TAN.UG.x.SIVagmTAN1           .........-G--CA--TA-TGG--A-C-AAG---G-GC-A-----GA----C-ATGC-CTAA-TGA-TGG---ATTA-AGA-----G---AGCA--TT-C--T--------C-C---GA-G--C-GATG---C--C-TT----TTG--T----G--G--A--TGTCTCC  164
VER.DE.x.AGM3                 ..................................................................................................................................  0
VER.KE.x.9063                 .........T---A-TTA-T-----G--A-G--GA-TAGA-----CA-C----ATGCTCTTA-TGA-TGG--AAT-A-TGA-----G-A-CGC--GGT--A----A-----CG-A--------C-GCTG------C-TT----T----------G---A---ATGCAT  164
VER.KE.x.AGM155               .........T---A-TTA-T-----G-TA-G---A-TA-G------A------ATGC-CT-A-TGA-TGG--AAT-A-TGA-----G-A-CGCA-GGT--AA---A--T-----A-----C--GAG-TG------C-TC----T----------G----CCCTGCAT  164
VER.KE.x.TYO1_patent          .........T---A-TTA-T-----G-TA-G---A-TA-G-----GA------ATGCTCTTA-TGA-TGG--GATAA-TGA-----G-A-TGC--GGT-GAAG--A-----A--A----TC--TAAATG------G-TC----TT--------G-----G--CTTACAT  164
WRC.CI.97.97CI14              .........G-ATA--T--TGG--AGCTA-----G----A--AG-GA-C--C-ATGCAC--A-TGA-TGG--GCTGA-AGA--GC-G---AGCT------GA---A--------A----GG---AA---A-----G-------GG--AT-G--T-----AC-CTTCAGT  164
WRC.CI.98.98CI04              .........GTATA--T--TGG--T-CTA-----G----A--AG-GA-C--T-ATGC--T-A-TGA-TGG--ACTGA-AGAA-G--G----GCA------GAG--A-----C--AC-C-GG---AA---A-----C-----T-GG--AT----G--T--G--ATTTAGT  164
WRC.GM.05.Pbt_05GM_X02        ..................................................................................................................................  0
```

Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

```
                                                                                                                                                                                     TCF-1 alpha
H1B.FR.83.HXB2              CCAGAGAAGTTAGAAGAAGCCAACAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGAATG......GATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCACATGGCCCGAGAGCTGCATCCGGAGTACTTCAA 322
H1O.CM.91.MVP5180           GA----GCAGAGAG-CTG-GT--T-C-AAT--AG-TG-T--TC-TC----T--A-CTT-T......AATCATGGAGCT--G--TG-AC-CG-G--GA-AC--A-A---CA------T--ATCAT---GC--AACA--T--A----TGC-AAA---C--A---CT----CC 325
H1N.CM.95.YBF30             ........................................................................................................................................................................   0
H1P.CM.06.U14788            ........................................................................................................................................................................   0
CPZ.CD.06.BF1167            GA----......................-AT--A-GG-ATCTAC-CA-G-------CTT--TCCT------GA.....--------TA-TG-----C----G-TA----A-------TCTGAA--G--TAGAAGA-----A---A-----AGA---------------- 313
CPZ.TZ.00.TAN1              ........................................................................................................................................................................   0
GOR.CM.07.SIVgorCP2135con   ........................................................................................................................................................................   0
                                                                                                                                                                                     TCF-1 alpha
MAC.US.x.239                GAT---............--AC-GG-G-AT---G-GCATTAT--AA-G--T--A-CTCAAACTTCCCAGTG-......--------TTG-G----G--TC---CA----A------TCCAACT--G--C-ACAC-T-TGA---ATAT-TTAGAT-C--A--AG-G--TGG 316
H2A.SN.86.ST_JSP4_27        ........................................................................................................................................................................   0
H2B.CI.x.EHO                G----A...........C-AG-GG-C-AG--A-C-CATT-TC--G-G--T--A-CACAGACCTCC-C-TG-......--------CC-CGA---GACCC-T-TC---CA-------TC--T--------A-G-CT-TG-----TTCAGCAG-TTC--A---G-G--TGG 316
H2G.CI.92.Abt96             ........................................................................................................................................................................   0
H2U.CI.07.07IC_TNP3         ........................................................................................................................................................................   0
H2U.FR.96.12034             ........................................................................................................................................................................   0
ASC.UG.10.RT03              ........................................................................................................................................................................   0
ASC.UG.10.RT08              ........................................................................................................................................................................   0
ASC.UG.10.RT11              ........................................................................................................................................................................   0
COL.CM.x.CGU1               ........................................................................................................................................................................   0
COL.UG.10.BWC01             ........................................................................................................................................................................   0
COL.UG.10.BWC07             ........................................................................................................................................................................   0
DEB.CM.04.SIVdeb04CMPF3061  ........................................................................................................................................................................   0
DEB.CM.99.CM40              ........................................................................................................................................................................   0
DEB.CM.99.CM5               ........................................................................................................................................................................   0
DEN.CD.x.CD1_CMO580407      ........................................................................................................................................................................   0
DRL.DE.11.D3                ........................................................................................................................................................................   0
DRL.x.x.FAO                 ........................................................................................................................................................................   0
GRV.ET.x.GRI_677_gri_1      GAG---............--ACG---CT-T----GACA-T-TC--A-G--T--A-CACAGA--GGG-A-GATCCTGAT-GAATAGATC-TG-------C--G-TC----A-------CCGAAGT-G--G..................-T-GA-T-C-GCCC-......G- 298
GSN.CM.99.CN166             ........................................................................................................................................................................   0
GSN.CM.99.CN71              ........................................................................................................................................................................   0
LST.CD.88.SIVlhoest447      ........................................................................................................................................................................   0
LST.CD.88.SIVlhoest485      ........................................................................................................................................................................   0
LST.CD.88.SIVlhoest524      ........................................................................................................................................................................   0
LST.KE.x.lho7               ...TGTG-TGAGT-TA---ATCC-TC--AT----CACAAT--C-T--G--TT-AAGCCAG---GGG-TCC--.......--A-----CTG-G----GAG-C-TATC---CAC--CA-TCCAATGT-G---..................-T-GACTT-ATA-CCCT-AAA-- 301
MAC.US.x.251_1A11           GAT---............--AC-GG-G-AT---G-GCATTAT--AA-G--T--A-CTCAAACTTCCCAGTG-......--------TTG-G----G--TCC--CA----A-------CCAACT--G--C-ACAC-T-TGA---ATAT-TTAGAT-C--A--AG-G--TGG 316
MND-1.GA.x.MNDGB1           ........................................................................................................................................................................   0
MND-2.CM.98.CM16            ........................................................................................................................................................................   0
MND-2.GA.x.M14              ........................................................................................................................................................................   0
MND-2.x.x.5440              ........................................................................................................................................................................   0
MNE.US.x.MNE027             ........................................................................................................................................................................   0
MON.CM.99.L1_99CML1         ........................................................................................................................................................................   0
MON.NG.x.NG1                ........................................................................................................................................................................   0
MUS-1.CM.01.CM1239          ........................................................................................................................................................................   0
MUS-1.CM.01.SIVmus_01CM1085 ........................................................................................................................................................................   0
MUS-2.CM.01.CM1246          ........................................................................................................................................................................   0
MUS-2.CM.01.CM2500          ........................................................................................................................................................................   0
MUS-3.GA.09.09GabOI81       ........................................................................................................................................................................   0
MUS-3.GA.11.11GabPts02      ........................................................................................................................................................................   0
OLC.CI.97.97CI12            --T---............A--CTTG---AT----GGGATGC------G--T---A-TCAGACC-C-CA--CA......CC---GGA-TTAG-G...--C-GG-T----GTC--CA-TCCA-ATAG----..................A-TAC--C--TTA----AAGG.. 229
RCM.CM.00.SIVagi_00CM312    ........................................................................................................................................................................   0
RCM.CM.02.SIVrcm_02CM8081   ........................................................................................................................................................................   0
RCM.GA.x.SIVRCMGAB1         ........................................................................................................................................................................   0
RCM.NG.x.NG411              ........................................................................................................................................................................   0
SAB.SN.x.SAB1               GAG---............--T--G--TA-T------CA-T--C--C-G--T----CTCAAG--GCCTATGA-......------G-CTG--A---GACCC-G-T-----A-------CCA-TA------..................-T-GACT--GTT-CT-GGAGGCT 298
SMM.CI.79.SIVsmCI2          ........................................................................................................................................................................   0
SMM.LR.89.SIVsmLIB1         ........................................................................................................................................................................   0
SMM.SL.92.SIVsmSL92A        ........................................................................................................................................................................   0
SMM.SL.92.SIVsmSL92B        ........................................................................................................................................................................   0
SMM.SL.92.SL92B             ........................................................................................................................................................................   0
SMM.US.04.G078              ........................................................................................................................................................................   0
SMM.US.05.D215              ........................................................................................................................................................................   0
SMM.US.06.FTq               ........................................................................................................................................................................   0
SMM.US.86.CFU212            ........................................................................................................................................................................   0
SMM.US.x.H9                 ........................................................................................................................................................................   0
SMM.US.x.PGM53              GAT--A............--TC-AG---AC----CACATTA---AG-G--T--A-CTCAGA-ATCCCAGTG-......-----T--CTG-G----G--TC-G-CA----A------TTCA-AAT----C-ACAGAT-TGA---TTTTATTAGA--C--A--AG-G--TGG 242
SMM.US.x.pE660.CG7G         GAT--A............---C-GG---AC----CACATT--C--G-G--T--G-CACAGACT---CAGTG-......--------CTG-G----G--AC-G-CA----A------TCCAGAAT----T-A-AGCT-T-A---ATTTATTAA-T-C--A--AG-G--TGG 316
STM.US.89.STM_37_16         ........................................................................................................................................................................  22
SUN.GA.98.L14               GA---C............AGGG-TCCTAATC-TCCTTGTCAAGCCC---TG-A-TCC----A---A--GG-TAAT...--A-----TTG-G----GAGAC-TATC----CA-----TCC-ACT--G--C..................T-TGATTT-AG--CCATACAG-- 301
SYK.KE.x.KE51               ........................................................................................................................................................................   0
SYK.KE.x.SYK173_COMGNM      GG-TGC............TTGG-GT-T-AG-----T--ACT-C-CC-C--T--A-CCTCT...GGGAC-GG-GAGTTC--AC-T......G-G--GAACC-CAT-----AC--CA--CC--A---G--C..................T-TACA-CAGGCTG-G-GA-GGC 192
TAL.CM.00.266               ........................................................................................................................................................................   0
TAL.CM.01.8023              ........................................................................................................................................................................   0
TAN.UG.x.SIVagmTAN1         -AG--A............---C-GG-T-AG...-GACA-T-T---C-C-------CCCAGA-TG-AT-GGAA......AGC-----TTG--A---GACA--G-T-----AA---A-TCCA-T-------..................-TT-A-T-CAACCC-......G- 289
VER.DE.x.AGM3               ........................................................................................................................................................................   0
VER.KE.x.9063               GA---.............--AG-G-CCT-T----GACATTA----GCG--T----CACAAA-AA-A-AGGACCCCGAT-GCAT-AGTC-TG----GACC--G-TC----A-------CCTATGT-----..................-T--A-T--GACCCAA--AGAC- 304
VER.KE.x.AGM155             GAG--A............--AG-A-C-T-T--A-GGCATT-----G-------A-CACAA-------A-GACCCTGAT-G-ATAAATC-TG-----A-A--G-CA----A------TCCAATGT-G--T..................-TT-A-T-CGACCCC-CAAGGG- 304
VER.KE.x.TYO1_patent        GAG--A............--AC-A-C-T-T--A-GACATT--C-AG-C--T--A-C-CAGA--GGA-A-GATCCAGAT-G-AT-AGCC-TG----GA-C--G-T-----A------TCCTATGT-G---..................ATA-A-T-CGACCCCA-TCGGG- 304
WRC.CI.97.97CI14            GA---ATCAGATA-TAC--A-CCAT-CTAT----GA-AT-AG--AC-G--T--A-CAT-TGCTG-GCC-..........--G--T--CTG-G-----CATC-G-TC----A-------CCTAAG--G--C......TGTGATTTTGT--CAGGAAGA-TCCCTGGACAGTC 319
WRC.CI.98.98CI04            GA----TCTAAT--CAGCT-TC-GT-CTAC-----G-A--AGC-------T--A-GCT---AAG-A...TCA.......--G--T--TTG-G-----CAC--G-TC----A-------CCT--A-----C......TGTGACTTTGTG-CAAGTAGA-TCCCTGGACAT-T 319
WRC.GM.05.Pbt_05GM_X02      ........................................................................................................................................................................   0
```

```
                                  TCF-1 alpha
H1B.FR.83.HXB2                    GAACTGC.....................................................................TGACA.....................................................TCGA  338
H1O.CM.91.MVP5180                 C--G........................................................................A-.....................................................  332
H1N.CM.95.YBF30                   ...........................................................................................................................  0
H1P.CM.06.U14788                  ...........................................................................................................................  0
CPZ.CD.06.BF1167                  A..........................................................................A-GGA.....................................CTGCTTCCGGCAAC-AC  337
CPZ.TZ.00.TAN1                    ...........................................................................................................................  0
GOR.CM.07.SIVgorCP2135con         ...........................................................................................................................  0
                                  TCF-1 alpha
MAC.US.x.239                      A-G-AAG...TCAGGCCTGTCAGAGGAAGAG...GTTAGAAGAAGGCTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAAGGAAACTCGC---.....................................  407
H2A.SN.86.ST_JSP4_27              ...........................................................................................................................  0
H2B.CI.x.EHO                      -T-TCAG...TCAGGAATGCCAGAGAAAGAG...TGGAAGGCTAAACTGAGAGCAAGAGGAATACCTACAGAG.....................-AGACAGGAAATAACAGCACAAAGAGGAACT.....................AGCTAACACTGCATAGAGAAGGAA-  439
H2G.CI.92.Abt96                   ...........................................................................................................................  0
H2U.CI.07.07IC_TNP3               ...........................................................................................................................  0
H2U.FR.96.12034                   ...........................................................................................................................  0
ASC.UG.10.RT03                    ...........................................................................................................................  0
ASC.UG.10.RT08                    ...........................................................................................................................  0
ASC.UG.10.RT11                    ...........................................................................................................................  0
COL.CM.x.CGU1                     ...........................................................................................................................  0
COL.UG.10.BWC01                   ...........................................................................................................................  0
COL.UG.10.BWC07                   ...........................................................................................................................  0
DEB.CM.04.SIVdeb04CMPF3061        ...........................................................................................................................  0
DEB.CM.99.CM40                    ...........................................................................................................................  0
DEB.CM.99.CM5                     ...........................................................................................................................  0
DEN.CD.x.CD1_CMO580407            ...........................................................................................................................  0
DRL.DE.11.D3                      ...........................................................................................................................  0
DRL.x.x.FAO                       ...........................................................................................................................  0
GRV.ET.x.GRI_677_gri_1            C-TG-TTAAGGACATGCACGAACAT.........GCAAAGCGC.....................................-AGTGTCAG.....................................  341
GSN.CM.99.CN166                   ...........................................................................................................................  0
GSN.CM.99.CN71                    ...........................................................................................................................  0
LST.CD.88.SIVlhoest447            ...........................................................................................................................  0
LST.CD.88.SIVlhoest485            ...........................................................................................................................  0
LST.CD.88.SIVlhoest524            ...........................................................................................................................  0
LST.KE.x.lho7                     -C-GCCTGCGAAGATACAAAATACT.....GCTTTCGCTTTTGACTGCAAGAGGAAG.....................-AG-CAAACCGCAAACCGCAGCACATCCTCTTGCAGCCCGGTTGCTAA...GGCAACCGGGCTAGCGCATGCG-A-  426
MAC.US.x.251_1A11                 A-G-AAG...TCAGGCCTGTCAGAGGAAGAG...GTTAGAAGAAGGCTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAAGGAAACTAGA---.....................................  407
MND-1.GA.x.MNDGB1                 ...........................................................................................................................  0
MND-2.CM.98.CM16                  ...........................................................................................................................  0
MND-2.GA.x.M14                    ...........................................................................................................................  0
MND-2.x.x.5440                    ...........................................................................................................................  0
MNE.US.x.MNE027                   ...........................................................................................................................  0
MON.CM.99.L1_99CML1               ...........................................................................................................................  0
MON.NG.x.NG1                      ...........................................................................................................................  0
MUS-1.CM.01.CM1239                ...........................................................................................................................  0
MUS-1.CM.01.SIVmus_01CM1085       ...........................................................................................................................  0
MUS-2.CM.01.CM1246                ...........................................................................................................................  0
MUS-2.CM.01.CM2500                ...........................................................................................................................  0
MUS-3.GA.09.09GabOI81             ...........................................................................................................................  0
MUS-3.GA.11.11GabPts02            ...........................................................................................................................  0
OLC.CI.97.97CI12                  -CA--TTCAGACTTGCATAGGATG.........TGGTTTGTGTGGTTAAGC.....................................-AGA-GACTTGGTTGCCTGGGAGACGGGGGTTGG....ACATGGGCGGGCT.....................  318
RCM.CM.00.SIVagi_00CM312          ...........................................................................................................................  0
RCM.CM.02.SIVrcm_02CM8081         ...........................................................................................................................  0
RCM.GA.x.SIVRCMGAB1               ...........................................................................................................................  0
RCM.NG.x.NG411                    ...........................................................................................................................  0
SAB.SN.x.SAB1                     -C--CCAGAGCAGGTGCCGTCAGCA.........CAAGGC.....................................A--CACAAAACCACATCCTATG.....................GAGTTGTCATGGTGATGACAT-AAG  377
SMM.CI.79.SIVsmCI2                ...........................................................................................................................  0
SMM.LR.89.SIVsmLIB1               ...........................................................................................................................  0
SMM.SL.92.SIVsmSL92A              ...........................................................................................................................  0
SMM.SL.92.SIVsmSL92B              ...........................................................................................................................  0
SMM.SL.92.SL92B                   ...........................................................................................................................  0
SMM.US.04.G078                    ...........................................................................................................................  0
SMM.US.05.D215                    ...........................................................................................................................  0
SMM.US.06.FTq                     ...........................................................................................................................  0
SMM.US.86.CFU212                  ...........................................................................................................................  0
SMM.US.x.H9                       ...........................................................................................................................  0
SMM.US.x.PGM53                    T-GTAAG...TCAGGCTTGTCAGGGGAAGAG...GTAAAGAGAAGGCTAACCGCAAGAGGCCTTTTAAAAATGGCTGACAAGAAGGAAACGAGC---.....................................  333
SMM.US.x.pE660.CG7G               T-GTAAG...TCAGACTTGTCAGAGGAAGAG...GTAAAGAGAAGGCTAACCGCAAGAGGCCTTATTAAAATGGCTGACAAGAAGGAAACAAGC---.....................GACAGC.....................................  413
STM.US.89.STM_37_16               ..........................GAGAAGGCTAACCGCAAGAGGCCTTCTAAAGATGGCTGACAAGAAGGAAACAAGC---.....................................  80
SUN.GA.98.L14                     -C--CCTGAAGAGTTCAAGCATGTT.........ACTAGTCTGCAGTGGGAGGTG.....................................-AG-TTAACCGCAAACCGCATCCTCTTGCATCGCCT.....................AGGCAACGGGGCTAGC.GCAT....  403
SYK.KE.x.KE51                     ...........................................................................................................................  0
SYK.KE.x.SYK173_COMGNM            CCG--TGCAGTTAGAGAGACAAACA.........GGAAAACCACAAGAGCTGCAGTCAGCACTTAGTAAGAACATCTCC.....................AGG-GACTCCATGGTGACAAGCTCGG.....................CCCAGGGGA.....................  298
TAL.CM.00.266                     ...........................................................................................................................  0
TAL.CM.01.8023                    ...........................................................................................................................  0
TAN.UG.x.SIVagmTAN1               C-G--TTAAGGACATGCATGGTCTT.........GTTAAGAGGAAG.....................................-AG-CTAACCGCAGGCTTGCGGTTAAGCTAACTAGGCCGTTGCCTAGGAGA.....................TGACATA-  385
VER.DE.x.AGM3                     ...........................................................................................................................  0
VER.KE.x.9063                     -T-T-TTGAAGACATGCATGCAATA.........GTGAAGAGGAAG.....................................-AG-TAACCGCAGGCTTGTGGTTAAG.....................CCGTTGCCAGGGAGATGACAT-T--  392
VER.KE.x.AGM155                   -T---TTACAGACTTATATTCAACA.........GTTGGTACAGGAAAC.....................................-AG-CGACCACAGGCTTGCGGTTTCC.....................TGGTTGCCTAGGAGATGACAT-AAG  395
VER.KE.x.TYO1_patent              -T---TTACTGACATGCATGGGCTG.........GTGAAGAGGAAG.....................................-AG-CAGACCGCAAGCCTGCGGTTAGA.....................ACATCACCATGGAGATGACAT-AA-  393
WRC.CI.97.97CI14                  A-T-GC-ACTGGGAAAGAGGCCATTGAG...AGTCTCTATAAGAACAAGAAGCAGTGCAAG.....................................-AGAGACATGTTGCAGTAGGAAGTGGCTTTGCGGTTAAGCTACAT.....CCTGAGACTGTGTACCAACGGA-TT  447
WRC.CI.98.98CI04                  A-TAGC-CAAGGAGAAGATGCTATTGAGACCCTCCATAAAGAGAAGCTGAAGCGGCACAAG.....................................-AGAGACATGTTGCAGTAGGATGTGGTTTTGTGGTTACGCTACATCCT...GGACCTGTGTACCAACGAACT-A-C  453
WRC.GM.05.Pbt_05GM_X02            ...........................................................................................................................  0
```

Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

+1 mRNA start site

5' LTR U3 end 5' LTR R repeat begin

TAR element end

```
H1B.FR.83.HXB2          .....GGGTCTCTCTGGTTAGACCAGA..TCTGAGCCTGGGAG..CTCTCT..GGCTA.ACTA..GGG.AACCCA..... 513
H1O.CM.91.MVP5180       .....TAG-TAGAG------G..--------C------.---C--------CT..------.CTG-----G..... 520
H1N.CM.95.YBF30         .....T-C-G-------..-TA----------.------.-------G---.----.------..... 73
H1P.CM.06.U14788        ..... 0
CPZ.CD.06.BF1167        .....-C---TGTCT-AGCT------.-----A---------.T----------GT-AGGGCATG-GA---..... 525
CPZ.TZ.00.TAN1          .....-C---TG-CTAA-CTG-----.----------.----------TAGTGGCTGGCTAG-GA--G..... 63
GOR.CM.07.SIVgorCP2135con .....------G--AC.---------.---A----------.----------TCTT-GC----.-------..... 56
```

+1 mRNA start site

TAR element start

5' LTR U3 end 5' LTR R repeat begin

TAR element end

```
MAC.US.x.239            C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGA.............--T------CA-CACT-G---G..-.A---------T-TT-C--GCTA-.ACTCT-A...CCAGC--TTGGCCG..GTGCTGGGCAGAGT..GACTC.CACGCTTG..... 651
H2A.SN.86.ST_JSP4_27    ....AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGA.............--T------CA-CACT-G---G..-.A---------T-TT-C--GCTA-.ACTCT-A...CCAGTG-TTGGCCG..GCACTGGGCAGAC..GGCTC.CACGCTTG..... 132
H2B.CI.x.EHO            C...AGTCGCTCTGCGGAGAGGCTGGCAGATCGAGCCCTGGGA.............--T------CA-CACT-G---G..-.A---------T-TT-C--GCTA-.ACTCT-A...CCAGC--TTGGCCG..GTGCTGGGCAGAC...GGCTC.CACGCTTG..... 690
H2G.CI.92.Abt96         ..........................................................--GCTA-.ACTCT-A...CCAGCG-TTGGCCG..GTGCTGGGCAGAGY..GGCTC.CACGCTTG..... 55
H2U.CI.07.07IC_TNP3     C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGA.............--T------CA-CACT-G---G..-.A---------T-TT-C--GCTA-.ACTCT-A...CCAGCG-TTGGCCG..GTGCTGGGCAGAC...GGCTC.CACGCTTG..... 160
H2U.FR.96.12034         C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGA.............--T------CA-CACTGG---G..-.A---------T-TT-C--GCTA-.ACTCT-A...CCAGT--TTGGCCG..GTACTGGGCAGAGT..GACTC.CACGCTTG..... 184
ASC.UG.10.RT03          ..... 0
ASC.UG.10.RT08          ..... 0
ASC.UG.10.RT11          ..... 0
COL.CM.x.CGU1           .....TGGAGTTCTGACTCTT.....G.CC-GG.CT.GCGCC-.GAGC----A-T-GCT-GC-AGCTATC-GC--TGCCT-AGT-TG-GTAGCTTG...CTTGGTTGAAGGGCTTGAGGGACTC..... 107
COL.UG.10.BWC01         ..... 0
COL.UG.10.BWC07         .....-TCCTGCCTGTCCTCGCCTGAG..CCCCGAGTGGCTTG..... 36
DEB.CM.04.SIVdeb04CMPF3061 .TGGAGTT.....GCATCCAGCGGG-CC-GGT-GGTA...AAGCG----------.-C--TGCCTAT-CAG-CCACCA-CCTG.AGCGGCCTGGGAGCCCTAGGTT..GGGG..... 228
DEB.CM.99.CM40          .....CGGGAGTC...GCGCCT-GCGAG-CC-GGTGAGTA...GAGCG----------.-C--TGCTCA--GGGA--GCCTAGCT--AGCGGCCTGGGAGCCCTGAGCTGGGG..... 100
DEB.CM.99.CM5           .....CGGGAGTC...GCGCCTAGCGAG-CC-GGT-GGTA...GAGCG---------CC--TG-CTGCTAGTG----CCTA-CTC-AGTGGCCTGGGAGCCCTGAGCTGGGGTTAGGTACTCCTCGGCACG 120
DEN.CD.x.CD1_CMO580407  .....A--TG--T-CG--T-TCTC..CAGCG----------.-C-------AG-GG-CTTA-C......AGCGGCCTGGGAGCCCTCTAAGGTATTG..... 80
DRL.DE.11.D3            ..... 0
DRL.x.x.FAO             ..... 0
GRV.ET.x.GRI_677_gri_1  TC.....A-TCTCT-ACTAGG---------.CTA--------T-.TTCGCTGGTTAGCCT-ACCCG-TTGGC-A-CGGGGGTA.AGGACTC..... 551
GSN.CM.99.CN166         .....--C--C-T-TTTCAC.AGCG---------CC--TG.....GAGCTAGGTAAAGCGGCCTGGGAGCCCTTT.....G 65
GSN.CM.99.CN71          .....--CC-CCGT-TCCAC.AGCG---------CC--TG.....GAGCTAGGTAGAGCGGCCTGGGAGCCCTTGC...CT 67
LST.CD.88.SIVlhoest447  ..... 0
LST.CD.88.SIVlhoest485  ..... 0
LST.CD.88.SIVlhoest524  ..... 0
LST.KE.x.lho7           G...GAG.....-CT-GCTC-GG-GA-CAGCAGTC-CT------T.GAGCACACTG-.TC-GCTA.AGCA-TC-G--CTGGGAGGCTACACTTAGCACCCACCTGGGCCTGATCA..... 619
MAC.US.x.251_1A11       C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGA.....-GT-----CCAGCACTAGC-GG-.A---------TTTT-C--GCTA-.ACTCT-A...CCAGC--TTGGCCG..GTGCTGGGCAGAGT..GACTC.CACGCTTG..... 650
MND-1.GA.x.MNDGB1       CAGATCCCCTTAGAGCAAGGACCA.....A---CTGAGT-ACT-GGTCTGAGCACCTCA--C--G-CTGATCACCTC-AGGT-G-G...AACTC-TTGCTTG..... 135
MND-2.CM.98.CM16        TCTGGTGAGTCTT.GGAAGGAGTGCCTGCTTGTAGCCCTGGGT.....-GT--G-A--CCCCTGGCTTG....AGCT-----CAGCTCG--AGGT-T-CCTGA...A-AGGCT-ATCGGGGGCTCCCTTGCTTGGTCTTGGTAGACCTCTAGCAGTCT 253
MND-2.GA.x.M14          TCTGGTGAG.....TCCTTGAAGGAGTGCCTGCTTGTAGCCCTGG-C-GT--G-A--CCCCTGGCTTG....AGCT-----TAGCTCG--AGGT-T-CTGGA-A.--TCTCG-T-AGTGAGCACTCTTGCTTGGTTTTGGTAGACCTCTAGCAATCT 179
MND-2.x.x.5440          ..... 0
MNE.US.x.MNE027         ....AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGA.....-GT-----CCAGCACTAGC-GG-.A---------T-TT-C--GCTA-.ACTCT-A...CCAGC--TTGGCCG..GTGCTGTGCAGAGT..GGCTC.CACGCTTG..... 133
MON.CM.99.L1_99CML1     .....A-C-GCTCG---CCA.-AGCGGC.CT---AGC--TTGGAGCTAGGTAAAGCGGCCTGGGAGCCCTTG.....C 66
MON.NG.x.NG1            ..... 0
MUS-1.CM.01.CM1239      .....CAG-TCTAGCTCCGCAGCG----------.C---GGAACTAGGTAA-GC-----.....CCTGAGAGCCCTTG.....C 65
MUS-1.CM.01.SIVmus_01CM1085 .....AG--CTAGCTCCGCAGCG----------.C---GGAACTAGGTA--GC-----.....CCTGGGAGCCCTTG.....C 64
MUS-2.CM.01.CM1246      .....TCTTC-GTCT-C-CAG...CG-------G-.C-CTCA.....GAACTAGGTAGAGCGGCCTGAGAGCC 59
MUS-2.CM.01.CM2500      .....G---TC-CCTGT-TCCGC.AGCG---------CC---G.....GAGCTGAGTAGAGCGGCCTGGGAGCCCTTG.....C 68
MUS-3.GA.09.09GabOI81   ..... 0
MUS-3.GA.11.11GabPts02  ..... 0
OLC.CI.97.97CI12        A.....-GTCT-GCTCGG-------CCT-GA-------A-T-CTT--GGCAGCCTGCTTTG-GC-CCTGGGTGTGCTTGGCAGGTGGTTAAAGAGTAAGAACCTCAG..... 475
RCM.CM.00.SIVagi_00CM312 ..... 0
RCM.CM.02.SIVrcm_02CM8081 ..... 0
RCM.GA.x.SIVRCMGAB1     ..... 0
RCM.NG.x.NG411          ..... 0
SAB.SN.x.SAB1           AC..AGTCTCTGCTTGGAG.G.CTGCCAGATTGAGCC.TGGGT.GTTCTCTG...GTAA-TCTGAA-CA-C-T--G-CTGGG-.GTTCG-----GAAGGCAAGCCTAAGAGC-.CTCT--TG-GGT-TTGGTA..GACCCTCGCCTGGCGAC.TG..GCCATTG.....C 627
SMM.CI.79.SIVsmCI2      ..... 0
SMM.LR.89.SIVsmLIB1     ..... 0
SMM.SL.92.SIVsmSL92A    ..... 0
SMM.SL.92.SIVsmSL92B    ..... 0
SMM.SL.92.SL92B         .....GTCGCTCTACGGAGAGGCTGCCAGATAGAGCCCCGAGA.....-GT------CAGCACTT--TAGGT.---------T-TT----ACTAA.ACTCT-A..ACCAGC-A-AGGCCA..TTGCTGGGTAGAC...GGCTC.CACGCTTG..... 132
SMM.US.04.G078          ..... 0
SMM.US.05.D215          ..... 0
SMM.US.06.FTq           ..... 0
SMM.US.86.CFU212        ..... 0
SMM.US.x.H9             C...AGTCACTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGA.....-GT-----CCAGCACTAGC-GG-.A---------T-TT-C--GCTG-.ACTCT-A...CCAGC--TTGGCCG..GTGCTGGGCAGAGT..RGCTC.CACGCTTG..... 134
SMM.US.x.PGM53          C...AGTCGCTCTGGGCAGAGGCTGGCAGATTGAGCCCTGGGA.....-GT-----CCAGCACTAGC-GGCA.---------T-TT-C--GCTA-.ACTCT-A...CCAGC--TTGGCCG..GTGCTGGGCAGAGC..GGCTC.CACGCTTG..... 577
SMM.US.x.pE660.CG7G     C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGA.....-GT-----CCAGCACTAGC-GG-.A---------T-TT-C--GCTA-.ACTCT-A...CCAGC--TTGGCCG..GTGCTGGGCAGAGT..GGCTC.CACGCTTG..... 650
STM.US.89.STM_37_16     C...AGTCACTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGA.....-GT-----CCAGCACTAGC-GG-.A-------A-T-TT-C--GCTA-.ACTCT-A...CCAGC--TTGGCCA..GTGCTGGGTAGAGT..GGCTC.CACGCTTG..... 324
SUN.GA.98.L14           GG..AGTCTCTGCTCAGAGG....CTAGGCGACGACCCT.....GA-AGGTC.....TCA----A---CTGG--TGTC-.--CCTGA.....GTGAGCACTTGCTGA.GCACCTCACCCGAGCCTGATCAGCTT 597
SYK.KE.x.KE51           GA..AGTTGCAGCTAGCTCTCCTCAGCGGCCTGGGAGCCC.TAGGAGAC.....TG-TAG.AGCG---------CC--TG-CAGT.....AA-GGCTAGT-AG..ACAAGGACTG.....ATCCACCTAGCCTCTACAATTCTCCGT 128
SYK.KE.x.SYK173_COMGNM  CGG.AGTTGCAAC.....T.....AG------CCAGC-G-----TGGGAGC----A---ACTGG-AG.A-CGGCCCGGGAGCCCTTG--AGT.....AAAGGCTAGTGAG.GTCCGGGTTTG.....GCCGC 465
TAL.CM.00.266           .....G-----C.---GTG-TCACAGCGT--------AC--TGGCCACCGGTA-AGC.....T--TGGG.....AGACCT..TGCCGGTAAACGCTCAGTAG 85
TAL.CM.01.8023          ..... 0
TAN.UG.x.SIVagmTAN1     TC..AGTCTCTTACTAGGAGA.....CA-C-T.....TTCGCTGGT-AGCCT-ACCTG-TCTGG--ATCCAGGGGTAAGAGCCTC..... 574
VER.DE.x.AGM3           C.....A-TCTCT-ACTAGG---------.CT---------T-.TTCGCTGGTTAGCCT-ACCTG-TTGGC-A-CAGGGGTA.AGGACTC..... 81
VER.KE.x.9063           TC..AGTCTCTTACTAGGAGA.....CA-C-T.....T-TT-G--GGTTA.GCCT-AC.CT--T.TGG---.CCAGGGGTAAGGACTC..... 585
VER.KE.x.AGM155         TC..AGTCTCTTACTAGGAGA.....CA-C-T.....T-TT-G--GGTTA.GCCT-AC.CT--T.TGG---.CCAGGGGTAAGGACTC..... 589
VER.KE.x.TY01_patent    TC..AGTCTCTTACTAGGAGA.....CA-C-T.....T-.TTCGCTGGTTAGCCT-ACCTG-T.TGG--A.CCAGGGGTAAGGACTC..... 588
WRC.CI.97.97CI14        ....GGTTGCAGTTTTCTGCCTGAGGCTAGGTGCCTGCGTAGT.....GTC--AC-TGGCCTAGCACTC-GA-T-----A-T-GTGGT.....CC..C--TTGGCTCAGAGACTCAGGTAGGAGGCTGTCACTCGCTTG..... 662
WRC.CI.98.98CI04        TTGCAGTCTTCTG.....CCTGAGGCTAGGTGCCTG...CGTA-T-GTC--AC-TGGCCTAGCTCTC-GA-T-----A-T-.GTGG-.....CCC--TTGGCTCAGAGACTCAGGTAGGAGTCTGTC.CTCGCTTG..... 663
WRC.GM.05.Pbt_05GM_X02  ..... 0
```


Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

```
                                                                                                                                      Extensive secondary structure
                                                                        Poly-A signal        5' LTR R repeat end  5' LTR U5 start
H1B.FR.83.HXB2                ................................CTG..CT.T......AAG...........CCT.........C.AATAAAGC..TTGCCTTG..AGT....GCTTCA.AGTA...GTGTGTGCCCGTCTG.TT.GT........GTGACTCT.GGTAACTAGAGATCC 597
H1O.CM.91.MVP5180             ................................---.--.-.......--C..........G--.........-.--------.--------.------....-AGAA...-C-..---------T-A----.--C............A--C--.---GT--------- 599
H1N.CM.95.YBF30               ................................---.--.-.......---..........---.........-.--------.--------.------.....--AA-..--G..----------A--CA.--C-...........-A-----.----C--------- 156
H1P.CM.06.U14788              ..............................................................................------.--------.-------AG-A-A.-C-..---------T-A----.--C.........AG.--------.---T--------- 67
CPZ.CD.06.BF1167              ................................---.--.-.......--C..........G--.........-.--------.------G....--AGTGT....T..-TC-GT.-----CC-ATC--A.T-CC-CGCCGGCCCTG-GG-AGA.-ATC-CTC-----T 618
CPZ.TZ.00.TAN1                ................................---.--.-.......--C..........G--.........-.--------..C-----G...--A....-TG-T..-AC-..-----------A-T-CA-ACCGCGT.....C--C-....CTGGGG--------- 147
GOR.CM.07.SIVgorCP2135con     ................................---.--.-.......--C..........G--.........-.--------.-----A-TG.------..-AA-T..--C-..---------T-A---A.--CA..........-A-------.---GT-------T- 139
                                                                        Poly-A signal     5' LTR R repeat end 5' LTR U5 start
MAC.US.x.239                  ................................CT--..---.-.....A---..........CCCT--T.........-.--------.-.---A-T..TTAGAAGTAAG-T.-.GT..-------T---A---.C-CCTAGCCGCCGCC--GTCAA.CTCGGTACTCA--AA 752
H2A.SN.86.ST_JSP4_27          ................................CT--..---.-....AA---..........ACCT--T.........-.--------.-.---AGT..TAGAAGCAAGTT.-AGT..------CT---A---.C-CCTAGTCGCCGCC--GTCA-..TCGGTGTTCATCTA 232
H2B.CI.x.EHO                  ................................CT--..---.-.....A--A........CCCT--T.........-.-------A-..-.----.A..TTAGAAGCAAGAC.-AGT..-------T---A---.C-CCTAGTCGCCGCC--GTCA-..TCGGTGCTCC-CTA 789
H2G.CI.92.Abt96               ................................CT--..---.-...ACA---.........ACCT--T.........-.---------..-GC-AAAT..TAGAAGTAAG--AGTGT..-------T---A---.C-CCTAGTCGCCGCC--GTCAC.CTCGGTGTTC-CGTA 159
H2U.CI.07.07IC_TNP3           ................................CT--..---.-.....A---.........ACCT--T.........-.--------.-.---AAT..TAGAAGCAAGTTA-.GT..-------T---A---.C-CCTAGTCGCCGCC--GTCA-..TCGGTGTTCATCTG 259
H2U.FR.96.12034               ................................CT--..---.-.....A--A........CCCT--..........-.--------.-.---AGT..TAGAAGCAAG-TA-.GT..-------T---A---.C-CCTAGTCGCCGCC--GTCA-..TCGGAG-TTC-C-T 283
ASC.UG.10.RT03                ........................................................................................................................................................................ 0
ASC.UG.10.RT08                ........................................................................................................................................................................ 0
ASC.UG.10.RT11                ........................................................................................................................................................................ 0
COL.CM.x.CGU1                 .CAGTCCCCTCGCGCTAAGGGTAGATAGGTG.CT-TGCA-.-....AT---AG..........................-.-------T.GA-TTTGAA..GC-TCAA--AG-G.TT-G...AGTAC-AGTTTCAC-TCCGCCTTCTC......................... 214
COL.UG.10.BWC01               ......................................................................................-.T-CT-TGGTCG---GCTCT-CC-T.--C-CT.C-CGTCTG--ACT-.CC-CACTGCGCGTCCGCCGCGC.-CCTTT--CTTTCA- 82
COL.UG.10.BWC07               ..................CAGGACACTCTGC.CT--..-CTGGGCTCTG--TAGCCTGCTTGGTAGAAGGGCAAGAG-CTCTC-AG-CC--GC-CGCTACGGGTAGCT-GTTGCTTT-CT--GAGGT-G-AATAAAAGATCCTTGCAA-GAA--GA--TCG-GTGAG-TA 185
DEB.CM.04.SIVdeb04CMPF3061    ............ATGCTAGCTC.TCGGCACG.CT--..---.-.....AGCTGCATTCT.ATTG-CCAG..............---------.TG--GTAA....CAGCAA.............C----C---T-TCA-CA--CAGATCCTTAGTGTGGGGGCCT............. 335
DEB.CM.99.CM40                ............TAGGCACTCCTCGGCACG.CT--..---.-......GCTACGCTTCTATAG-CCAG.............---------.TG--GTAA....C-GCAA.................------CTG-CCTCATCATTGACTCCAGG-.......................... 196
DEB.CM.99.CM5                 ................................CT--..---.-......GCTGCATTTT.ATTG-CCAG.............---------.TG--GTAA....C-GCAA.................------CTG-CCTCATCATTGACTCCAGG-.......................... 197
DEN.CD.x.CD1_CMO580407        ............CAACAAGTACCCAGACAAG.CT--..--CG.....TGCTGCTTTTG.ATGA-CATA..............---------.TCA-TTAA....C-GGCA.............TCT---CTTT-CTCCTCATTAGTGGGTTCC...................... 178
DRL.DE.11.D3                  ........................................................................................................................................................................ 0
DRL.x.x.FAO                   ........................................................................................................................................................................ 0
GRV.ET.x.GRI_677_gri_1        ................................CT--G.--.-.....CAT-T..............AGCT...-.------C-..---TCGCT.TAGTCGCTA-ATTGGAGTCAA---CTCATT...............GCTGC-CCGAG-CTCTAG-GGT--ACCT- 644
GSN.CM.99.CN166               CTAG.........AAGACGGTAGAATTTCG.CT--..---.-...GCCT-AGGCT...TATGA--..........-.--------.CT-CCAAAG-..-TAA.CGAG-CTGCTC-C....AT-CATTGG-.TC-AGGTCGGGCTTG.......................... 168
GSN.CM.99.CN71                AGAG..........GACGGTAGATTCTCGG.CT--..---.-...GCCCTAGGGC...TTTAGTGGCT.......-.--------.CT-C-AAAG-..TACAGCAAG-CTGCTC-C....AT-CATTGG---G-GA-C-GGCTTGG.......................... 176
LST.CD.88.SIVlhoest447        ........................................................................................................................................................................ 0
LST.CD.88.SIVlhoest485        ........................................................................................................................................................................ 0
LST.CD.88.SIVlhoest524        ........................................................................................................................................................................ 0
LST.KE.x.lho7                 GCCTAGGGAGC.TAAGGGCTCCTAG.......CT--..---.....ACTTTAAA...G.......---T..........------AA.G----TCAT..T--TAGAAAG-AC.-AGC...TA----CT-ATAG-.AG-CTCTTTCTCCTTC-C-CTG.-T-C-G-GATCTCT- 748
MAC.US.x.251_1A11             ................................CT--..---.-.....A---..........CCCT--T.........-.--------.-.---A-T..TTAGAAGTAAG-T.-.GT..-------T---A---.C-CCTAGTCGCCGCC--GTCAA.CTCGGTACTC-G-AA 751
MND-1.GA.x.MNDGB1             ................................CT--..--A-...................TGT--T........-.-------TAAC-TAGAA...TTAGAGCAAG-G-.GTA-..-----TAT--AT-G--CGCCTCTCTTCT........................... 211
MND-2.CM.98.CM16              CGGTGGCCAAGAGGCTGTGGGATTGACTACCGCT--..---.-...GCTTT......TAAGTGCT..........-.-------AG.C--A--CAAGTTAAGAA.AAGCA-.GTG-..-----ACT-G-C-A-.T--GTCTTTGGT.............AG-AACTCT.GGT 379
MND-2.GA.x.M14                CT.AGGCCAGGAGATTGTGGGTTAGACTACCGCT--..---.-...GCT......ATTGATGCT..........-.-------G.C--A---GA..TTAGAA.TAGCA-.GTA-..-----ACT-GTCCA-.T--GTCCTTGAT.............AG-AACTCT.GGT 301
MND-2.x.x.5440                ........................................................................................................................................................................ 0
MNE.US.x.MNE027               ................................CT--..---.-.....A---.........ACCT--T.........-.--------.-.---A-T..TAGAAGTAAG-C.-.GC..-------T---A---.C-CCTAGTCGCCGCC--GTCAA.CTCGGTACTC..AAT 231
MON.CM.99.L1_99CML1           CTAGAGAGTTGGTGAAGCTGACCCCGGCTCG.CT--..---.-.......T--TA....AGGC-T-ACAGCC...-.--------.......................................................-GCTT-A--GAGACTTG-G 149
MON.NG.x.NG1                  ........................................................................................................................................................................ 0
MUS-1.CM.01.CM1239            CTAG..........GAGAACTGGTTCTCGG.CT--..---.-..TGCTTC-AAGC...TTATTG..AGCCCA..-.-------TAG-A-AACC.AAG-CAAGTCT................................................................ 144
MUS-1.CM.01.SIVmus_01CM1085   CTAG..........GAGTACTGATTCTCGG.CT--..---.-..TGCTTC-AAGC...TTATTG.AAGCCCA..-.-------G..................................................................TTG----AC-AA 135
MUS-2.CM.01.CM1246            CTTGCCTAG...GGGTTGAGACCCCGGCTCG.CT--..---.-...GCAG-AA...GGCTTTTA-ATAGCC...-.-------TCC---AA.....ACAACCTGCGA.........--C-A.----TCA-.C-GT--GGCGTCTAAGAGTAGAC---TG-CG-CACAG- 194
MUS-2.CM.01.CM2500            TCAG..........AAGGAATAGACCCCGG.CT--..---.-...GCTT-CTTCAGGCTTTAT-..AGCC...-.--------.CT---AAC-.AACT-GCGA-TC-A...............----TCA-.CA-T-GCGGCGC....CTAGAAGT--............ 179
MUS-3.GA.09.09GabOI81         ........................................................................................................................................................................ 0
MUS-3.GA.11.11GabPts02        ........................................................................................................................................................................ 0
OLC.CI.97.97CI12              ................................CT-TG.---.GTGT.GAG--A...................-.------......................GCTGAG-AA-AGAAGACTC-C---CT-G-GTGTT--AC-CATCCAGGT-A..................... 549
RCM.CM.00.SIVagi_00CM312      ........................................................................................................................................................................ 0
RCM.CM.02.SIVrcm_02CM8081     ........................................................................................................................................................................ 0
RCM.GA.x.SIVRCMGAB1           ........................................................................................................................................................................ 0
RCM.NG.x.NG411                ........................................................................................................................................................................ 0
SAB.SN.x.SAB1                 CAGTAGCAGAGACTCCG...............CT--..---.-...GCTT.-A......TTG---..........-.-------T.C-------A..GT-TACAAG-AA-.GCA-..----------A-T-.A--CC-CAGGACAACCCT.................... 730
SMM.CI.79.SIVsmCI2            ........................................................................................................................................................................ 0
SMM.LR.89.SIVsmLIB1           ........................................................................................................................................................................ 0
SMM.SL.92.SIVsmSL92A          ........................................................................................................................................................................ 0
SMM.SL.92.SIVsmSL92B          ........................................................................................................................................................................ 0
SMM.SL.92.SL92B               ................................CT--..---.-.................GACT--..........-.-------G.C-.--AACT...TAGAAGCAAGTC.-AG-..------T-T-A---.C-CCTAGTCGCCGCC--GTCA-..TCGGGGTCCCGGAT 228
SMM.US.04.G078                ..........................................................................................................................................-C.TCGGTACTCTTGTG 16
SMM.US.05.D215                ..........................................................................................................................................-C.TCGGTACTC-TCA. 15
SMM.US.06.FTq                 ..........................................................................................................................................-C.TCGGTACTCAC-.. 14
SMM.US.86.CFU212              ..........................................................................................................................................-C.TCGGTACTC-.... 12
SMM.US.x.H9                   ................................CT--..---.-....AA---.........ACCT--T.........-.-------.C-.---AGT..TAGAAGTAAG-C.-.GT..------T---A---.C-CCTAGMCGCCGCC--GTCA-.CTCGGTACTC--A-A 235
SMM.US.x.PGM53                ................................CT--..---.-.....A---.........ACCT--T.........-.--------.-.---GCT..TAGAAGTAAG.C.-AGT..------T---A---.C-CCTAGTCGCCGCC--GTCA-.CTCGGTACTC--A-A 677
SMM.US.x.pE660.CG7G           ................................CT--..---.-.....A---.........ACCT--T.........-.--------.-.---A-..TTAGAAGTAAG-..-A..GT.------T---A---.C-CCTAGTCGCCGCC--GTCA-C.TCGGTACTC-.A-A 749
STM.US.89.STM_37_16           ................................CT--..---.-.....AT-T........ATCT--T.........-.--------.-.---AAT..TTAGAAGTAAG.C.-AGT..------T---A---.C-CCTAGTCGCCGCC--GTCAA.CTCGGTACTC..... 420
SUN.GA.98.L14                 G...........GGGAGCTAGAGGCTC....CT--..---.-...TGC-GTACA..AG......---T........-.--------..CC------A..GT-TTACTGCAAG.CAA.........----T-----G-TC-C.TCTTTCTCCTTCT....A.AACCCTGTTC-.... 709
SYK.KE.x.KE51                 TA..............................CT-A..---.-...TACT-CGCG..CTGATCAGCC..........-.-------T.G-GTTAACT..-CAAGTAA-GAGT...GC...TCTCA-CTGTTCTGCCGGC--CTCCCAGAAG.--G-AC--G--AC........ 233
SYK.KE.x.SYK173_COMGNM        CCGGTCTTCGGCAGCTCTCCGTTG........CT--..---.-...TGCTTACGCT..GATCAG-CCA..........-.-------T.G-G-TAAAC..GCAAGTAA-G--T.GC-C...TCA-C--TT-TG---GG-CACCTCCCGAA-GTG---A.--G-G-GC-...... 596
TAL.CM.00.266                 ACATCCCGCTGCTTGCTTCGCGAAG..........C-..---.-T......--A.........CATAGCC..........-.---------..-GTGT-CA..-AG......AG-..--C-..---------AGCC-ATC-GTTGG.........AG-CGC.--CGGG-GC...... 189
TAL.CM.01.8023                ........................................................................................................................................................................ 0
TAN.UG.x.SIVagmTAN1           ................................CT--..---.-....AG--.........GGCTA-TG.......-T-------..------GC..TTAAGAGT--ATC.T-AGCAA---CC.T-.AT-GC...GC-CCC.TTCCTC..-GGGAACCCTC.GTT-CTGG. 673
VER.DE.x.AGM3                 ................................CT--G.--.-...AGA---.......................-T------T..C--CG--GC..-T-AGAGCT-CTC.T-AG...TCAA------TCA-.TGAC-CCTCACTCTTGA--GGG.TAA---.TTCCT-A- 180
VER.KE.x.9063                 ................................CT--G.--.-...GGA---.......................-T------A-.T------GC..-T-AGAGCT-ATC.T-AG...TCAA------TCA-.TGAC-CCTCACTC..AAG-AGG.--A-C-GTTCCT-A- 684
VER.KE.x.AGM155               ................................CT--G.--.-...AGA---.......................-T-------..C------GC..-T-AGAGCT-ATC.T-AG...TCAA------TCA-.TAAT-CCTCACTC..-TGAA-G.--AG-AGTTCCT-A- 686
VER.KE.x.TYO1_patent          ................................CT--G.--.-...AGA---.......................-T------.-.------GC..-T-AGAGCT-ATC.T-AG...TCAA---T--TCA-.TGAC-CCTCACTCTC-TGAA-G.--A-T-.TTCCT-A- 686
WRC.CI.97.97CI14              ................................CT--.--.A.............GG.GTTT.GCAGTT.......-------AGTC-TTGACTTAGACAAGCAG-GAGTCAGTAGTC-TAT-TGGGCAAA-TAAG-AACCCTGGTTC-CT-GAG-T-G-TCTTCAG.AG 774
WRC.CI.98.98CI04              ................................CT--..T-.A.............GGAGTTTTGCA.TT......-------AGCC-TTTGCTTG-CAAGCA-TGAGTTT-GTAGTT-CAT-TA-A--ACAG-AG-AACCCTGGTTA-CT-GAGCT-GGACTTCAGAAG 777
WRC.GM.05.Pbt_05GM_X02        ........................................................................................................................................................................ 0
```

Extensive secondary structure

5' LTR U5 end

Lys tRNA primer binding site

```
H1B.FR.83.HXB2            CTCAGACCCT.TTTAGTCA.GTGTGG.....................................AAAA.TCTCTAGCAG....TGGCGCCCGAACAGGG..ACCTGA..... 657
H1O.CM.91.MVP5180         ------T-AC.--AGAC.T.-AAGCAG.....................................-----.---------.---------.---GC--..... 659
H1N.CM.95.YBF30           -------A-.C-AGAC.T.-A---A.....................................-----.---------.---------.---T-----..... 215
H1P.CM.06.U14788          ------G-AC.---TAG-C.-A---.....................................A----.---------.---------.---T-----..... 127
CPZ.CD.06.BF1167          --G-AGG-TAA.....................................GCA----.-------C---.---------.---------.---T-----..... 668
CPZ.TZ.00.TAN1            ------T.-----------------------------TTGTA.GTGGCTAAGTA.----.-------C---.---------.---------.---T-----..... 206
GOR.CM.07.SIVgorCP2135con ------GTAA.AG---CTGA-CAG.....................................-----.---------.---------.---T-----..... 198
```

5' LTR U5 end Lys tRNA primer binding site

```
MAC.US.x.239              TAAG...AAG.ACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCGAAGCAGG------.--C-------A..T-------T---------.---T-----..... 843
H2A.SN.86.ST_JSP4_27      AAGTA--AAG.ACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCAAGGCAGG------.--C-------G..T---------.---T-----..... 326
H2B.CI.x.EHO              G.......AA.ACCCTGGT.C---TAGGACCCCTTCTGCTTTGGGA..AACCAAGGCAGG------.--C-------A..T---------.---------..... 876
H2G.CI.92.Abt96           GC.....AAG.ACCCTGGT.C---TAGGACCCTTTCTGCTTTGGRG..AACCAGAGCGGG------.--C-------A..T---------.---T-----..... 248
H2U.CI.07.07IC_TNP3       GCT-ACAA.G.ACCCTGGT.C---TAGGACCCTTTCTGTTTTGGGA..AACCGAAACGGG------.--C-------A..T---------.---T-----..... 352
H2U.FR.96.12034           TAT------.A.ACCCTGGTAC---TAGGACCCTTCTTGCTTTGGGA..AACCAGAGCAGG------.--C-------G..T---------.---T-----..... 373
ASC.UG.10.RT03            ..........................................................................................................  0
ASC.UG.10.RT08            ..........................................................................................................  0
ASC.UG.10.RT11            ..........................................................................................................  0
COL.CM.x.CGU1             ..........GCCCT....C....AGAGCGGGGCCTTTTTGCTG......TCCGGGAAGG---.GGC--T---------------G---------.---T-----..... 286
COL.UG.10.BWC01           T-TT-CG-AG.ACCGTAG-------------.TTCTCTCCTTAG.....................--TCCA-CACGTG--AGA.G---TTA--TGGT------.T-CT..... 150
COL.UG.10.BWC07           ACGCCCG---TCCCTAC-CTCA-ATTGCTGGGCTGAGGACTTGACCAAGCAGGTCTTGTCC--GAG-G-CTTGCA...G---------.---T-----..... 283
DEB.CM.04.SIVdeb04CMPF3061 ..........................GAACGGCTTGGGTTGATCTGTAAGAGACAGA--G-CA-C--C----.---------.---------.---T-----..... 403
DEB.CM.99.CM40            ..........................AATTCGGGTCGAGCGGGCTAGGCTCAAG.CG-G-G.GTC--------...T---------.---------.---T-----..... 263
DEB.CM.99.CM5             ..........................GTCGGGAGCTCAAGCGGCTAGGCTCTGGGCAGG-G.AAC--------...T---------.---------.---T-----..... 265
DEN.CD.x.CD1_CMO580407    .......................AGGAGGTTCGGGTTGTGAGCCTTACGCAGGGTGA-----.C-C--C----.---------.---------.---T-----..... 248
DRL.DE.11.D3              ..........................................................................................................  0
DRL.x.x.FAO               ..........................................................................................................  22
GRV.ET.x.GRI_677_gri_1    TCTTACTG.........................GGTT..CTCCTG..TACCCAGGTGGG-G--A.---C----.---------.---------.---T-----..... 710
GSN.CM.99.CN166           ..................................AGTCCG.GGTGGACACC-G-...-CA-C----.---------.---------.---T-----..... 219
GSN.CM.99.CN71            ..................................AGTCGC..GGTGGATACC-G-...-CA-C----.---------.---------.G-T-----..... 227
LST.CD.88.SIVlhoest447    ..........................................................................................................  0
LST.CD.88.SIVlhoest485    ..........................................................................................................  0
LST.CD.88.SIVlhoest524    ..........................................................................................................  0
LST.KE.x.lho7             ACT---GAGA.-GGAG-C.T-----------------------------.GATCCGGG-C-.GGG--TC---.---------.---------.---T-----..... 813
MAC.US.x.251_1A11         TAAG...AAG.ACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCGAAGCAGG------.--C-------A..T---------.---------.---T-----..... 842
MND-1.GA.x.MNDGB1         ..............AAACCTGT--TGTTCTCATTTAGAGAACAGAAGGACTTCTAGTTA-CCCTAGAAGCCTTTCAG.---------.---------.---T-----..... 295
MND-2.CM.98.CM16          TA-T-G.AGA.-CCCTCAGATC-GT-GCAGAGATTTGATATCTAGTGGC..AGTAGAGAA------.--C-----.---------.---------.---T-----..... 471
MND-2.GA.x.M14            TA-T-G.AGA.-CCCTCAGATC-GT-GCAGAGATTTGATATCTAGTGAG..AGTAGAGAA----.----C-----.---------.---------.---T-----..... 393
MND-2.x.x.5440            ..........................................................................................................  22
MNE.US.x.MNE027           AAT-AG.AAG.ACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCGAAGCAGG------.--C-------A..T---------.---------.---T-----..... 324
MON.CM.99.L1_99CML1       AGTCTG----.CA-CTGTG.CG-ATCAGGAGCGGGCTTGAGCCGGGGAG..GGGACAGCAG-C-.G-C--TCTCCA..G---------.---------.---T-----..... 243
MON.NG.x.NG1              ..........................................................................................................  0
MUS-1.CM.01.CM1239        .........G.C-CCTGTT.TAT-CACACTGAGAGCCTACCCCGGCGTGGTGTGGTTCCAG----.CAC--C----.---------.---------.---T-----..... 229
MUS-1.CM.01.SIVmus_01CM1085 GG--AGT-TG.C-CCTATT.CAT-CACACTGAGAGCCTATCTCAGCGTGGTGTGAGTCTGG----.CAC--C----.---------.---------.---T-----..... 229
MUS-2.CM.01.CM1246        -CG-.....................ACCTCGCGGG.....................GGGA--T-----CTC----.---------.---------.G-T-----..... 248
MUS-2.CM.01.CM2500        .........................ACC.....................GATCCAGGGCT-C--.AC--CC---.---------.---------.---T-----..... 228
MUS-3.GA.09.09GabOI81     ..........................................................................................................  0
MUS-3.GA.11.11GabPts02    ..........................................................................................................  0
OLC.CI.97.97CI12          ..................................................................T--C-A----.---------.---------.---T-----..... 584
RCM.CM.00.SIVagi_00CM312  ..........................................................................................................  0
RCM.CM.02.SIVrcm_02CM8081 ..........................................................................................................  0
RCM.GA.x.SIVRCMGAB1       ..........................................................................................................  22
RCM.NG.x.NG411            ..........................................................................................................  22
SAB.SN.x.SAB1             ..................................GGTTACTAAGGA..TCCCTGATAGAA------.--CT-----.---------.---------.---T-----..... 790
SMM.CI.79.SIVsmCI2        ..........................................................................................................  0
SMM.LR.89.SIVsmLIB1       ..........................................................................................................  0
SMM.SL.92.SIVsmSL92A      ..........................................................................................................  0
SMM.SL.92.SIVsmSL92B      ..........................................................................................................  0
SMM.SL.92.SL92B           AA....GAG..CCCTGGTAC---CAGAACCCTCTCACTCTTGGGGGCAACTCCTGAGTG----A-----G----T..T---------.---------.---T-----..... 321
SMM.US.04.G078            A-ATA-GAAG.ACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCAGAGCAGG------.--C-------A..T---------.G-------.---T-----..... 110
SMM.US.05.D215            ....-TAAAA.ACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACC.AAACAGG------.-AC-------A..T---------.---------.---C-----..... 104
SMM.US.06.FTq             ...AG-AAG.ACCCTGGT.C---TAGGACCCTTTCTGCTTCGGGA..AACCGGAGTAGG------.--C-------A..T---------.---------.---T-----..... 104
SMM.US.86.CFU212          --A-C-GAAG.ACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGC..AACCGAGGCAGG------.--C-------A..T---------.---------.---T-----..... 106
SMM.US.x.H9               TAAG...AAG.SCCCTGGT.C---TAGGACYCTTTCTGCTTTGGGAAAC..CGAGGCAGG------.--C-------G..T---------.---------.---T-----..... 326
SMM.US.x.PGM53            TAA---.AAG.ACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGC..AACCGAGGCAGG------.--C-------A..T---------.---------.---T-----..... 769
SMM.US.x.pE660.CG7G       -AT-AGAAG..ACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCGAGGCAGG------.--C-------G..T---------.---------.---T-----..... 842
STM.US.89.STM_37_16       .......AAG.ACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGC..TACTGAGGCAGA------.--C-------A..T---------.---------.---T-----..... 507
SUN.GA.98.L14             ........AC.-GCTC-T-.ACTA--GAGCAATTACTGGCCGCCTA..GTGATCCGGGGCTG-C.GGC.----------.---------.---------.---A-----..... 792
SYK.KE.x.KE51             ..........................................................G-G-CTT----.---------.---------.---T-----..... 265
SYK.KE.x.SYK173_COMGNM    ..................................................................CTT----.---------.---------.---T-----..... 624
TAL.CM.00.266             ...............TGGTCCC-C-CCCTCGGGTCGAGGGAACTTGGTGGGTGAAGAAATTT-CCCAC-AGCTG-CGCCC-AACAGG--CTT-A-AACA-CC-GGTTCTCTGTCCTGAGGGCCAAGCGGCTTCGGGGACCGCGGGTTTTGGGTCCCAAAGAAGAGAGAGA 344
TAL.CM.01.8023            ..........................................................................................................  0
TAN.UG.x.SIVagmTAN1       ..............................GGTT..TCTTA..T..CCAGGCTGAG-------.C-----.---------.---------.---T-----..... 728
VER.DE.x.AGM3             TG.............................GGTTCTCTCTCA...ACCCAGGCGAG-G---.A---C-----.---------.---------.---T-----..... 241
VER.KE.x.9063             TG.............................GGTTCTCTCTCT..GACCCAGGCGGG-G---.A---C-----.---------.---------.---T-----..... 746
VER.KE.x.AGM155           TG.............................GGTTCTCTCTCA..AACCCAGGCGAG-G---.A---C-----.---------.---------.---T-----..... 745
VER.KE.x.TYO1_patent      TG.............................GGTTCTCTCTCT..GACCCAGGCGAG-G---.A---C-----.---------.---------.---T-----..... 748
WRC.CI.97.97CI14          GAGCTGAAGA.CCCC.....................................................A-----T----.---------.---------.-------..... 822
WRC.CI.98.98CI04          GAGCTGAAGA.CCCC.....................................................A-----T----.---------.---------.---AA----..... 825
WRC.GM.05.Pbt_05GM_X02    ..........................................................................................................  0
```


Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

```
H1B.FR.83.HXB2                  ..........................................................................................................AAGCGAAAGGGAA 670
H1O.CM.91.MVP5180               ..........................................................................................................---T-----T-G- 672
H1N.CM.95.YBF30                 ..........................................................................................................--A------TAG- 228
H1P.CM.06.U14788                ..........................................................................................................------T--- 134
CPZ.CD.06.BF1167                ......................................................................................................................  668
CPZ.TZ.00.TAN1                  ..........................................................................................................G-AGC-GG-....  215
GOR.CM.07.SIVgorCP2135con       ..........................................................................................................---T----CCA-.  210

MAC.US.x.239                    ..........................................................................AGGAGAGTGAGAGA...CTCCT.GAGTACG.---TG-GT-A-G  880
H2A.SN.86.ST_JSP4_27            ..........................................................................AGAAGACTGAGAA.G..CCTTG.GAACACG.---TG-GT-A-G  363
H2B.CI.x.EHO                    ..........................................................................GGAAGAGTGAGAG.T..CTTCG.GGACACG.---TG-GT-A-G  913
H2G.CI.92.Abt96                 ..........................................................................GAAGGAGTGAGAG.T..CCTTT.GAACACG.---AG-GT-A-G  285
H2U.CI.07.07IC_TNP3             ..........................................................................AGGAGAGTGAGAG.G..CTCCT.GAACACG.---TG-GT-A-G  389
H2U.FR.96.12034                 ..............................................................................AGAGTGAGAG.T..CTTCTGAAACTCG.---TG-GT-AGG  408
ASC.UG.10.RT03                  ..........................................................................................................................  0
ASC.UG.10.RT08                  ..........................................................................................................................  0
ASC.UG.10.RT11                  ..........................................................................................................................  0
COL.CM.x.CGU1                   ................AAGAGGTCCCTGAGGGTCGCGGAAG.AAGAAGAGGAGACTGATC.GCCTCCTCA...GAAGAGAGGTAAGTCGAAAAAGGGGACTTACG.GGTGGA..GTCA.AGGAGGACTCTACGGTCTGCGCAAAAGTGAAAGTAAA-GTGCGCGCTCGT  430
COL.UG.10.BWC01                 ..........................................................................AAGGAGAGA...ATCTACGGTCTGC..---A----T---  183
COL.UG.10.BWC07                 ...................CTAGGACGGACACTAGGAGT.CGCGGGGCCACAAGAGAG.AGAGAAAAA...GAGAAGAAAGGAGCCTGATCAACTCCTGAGGTAAGTGGG.TCTAG.GAGAGACCTACGGTCTGCGCCCTGAAGAGAAGTTCCCT-A.....T---  420
DEB.CM.04.SIVdeb04CMPF3061      ..........................................................................GCATC.CCAGTGTTCAGGCCCGAAGACGGTAGGGACGGAGAGG-TCT-TGGA--CC  458
DEB.CM.99.CM40                  ..........................................................................GCTAC.CCAGAGCTC.GGCCCG........CAGGCGGTGTGCGG--TCTGGAAACC  309
DEB.CM.99.CM5                   ..........................................................................AGCATC.CCAGTGTTTGAGCCCG........AAGGCGGAGTGTGG--TCTGGAAACC  313
DEN.CD.x.CD1_CMO580407          ..................................................................AAGAGAGTCCTGAAACAGG.GACTTGAGTACCCAGAAGGCTCAGGCCCGAAGGCGGTTG-AACGTACC-C  317
DRL.DE.11.D3                    ..........................................................................GC..TAAGAAGCCCCGACGCGAGACCCGC-TCGGTAAGT   37
DRL.x.x.FAO                     ..........................AGACAGT.CGACCCCAC...CTGTGAGGGACGAAGGCGGCAGCCATCCGGACCGACCCACCCGG...AAAAGAGAGTGC..TAAGAAGCCCCGACGCGAGACCCGC-TCGGTAAGT  129
GRV.ET.x.GRI_677_gri_1          ..........................................................................GAAGAGGCATCGGCA-CG-CC-CTG  735
GSN.CM.99.CN166                 ..........................................................AACAGCCCAG.GACTTCAA.GCCCGAAGGCGGGACCCAAGGATTGCCCCTGGCCGAGTG-TCCGG-CTCC  287
GSN.CM.99.CN71                  ..........................................................AGCATCCCAGTGACTTCAA.GCTCGCAGGCGGGACCCAAGGATTGCCCCTGGCCGAGTG-TCC.G-CTCC  295
LST.CD.88.SIVlhoest447          ..........................................................................................................................  0
LST.CD.88.SIVlhoest485          ..........................................................................................................................  0
LST.CD.88.SIVlhoest524          ..........................................................................................................................  0
LST.KE.x.lho7                   ..................................AGGAGA...AGAGGCGCTGGCAAGAACCTCTGGAGGGAGGAAAAGCGCGCAG.........AAGAGGCCAGGAAGCTAGGACAGAGAAAGTCTGCTCGGG-CAGG  906
MAC.US.x.251_1A11               ..........................................................................AGGAGAGTGAGAGA...CTCCT.GAGTACG.---TG-GT-A-G  879
MND-1.GA.x.MNDGB1               ..........................................................................................................................  295
MND-2.CM.98.CM16                ..............................CAAGGGTTC...GAGTCATTCCAACCTGTAAGGGACGAAGGCGGCAGCCAGCCGGAC.CGACCCACCCGGCGAAGTGAGTTAACCAAGGAGCCCCG-CGCGC---ACT  575
MND-2.GA.x.M14                  ..................................CAAGGGTTCGAG.TCGCTCCGA...CCTGTAAGGGACAAGCGGCGGCAGCCGGTCGGACAGACCCACCCG.ACGAAGAGAGTTAATCAAGGAGCCCCGACGCGCAGGACTC.A-GGTAAGC  508
MND-2.x.x.5440                  ..................................CAAGGGTTCAGG.TCGCACCCA...CCTGTAAGGGACGAGCGGCGGCAGC.AGTCGGACAGACCCACCCG.GCG.AAAGAGTTAATTGAAGAGCCCCGACGCGCAGGACTCAA-GGTAAGC  136
MNE.US.x.MNE027                 ..........................................................................AGGAGAGTGAGAGA...CTCCT.GAGTACG.---TG-GT-A-G  361
MON.CM.99.L1_99CML1             ..........................................................................GTAACGACCCAGCGACTCGAGCCCGCAG-CG-ACAG-  280
MON.NG.x.NG1                    ..........................................................................................................................  0
MUS-1.CM.01.CM1239              ..........................................................TTAGCCCGCGAGGGGGT.AGGCGGTGTGACAGAGAGAAGG...CCCC...AGCCG..AGTGCGT-CC-  288
MUS-1.CM.01.SIVmus_01CM1085     ..........................................................TTAAGCCCGCGAGGAGGT.AGGCGGTGTGACGGAGAGAAGG...CCCC...AGCCG..AGTGCGT-CC-  289
MUS-2.CM.01.CM1246              ..................CTACATATCAA.CCCGTTCGG...AGGTTCGCGGCCGGCAGTCAACAGGCGGTGGAGCAGTCCCAGAAG.GCTCCATCCCGAAGG..CGAGACGCCTGAGGATCGCCT---CG-GT-..G  359
MUS-2.CM.01.CM2500              ..........................CGA...ACGTCAGGTAGGATCGGTGAGCAGGCAGTAGAG.CCTCCCAGGGA.CTCTGCGCTGAAGGCGACAGGCGCCGAGAGAACGCCCC---CG-GT-...  322
MUS-3.GA.09.09GabOI81           ..........................................................................................................................  0
MUS-3.GA.11.11GabPts02          ..........................................................................................................................  0
OLC.CI.97.97CI12                ..........................................................................................................................  584
RCM.CM.00.SIVagi_00CM312        ..........................................................................................................................  0
RCM.CM.02.SIVrcm_02CM8081       ..........................................................................................................................  0
RCM.GA.x.SIVRCMGAB1             ..........................................................................GAGTGGCTGAGAGACCTC---GGCTA-G  50
RCM.NG.x.NG411                  ..........................................................................GAGTGACTGAGAGACCTC---CGC--GT  50
SAB.SN.x.SAB1                   ..............................................................CTAAAGACGTCCCGGCCGCAGAGTAGACCAGCCTGAGG-CTCCTTT-ACT  840
SMM.CI.79.SIVsmCI2              ..........................................................................................................................  0
SMM.LR.89.SIVsmLIB1             ..........................................................................................................................  0
SMM.SL.92.SIVsmSL92A            ..........................................................................................................................  0
SMM.SL.92.SIVsmSL92B            ..........................................................................................................................  0
SMM.SL.92.SL92B                 ..........................................................................A.GACG.....................GACT.CG.---TG-GT-A-G  343
SMM.US.04.G078                  ..........................................................................AGGAGAGTGAGAGA...CTCC.TGAGTACG.---TG-GT-A-G  147
SMM.US.05.D215                  ..........................................................................AGGAGAGTGAGAGA...CTCC.TGAGTACG.---TG-GT-A-G  141
SMM.US.06.FTq                   ..........................................................................AGGAGAGTGAGAGA...CTCCAAGAGTACG.---TG-GT-A-G  142
SMM.US.86.CFU212                ..........................................................................AGGAGAGTGAGAGA...CTCC.TGAGTACG.---TG-GT-A-G  143
SMM.US.x.H9                     ..........................................................................AGGAGAGTGAGAG....CTCCT.GAGTACG.---TG-GT-A-G  362
SMM.US.x.PGM53                  ..........................................................................AGGAGAGTGAGAG....CTCCT.GAGTACG.---TG-GT-A-G  805
SMM.US.x.pE660.CG7G             ..........................................................................AGGAGAGTGAGAG....CTCC.TGAGTACG.---TG-GT-A-G  878
STM.US.89.STM_37_16             ..........................................................................AGGAGAGTGAGAGA...CTCCT.GAGTACG.---TG-GT-A-G  544
SUN.GA.98.L14                   ..................................GGGAG.AGAGAGTGA...CGAAGAGACGTGGAGTGAAACCTTCAGAAGAAGTGAGCAGCTGCT.GAGAAGAAGGAAGGTTTCGACCTGGTGGAAGGCGCCGGGAA-TT-A--T-  901
SYK.KE.x.KE51                   ..........................................................................................................CT--G  270
SYK.KE.x.SYK173_COMGNM          ..........................................................................................................CT-G-CC-G  633
TAL.CM.00.266                   AACTCCGAGGGGTCGTGAGAGGACGACAGTTCCGGAGTCGAGCCCCGGATCGGCGAAGACGACGCGCGTCGCAAAGTGAAAGTAAGGTAAGGAGCCGATCGGTCACCCCGGGCGAGTAGCCGCCGTGAGCGAGTGCGGTGCTAGGAAGAGTGGCAGAG--GCCT--CTT-  514
TAL.CM.01.8023                  ..........................................................................................................................  0
TAN.UG.x.SIVagmTAN1             ..........................................................................................................GTCA-----ACTT  741
VER.DE.x.AGM3                   ..........................................................................................CTTGAGTGAAGGCACGT-CAGCTG--AAG-  271
VER.KE.x.9063                   ..........................................................................GGTCCGAGGA.CTTGAGTGAAGGCACGT-CAGCTG--AAG-  786
VER.KE.x.AGM155                 ..........................................................................GAGTGAAGGCACGT-CAGCTG--AAG-  772
VER.KE.x.TYO1_patent            ..........................................................................GTGA.GAGTGTAGGCACGT-CAGCTG--AAGG  779
WRC.CI.97.97CI14                ..........................................................GGGG.CGTCTAGAGTGAAAGAAGACCGGCCGGCGAGGGATTC-CT------AAGT  876
WRC.CI.98.98CI04                ..........................................................GGGG.CGACTAGAGTGAAAGAAGACCGGCCGGCGAGGGATTC-CT---G--A-GT  879
WRC.GM.05.Pbt_05GM_X02          ..........................................................................................................................  0
```


Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

```
                                  ┌Packaging loops begin
H1B.FR.83.HXB2              A......CCAGA.....GGAGC....TCTCTCGACGC..AGGACTCGGCTTGCTGAA...GCGCGCAC.GGCAAGAGGCGAGGGGCGGC.....GACTGGTGAGTACGCC....AAAAATTTTG.ACTAGCGGAGGC..................................  776
H1O.CM.91.MVP5180           -.........---G.....-A--AAAACC..TC------AAC--G--------AGC-G-...-T--A-CT.-CT---------A--..AACTCACA-G...A-G--GA-TA........-A----.---G----T--.CC....................AGACCTAGGG  788
H1N.CM.95.YBF30             -.......--G--.....--CTGAA..------------.-----------C-T--.....-T--A---.A--G-------------.----G.....--AGT-------------.........-.--.---------G----T--.CC..............................  330
H1P.CM.06.U14788            -.......-----....TTTCTGAAACC..TC------A.C--G-------CAGC-G-...-T--A-CT.-CTG---------A--..AACTCAC-G.C---------.........-------..TC--T--T-A-T.....................GACCCTAGGG  248
CPZ.CD.06.BF1167            .GAAGCAGGG--......CGCGGAACC-.T-GACGC-A..-------.-GC-TG--.....CT-----A.CA-------------TAAGTT.........................G------C.TT-GCT-..............TTAGT.GTACAGCCCTAGGGGAAG  780
CPZ.TZ.00.TAN1              ...........-.....ACGCGG.CCC..--G------..----------------.---C..A----AAT-..A-----------C-GACTC....CGG-----------......A--TTT-G--.....T---T--GC.....................AACCCTAGAG  322
GOR.CM.07.SIVgorCP2135con   ..........G.....TTCTGA...CTC---CGACGCAC--G-------CAGC-G-...-T--A-C-.ACTG---------A--AA-ACTCACAGAC---------A..............-------.-C-----T--..CTG.................GACCTAGAGG  323

MAC.US.x.239                GCAGTAAGGGCGGCA.G-A-C-AACCA-GACG--GTGCTCCT-TAAA-GCGCGG-TC..G-TA-C.....AG-C.G-CGTGA--AGC-GG...AGAG-AA---G.-CT-CGGTTGC...AGGTA-G-GC........AACACAAAAAAGAAAT..........AGCTGT 1015
H2A.SN.86.ST_JSP4_27        GCAGTAAGGGCGGCA.G-A-CAAACCA-GACG--GTGCTCCT-GAAAAGCGCAG-CCGAG-TA-C..A.A-GGC.G-CGTGT--AGC-GG...AGTGAAA---G.-CT-CGGGTG-...AGGTA-G-GC.......CTACACCAAATACAGTAGCC..AGAAGGGCTT  510
H2B.CI.x.EHO                GCAGTAAGGGCGGCA.G-A-T-AACCA-GACG--GTGCTCCTGTAAAAACGCAG-CC..G-TA-C.....AGGC.--CGTGA--AGC-GG...AGGA-AA---G..CT-CGGG-GT--AAGGTA-G--C........TCCAAA...................TTTGCAA 1041
H2G.CI.92.Abt96             GCAGTAAGGGCGGCA.G-ATC-AACCA-GACG--GTGTTCCT-GAAA-GCGCGG-TC..G-TA-C.....AG-C.--CGTGA--AAC-GG...AGTG.AA---G.-CT-CGGGTGCGC.AGGTA-G-GC........ACCAAAAATTTGTAGCCAAAAA...GGGCTGT  428
H2U.CI.07.07IC_TNP3         GCAGTAAGGGCGGCA.G-A-CAAACCA-GACG--GTGCTCCT-GAAA-GCGCGG-CC..G-TA-C.....AGGC.G-CGTGA--AGC-GG...AGTG.AA---GC-T--GGGTG--GG-AAGTAC--TC........ACCAAA.ACTTGTAGCCAGAAA...GGGCT.T  532
H2U.FR.96.12034             GCAGTAAGGGCGGCA.G-A-CAAACCA-GACG--GTGCTCCTGTAAA-GCGCAG-CT..G-TA-C.....AAGC.--CGTGA--AGC-GG...AGTG..A---G.TCT-CGGTTGC...AGGTA-G-GC........AACAAAAAAATAGAACCAAA......GGGTTGT  545
ASC.UG.10.RT03              ..................-G-TCTAGGACAGT-G-AGGGTCCTAGGTCGG--GGTTGAGA-AAG-GTA............CCCACACAGC.........-G-GAG-AA--TTAGT-GTGGGAA-ACTC..............................ACGAGGCTAC   99
ASC.UG.10.RT08              ...GTCCTAGACCCGACC-G-TCTAGGACAGT-G-AGGGTCCTAGGTCGG--GGTTGAGA-AAG-GTA............CCCACACAGA.........-G-GAG-AA--TTAGC-GTGGGGA-ACTC..............................ACGAGGCTAC  114
ASC.UG.10.RT11              ..................-G-TCTAGGACAGT-G-AGGGTCCTAGGTCGG--GGTTGAGA-AAG-GTA............CCCACACAGA.........-G-GAG-AA--TTAGC-GTGGGAA-ACTC..............................ACGAGGCTAC   99
COL.CM.x.CGU1               CGCGCAGT---G.AA.GT-GCAGACGAGAGAGC--AGAG---GAGT-CTCGC-A--GAAG-A-AAG-GCTCATTAG-AG-TCTTCA---A....-GGA-TA-C-CC-CTGGGCAGGTCGAGGG-GG-CT........GCAAGTGGAGTAAGCTGCGGCGCTCCAGGACCG  586
COL.UG.10.BWC01             -GTAAAAGGC-CGCGGC---CGCGCAGTGAGGA-GTGGC--ACGAGA-TGCTAG-GGA.GAGCACTCGACCAG---A-GAGAA-AGCT-ACTAG--G-TC-TCAGG-AGGGAGT-GCGCCCC--GG-A---CAGT-AAGGGCTTGTAGTGGGAGGTAAGCTGCGGCGCCT  352
COL.UG.10.BWC07             -GTGAAAGAG-TAGT.G-C--ACACAGTG-G-CTAAGGG--AGGA--CACGCTA-GGAAG-A-A.GGT.CCGCTAG--ACCTCT-GT--A....-GGA-TA-C-CC-CTGGGCT-GTCGCC-CAG-GGGAA-AG--ACAGGACTAGCAACGGAAGCTAAGCTGCGGCGGC  583
DEB.CM.04.SIVdeb04CMPF3061  CGTGCTGATCTCTTC.CT-C-AGACAA.GGCT-CGCAGCCAA-GGTA..AGCGGACGCCGAGTA.AGA.TC-GC---TGAG-CA-.-TG.....AGTG---TCTGCA-G-AGAATTCC-C-G-TGGGGG..AA-TAGAGAGTA.CCCCTGGTCCACTAGGGACG.GTGT.  611
DEB.CM.99.CM40              C.TGAGC-G-TCTCT.TCCC--CGGA-.AGCT-CGCAGC-AAGG-AA--GGA-GCCGGAGAATACAGG.-C-GC--ATAAG--C-G-TGT....-TG.-TCTCTGCA-G-AGAGT-CC-C-G--G-GC-........AGAGAGGCTCCTGGTCCACTAGGGACGTTTGTG  462
DEB.CM.99.CM5               C.TGAGC-G-TCTCT.TCCC-TCGAG-GAGCT-CGCAGC-TAGG-AA--GGA-GCCGGAGAATACAGG.-C-GC---TG-G-AC-G-TGT....-TG.ACCTCTGCT-G-AGAGTGTC-C-G--G-GC-........AGAGG..CTCCTGGTCCACTAGGGACGTTTGTG  465
DEN.CD.x.CD1_CMO580407      CCTCTTA--GAGCAA.G--C-AGGAAGACGCT-CGCAGCGCA-AGGTAAGGATCTCGCCG-A--A---.ACT-GTG-CA-TA-T-T--GG...CTGGAAGATAGCAAGGAGGGCT-TC-GGAAACAA-........GGGCTAGATAAGGGGTTGAAAAAAATCTAGT..  471
DRL.DE.11.D3                -AAAAGGAA--TAAA.ATC--TGCAC-G-GAG-CGAAGGGTCTAGT--GCC-TGTG-GCA---GCT-G.-CA---GCAACTCC-TA---A.....-AA-AAAGAA-AAGGGAGA-GT--GGCCTGTTGC........TGGCAAGTGAGGCAGAAGAAGCCTAGG......  186
DRL.x.x.FAO                 -AAAAGGAA--TAAA.ATC--TGCAC-G-GAG-CGAAGGGTCTAGT--GCC-TGTG-GCA---GCT-G.-CA---GCAACTCC-TA---A....A-AA-AAAGAA-AAGGGAGA-GT--GGCCTGTTGC........TGATAAGTGAGGCAAAAGAAGCCTAGG......  279
GRV.ET.x.GRI_677_gri_1      -GTTGCTGAGCGTCG.GA---GGACGA---AG-TA-GGTGA--GC-TA-GA-T-TTTTGCTAC-TAGT.CAGCGAGAAA-GCTA-GCCGC....---A--G-C-CGG-T-CCATT-GTGGCAACC-ACC........CAGTTGGACGAAGGGTTGGTAGGGGACGGGTCG  891
GSN.CM.99.CN166             -CGAAATTTGCTGCG.CA-CC-CTTG...AAGAGAAGGT-A-GAGACCGGACGG-TCTA.-AAG-GTG.-CAG-CTTCTAG-CCAG-TGG.....GT--.AAGAAGG-TACCCAC-C-G..AG-GGAG-AACTGT-............................AAAGAG  415
GSN.CM.99.CN71              -C..AACTTGCTGCG.CA-CA-CTAG...AAGAGAAGGT-A-GAGACCGGACGG-TCTG.-AAG-GTG.-CAG-CTTC-AG-CC.G-TGG.....GT--AGAGAAGG-TACCCAC-C-GAGGG-AGAA-CT-TGA..............................AAAAG  421
LST.CD.88.SIVlhoest447      ...............................................................................................................................................................................    0
LST.CD.88.SIVlhoest485      ...............................................................................................................................................................................    0
LST.CD.88.SIVlhoest524      ...............................................................................................................................................................................    0
LST.KE.x.lho7               -AGAGAAGACCTGCT.T-AGA-GAGGATAAGAAGGA-TTGACTGCTA-AAG--G-C-GAA-GTAAGGA.-AGT-AGAC-TCACC-G-AGA....-TGGCAGC--GC--GGGGCAGCCTTG-GA-A-G-G........TGTAATTCTGATCTAGTGTAACTGGGACACTAG 1062
MAC.US.x.251_1A11           GCAGTAAGGGCGGCA.G-A-C-AACCA-GACG--GTGCTCCT-TAAA-GCGCGG-TC..G-TA-C.....AG-C.G-CGTGA--AGC-GG...AGAG-AA---G.-CT-CGGTTGC...AGGTA-G-GC........AACACAAAAAAGAAAT..........AGCTGT 1014
MND-1.GA.x.MNDGB1           .......AG---GGC.ACTGA-ACTTGAGGCA--GCACTCC-C--G-AAGAAGCAGGTTGAA-GAG-G.T-G-CTG-T-TGAA-A--C-A....-GAG-TGAGTC-GTGGGACTG-CTT-ACAAG-A-T........AGTTGTACCCTAGTGTAAGGGGCAGCATAGTCA  444
MND-2.CM.98.CM16            CAAGGTAAGC-GTGC.ACCGAACAGT-G-G-GTGT-AGTCCAGGGTCCGC-AGA-C-AGC-A-AC-G-.C-AGGC-AC-CCA-TAG-CAA....A-GGAAAAGAGGG-AAGTAAGGCCGAGGCAA-G-G........AAAAGAGGTCCTTGTGAAAGAGGACAGGGGGCA  731
MND-2.GA.x.M14              GGTGCAC-........GTATTGTGGTAG-GAGTC-.AGG-TCCGCTT-AGCAGGCG-GATCGCTAAGG.CAACCCCA-TAG-T.AAAAGA....A-AGA-G-GAAGTAAGGCCG--GC-AAG--A-AGT........AAAAGGGCCTTCCGAAGGGAAGGGCAGGGGGCA  655
MND-2.x.x.5440              -GTGCACTGTA.....ATTGATTGAGAGAGAGTC-AGGG-TCCGCTT-AGCAGGCG-GATCGCGAAGG.CAACCCCA-TAG-CAAAA-AA....A-GAA-G-GAAGTAAGGCCG-CGC-AAG--A-AGT........AAGGGCTTCCTCCGAAGGGAGAAGCAGGGGGCA  288
MNE.US.x.MNE027             GCAGTAAGGGCGGCA.G-A-C-AACCA-GGCG--GTGCTCCT-GAAA-GCGCGG-TC..G-TA-C.....AG-C.G-CGTGA--AGC-GG...AGAG.AA---G.-CT-CGGTTGC...AGGTA-G-GC........AACACAAAAAAGAAAT..........AGCTGT  495
MON.CM.99.L1_99CML1         TCCGCGAGA--CCTT.GCCGAGTGCGC-CGA-A-A--TGC-C-GCAA-AGGTAA-G-GAGATCG-GCG.-T-CCCTA-G-T--CAGACTA....-CG---GAC-GGAA--TATCC--G-GGGAAGGT-C........CAGAGGTGCTCCAAGAGGGAGCATAAAATAAGG  436
MON.NG.x.NG1                ...............................................................................................................................................................................    0
MUS-1.CM.01.CM1239          -G.ACCA--G--GGA.CAAGTGCTGCG-AGC-A-GAGGT---TGG-TTGGACGG-TCTGG-AAG-GTG.-CAG-CTTC-TCA-AT--AGC...AGG-ACGAGAA--C-TGCGTG-G-GGAGAAACTCC..............................TGAGA  415
MUS-1.CM.01.SIVmus_01CM1085 -G..CCA--G--GAA.CAAGTGCTGCG-AGC-ATAAGGT---TGA-TTGGACGG-TCTGG-AAG-GTG.-CAG-CTTC-TCA-AT--AGCA...-G-ATAGAGAG-TC-TGCGGG-GGGGAGAAACTC...............................GGAGA  415
MUS-2.CM.01.CM1246          CCCGGCA-GGAGAAA.AAAGAGCTGCG-AGCAA-A-AGGTAAGG-GTCGGGCGA-CCTA.-A-G-GTG.-CAG-CCTCTAG-T--GC-GGT...AGAC-AG---.GTC--CGCG-GG-GGAGAAAC-GCT..............................AGTAG  487
MUS-2.CM.01.CM2500          -GCACCGA--CGAGG.AAATCGCTGCAGC.GT--A-AGGTAAGG-GTCGGGCGA-TCTA.-A-G-GTG.-CAG-CCTCTAG-C--GC-GT.....T-A-AGA--.GGA--GCGAG-G-GGAGAAAC.................................GCGAG  443
MUS-3.GA.09.09GabOI81       ...............................................................................................................................................................................    0
MUS-3.GA.11.11GabPts02      ...............................................................................................................................................................................    0
OLC.CI.97.97CI12            .............................................-AA.AG-GTTCG.A-TCCCT-AA--A-AAGTTGGTC.AT-AG-TGAGTG--CGGGGAAC-CG-GG--TAGG-CCT...........................ACAGGATAGG  666
RCM.CM.00.SIVagi_00CM312    ...............................................................................................................................................................................    0
RCM.CM.02.SIVrcm_02CM8081   ...............................................................................................................................................................................    0
RCM.GA.x.SIVRCMGAB1         GTTCGGCG-G-CAGG.TCACCGCGGGAG-GGAAC-TTGACCAGG-AA-AGCTGCCTGGTG--TTCA-A.-TGCTAGA-AA--T-AGCTAG....CGAA--AAGCAGG--AACCCGGTCCGG-ACTGGGC........CTCTAGGGAAGGAGCGAAGTCCT..........  196
RCM.NG.x.NG411              CCGGACCAG-AGAGT.GACT-GTAGAA-G-GGAC-CAGTCCAGG-AA-AGCAGCCTGGTG--T--T-A.AT-T--C-AGA-AA-TGA-GG....CTAGTAGAGCA-G--GACCCGGTC-GG-ACTGGGC........CTCTAGGGAAGGGTCGAAGTCCT...........  196
SAB.SN.x.SAB1               GAGTAAGA-C-GACG.GAAGAAAGGCGGA-CGTG-CGGAGT-CAGGTAAGA-GC-CCTGC-GA-C-GA.-A--CCCTCG-GCT-TGT-TG....TGT.....---A-GAGAGGAA---G-GAAGTC-GAC........AGGGCTGACAAGGCTGGCTTAGCAGTGGCAGAA  992
SMM.CI.79.SIVsmCI2          ..............................ACG--GTGCTCCAGTAAATAGC-GG-GCC.G-TA-C.....AG...GC--CGT-T-GA--GGGA..GTG.AA---G.-CT-CGACTG-C..CGGTA-G-G--A-TAAAACTTGAGCCAAGAAGGCTGTTAGTCTACCCTT.  125
SMM.LR.89.SIVsmLIB1         ..............................ACG--GTGCTCCT-GAAA-GCGCGG-CC..G-TA-C.....AG...GC-GCGT-A-GA--AGGA..GTCAAA---G.-CT-CGGTTGC...GGGTA-G-GCGCA-CAAAAACAG.GTAA..........TCGTACTCTGAA  114
SMM.SL.92.SIVsmSL92A        ..............................ACG--GTGCTCCT-GAAAAGCG-GGACCG...A--C..GGTA-GGTGCCGTGT--AGC-AGAG...CTG..A---G.-CT-TGGTTGC...AGGTA-G-GCAAC-A..CTTGGT.AGCCGAAAGGGCT.GTCTATCCTTAG  123
SMM.SL.92.SIVsmSL92B        ..............................ACG--G....TTTGA....-CC..........G-T.....-C--GTG---TGA-C-G-AGTACAG.GTAA-A---G.-CTGTGGG-.C...AG-AAGG-CCA.............AAAGGTGTGAATA.TCCTACCTACGG   98
SMM.SL.92.SL92B             GCAGTAAGGGCGGCA.G-A-T-AACCG-GACG--G....TTTGA....-CC.........G-T--C.....AGTG.G-C.TGA...GC-GG....AGTACAG-TAAGA-AGGCCTGTGGGACAGTA-GGTC..........CAAAAGG...................TGTGA  454
SMM.US.04.G078              GCAGTAAGGGCGGCA.G-A-CAAACCA-GACG--GCGCTCCT-GAAA-GCGCGG-CC..G-TA-C.....AGGC.G-CGTGT--AGC-GG...AG-C-AA---G.-CT-CGGTTGC...AGGTA-G-GCA.ACA...AGTAGCC.AGAGGGCTGATAG.TTCTGCTCTT  294
SMM.US.05.D215              GCAGTAAGGGCGGCA.G-A-CAAACCA-GACG--GTGCTCCT-GAAA-GCGCGG-CC..G-TA-C.....AGGC.G-CGTGT--AGC-GG...AG--AAA---G.-CT-CGGTTGC...AGGTA-G-GCA.ACA...AATAGCC.GTAGGGCAAGCAG.CATCACTCTG  288
SMM.US.06.FTq               GCAGTAAGGGCGGCA.G-A-CAAACCA-GACG--GTGCTCCT-GAAA-GCGCGG-CC..G-TA-C.....AGGC.G-CGTGT--AGC-GG...AGGC-AA---G.TCT-CGGTTGC...AGGTA-G-GCA.ACA.AAAAGAGCC.AGAGGGCTGTAG..TTCTACTCTC  290
SMM.US.86.CFU212            GCAGTAAGGGCGGCA.G-A-CAAACCA-GACG--GTGCTCCT-GAAA-GCGCAG-CC..G-TA-C.....AGGC.--CGTGT--AGC-GG...AGTC--A---G.-CT-TGGTTGC...AGGTA-G-GCA.ACAA-TAGTAGTT...............GACAGTACTC  280
SMM.US.x.H9                 GCAGCAAGGGCGGCA.G-A-CAAACCA-GACG--GTGCTCCT-TAAA-GCGCGG-CC..G-TA-C.....AGGC.G-CGTGA--AGC-GG...AGTC--A---G.-CT-CGGTTGC...AGGTA-G-GC........AACAAAAGAGTCATA..........GGACTGA  497
SMM.US.x.PGM53              GCAGTAAGGGCGGCA.G-A-CAAACCA-GACG--GTGCTCCT-TAAA-GCGCGG-CC..G-TA-C.....AGGC.G-CGTGA--AGC-GG...AGTC--A---G.-CT-CGGTTGC...GGGTA-G-GC........AACAAGAAGAG.........TCATAGGACTGA  941
SMM.US.x.pE660.CG7G         GCAGTAAGGGCGGCA.G-A-CAAACCA-GACG--GAGCTCCT-GAAA-GCGCGG-CC..G-TA-C.....AGGC.G-CGTGA--AGC-GG...AGTC--A---G.-CT-CGGTTGC...AGGTA-G-GCA.AC..AGAAAAGTC.ATAGGACTGAGTT.CCCTACTTTT 1026
STM.US.89.STM_37_16         GCAGTAAGGGCGGCA.G-A-CAAACCA-GACG--GTGCTCCTGTAAA-GCGCGG-C...........-.....AAGC.G-CGTGA--AGC-GG...AGGC-AA---A..CT-CGGTTGC...AAGTG-G-GC........AACAAGTAGCCAGA...........GGGCTGA  671
SUN.GA.98.L14               -GTGAGGG--CTTAC.CA-GCACTGTCAGGGGACA-GCC--A-AGTCCTAG----GGGGGAG-TC-CG.CTAGGTTA-TACT-TCAA-TA....-G-A-ACT----GTA-GTGG--T--AA-AAG-AGT........GTGTGTATAACTCTGTGTCTCGGAAGAGTGTCT 1057
SYK.KE.x.KE51               CTAGGGA--C--GAG.A----ACGCCAG--AAAGG-ACCGCAGAGGAA-G---GC-CGAA-AAGAG-A.-CGGT-TCCGCCCTATGCCGA....TCGGC-ACC-G-TTGAGTAGGTG-G--C-CCGGCG........TCAGACTGCACGCGGCGGTCGGGGATAAATAGG  426
SYK.KE.x.SYK173_COMGNM      GGACCCGAG--GAGG.T-AC--CAGTCAAAGG--GT-CGGATC-AA-AGGGATC-GCTGTC--TC-TA.---CGACC-GTGCT--T---T....CGTG-TGT-CACACGAGACG-GGT-AG-GATC-AG........CGCCAGACTGCCCACAGCAGTCGGGGAAATAGG  789
TAL.CM.00.266               -CCCAGGGGG--GGACG-AC-GAGGACT-GCA-G-AGGGGTCTGC...AGGAAGTTC...-T--.......-T-C.CTC-AGACAAGCA-C....CGG-A-AA--G..A........G-C-GC-C.-GG--GA-CACAG.................GGACCTGCGGGTGG  638
TAL.CM.01.8023              .............CA.G--G-GAAGG--G--T--G-GACGC-CAGGCAGGG-TCTGCGGG-AAGT--G.C-AC-CTTC-AGACAAGCA-C....CGG-A-AA--GTGA-AGCTC-G-GGAGGA-CT-AG........GGACCTGCGGGTGGTGCTAAGTGCGCACACGAG  143
TAN.UG.x.SIVagmTAN1         CACTGAAGGG--GGA.A--GCATCACCAGTCGAGTT-CCC----G---TAA-TC-GGACGCG-G-AGG.C-A-GAG-ATCC-CCAG-AAA....A-AGTACTGCC-G-GAGGCAGGGGTGG-GCAGG--........GATCTAGGTGGGCGACAGATCAGCA.CCTGGGT  896
VER.DE.x.AGM3               CGTCGGA-GC--AGG.AA-GCGCGGGGTG-GACGT-ACC..A-GAA--GC-TGGTG-GTA-GCTT-T-.-AGTGCC--GA-AAA--TCGA....-C--A--T--AGGA-TAGGA-GGGCCG-A-CCA--........ACTACTCTGGGCAAGTAGGGCAGGCGGACGGGT  425
VER.KE.x.9063               CGTCGGA-GC-TAGG.AA.CCGCGGGGTG-GACGT-ACC..A-GAA--GC-TGGTG-GTG-GCTT-T-.-AGTGCC-.AA--AA--TCGA....-C--A--T--AGGA-TAGGA-GGGCC.-A-CCA--........ACTACTCTGGGCAAGTAGGGCAGGC.GGACGGT  936
VER.KE.x.AGM155             CGTCGGA-GC--AGG.AA.CCGCGGGGTG-GACGT-ACC..--GAA--GC-CGGTG-GTA-GCTT-T-.-AGTGCC--GA-AAA--TCGA....-C--A--T--AGGA-TAGGA-GGGCCG-A-CCG--........ACTACTCTGGGCAAGTAGGGCAGGCGGACGGGT  925
VER.KE.x.TYO1_patent        CGTCGGA-GC--AGG.AA.GCGCGGGGTG-GACG--ACC-A--AGGA-AC-TGGTG-GTA-GCTT-T-.-AGTGCC--GA-AAA--TCGA....-C--A--T--AGGA-TAGGAG-GGCCG-A-CCG--........ACTACTCTGGGCAAGTAGGGCAGGC.GGTGGGT  933
WRC.CI.97.97CI14            CGTCC...AGA-.GG.TAAGAGGGAGGACT--TGGT-CCT-TTGAGAC--CTAGA-GGAATTAGAGG-.TAGCCTT--AAC-TA-A---TGAGTAGT-CA-A--GGACTACT...........................GGGAGACAGGAGAAAGGACCAGGGGGGAGGC. 1013
WRC.CI.98.98CI04            CAGCCTC-AGA-AGG.TAAG-AGGAGGGCT--TGGT-CCT-TTGAGAC--CTAGA-GGACTTAGAGG-.TAGCCTT-AGAC-CATA---TGAGTAGT-CA-A--G-GCTACT...........................GGGAGACAGGAGA.GTAACCAGGGGGGAAGC. 1019
WRC.GM.05.Pbt_05GM_X02      ...............................................................................................................................................................................    0
```

```
H1B.FR.83.HXB2               ACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGG 1084
Gag                           R  F  A  V  N  P  G  L  L  E  T  S  E  G  C  R  Q  I  L  G  Q  L  Q  P  S  L  Q  T  G  S  E  E  L  R  S  L  Y  N  T  V  A  T  L  Y  C  V  H  Q  R  I  E  I  K  D  T  K
H1O.CM.91.MVP5180            -A---A----TG---------T--A--------TG-------TAC-GAG---C-G--ACAG---T--G-G---G-T--CA-------G----G--C--G-A---TC-C-GG--CG--A-----GTA----GG--C--T--CA-C--AT-T--C--CCG---T--AC--C 1110
H1N.CM.95.YBF30              -A-------TTG---C--A--T--CA-G--C---G-G--C------GCCA-GT----AAAT--AT--G-----G-T--CA-------G----------GC-C--T--------CG-TC-----GTT--T--------C---AGT------C-----C-CA----AC---  642
H1P.CM.06.U14788             -A----T---TG---------G--TA-G------G-G--------CGA-A--T--T-A-AG---T--G-----G-T--CAGA-----C--Y--T-GC--GCAG--TC-T-GG--C--CC---T-GTTY-A-GG-----T--CA----A----------GT-----AC--C  570
CPZ.CD.06.BF1167             G--------TT-G-----C------A-G------G----T--T--C---A-G--CA-TATG--A--CTC------AG-AG-A-T---G---CC----A-A-TC--C----T---C--CA-T---GTGT-AC----CA-A---G--G-AGA-A-A--TC-G-----TG-A- 1099
CPZ.TZ.00.TAN1               G--T-----GA-G-----C--T--CA-G--G-ACGT---------CT-GA---CA-CCTC-----G--G--T--GG-AG-C-TT--T--TCC----A-C-TT--T--G-T------CA-CTGTGTA-----C--C--A--CGC-G-AGA-AGAG-CC----T--GG-A-  647
GOR.CM.07.SIVgorCP2135con    -A-G--T---TG--------AG--TA-G------G-Y--G------GA-A--T--T-A--G--A---G-----G-T--CA-A-----G-----T-GCT-AC----TC-G-GG--C--TT-G-T-GTGT-G-GG-----C--C-GG--A--T--TG--TC---T--TC--C  644

MAC.US.x.239                 TA----T-G-T-AGCAGAAA-------G--G-ACAA------A---CA-A-------TTCGGTCT--GCT----TAG-G-CA-----C------A-TT-A-A-AGCC-T--------T--CTGCGT-A---GG--CA-T--CGC-GAAGAGA-AG-G---C----TG--- 1347
Gag                           R  F  G  L  A  E  S  L  L  E  N  K  E  G  C  Q  K  I  L  S  V  L  A  P  L  V  P  T  G  S  E  N  L  K  S  L  Y  N  T  V  C  V  I  W  C  I  H  A  E  E  K  V  K  H  T  E
H2A.SN.86.ST_JSP4_27         CA------G-T-GGCAGAGA-------G--GT--AA---G--T--CCA-A----T--TAC-GTTT--G-T----TAG-A-C------G-----A-TT-A-A-AGCC-T-T------T--CTGCGT-A-T-GG---A-A--CGC-GAAGAGA-AGCG-----T--TG-A-  842
H2B.CI.x.EHO                 GA----T-G-T-GGCAGAGA---G---G-G-T--AA------A--CC-GA----T-G-AA-GTCT--GG-----TAG-A-CA-----T------A-TT-A-A-AGTC-T--------C--CTGCGT-A---T---CC----CGC-GAAGAGA-AG-G-----T--TG--- 1374
H2G.CI.92.Abt96              CA------G-T-AGCAGAGA-------G---T--AA------T--CCA-A----T--TTC-GTAT--GCT----TAG-A-CA-----T------A-TT-A-AGAGCC-T--------T--CTGCGT-A-T-GG---T----CGCCGAACAGA-AG-----C-T--TG-A-  760
H2U.CI.07.07IC_TNP3          TA----T-G-T-AGCAGAGA-------G---T--AA------T---CA-A----T--TTC-GTTT--GCT----TAG-G-C------G------A-TT-A-AGAGCC-T--------TACCTGCGT-A-T-GG--CT----CGC-GAAGAGA-AG-G---C-T--TG---  865
H2U.FR.96.12034              CA----T-G-T-AGCAGAAA-------G--GG-CAA------A---CA-A----TT-ATCGGT-T--GCT----TAG---CA-----T------A-TT-A-AGAGCC-T--------T--CTGCGT-A-T-GG---T----CGC-GAAGAGA-AG-G---C-T--TG---  876
ASC.UG.10.RT03               ------T-G-C-ATCAGACTC-T-A--G----ATAAG--G-----CCA-A-G--CT-A-AGGTCA-TGTT----TGGAAG-A-----G--C--GTCT--A-A---CC-C-T-GGATT---TT--GTTT-G-GG-----A---GC-GATG-TACAG--G-------AG---  433
ASC.UG.10.RT08               ------T-G-C-ATCAGACTC-T-A--G----GTAAG--G-----CCA-A-G--CT-A-AGGTCA-TATC----TGGAAG-A-----G--C--GTCT--A-A---CC-C-T-GGATT---TT--GT-T-A-GG-----A---GC-GATG-TACAG--G-G-----AG---  448
ASC.UG.10.RT11               ------T-G-C-ATCAGACTC-T-A--G----ATAAG--G-----CCA-A-G--CT-A-AGGTCA-TGTT----TGGAAG-C-----G--C--GTCT--A-A---CC-C-T-GGATT---TT-GGTTT-G-GG-----A---GC-GATG-TACAG--G-G-----AG---  433
COL.CM.x.CGU1                C--C-GT-TTT-A-TATT--AG------A-GT--G--AC---AGTGGCT------T---A-A--G-GACG----TAG-GG-C------------TGT-G--G--CC-----GGA...TGTT-TGTGTGT--C--CC----CAG--AATGGA-T--TG-G-----TC-A-  895
COL.UG.10.BWC01              C--GC--TGCC-C--CATA-AGT-----A--T----TC-G--AATAGC-GT----T-A-A-A--G-GAGG--C-TGAAAG-AG---G--T----TG--A-A---CC-C-GGGGA-TGC--TGTGTGTG------CT-----AG---TTGG--C--C--------AC-A-  673
COL.UG.10.BWC07              GA----GCA-C-A--CATC-AGT--C--A--T--G--C----AATAG--A-G--T--T-AG--AG-G--G--TCTGG-AGTT-----T-----G-TG--G-A-AGCC-T-GGGGA-TGC-CTGCGTATGT--C--CT-----AG-G-ATGG--CG--G-GC-T--TC-A-  907
DEB.CM.04.SIVdeb04CMPF3061   TA-G----GCC-CTCAGAG-CAT---------ACAAG--G--A---CAGG-G--T--AACCAG-T--AC---T-TGGAAGCA-----CAGT--GA-T--C-AG--TC-T---GGG---TGTTGTGT-A-A-GGGCCTGC---GC-G-TC-CC-A---G-G-----AG-A-  930
DEB.CM.99.CM40               CA----T-GCC--GCAG-A-CAT-A-------ATAAG----------CAGG----TT--ACTAG-T----G--T-TAGAAGCA-----CAGT--GA-TT-G-A---TC-----GGC-T---CTGTGT-A---GGGCATGC---GCC-ACC-CTC-G-GG-T------G---  782
DEB.CM.99.CM5                CA----T-GCC--GCAGGA-CAT-A--G-----CAA------A---CA-G-G--CT-AACTAGAT-------T-TAGAGGCAT----CAGT----GC----AG--CC----CGGC-T----TGTGT-G-T-GGGCATGC--CGCC-ACC-CTC-G--G----T---G---  785
DEN.CD.x.CD1_CMO580407       G-----T-GCT-AGGAGAACAG---C-------C-AT-----A--CCA-ACT--CA-A-AGACCT--GCT----TAGAA-CACA---CAGT--GAGT--G-A-AGCC---TC-GA---TGCTGCGTTG---GG---TGC--CAGGG-AT-G--AG-GGC---T--AG---  776
DRL.DE.11.D3                 T-------GCC-AC--GAGA-AT-A-------GTCAG-----G--CCA-A----TT-A---GTA--CTTT--TCTAG---CT-----G-----GA-T--A-TT--TC-C-T------CTGCTGCGTAG-T-GG--C--T--CTCG-AAG-GA-AG---C------AG---  492
DRL.x.x.FAO                  T-------GCC-GC--GAGA-AT-A-------GTCAG-----G--CCA-A----TT-A---GTA--CTTT--TCTAG---CT-----G-----GA-TT-A-TT--TC-C-T------CTGCTGCGTAG-T-GG--C--T--CTCG-AAG-GA-AG---C------AG---  585
GRV.ET.x.GRI_677_gri_1       ------T-GGT-AC--GAGAAA--T---------AA---------CA-A----CA-A-A-GTTT--ACC--G-TGGAA-C-------C--C--G-GG--A-A-G-TC-G-T----TTGTGCTGCGT-A-T-GG--CA-T--CGC-GAACAGA-AG-G--------AG--- 1191
GSN.CM.99.CN166              GA---AT--GC-GTCCGAC-CT--TC-T----ATCA---G--G----C----GG---T-A-A-ATTG-CTTC--CCTAGAA-CA-----G-------CT--C-A---TC-C-T-GGA-TC--GT--GTG---------A-C---GC-GAC---C--G-GG-G--T--AG--C  746
GSN.CM.99.CN71               GA---A---GC-ATCAGGCA-T--TC-T---T-TAA------G---CA---GG--T----A-GTCACCTCC----TAGAA-CA-AT--G--C---TCT--C-A-AGCC-C---GGA-TC--GT--GTG---------A-C---GC-GAC---TG-G--G-----T--AG--C  752
LST.CD.88.SIVlhoest447       TA-G--T-G-T-AGGAG-ACA-T--C----G---G-----T-----C-AGA----T--T---GTCTGTTGG---CTATACTCC-----C--GA-GA-TT-A-A-G--C-TGTAGGA--T---TGTGTG--G-G---CTGC---TT-G-----A-AG--GC-------C-A-  295
LST.CD.88.SIVlhoest485       TA-G--T-GTT-AGGAG-ACA-T--C-G--G---G-----T-----C-AGAG---T--T---GTCTGTTGG---CTATACTCC-----C---A-GA-TT-A-A-G--C-TGTAGGA--T---TGTGTG--G-G---CTGC---TT-G-----A-A---GC-A-----GCA-  295
LST.CD.88.SIVlhoest524       TA-G--T-GGC-AGGAG-AAA---TC---------G----G-------AGA-G--TT-----GTC-GTTGG--T-TATA--CT-----C--TA--A-T--G-A-G-TC--GTAGGG-----GTGTGT-A---G---CTGT--C-T-G-A---AGAG-C--T--T--AC---  295
LST.KE.x.lho7                CA----T-G-T-GGGAT-ACAA--C--G-------G-------G----AGA---C--ATCTGTCTGCTGG--T-TATA-GCA--T--T--TA-GA-C--A-AGG-C---GTAGGG------TGTGT-A---GC--CTGT--C-T-G-AG--AGA---TC------TC--- 1370
MAC.US.x.251_1A11            TA----T-G-T-AGCAGAAA-------G--G-ACAA------A---CA-A-------TTCGGTCT--GCT----TAG-G-CA-----C------A-TT-A-A-AGCC-T--------T--CTGCGT-A---GG--CA-T--CGC-GAAGAGA-AG-G---C----TG--- 1346
MND-1.GA.x.MNDGB1            TA----T-GCT-ATCGGA-AAA--CC-T------CAGC----A---GA-A----C--CTC-GTATGTTGG----TATA-G-CCA---------TA-T--A-A-G-T--GGTAGGG-----CTGTGTTG-AGCC--CA-A--CGC-G-T-----A--T--GAG---AC-A-  744
MND-2.CM.98.CM16             CA----T-GCC--C--GAAAAG----------GTAAG-----A--CGAGA----T--TA-CGTA--CTTT---CTAG---CT-----G------A-TT-A-TC--GC-G--C--C--CTGCTGCTG-G-A-GG--C--A---GCG-AAGAGA-AG---C---T--AG--- 1031
MND-2.GA.x.M14               T-----T-GCC--C--GAAAAA--A--------TCAG-----A--CGAGA----T--TA-CGTA--CTTT---CTAG---CT-----G-----GA-T---CTTG-GC-G--C--C--CTGTTGTTG-A---GG--C--A---GCG-AAG-GA-AG---CT--T--AG-A-  955
MND-2.x.x.5440               CA----T-GCC-AC--GAGAAGT---------GTCAG-----A--CGA-A----T--TA-CGTA--CTTT---CTAG---CT-----G------A-TT-A-TT--GC-G--C--C--CTGTTGTTG-A-T-GG--C--A---GCG-AAG-GA-AG---C---T--TG---  588
MNE.US.x.MNE027              TA----T-G-T-AGCAGACA-------G--G-ACAA------A---CA-A-------TTCGGCCT--GCT----TAG-G-CA-----C------A-TT-A-AGAGCC-T--------T--CTGCGT-A---GG--CA-T--CGC-GAAGAGA-AG-G---C----TG---  827
MON.CM.99.L1_99CML1          TA------GTC-CTCGGACTCG--CC-G--G---CAG--C-------AGA-G--TT---A-GTTA-T-TT--TCTA-AG-C---------C--GTCTA-C-A---TC-C-T-GG--TC-CTT-GGTAT-G-----CA-A---GCCG-T-----A--TG-G-----AG-AC  739
MON.NG.x.NG1                 ..................................................................................................................................................................    0
MUS-1.CM.01.CM1239           -A-G--T-G-C-CTCGGACTC---T--G--G---AA---G--A---CA-A-G--CT-A-A-GTT--CAT-----TAGAA-CA-----G--C---TCT--C-A----C-C-T-GGA-TCACGT--GTG----GG---A-C---GC-GA-G-G---G-GG-------AG--C  746
MUS-1.CM.01.SIVmus_01CM1085  TA----T-G-C-CTC-GA-TC---T--G---T--AGG--T--G---CA-A-G--CA-A-AGGTCA-C-TC--C-TAGAA-CAG----G--C--GTCC--C-A----C-C-T-GGA-TC--GT--GTG----GG---A-C--CGC-GA-G-GC-TG-GG-------AG---  744
MUS-2.CM.01.CM1246           G--G----GTC-CTCAGA-ACG--CC-----G-GCAAG--C--A--CCA--GC--C--A-A-GTAA-CTGC--C-TGGAAG-A--T--GAGC--G-CG--C-A----C-C-T-GGA-TC--GT--GTA----GG---A-C---GC-GACG--ACAG-G--------AG---  815
MUS-2.CM.01.CM2500           CA------G-C-CTCAGACTC---T--------TAAG-----G--CCAG-GC--C--A-A-GT-A---T------TGGAG-C---C--G--C--GTCT--C-A----C-C-T-GGA-TC--GT--GTA---------A-C--CGC-GA-G-GAGCG-GG-------AG---  774
MUS-3.GA.09.09GabOI81        TA-G--T-G-C-ATCAGAC-CT--CC----G-GCAA------A---CA-AG---C--T-AGGTAA-CAT---TCTAGAATCA-----G--C--GA-C----A----C-C-T-GGA-TC--GT--GTG----GG-----T---GCTGATG-GACAG--G-------AG---  298
MUS-3.GA.11.11GabPts02       GA----T-GGC--TCAGA-TC---A--G-----CAAG-----G--CCA-A----CT-A-AGGT-A--AT---TCTTGAA-CC--T--G--T----CT--C-A----C-C-T-GGA-TC--GT--GTG----GG---A-T---GCTGAAG---CAG-GG-------AG---  298
OLC.CI.97.97CI12             --ATC-ATGCA-AG----A-ATGATC---G---TAAG----------A-AG---TG-A-A-A-AT---TG--T-TGGAAG--GG----A-TC-GA-GT-G-A-AGT--GCTG--C-TCTGTAT-GTTA-T-TG---T----AG--G-CC-TCCAG---G-A------TA- 1000
RCM.CM.00.SIVagi_00CM312     GA----T-GCT-A-G-GA-ACAT-------TT--G---GG--A---GA-A----CT-A---GTC--TTTG--TCTGG---C------G-----G-GG--G-A-AGCC-C-T----TTGTGCTGCG-A--T-GG--C--T--CA-GGAAG-GA-AG-G--------AG-A-  295
RCM.CM.02.SIVrcm_02CM8081    GA----T-GCC-A-G-GA-ACGT-------TT--G---T---A---CAGA----TT-A---GTC--GTTG--TCTAG---C------G-----G-GG--A-A-AGCC-C-T----TT-TGCTGCG-A--T-GG--C--A---TCGGAAG-GA-AG-G-G------AG---  295
RCM.GA.x.SIVRCMGAB1          TAA---T-GCT-G-G-GA-CATT-----------G--AC---A---GA-A-----T-A---GTC--G-TG--TCTAG---C------G-----G-GG--A-A-AGCC-C-T----TTGTGCTGCGTA----GG--C--A--CA-GGAAG-GA-AG-G--------AG---  506
RCM.NG.x.NG411               GA----T-GCT-G-G-GA-ACTT---------GTG-------AATCCAGA----TA-A--GGTTA-T-TG--TCTAG---C------G--T--G-GCT-A-A---CC-C-T----TTGTGCTGCGTA--T-GG--C-----CA-GGAAG-GA-AG-G-----T--AG---  506
SAB.SN.x.SAB1                T-----TAGCC--TCAG-AAAT--C-----G---AAG-----GGTCGTCA----T--TA-TGTA--C-TT--CCTAG---CT-----G-----GA-TT-A-TTG-TC-C-T----TT-TGCTGCGTGT-AGC----A-T--CGC-GAA---A-AG-G-----T--AG--- 1311
SMM.CI.79.SIVsmCI2           TA----T-G-T-AGCAGA-A-------G--GT--AA------A---CACA----TT--TC-GTTT--GCT----TAG-A-C------T------A-TT-A-A-AGCC-T-T------T--GTGCGT-A-T-GG--CA-A---GC-GAAGAGA-AG-G---C-----G---  443
SMM.LR.89.SIVsmLIB1          CA----T-G-T-AGCAGA-A-------G--GGATAA------T---CA-A----T--TTCGGTCT--GTT----TAG---CA-----T------A-TT-A-AGAGCC-T--C-----T--CTGCGT-A-T-GG--CA-A--CGC-GAAGAGA-AG-G---C-T--AG---  432
SMM.SL.92.SIVsmSL92A         CA------G-T-AGCAGAAA-------G---T--AA------T---CA-AG---TT--TC-GT-T--GTT----TAG-C-CA-----C------A-TT-A-AGAGTC-T--------T--CTGCGT-A-T-GG--CA-A--CGC-GAA--GA-AG-G---A-T--TG---  440
SMM.SL.92.SIVsmSL92B         CA----T-G-T-GGCAGAAA-------G----GCAA------A---CA-AG----T-AAC-GTAT--GCC----TAA-G-CA-----C--G---A-TT-A-A-AGC--G-T-------T--CTGCGTAG-T-GG---C----CGC-GAA--GA-AG-G--------AG---  416
SMM.SL.92.SL92B              CA----T-G-TCGGCAGAAA-------G----GCAA------A---CA-AG----T-A-C-GTA---GCT----TAA-G-CA-----C--G---A-TT-A-A-AGC--G-T--G---T--CTGCGTAG-T-GG---C----CGC-GAA--GA-AG-G--------AG---  787
SMM.US.04.G078               TA----T-G-T-AGCAGAAA-------G--T-ACAA------A--CCA-A----T--CTCGGTCT--GCC----TAG-A-C------T------A-TT-A-AGAGC--G--------T--CTGCGTTA-T-GG--CA-T--CGC-GAAGAGA-AG-G---C-T--TG---  613
SMM.US.05.D215               TA----T-G-T-AGCAGAGA-------G----ATAA------T---CA-AG---TA-TTCGGTCT--GCT----TAG-A-C------T------A-TT-A-AGAGCC-T--C-----T--CTGCGT-A-T-GG--CA-T--CGC-GAAGAGA-AG-G---C-T--TG---  607
SMM.US.06.FTq                TA----T-G-T-GGCAGAAA-------G--G-ACAA------T---CA-A----T--CTCGGTCT--GCT----TAG-A-CA-----T------A-TT-A-AGAGCC-T--------T--CTGCGT-A-T-GG--CA-C--CGC-GAACAGA-AG-G---C-T--AG---  620
SMM.US.86.CFU212             CA------G-T-GGCAGAGA-------G----RCAA------T---CA-A-R--T--TTC-GTTT--GCT----TAACA-CA-----G------A-TT-A-AGAGCC-C--C-----T--CTGCGT-A-Y-GG--CA-T--CGC-GAAGAGA-AG-G---C-T--YG---  599
SMM.US.x.H9                  CA----T-G-T-AGCAGAAA-------G--G-ACAA------T---CA-A----T--TTC-GTTT--GCT----TAG---C------T------A-TT-A-AGAGCC-T-----C--T--CCGCGT---T-GG--CA-T--CGC-GAAGAGA-AG-G---C-T--TG-A-  829
SMM.US.x.PGM53               CA----T-G-T-AGCAGAAA-------G--G--CAA------T---CA-A----TA-TTC-GTTT--GCT----TAG---C-G----T------A-TT-A-AGAGCC-T--------T--CTGCGT-A-T-GG--CA-T--CGC-GAAGAGA-AG-G---C-T--TG--- 1273
SMM.US.x.pE660.CG7G          CA----T-G-T-AGCAGAAA-------G--T-ACAA------T---CA-A----T--TTCGGTTT--GCT----TAG---C------T------A-TT-A-AGAGCC-T--------T--CTGCGT-A-T-GG--CA-T--CGC-GAAGAGA-AG-G---C-T--TG--- 1345
STM.US.89.STM_37_16          CA----T-G-T-GGCAGAAA-------G--GT--AA------G---CA-A----TA-AACGGTCT--G------TAG-G-CA-----T------A-TT-A-AGAGTC-T-T------T--CTGCGT-A-T-GG---A-C--CGC-GAAGAGA-AG-G---C-T--TG--- 1003
SUN.GA.98.L14                TA-G--T-G-C----CAAG-AA------AG----GTG---------CAGA-G--T---TC-GTCTGTTGG--T-TGTA-GCAT-T--T--TA-GA-C--G-AGG-C---GTAGG-------TGTGTG--TGCC---TGT------G-C--TCCAG----------AC--- 1357
SYK.KE.x.KE51                CA------GCC-CTCAGACCAG--TC-T----GTAA------T---GA------CA-AAA-ACCA-CTT-----TAGAAA-ACAT--C------A-TT-A-A--------T-GGC-T-AC----GTAG---GGGCC--C--CGC-GAA-A---CG-GG----T--GG-AC  729
SYK.KE.x.SYK173_COMGNM       CA------GCC--TCGGACCAA--CA-G--GT-TAA----------GA-A----T--CAC-GTC----TT----TAGAAGCT-AT--G------A-TT-A-A-AGCC-T-T-GGC-T-A-TT-GGT-G-A-GGGCC--T---GCC-AA-AG--AG--G-------TG--C 1092
TAL.CM.00.266                GA-G--T-GCC-G-G-GAC-C---------G--CAA------G--CGTTA-G--CA-A-A-GCCA-TT-CATG-TAGAA-C--AT--C-------GGA-A-AG--TC-C-T-GGA-TC--GTGTGTT--A--C--CA-T---GCTG-T-----C--TG-G--T--AG--C  960
TAL.CM.01.8023               -A-G--T-GCC-G-G-GAC-C---T--G--G--CAA------G--CGTTA-G--CA-A-A-GCCA-TTCT--G-TAGAA-CA-AT--C-------G-A---AG--TC-C-T-GGA-TC--GTGTGTT--A--C--CA-T---GC-G-T-----A---G-G--T--AG-AC  446
TAN.UG.x.SIVagmTAN1          G--C--T-GGC-CC-CGAGAAA--C----------AA---G------CA-A-G--CA-A-A-GTAT-GAGC--GCTAGAA-CA------------G---A-A-AGTC-G-T----TTGTGCTGCGT-A-T-GG---A-T--CGC-GAACAGA-AG-G-----T--AG--- 1196
VER.DE.x.AGM3                G--C----GCC-CC--GAGA-AT-AC-----T--GA------A----AGA-G--CA-A-A-GTA--CT-C--GCTAGAA-CA-----G--G--G-GCT-A-A-AGTC-G-T---CCTT--GTGCGTAT-G-T---C--A--CA--GAT-AG--AG-G--------AG-A-  725
VER.KE.x.9063                G--C----GCC-CC--GA-A-GT-----------GA---T--A----A-A-G--CA-A-A-GTC--TTTT--T-TAGAA-CA-----C--G--G-GCT-A-A-AGTC-G-TC---CTT--GTGCGTG--A-----CA-C--C-GCGAACAGA-AG-G--------AG--- 1236
VER.KE.x.AGM155              C--C----GCC-CC--GAGAAGT-A--G--G---GAG-----T----A-A-G--CA-A-A-GTT--CTCT--CCTAGAA-CA-----G--G----G-A-G-A-AGTC-G------CTG--GTGCGTAT-GCT---C--C--C---GAA-AGA-AG-G--------AG--- 1225
VER.KE.x.TYO1_patent         G--C----GCC-CC--GAGA-GT----G--G---GAG--G--G----A-AG---CA-A-A-GTC--CT-C--CCTAGAA-CA--------G--G-GCT-A-A-AGTC-G-TC---CTT--GTGCGTA--A-----CT----CA-GGAACAGA-AG-G--------AG--- 1233
WRC.CI.97.97CI14             GAAG--GTGCT-AG--T-A-CAT-AA-CAGT---G-------GGTG-AGA-GC-TA--TC---T---GCT--T-TA--G-CTGAT-CC--TA-G---G-A-AGAGCA-T-GGGGA-TGT--TGTGTGT-AGCA---C-C---G-G--A--TCCATGT-G-A-T---C-A- 1342
WRC.CI.98.98CI04             GAA---GTGTT-AG-CT--ACAT-AA---C----GA-AC---GAC--A-A--C-GA---A---TT--GC-----TAG---CAGAT------A-----G-G-A-AGCC-T-GGGGA-TGA-GTGTGTG--TGCA---T-----GC-C-T--TCCCTGC-CCA-T---C--- 1348
WRC.GM.05.Pbt_05GM_X02       GAAG--ATGTC-AG--GGGAAAT--A--TCT-GCCA------AGTG-AGA--C-CTGT-AG--TA--GCT--T-TGG-A-CAGA---G---A--AC-A-A---AGTA-T-GGGGA-TGA--AGTGTG--AGCA--CT-AT--A-T-AC--TC-AG---GT--T--TC-A-  298
```

Draft of 30 September 2015, not for public release
plvdna_main.tex r8876–10546

Draft of 30 September 2015, not for public release
plvdna_main.tex r8876–10546

```
                            Gag p17 Matrix end Gag p24 Capsid start
H1B.FR.83.HXB2              CAGGTCAGCCAAAATTACCCTATAGTGCAGAACATCCAGGGGCAAATGGTACATCAGGCCATATCACCTAGAACTTTAAATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTGATACCCATGTTTTCAGCATTATCAGAAGGAGCCACCCCACAAGA 1337
Gag                          Q  V  S  Q  N  Y  P  I  V  Q  N  I  Q  G  Q  M  V  H  Q  A  I  S  P  R  T  L  N  A  W  V  K  V  V  E  E  K  A  F  S  P  E  V  I  P  M  F  S  A  L  S  E  G  A  T  P  Q  D
H1O.CM.91.MVP5180           ---ACA--T-----------------AACA--TGCA----A-------------A----C--C--C--G-------------------G-C--------------C--T-A---T---A-T--T--T-----ATG-------------G--TGT---CT-T-- 1357
H1N.CM.95.YBF30             A-TA----TAGG-----T---C----C----CTGCT--A--A--------------C-GC-GA---C-----C-------T-----G-----GA---G-------C--T--T--------A-----A------ATG--C--G--------G--A--G--CTC--- 901
H1P.CM.06.U14788            GGAACA--T-----C--T-----C--------TGC-------------ACC---AT-C-GC-G--C--C--G--GC-G-----C-----G--G-C---------A--A--C----A---T---A-T--C--T------ATG--C--G--------G--T-TA--TG-T-- 802
CPZ.CD.06.BF1167            ---ACAG-GAG---C--------T--TTTTG-TGCTGC-CAA--GCCAAGG------C-A--T-----C------C-T-----T---------AC------G-----AAA---T-AG--------A--C--A-----C-----C---G----G-------T------T-- 1415
CPZ.TZ.00.TAN1              AGT-GG--TAG-CTA-----AG-GA-AACTG-TGCA-----AGTTGCAAGG------C-T--T-----------C-------C-------GG--A-------A--A-GG--T-AT--------A-----A-----C--------G--T--G-----A-----TT-T-- 969
GOR.CM.07.SIVgorCP2135con   G--ACA--T--G--C--T--AG-G--AGCA--TGCT--A-----G-----C------C-GC-G--T--A--G---C-------C-----G----C---------A-----C----AT------A-C--C--T-----CATG-------------GT-T-T---TT-T-- 870
                            Gag p17 Matrix end Gag p24 Capsid start

MAC.US.x.239                GGCAGAG-AGG---------AG--CAA--A-TAGGT...--TA-CTAT--C--C-T-C-AT--AGC--G-----A-------C--------T-GA---G--A---AAA--TG-AG-------AG-G--AGGA---CAG---C-G--------TTG------CT-T-- 1606
Gag                           G  R  G  G  N  Y  P  V  Q  Q  I  G  .  G  N  Y  V  H  L  P  L  S  P  R  T  L  N  A  W  V  K  L  I  E  E  K  K  F  G  A  E  V  V  P  G  F  Q  A  L  S  E  G  C  T  P  Y  D
H2A.SN.86.ST_JSP4_27        GG-AA-G-AGG---C-T---CG--CAA--AGTGGC-...--CA-CTATACC---GT-C-AC--AGT--CC----CC-------T---------C-------G--A---AAG---G-GG-------AG-G--AGGA---CAG---C-C--------CTG---G--CT-T-- 1101
H2B.CI.x.EHO                -GCC-A......GC---T--AG-GCA---A-TAGCT...--CA-TTATTCC--C-T-C-GC--AGC--CC----C-------C--------T--------------AAG---G-GG-------AG----AGGA---CAG---C----------GTG------TT-T-- 1618
H2G.CI.92.Abt96             GGCAGAG-AGG---C-----AG-GCA---AGTAGG-...--CA-CTAT--T----T-C-AT--AGT---C-----------------------T-G--------A--AAG---G-GG--------G-G--AGGA---CAG--GC-G--------CTG-CTA---T-T-- 1019
H2U.CI.07.07IC_TNP3         GGCAGA--AGG---------AG--CA---AGTGGGA...--CA-TTAT-------T-C-TC-GAGT---C-G---C-C-----T-----G---T----G-----A---AAA--TG-AG-------AG-G--AGGA---CAG--------------CTG---A--CT-C-- 1124
H2U.FR.96.12034             GGC-G---AGG---------GG-GCA-----TAGGT...--CA-TTAT-CC--C-TAC-TC--AGT--------AC-C-----C---------C-G-----G-----AAAA--TG-AG-------AG-G--AGGA---CAG---C-C-----G--CTG---G--CT-T-- 1135
ASC.UG.10.RT03              GCCTCAG-A-----C--T--A--T--TAG---TGCA-----A--GCATCAG--------TT-G--C--A---TCC-C--GA-----A---C-ACCA-T--G-----AAAA--TTCAG-------AG-TG--C-A--CCAG---C-G--------CTGT-TA--CT-T-- 674
ASC.UG.10.RT08              GCCTCAG-G--------T--AG-T--TAG---TGCA-----A--GCATCAG--------TT-G--C--A---TCC-C--GA-----A---C-ACCA-T--G-----AAAA--TTCAG----G--AG-TG--T-A--CCAG---C----------CTGT-TA--CT-T-- 689
ASC.UG.10.RT11              GCCTCAG-G--------T--A--T--TAG---TGCA-----A--GCATCAG--------TT-G--C--A---TCC-C--GA-----A---C-ACCA-T--G-----AAAA--TTCAG-------AG-TG-TT-----CAG---C----------CTGT-TA--CT-T-- 674
COL.CM.x.CGU1               ......TTGG-C-TG...--A-----AACAGGGCCT...CAAGGGCCA--G...--TCAACCCCTGAGCCCC-GGACGTTA-G-GCCTGGGTTAAGTGT-T----GGA-GAA-TGCT--TTCCT--GCT--------C-T---T-AT--TACG--G--A-TAG-TT-T-- 1145
COL.UG.10.BWC01             ......TTGG-T-TG...--A-----AA-T-C-CCA...CAAGG-CCT-CC...--TCAACCT-T-AG-CC--GGACTTTA---GCATGGGT-AAGTGT-TG---GGA-GAA-AGCA--CTC-C-TGCC----------TG--T-AT---AC---G--T-TAG-TT-T-- 920
COL.UG.10.BWC07             .........G---TG...--C-----AACA...GGG...CCA---GG-CCTATG--TAG-CCCCT-AG-CC--GGACTTTG--GGCATGGGTGAAGTGT-TT---C-A-GGA---CTG-CTCCC-TGCT--A------CT-----AC--CACG-----T-TAG--T-T-- 1163
DEB.CM.04.SIVdeb04CMPF3061  GTA--A--GGGC--C-----A--CC-CAGAC-AGGG...--A--GTAT--G---AT-C-TC--AGT--A--G-TAG----AA-T---A----T-C------G-----AAAG---TC------GA-AG-C--TC-T--CCA--TTC--G----G--CTGT------T-T-- 1267
DEB.CM.99.CM40              ACACCATCAGTT--C--------CT-AAGAC-GGGG...CAA--GT-T--G---AT-C-AT--AGC--C--G-TAG----AA-C---A----T-CT-----G--A---AAA--TTCT-----GA-AG-T--TC-T---CA--TC--GG----G--CTGT-----CT-T-- 1107
DEB.CM.99.CM5               A-CCAACCAGCT----TT--A--CC--AGAC-AGGA...CA----TAT------AT-C-TT--AGT-----G-T-G-G--AA-C---A----T-C---G--G--A---AAA--TTC------GA-AG-G---C-A---CA--T---GG-------GTGT--T--CT-T-- 1098
DEN.CD.x.CD1_CMO580407      -CAAATTCT--------T--A--T--CAG--CAGCT-----A---TATCAG-----A---T--AGTA-C--GGTG-----AA-----A-T----AG-----G--C-GAAAA--TTCG-----GA-AG-G--AC-A------G-TCC-CG-C------GCTG-TA--TT-T-- 1104
DRL.DE.11.D3                GGCAGGTCAAG---------C--TCA-GTTGTT-AT...CAA--GGCA-----C-----T--T-----A--------G-----C-----G--G---A-T--G--A---AAG-----TG--------G-G--T-------TG--CC-G--------CTG--T-AGCT-T-- 754
DRL.x.x.FAO                 GGCAGGTCAAG---C-----C--TCA-GTTGTT-AT...CAA--GGCA-----C-----T--T-----A-----C-------C-----G--G---A-T--G--A---AAG-----TG--------G-G--T-----C-TG--CC-G--------CTG--T-AGCT-T-- 847
GRV.ET.x.GRI_677_gri_1      GGC-AATCAAG---------AG----AA-TC-G-AT...AATGCCTG------C---C-TT-G--T--GC-C--G--------G-----C---TGC--G--G--A--AAGG-GGG-AG-------AG-C--------CCA----C-C-----G---TGTCT-T-CT-T-- 1450
GSN.CM.99.CN166             GTA--T--T-----C------G-GA-TAG---TGC---A--A--GTATCA----------C-CAATAG-----T-C-T--GA-T-------GCA----G--G-----AAAG--TGCT-----GACAG-TG--T-----CAG--T-----------CTGT-TA--CT-T-- 1014
GSN.CM.99.CN71              GTA--T--T--G--C-----C-----TAGA--TGCT-----A--GT-TCA---C-------C-CAACAGC---GTCC-T--GA-T-------GCA----G-----A--AAG--TGCT-----GACAG-CG-TC-----CAG--TC--A-------CTGT-TA--TT-T-- 1020
LST.CD.88.SIVlhoest447      -CTAAA--TGGC-----T--AC-C--CAGAG-A-AT...CA-AG-TG-----CACAC-AT-G-----------AA--C-AA-T---------A---------T-GA-GA-GG-AG------ACAG-GG-A--------T--CC-CA-----AA---TTTG---G-T-- 509
LST.CD.88.SIVlhoest485      ---AAA--TGGC-----T--AC-CA-CAGAG-A-AT...CA-AG-TG-----CAC-C-AT-G-----------AA--C-AA-T---------A---------T-GA-GA-GG-AA------ACAG-GG-A--------T--CC-CA-----AA---TTTG---G-T-- 509
LST.CD.88.SIVlhoest524      -CA-AA--TA------TT---T-GA-TAGAG-A-AT...CAAAG-TG---C---ACCC-TC-------C-----AA-TC-AA------G---A---------T-G--GA-GG-AG------ACAG-GG-T-----------TC-GA-----AA----TTG---G---- 503
LST.KE.x.lho7               ---ACAG-GGG-----T--CC-GA-AAGAG-A-AT...CA-AG-TG------ACAC-TT----G--C-----AA--C-AA------G---A---G-----C-GA-G--GG-AA------ACAG--G-A------------C---A-----AA---TTTG--TG-T-- 1593
MAC.US.x.251_1A11           GGCAGAG-AGG--------AG--CAA--A-TAGGT...--TA-CTAT--C--C-T-C-AT--AGC--G-----A-------C--------T-GA---G--A---AAA--TG-AG-------AG-G--AGGA---CAG---C-G--------TTG------CT-T-- 1605
MND-1.GA.x.MNDGB1           ......--TA-G----T----G--CAAAG-G-TGCAGCA--A--GTATCAGT--ACTC-A---AGT-----G-T-A--C-GA-----------AC---G-----A---AAG-GG-AA--G--G--C--C--TC-A--C--------GA----------A-T-AGT--T-- 964
MND-2.CM.98.CM16            GGCCG-TCAA----------A---CA-ATA-TA-AT...CA-ACCCCA--C-------GA--T-----GC-C--GC-G-----G----------TGCA------------AAG--T---------A-AG-G--------CAT---C--G--T-----GTG-CT---CT-T-- 1293
MND-2.GA.x.M14              GGCCG-TCAA-------T-----TCA-ATA-TA-AT...CAAAC-CCA--G-----A-GA--T-----A--------------C-----C---TGTA-T--G--A--AAA-----T------A-AG----T-----CAT---CC----------TG-CT----T-T-- 1217
MND-2.x.x.5440              GTTCG-TCAA----------C--TCA-GTA-TA-AT...CAAAC-CCA---------GA--T-----A----C--G--------------TGTA---G----AAAG-----------A-AG-G--T-----ATT--C--G----------TG--T---GT-T-- 850
MNE.US.x.MNE027             GGCAAA--AGG---------AG--CAA--A-TAGGT...--TA-CTATACC--C-TAC-AT--AGC--G-----A-------C--------T-GA---G-----AAA--TG-AG-------AG-G--AGGA---CAG---C-G--------CTG--T---CT-T-- 1086
MON.CM.99.L1_99CML1         GTACC---TGG---C------G-C---AGA-CACAGGG---AGGCT-CCA---------AG-GGAG---C-GCTCC-C--GA-------C-G--CA-T--G------AAA--TGC----------G-GG--C-C--CCAG--TC----------CATG-TT--CT-T-- 1001
MON.NG.x.NG1                ........................................................................................................................................................................ 0
MUS-1.CM.01.CM1239          --AACT--TGCT--------C--CA-AAG----CCA-----A---T-TCA---C---T--C-TAATG------TC-----GA-------GTC-A----G----A---AAA--TGCAG-------AG-GG-A------CA----C--G--------A-A-TA--CT-T-- 987
MUS-1.CM.01.SIVmus_01CM1085 GCT...--TGCC--C-----AG-CA-AAG----GCA--A-----GT-CCA---------C-CAATG-----TA-----AA-------GTC-A----G--G--A---AAA--TGCTG-------TG-GG-A------CAG-----GG----G--G--A-TA---T-T-- 985
MUS-2.CM.01.CM1246          ACCAGTG----G--------AG-CA-AAGA-CAGCA-----A--GTATCAG------T-TC-TAGC-----GCTA-----AA-C------TC-AC------G-----AAAA--TGC---------AG-GG--C----CCAG--CC-----------TG--T---CT-T-- 1068
MUS-2.CM.01.CM2500          GCAAG---------C-----AG-GA-TAG--CAGCT--A---ACCTATCA-------T--C-CAGC--C----T------GA-C-----GTCCAC-A-T---------AAA--TGC---------AG-GG-TT----CCA---C------------TGT-TT--TT-T-- 1030
MUS-3.GA.09.09GabOI81       ACAACA--T--G--------A--TA-TAGA-CA-CAA-T-----GT-TCAG----------C--GA---C---TTA-----GAGT-------GCAC------G------AAA--TGC---------AG-GG--T-----CAG--C---A----G---CT--T---TT-T-- 566
MUS-3.GA.11.11GabPts02      GCCAG---TT-C--C-----GG-CA-TAGAGCAGCTA-T------TATCAG--------AC-C--CA-----GTGC---AAGT-----T--CACCA----------AAAG---GC------GACAG-GG-TC-T---CAG--CC-G---------TTA-TAG-TT-T-- 578
OLC.CI.97.97CI12            ......CAGTT-GCCATG--A---A-CAGA--TGCT-----AATCT-T-AG---GT-C-GC-GG-TA-------CC-GC-AAGT----------C-T-G-----C-G--G---T-AAT-----A-AG-G--TC-A-----T--G--TG-T--G---TTGCTG---G---- 1205
RCM.CM.00.SIVagi_00CM312    GGTCAA--AGG---------C-----AACA-TT-AT...CA---GCCA-A---C---C-T--T--T--C--G--G--G-----C-----G--G--G--T--G--A---AAG--TG-AG-------AG-G-----------G---C-T-----G--CTG--TA--TT-T-- 554
RCM.CM.02.SIVrcm_02CM8081   GGCCAA-AAGG---------A---A-AACAGTA-AT...CA---GCCA-AG--C---C-----T-----AC-G--AC-T-----T-----G-----G-----------AAAG--TG-GG-T------G-G-----------G---C-T----------TG--TA--TT-T-- 554
RCM.GA.x.SIVRCMGAB1         GGACAA--AGG------T------A-TACT-TA-AT...CA---GCCT-AG---A-TC-T---------A--G---C-------C-----C--G--G-----G-----AAAA---TCAG-------AGCG--------C--G---C----------CTG--TA--CT-T-- 765
RCM.NG.x.NG411              GGACAA--AGG---------A-----AACA-TT-AT...CA---GCCT-A-------C-A-----T--------A-------T-----------------G--A---AAA--TG-AG-------AG-G--T--------G---C-G--------CTG--TA--CT-T-- 765
SAB.SN.x.SAB1               GGT-G-TCG-G---------C--T--CAGTGTT-AT...AAT---TG---C--C---C-AT-G--C--A--G--AC-T-----T----------GA-T--G--A---AAG-----TG----G--AG-----------AGC--CC--G----------T-T---TT-T-- 1603
SMM.CI.79.SIVsmCI2          GGCAAAG-AGG---------AG--CA---AGTGGGA...--TA-TTACACC--CTTAC-AT--AGT---C----C--------C--------T--A---G--A---AAG---G-GG-------AG-G--AGGA---CA----C----------CTG---T--CT-T-- 702
SMM.LR.89.SIVsmLIB1         G-CAGG--AGG---------AG--CA---A-TGGG-...--CA-TTAT--C--CTT-C-GC--AGC--C--G--AC---------A-----GC---G-----A--AAG--TG-GG-------AG-G--AGGA---CAG---C----------CTG---T--GT-T-- 691
SMM.SL.92.SIVsmSL92A        GGCAAG--AGG---------AG-TCAA--AGTAGGA...--CA-TTAT--T--C-TAC-TT--AGC--CC-C--G--G-----T-----T--GC---G-----A--AAG--TG-AG-------AG-C--AGGT---CAG--G--G---------TG------CT-T-- 699
SMM.SL.92.SIVsmSL92B        ......G-AGG---C--T--AG-GCA---AGTAGGA...AATA-TTAT--G--CACTC-AC-G--C--AC-G--AC-C-----G-----C---C-G--------A--AAA--TG-AG-------AG-G--AGGA---CAG---C----------CTGT--A--CT-T-- 669
SMM.SL.92.SL92B             GGA-GA......--C--T--AG-GCA---AGTAGGA...AATA-TTAT--G--CACTC-AC-G--C--AC-G--AC-C-----G-----C---C-G--------A--AAA--TG-AG-------AG-G--AGGA---CAG---C----------CTGT--A--CT-T-- 1040
SMM.US.04.G078              GGCAAAG-AGG---------AG--CA---AGT-GGT...--CA-TTACACC--CTT-C-GC--AGT--A--------G-------------GT-G--------A--AGG--TG-AG-------AG-G--AGGA---CAG---C-G--------CTG------TT-T-- 872
SMM.US.05.D215              GGCA-G--AGG---C-----AG-GCA----GTAGGA...--TA-TTATA-C--C--AC-TC--AGC--A--G-----------T-------GT---T-----A--AAA--TG-GG-------AG-G--AGGA---CAG--------------CTG------GT-T-- 866
SMM.US.06.FTq               GGCAGAG-AGG----------G-GCA---A-TAGG-...--CA-TTAT--C--C-TAC-TT--AGT--A-----AC-G-------------T-G--G--------AAA--TG-AG-G--G--AG-G--AGGA--CCAG---C----------CTGT------T-T-- 879
SMM.US.86.CFU212            GGCA-G-AGGG---------AG-GCA---AGTRGGT...--CA-TTAT--Y--C-TAC-AT--AGY--C--G--A-----C--C---------T----G--R--A--AAA--TG-GG----G--AG-G--AGG----CAG---C-R--------CTGT--T--TT-T-- 858
SMM.US.x.H9                 GGCAGAG-AGG---------AG-GCA---AGTAGGT...--CA-TTATACC--C-TAC-TT--AGT--A-----A--------T---------T-G-----------AAAA--TG-GG----G--AG-G--AGGA---CA---GC----------CTG---T--CT-Y-- 1088
SMM.US.x.PGM53              GGCAGAG-AGG---------AG-GCA---A-TAGGT...--CA-TTACACC--C-TAC-TT--AGT--A-----A--------T---------T-G-----------AAAA--TG-GG----G--AG-G--AGGA---CA---GC----------CTG---T--CT-T-- 1532
SMM.US.x.pE660.CG7G         GGCAGAG-AGG---------AG-GCA---A-TAGGT...--CA-TTAT--C--C-TAC--T--AGT--A-----A--------T---------T-------------AAAA--TG-GG----G--AG-G--AGGA---CA---GC----------CTG---T--CT-T-- 1604
STM.US.89.STM_37_16         GGCAGAG-AGG---------AG--CA---AGTAGG-...--CA-TTAT--C--C-TAC-AT--AGC--A-----C--G-------------GT----G--G--A---AAA--TG-TG--------G-G--AGGA---CAG---C----------CTGT-----CT-T-- 1262
SUN.GA.98.L14               ......TCT-------TT--AG-GCA-AGAG-GGGA...CAAA-TTACA-T------C-TC-C--G--------AG-GC-AA-------G--GA----G----------GA-GG-AA------ACAG-GG-A------G----T---ATGAC------A-TT---G---- 1571
SYK.KE.x.KE51               GGC-G-TCAGG----------C--A-TAGA---CCA--AAAT---TG-A-T---GTA-GGG-CAACA-C--G--------AA-C------G-G-C---GA-CAGC--AAAA--TGATG-CTC-C--G-C--AC-A--CCA-ATT---A-------TTTT-TA--TT-T-- 1045
SYK.KE.x.SYK173_COMGNM      -CACCT--AGG---C--T--AC-GC-CAGA--TCCT--AAAT--GTG-A-T---ACA-GAG-CC--GT---G---C----AA------GG-G-CT--GA-CTCC--AAAA--TGATG--TCCA-AG-G--AT-A---CAG--TC--A--------TTT-TT--TT-T-- 1417
TAL.CM.00.266               GGCCAA-CAGG---------C---A-AAGA--TGCA---CA---GTATCA---------AC-G-------GCTGC----GA-----A-TGCCACTA----G-----AAAA--TGC---T---ACAG-GG-GC-C--C--G---C--G-G-----GTG--T---CT-T-- 1225
TAL.CM.01.8023              GGCCAT--TGGG--------C--TA-AAGA--TGCA--ACAA--GTATCAG---------T-G-----C---CTGC-G--AA-----A--GCCACTA-------A---AAA--TGC------GACAG-GG-GC-T--C--T---C--G--------TG--TT--TT-T-- 723
TAN.UG.x.SIVagmTAN1         AGA...CATGGT---------G----A--AC-G-AT...AAT---TG-------ACTC-TT-G-----AC-T--CC-------------T---AC---T--G--A--AAGG---G-GG-----A-AG----------CCA---CC-C--------GTG-TTAAGTT-T-- 1473
VER.DE.x.AGM3               GGT-G---T-------T---AGC-CAA--AC-GGGA...AATGC-TG-A-----GT-C--T-G-----AC-C--C---------G----------C------G-----AAAA---G-AG-----A-AG-G--------CCAG--T-----------GTG---A--CT-T-- 1014
VER.KE.x.9063               GGC...--T-------T---AGCGCAA--AC-AGGA...AATGCCTG-------GTTC-AC----C---C-C--A--------G----------C------G--A---AAA--TG-AG----GA-AG-G--A-----CCA----C-C--------GTGT-----CT-T-- 1513
VER.KE.x.AGM155             GGC...-----------TT--CGC-CAA---C-AGGG...AATGC-TG---G---GTAC-AC-T------C-C--CC-------------------C-------------AAAA--TG-GG-----A-AG----------CCA---CC-C--------CTG-------T-T-- 1502
VER.KE.x.TYO1_patent        GGC...--T--G----TT--AGCGCAA--AC-AGGA...AATGCCTG-------GTAC--T-G-----GC-C--C--------G----------C------G-----AAAA--TG-AG-----A-AG----------CA---CC-----------CTG---A--CT-T-- 1510
WRC.CI.97.97CI14            ......TCT--G---ATG--AG-GC--AGA-CAGCA...CAAGGGCCTCA-TGGG-CC--T-G-----------AA--C-AA-------G---CATA-T--G--T-G--GG--TG-AG-----AGTG-T--AT-C---G----TC-TA----CA-G--ATTAT-T--T-- 1541
WRC.CI.98.98CI04            ......TCT-----CATG--AG--T--AG--CAGCT...CAAGG-CCTCA-TGGG-TC-TC-G--T--A-----CA--C-GA-T-----G---CACA----G--T-GA-GC--TG-AG-----TGTG-T--AT-C-----T--T--GA----TA----TTTGG-T--T-- 1559
WRC.GM.05.Pbt_05GM_X02      ......TCAGTG---ATG--AG-CC-TAGAGGACCA...CAAGG-CCACA-TGGG--C-TC----C--G--G--CA-TC-AA-----------CATA-------T-G--GA--TG-AG-----TGTG-T--AT-C-----T--T--GA----TA----ATTGA----T-- 503
```

```
H1B.FR.83.HXB2                TTTAAACACCATGCTAAACACAGTGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCTGCAGAATGGGATAGAGTGCATCCAGTGCAT...GCAGGGCCTATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAA 1504
Gag                            L  N  T  M  L  N  T  V  G  G  H  Q  A  A  M  Q  M  L  K  E  T  I  N  E  E  A  A  E  W  D  R  V  H  P  V  H  .  A  G  P  I  A  P  G  Q  M  R  E  P  R  G  S  D  I  A  G
H1O.CM.91.MVP5180             -A-C--T--------G--TG-CA-A--------C----GG--TT-A---G----G--G--AGTA--------------A---------------ACT------CCAGCA...ATG-----GT-AC-------G-----A--G-------CA-----------T--T---- 1524
H1N.CM.95.YBF30               -C----T--T---T----T-----A-----------G-----A-----G---C-G--G--AGT---------------A-----C---------GACA---------C-C-...-TG--A--AC-AC-C-----G--AC-------C--T--A---------T---------- 1068
H1P.CM.06.U14788              CA-C--T--------T--TG----A--C--------G-G---GT----GG----G------GTA--T-----A--G------------------ACA-----C--T-CA...-T---A--AT-AC-------G---C-T----------CA---G-G--T--T------- 969
CPZ.CD.06.BF1167              -C-T--T--A--------TG--A-T----AC-------GG--A------G--C-G--G--AGTT---------A-----A--T--T-----C--GT-A--C---ACT---..---------GC-A-TC-----GACTC-A-----G--C-CA--G--------T------- 1582
CPZ.TZ.00.TAN1                -C----T-GT-----C--TG-T--T----A------------A-----------G--G--AGT---------------A-----G-----C----CA-----C-CT---..-----A--CCAGCA-G----GAT-C-A-----G--C-CA--GGCA--T--T-----G- 1136
GOR.CM.07.SIVgorCP2135con     -A-T--T--T-----G--TG-CA-T--T--------G-G---AT-A--GG-A-----G--AGT---------------A-----------C--GC----C--G-CAGTA...-T---A---C-GC--GT---A---C----G--G--GTCA--GG-A--T--T-----G- 1037

MAC.US.x.239                  CA-T--TCAG---T----TTGT-----A-AC--------G--T-----G--TA-C-G---T-TT--A--C-----G--------T-----CTTGCA---C---CAA-CA......-CT--A...CA--A---A--AC-T--G--G--GTCA---TCA--T--T------- 1767
Gag                            I  N  Q  M  L  N  C  V  G  D  H  Q  A  A  M  Q  I  I  R  D  I  I  N  E  E  A  A  D  W  D  L  Q  H  P  Q  P  .  .  A  P  .  Q  Q  G  Q  L  R  E  P  S  G  S  D  I  A  G
H2A.SN.86.ST_JSP4_27          -A-T--TCAA-----T--TTGT-----C-AC----------T--------AA-C-GG--A-TT--T-----A-----A-----T-----CGC-CAA--C---A-A-CA........--C--CT-AC--G-G--G---C-C--G--G---------TC-----------G- 1265
H2B.CI.x.EHO                  -A-T--TCAG---T----TTGT--A--A-A------G--------------TA-T-GG----TA--T-----------A-----C-----CCA-CA-------TC--CA.........--C--A--GC--G----G--AC-C--G--------A--GTCA--------G---- 1782
H2G.CI.92.Abt96               -A-C--TCAA--------TTGT--A--A-A---------------------TA-T-G---A-TA--T---------G--A-----K------CTTCAA------CAA--A..........-----GCCGC--G----A--AM-C--G-----G-CA---TCA--T--------G- 1183
H2U.CI.07.07IC_TNP3           -A-C--TCAA---------TGT--A--R-A------G--------------CA-T-G---A-TA--T-----------A-----T-----C-T-CA-A----TCAA-C-........--C--AC-AC--G----G---C-T--------T-----GTCA------------- 1288
H2U.FR.96.12034               -A-T--TCAA---T----TTGC-----A-A---------T--T--------TA-C-GG--A-TA--T-----T-----A-----T-----C-C-CA---C---AAC--A.........--C--ACCAC--G----G--AC-T-----G-----A--TTC---T--T------- 1299
ASC.UG.10.RT03                -A----TCAGC-CA-G--TG-CA-A----A--------G---GG-A--GC-TA-C-----TGTAG-T-----------A--C-CT------CAGT----C---CA--CA...---CA---GCAGC-CAAC-CAGGCT-----T-C---TCA--G---------T-----GG 841
ASC.UG.10.RT08                CA----TCAGC-CA-G--TG-CA-A----A--------G---GG-A--GC-TA-T-----TGTAG-T-----------ACACACT-----CCAGT----C---CA--CA...---CA---GCAGC-CAAC-CAGGCT-----T-C---TCA--G--------------GG 856
ASC.UG.10.RT11                -A----TCAGC-CA-G--TG-CA-A----A--------G---AG-A--GC-TA-C-----TGTAG-T-----------A--CA-T-----CCAGC----C---CA--CA...-T-CA---GCAGC-CAAC-CAGGCT-A---T-C---TCA--G--------------GG 841
COL.CM.x.CGU1                 -A-G---TTG-----C--T-TTT---ATACT-------G-TT-CCT--GG-C--G--G--TGAG--T---A-AA-G--A-A---G-AT--CCT-C-A-----G--C--G.....CA---ACAACAG-A---AGCAT-A--GC-G--T-CA-C-TCA--T--CA----C- 1309
COL.UG.10.BWC01               CA-G--TTTG---------TCT-A-ATAC--------GGTT-C-T--GG-TC-C-----CGAG------A-AA-G--A-A---G-AT---CTCCAAA----------CG.....CA---ACAACAG-A---G---C-A---C-G----CA-CTTCA-----CA----G- 1084
COL.UG.10.BWC07               CA-G---TTG---T------TCT---ATACC-------GGTT-C-C--GG-C-----G---GAAC-A---A--A-G--A-AG--G-AT---TTGACA--C--------CA.....CAA---CAACAG-A---GGCAC-A--GC-G--C-C--CTTCA-----CATG--G- 1327
DEB.CM.04.SIVdeb04CMPF3061    CA-C--TGGGT-A--C--TG-CA-A--A-ATTTA----GG--A------G--AA----G--TGTA---------------G--C--G------TT-CAA--C---CAA--G.....CA---A...C-C-A---A---T-----------CTCA---GCA--T--T-----G- 1428
DEB.CM.99.CM40                CA-T--TGGTC-A--T--TG--A-T--A-AT-T---G-GG--T------G--CA-T-----TGTA--T-----------A-----G------TT-CAA--C---CAA--A.....CA---G...C---AG--G---C-T--G-------TCA---GCA-------------- 1268
DEB.CM.99.CM5                 -A-T--TGGGC-------TG-CA-A--A-AT-TA----GG--A---------TA-------TGTG--T-----------A-----G-----CTT-CAA--C---CA---A.....CA---A...C---A---T---C-T-----G---TCA---GC-------------- 1259
DEN.CD.x.CD1_CMO580407        -A----TCAAC-C-----TG--A-T--A-A---C----G-----------G----------TGT---T-----A------CAT--T------TTG-CT-----G-CTA--TTACA-CA---GC--C------CAGCACAA--G--C-----A--TTCA--T---------G- 1274
DRL.DE.11.D3                  CA-C--TCAG---T----TG-TA-T--A-AT--C----GT--TC-C-----AG-------C-T---T-----------A--T--C---------GAGA-----TCAAGTA.......--A---C-CC-G-A---GGT-C-T--GA-C--CTCA--GTCA-----T------- 918
DRL.x.x.FAO                   CA-C---CAG---T----TG-TA-T---AT--C----GT---C-C-----AG-------C-T-----------------A--T--T---------GAGA-----TCAAGTA.......--A---C-CC---A---GGT-C-T---A-C--CTCA--GTCA-----T------- 1011
GRV.ET.x.GRI_677_gri_1        -G----TCAG-----C--TGT-A-A--A-AC-----G-GG--AT-A-----TC-T--G--AGT---T-----A-----A-----G-----C--GACA--CAG-CCA-CA...-T--C--GT-AC--G----G---C-A-----C--G-CA--GTCA--T----------- 1617
GSN.CM.99.CN166               -A-G---CAA-----C--TG-TA-A--A-ATT------GG--AG-A--G--TA-T-----TGT-----------AC-G--A-----G-----CCTTT-A-----G--A-CA...----CA---CAAC-GGTG-CAGG-C-T-----C--T--C---GCA-----------G 1181
GSN.CM.99.CN71                -A-G---CAA-----C---G-T--T--A-ACT------G---AG-A--G--TA-T-----TGT-----------AC-G--A-----G-----CTCT-A-----GCA--CC....-G-CA--CCAAC--GT--CAGG-C-T-----C--T--C---GCA-----------G 1187
LST.CD.88.SIVlhoest447        C-----TGT----T----TG-CA-A--A-AC-------GG--A-----G--AA-T-----TCAT--TGTG--A----GG-----------C--GAAA-----GCAA--A...CA-CCTG-ACAAC--GG---AGG-C-T---ACC--T-AT-C--CA--TG-C------G 676
LST.CD.88.SIVlhoest485        C-----TGT----T----TG-CA-A--A-AC-------TG--A-----G--AA-T-----TCAT--TGTG--A----GG-----------C--GACA-----GCAA--A...CA-CCTG-ACAAC--GG---AGG-C-T---ACC--T-AT-C--CA--TG-T------G 676
LST.CD.88.SIVlhoest524        C-----TGT----T----TG-TA-A--A-AC-------GG--A-----G--AA-------TCAT--AGCG--A----GA--T-------------CA------CCA---G...CA-CCAG-ACAAC--GG---AGG-T-A---ACT----AT-C--CA--TG-G-----G- 670
LST.KE.x.lho7                 ------TGTT---T----TG--A-T--A-AT-----G-GC-----------AA-------TCA---AGTG--------G------G------------CA---C--GCAA--A...CAGCCAG--CAAC--GG---GGGTT-A---ACT--T-AT-C--CA--TG-G-----GG 1760
MAC.US.x.251_1A11             CA-T--TCAG---T----TTGT-----A-AC--------G--T-----G--TA-C-G---C-T---A--C-----G--------T-----CTTGCA---C---CAA-CA.........G--CCACA--A---A---C-T--G--G--GTCA---TCA--T--T------- 1766
MND-1.GA.x.MNDGB1             ---G--T-T------G--TG----A--A-AT-----G-G---A-------G-C--------TGTA--T------C---A-------------CT-ACA-----TCAA--A...CA-CCAG-ACAAC--GG---AGGAT-A--GACC--TTCA--CTC---T---------- 1131
MND-2.CM.98.CM16              -C-C--TGG------C--TG-TA-T--A-AC-----G-G---GC-T-----AG-G--G--TGT--------------------C---A-CTTGAGA--C---CAAGTG.......--T-----GC-C-A---TGTCT-----A-C----C---C-----T--T--T---- 1457
MND-2.GA.x.M14                CC-C--TGGT---T----TG-TA-T----A------G-G---TC-C-----AG-G--G--TGT--------------------C------CTTAGA--C--CCAAGTA.......--A---T-GC-T-A---GGT-C-----A-C----CA------------T------- 1381
MND-2.x.x.5440                CC-C--TGG------C--TG-CA-T--A-AT-----G-G---TC-C-----AG-G--G--TGT--------------------C------CTTAGA--C--CCCTGTG.......--A-----GC-C-A---AGT-C-A---A-C----CA----A------T------- 1014
MNE.US.x.MNE027               CA-T--TCAG---T----TTGT-----A-A-------------T-----G--TA-C-G---A-TT--A--C-----G--------T-----CCTGCA---C---CAA--A.........G--CCACA--A---A---C-T--G--G--GTCA---TCA-----T------- 1247
MON.CM.99.L1_99CML1           CA-C--TCAGC-A------G-CA-A--T-AT-----G-GG--AC-C-----TA-T--G--TGT---------------G--CA-T-----CTT------C---CA--CG...C--CAA--GCAAC-CAAT-CAGG-C--G----C----C---GGCC--T---------G 1168
MON.NG.x.NG1                  ............................................................................................................................................................. 0
MUS-1.CM.01.CM1239            CA-G--TCAGT-------TG-CA-A---AT-----G-GG--T--A--G--CA-------TGT---A-----C----A------------CTGC-A--C---CA--C-...CAGCA---GCAGC-CAAT-CAGGCC--C-GAGC---TCA---TCA-----------GG 1154
MUS-1.CM.01.SIVmus_01CM1085   CA-C--TCAAC----G--TG-CA-A--A-A------G-G---T--A-----TA-------TGTA--T-----AC----A------------CCTGT--A-C---CA--CC...CA-CA---GCAGC-CAAT-CAGGCC-C--GA-C----CA---TCA-----------G 1152
MUS-2.CM.01.CM1246            -G----TCAG-----C---G-CA-A--A-AC--C--G-GG--A--A---C-AA-------TGTGG---------C-G--A-----CA-A--CGTTT-A----CCAA-CA...A--CA---ACAAC--AAT-CAGG-C-A--GT-C---TCA--GGC---TG-G--G---G 1235
MUS-2.CM.01.CM2500            -G-G--TCAA-----T--TG-------A-AT--C----GT--AG----GC-AA-T--G--TGTGG---------C-G-----T--G-----CTTT----C--CCAA-CA...---CA---CCAAC--AAT-CAGG-C-TC--T-T--CTCA---GCA--T---------GG 1197
MUS-3.GA.09.09GabOI81         CA-G--TCAG---T----TG-CA-A--A-A------G-GG--AT-A--R--CA-------TGT---------AC-------T--T------TTGT----C--GCA--CA...---CAA--ACAGC-CAAT-CAGGAT-A---TC----TCA--CTCA-----------GG 733
MUS-3.GA.11.11GabPts02        CA-G--TCAA-----G--TG-TA-A----A------G-GG--GT----G--AA-C-----TGTG--A-----AC-G--A-----C-----C-TGT---------CAA-C-...---CA---ACAAC-GAAT-CAGG-T-A---TC---TTCA--GTCA------------G 745
OLC.CI.97.97CI12              GA-C--T---C--T-G--TGT---T----ACT------GG--T------G----G--G--AG---------------A-A---G-AT--C--GCA-------CA--CA...AT------AT-GC-GGTG--A--A--TCC-CCG-------CT-----T----T----- 1372
RCM.CM.00.SIVagi_00CM312      -G----TCAA--------TG-CA-A--A-A--------G---AT-A-----AG-G-----AGTA--A-----T-----A-----C------CTTAG------CAAT-CA...........------AC-AC--G----A---C-A-----G--G-CA--T-----T--T------- 718
RCM.CM.02.SIVrcm_02CM8081     -A----TCAA---T----TG--A-A--A-A--------G---TT-------AG-G-----AGT---------------A-----T------GCTAGA----TAAT-TA.........--A--AC-AC-TG----G---C-T-----G--T-CA-----C--T--T------- 718
RCM.GA.x.SIVRCMGAB1           -A----TCAA--------TG-CA-A----A---C--G-GT--GC----G--AG----G--AGTG---------------A-----C------GCTAGA--------A-CA.........--C--G--AC--G----G--AC-T-----------CA------------------G- 929
RCM.NG.x.NG411                -G----TCAG---T----TG-CA-A----A--------GT---T-A-----AG----G--AGTA--------T--G--A-----T-----CTTGAGA----CAAT-CA.........--A--AC-AC--G----T--AT-A--G-------CA-----C--T--T------- 929
SAB.SN.x.SAB1                 -A-C--TCAA-----C--TG----A--A-A---C--G-GG--AC-A-----AG-G-----TGTT---------A-----G-----T-----CCTTAGA--C--GCC--CG...CA-CA---CCCA--T-AG--AGT-C-A--G--C---CAA--GTCA--------G--C- 1770
SMM.CI.79.SIVsmCI2            -A-C--TCAA---T----TTGT-----A-A-------------A-----G--AA-T--G--A-T---T-----------A-----T-----CTTGCA------TAATGTA.............-----AT-GC-GG----A---T----------G--T--CTCA-----T------- 866
SMM.LR.89.SIVsmLIB1           CA-C--TCAG---T-----TGT-----A-A------G--------------TA---G---A-TA--T-----A-----A-----T--A---TTGCA------CAA-CA......A-C--A--AC--G----G--AC-T--G-----GC-A--TTCC--T-----T--C- 855
SMM.SL.92.SIVsmSL92A          CA-T--TCAA-----T--TTGT--T----AG--C--------A--------AA-T-G---A-T---------A--G--A-----C-----CTTGCA------GCA--CA.........--T--AC-CC--G----G--AC-C--G--G-----A---TC---T---------- 863
SMM.SL.92.SIVsmSL92B          CA-T--TCAG-----T--TTGT--A--T-A------G--------------TA-C-G---A-TG--T-----------A-----C-----CCT-CAA--C---AGAGGACAACAGCC-G-GCAAC-GG-T--GGG-C-A--G------TCA---TCA-----------G- 839
SMM.SL.92.SL92B               CA-T--TCAG-----T--TTGT--A--T-A------G--------------TA-C-G---A-TA--T-----------A-----C-----CCT-CAA--C---AGAGGACAACAGCC-G-GCAAC-GG-T--AGG-C-A--G------TCA---TCA-----------G- 1210
SMM.US.04.G078                CA-C---CAA--------TTGT--A----A---------------------TA---G---A-TT--A-----A-----A-----C------TTGCAA-----CA--CA.......--C--AC-AC--G----G--AC-C---------------CTC---T---------- 1036
SMM.US.05.D215                CG-C---CAG--------TTGT--A--A-A------G--------------CA-T-GG--A-TA---------A---------T--T------CTGCAA-----CA--CA.......--C--GC-AC--G----A---C-C--G-------CA--GTC-------T------- 1030
SMM.US.06.FTq                 -A-T--TCAA---T----TTGT-------AT-----G--------------TA---G---A-T---A-----------G--T--T-----CCT-CAA--C---CA--CG...ATG--C--A--AC--G----G---C-C-----C--G-------TCA-----T------- 1046
SMM.US.86.CFU212              -A-C--TCAG--------TTGT-----A-A---C--R-----------G--CA---GG--A-T---------A--G--A--T--T-----CCTGCAA------CAA-CA.......--T--A--AC--G----A---C-T--G--C-----A--RTCA--T---------- 1022
SMM.US.x.H9                   -A-C--TCAA--------TTGT--A--A-A------G--------------TA-T-G---A-TT--A-----A--G-----T--T------TT-CAA--C--RCA--CA.......--T--A--AC--G----A--AC-T-----G-----R---TCC-----T------- 1252
SMM.US.x.PGM53                CA-C--TCAA--------TTGT--A--A-A------G--------------TA-T-G---A-TT--A-----A--------T--T------TTGCAA--C--GCA--CA.......--T--A--AC--G----A--AC-C-----T-----A---TCA--T--T------- 1696
SMM.US.x.pE660.CG7G           -A-C--TCAA--------TTGT--A--A-A------G--------------TA-T-G-----TT--A-----A--------C--T------TT-CAA--C--GCAA-CA.......--T--AC-AC--GT---G--AC-T-----G-----A---TCA-----T------- 1768
STM.US.89.STM_37_16           CA-C--TCAAC--T----TTGT-------AG--------------------TA---------T---A---------------T--T-----C-TGCA-------CAA-CA...C-G--T--GC-GC--G----G--AC-T------------C--GTCA--T---------G- 1429
SUN.GA.98.L14                 CA-C--TGTG-----G--TG-CA-A----AG-------G---T-----G---A------TCAT--AGTG------AG-----------C---A-----C--TCA---G...CAGCCTG-ACAGC--GG---AGGCC-A--GACT--T-AT--T-CA-----------GG 1738
SYK.KE.x.KE51                 CC-C--TGA------T--TG-CA-T--A-AT-------G---A------G--A-C-----TGTG--T-----------GGT-T---------CCT-CAA--C---CAA-C-...CA-CA--AGCAGC-TGT--CAGG-C-A-----T--T--T-CCTCA------------- 1212
SYK.KE.x.SYK173_COMGNM        -C-C--TGG------T--TG-T--A--A-AT-----G-G-----------GG-AA-T-----C-T--------------GG--G-C--------CTT-AA-----CCAA-CG...CA-CAA---GC-C---A--CAGGAT----G--T--CTCC-CCTCA------------- 1584
TAL.CM.00.266                 CA-C--TCAA-----T--TG----A--A-AG-----G-GG--A-----G--CA----G--C-T---T-------C-G--AT----G-----CAGA-------CCA-......AT---T--AC-AC--G----A---T-A--G------GCA--G-------------G- 1389
TAL.CM.01.8023                -A-C--TCAG-----T--TG----A--A-A------G-GG-----------G--CA-T-----T-T---A-------C-G-----C---------CAGA----C--CCA-......AT---T--AT-AC--G----A---C----G--G---GCA--C--------------- 887
TAN.UG.x.SIVagmTAN1           CA-T--TCAG---T----TGTCA-A--A-AC-------GG--A-----G--TA-C--G--AGTA-------T-----A--CC-G-------T-ACA--C--CCCA-CA.......---A--GC-AC--G----G---C-A--G--C-----A--TTCA--------G---- 1640
VER.DE.x.AGM3                 CA-C--TCAA-----T--TGTCC----A-AC-------GG--GC-A-----AG-------A-T---------------A--CC-G-------T--CT--C---CCA-CA.......---A--AT-AC--G----A--AC-C-----C--T--A--CTC-------------- 1181
VER.KE.x.9063                 CA-C--TCAA-----T--TGTGT-A--A-AC-------GG--AC-A-----AG-------A-T---------------A--CC-G------GT-ACA--C---CCA-CA.......---A---T-AC-------A---C-T-----C--T-----TTCA------------- 1680
VER.KE.x.AGM155               CA-C--TCAG-----T--TGTCT-A--A-AT-----G-GG---T-------AG-G-----A-TA--T-----------A--CC-G------GT-ACC--C---CC--CG.......----C--CT-GC--G-G--A---C-C--G--T--GG-----TCA--T--------G- 1669
VER.KE.x.TYO1_patent          CA-T--TCAG-----T--TGTGC-A--A-AT-------GG--AT-A-----AG-G-------T---T-----A-----A--CC-G------GT-ACA--C---C-A-CC.......----A--CC-AC--G----A---C-C--G--C--TC-C--CTCA--T--------G- 1677
WRC.CI.97.97CI14              -A----T---C--T------ACA-A--A-AT--------T--A--------AA-------C-TAG-A-----A---ATACA---G----------ACA------CA-GCA...------------A---------A--A------C-G----CA---TCA--T----------G 1708
WRC.CI.98.98CI04              CA-T------C--T----T-ATA-A--T-AT-----G--T-----------TA-------A-TAG-A---------AT-CA---C----------ACT--C---CAA-CA...---------A--A---------A--A-----GC-G----CA--GTCA--T---------GG 1726
WRC.GM.05.Pbt_05GM_X02        CA-C--T--TC-AT-G--T-ATA-A--A-AT-----------------G--AG-T-----A-TAG-G-----A---ATAA----------C---ACA------CCA-CA...-T---A---C-GC-C-----GGCT------A-T----CA---TCA--T---------G 670
```

```
H1B.FR.83.HXB2              CTACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAAT..............AATCCACCTATCCCAGTAGGAGAAATTTATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAATGTATAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAAG 1659
Gag                         T  T  S  T  L  Q  E  Q  I  G  W  M  T  N  .  .  .  .  .  N  P  P  I  P  V  G  E  I  Y  K  R  W  I  I  L  G  L  N  K  I  V  R  M  Y  S  P  T  S  I  L  D  I  R  Q  G  P  K
H1O.CM.91.MVP5180           -A-----C--A-AG--A--G-----TAT-----CT--T-GA...........GGGGC-AACT----------------C--C----G--A-------G-G--A---C----C-----G----A------C--T--AGTG-----CT-A--T--T--G--G--------A 1682
H1N.CM.95.YBF30             -A-----C-----GGCA-----GG-G-CT--------TGC-...............-----T---G-T-----------T-------G----------G-------G-----C-G---T--G-----------T---GT---------A--G--C-A------------A 1223
H1P.CM.06.U14788            -C---------AAA--------G---ACC----------GG...............--CAAT---G-A---------G--C--C----G--AR-------G-GT----GC-C------G-G----A------CT----CGTT----------------AG--G-----T--- 1124
CPZ.CD.06.BF1167            -A--A------A--GT---G--G---CA----G-A--T-C-......CCTCAAAATC-AGG-GGAG-T-----------T--C---------------C-----A--GC-G------T----G-A------C------GTA---------A--T---------G--------A 1746
CPZ.TZ.00.TAN1              -C--------G--A--A------G--CTG----------CC......CCACAGGCAC-AGG-GGAG-G-----------C--C--------G--------TT-A------------T----C-------C------GTT------T----------A---G--------A 1300
GOR.CM.07.SIVgorCP2135con   -A--A-----T-AA--A------G---CT-----A----GG...........AAT---AAC---G-G-----G--T--C--A--C-GG-A-------G-G--A--G-----C---G------A------CT-T--AGT---T--C--A--T--T-AG--G-----T--- 1195

MAC.US.x.239                -A------T-AG-AG-T--------CCAG------TAC-GA...........CAAC-GAAC--C--A---------CA-C-----C-GG--------CCAA-----G--GC-A---TGT--C----------A---A--A-A------A--TG---A------G-----A 1925
Gag                         T  T  S  S  V  D  E  Q  I  Q  W  M  Y  R  .  .  .  .  .  Q  Q  N  P  I  P  V  G  N  I  Y  R  R  W  I  Q  L  G  L  Q  K  C  V  R  M  Y  N  P  T  N  I  L  D  V  K  Q  G  P  K
H2A.SN.86.ST_JSP4_27        -A--A--C--AG-AG-A--G--G--CCAG------TTT-GG...........CCAC-AAAT---G-A---------AGC--C----G---------CCAGA-A--GC--C-G--GTGC--C--G-----C-A---A----A---C--A-------A---G--------- 1423
H2B.CI.x.EHO                -C--C--C--AG-AG-A-----G---CAG------TAC-GA...........CCCC-AAAT---G-------G---A-C--C----G---G-----TCAGT-A--GC--C-G---TGT--CC-G-----C-AT-----T-AT--A-----------AG-----G-----A 1940
H2G.CI.92.Abt96             -C---------AA-AG-A---------TCAG----CACAC-GG...........CAGC-AAAC-----A---------A-T--Y----G---------CCART-A---C-TC-A---TGT--C--G-----Y-AT--A--A-A---C--A--TG---A------G-----A 1341
H2U.CI.07.07IC_TNP3         -C--C--C--GG-GG-A-----G--CCA--------TAT-GG...........CAAC-AAAC-----A--G-------A-C--C--C-GG--------TCAAT-A---C-CC-A---TGT--C--G-------AT--A--A-A---C------G---A------G-----A 1446
H2U.FR.96.12034             -A-----C--AG-GG-A--G--G---CAG------TAC-GG...........CCGC-AAAT--A--A--G--G--TA-C--C---CGG--------CCAAT----CC--C-A---TGT--T--------C-AT--A--T-AT--CT-A--T----AG-----G------ 1457
ASC.UG.10.RT03              TAT-ATCAT--A-A--A-----GC---A-----------GA...........GCAGTAGAC--AG-AAAT-----------A--C-G--A-------G-AA----TC-GC-G-G-TGT--C-A------C-A---AGTT-A---------------------G-----A  999
ASC.UG.10.RT08              TAT-ATCAT--A-A--A-----GC---A-----------GG...........GCAGTAGAC-AAG-AAAT--G---------A---G--A-------G-AA----TC-GC-G-G-TGT--C-A------C-A---AGTT-A---------------------G-----A 1014
ASC.UG.10.RT11              TAT-ATCAT--A-A---------GC---A---------C-GA...........ACAGTAGAC-AAG-AAAT--G---------A--C-G--A-------G-AA----TC-GC-G-G-TGT--C-A------C-AT--AGTT-A---------------------G-----A  999
COL.CM.x.CGU1               AC--A--CT-AG-AGCA------G-G-CT---GGAGAG..............................--CA---CCA-C-------GGG-C------G-A-A-A-TC-GG-A--GG--A-CCAG--TGCA--A--AT--TCTG-A--A--T--T-----G---AGT--A 1449
COL.UG.10.BWC01             AC--A--CT-AG--GCA-----GG---CT---GGAGAG..............................-------C-A-C--C--C-G-G-------TG-T-A-A-CC--G-A--GG-GA-CCAG--TGC---G-----ATCAG-G--A-----C--G-----GAGC--A 1224
COL.UG.10.BWC07             -A--A--CT-AG-GGCA----------C----GGAGAG..............................-------C-A-C-------G-G-------G-G-AAA-C---G-A--GG--A-TCAG--AGCA--A----G-TCAG--T-------C---------G----A 1467
DEB.CM.04.SIVdeb04CMPF3061  -C-AC--C--AG-AG-A--------T-C-----------GAGCCCAGGGAGTGGGAGGAGG---C--AGAT------C---------G-------G----TT----TC--C-A-G-TGT----A---------AT-------A---CT----TG-G-A------G--C--A 1598
DEB.CM.99.CM40              -A-AC--C--AG-GG-A-----G----C---------C-GACCAGCAGGTCAAGGAC-AGG---C--AGAT------C----A----G-------G-C--A--A--TC-GC-A-G-TGT--G-A------C-AT--A----A---CT----TG-T-A------G--C--- 1438
DEB.CM.99.CM5               -A-AC--C--AG-GG-A--G--G----CT--------C-GGCCAGCAGGACAAGGAGCAGGC--A--AGAT------C-G--A--C-GG------G-C-----A--C---C-G-G-TGT----A------C-AT--A--T-A----T----TG-T-AG--G--T--T--- 1429
DEN.CD.x.CD1_CMO580407      -C--CTCCT-TA-AG-A--T--G---AC---------T-GA...........CAAC-AGGC--A--AAAT--------T--A-----G--G---G-T--TA--------C-A--GG-G--C-AG------CCAA--AGT-TC----T----------AG-----G--C--A 1432
DRL.DE.11.D3                -C--C--CT-TA-AG-A--G--G----AG----CC--C-GA...........GCCC-GGACT-C---AAT--G----C--------G-CA----G-GG--T-A--G--GC-A-G-TGT--G-C------C-AT--AGT--A---CT----TG-G-AG-----T--T--A 1076
DRL.x.x.FAO                 -C--C--CT-TA-AG-A--G-------AG----CC--C-GA...........GCCC-GGATT-C---AAT--G----C--------G-CA----G-GG--T----G---C-A-G-TGT----C------C-AT--AGT--A---CT----TG-G-AG-----T--T--A 1169
GRV.ET.x.GRI_677_gri_1      ----C--CT-AA---------------AG----CCTTC--............GCC------AGA--AGAC-----G-C-CAA--C-G--A----G-T--TT----C---C-A--GG----GCAG-----C-AT--CCAA-AGG-C--A-----TC-----G-----T--A 1775
GSN.CM.99.CN166             T----TCA--T-CCA-T--G--G----AG-----C----GG...........CAG--C-A----G--AAT--GTC---T--A----G--AG-----C--TT-----CC-GC-GCG-TGT----A---------A-----GT--AT---T--------C-A------C--T--A 1339
GSN.CM.99.CN71              TA--ATCA--T-C-A-T--G--G----AG-----T----GG...........CAG--CAAT--AG-TAAT--G-C---T--A----G--A------C---T-----C--GC-GCG-TGT----A---------A-----GT--AT---T---G---C-A---G--C--T--- 1345
LST.CD.88.SIVlhoest447      TC--CTCA--AG-AG----G--GC---CT-----CC--T.............GCAG-CA-C---G-AGAT------A-G---------GAG---G----T-AA-CTA-GG-A---G----T-----CC--CAA--AGTT--TG-AA-A-----C-AG--G--G-----A  831
LST.CD.88.SIVlhoest485      TC--CTCA--AG-AG----G--GC---CC-----CC--T.............GCAG-CA-C--AG-GGAT------A-----------GA----G----T-AA-CTA-GG-A---G----T-----TC--CAA--AGTT--TG-AA--------C-A---G--G-----A  831
LST.CD.88.SIVlhoest524      TC--ATCA--AG-AG-A-----GCA--CC-----CA---.............GCAG-CA----C--AGAT-----GA----C------GA----G------A--CTA-GG-G---G-G---------AC--CAG--AGT---TG-AA----------A---G--T-----A  825
LST.KE.x.lho7               TC--CTCA--AG-AG-A------C---CC-----CT---.............GCAG-CA-T---G-AGAT------A-----------GA----G-C--T-A--CTA-GG-A---G----C-----CC--CAA--AGTA--TG-CA-------C-A----------C--A 1915
MAC.US.x.251_1A11           -A------T-AG-AG-T--------CCAG------TAC-GA...........CAAC-GAAC--C--A---------CA-C-----C-GG--------CCAA-----G--GC-A---TGT--C----------A---A--A-A------A--TG---A------G-----A 1924
MND-1.GA.x.MNDGB1           -----TC---AG-GG-A------T-G-C---------AT-TG...........CAAC-AAATG-A---AAT-----AC---C-----G--T-----T--A-----CA-G----G-T-G----A--GTC--T-T--A-TA--T--AACA--TG-------G--------- 1289
MND-2.CM.98.CM16            -A--A--CT--A-AG-A--G--G----A-----CT----GG...........GAGC-AGATG-AG-AAAT--G----G---A-----CA-------G-TT-A--GC-TC-G--GTGT-----C-------AT--AGTA-A---C-----------AG----------- 1615
MND-2.GA.x.M14              -A--A--C-G-A-AG-A-----G----A-----CA--T-GA...........CAGC-GGA--AGG-AAAT-----G-C---C-----GCAG------G----A--GC-GC-A---TGT--G--C-----C-A---AGTG-A--------------A------G-----A 1539
MND-2.x.x.5440              -A--A--CT--A-AG-A--G--G----A-----CA--C-GG...........CAGC-AGAC-AGG-AAAC-------GT--A--C---CA-------G-A--A--GC-GC-A---TGT-----C-------A----GTA-AT---T-A-----C-A---G--------A 1172
MNE.US.x.MNE027             -A--------AG-AG-T--------CCAG------TAC-GA...........CAAC-GAAC--C--A---------CA-C-------GG--------CCAA-----G--GC-A---TGT--C-A-------A---A--A-A------A--T-----------G-----A 1405
MON.CM.99.L1_99CML1         TCT-GTCC--A-CC--AC-G--G----AG-----A--C-GG...........GCT--CAAC--CG-T-AG--CTC---C--C--C-G--A----G-C---T----CC-GC-GCG-TGC--C-A------C-A----GT--A----T----T--C-A---G-----T--A 1326
MON.NG.x.NG1                ..........................................................................................................................................................................    0
MUS-1.CM.01.CM1239          TA--ATCC----CCA-T----------A------A--T-GG...........CAA--CGAT-----AAAT--G-C---C--C-----G-A-----------A----G---C-G-GGTGT--C-A--------A----GTT--T--AT----T--C-AG--G--------A 1312
MUS-1.CM.01.SIVmus_01CM1085 TAT-ATC-----CA--A-----G----A------C--T-GG...........GCA--CAAC----AAAT----C------C--C--G-AG---G-G---A----GC-GC-G-G-TGT--C-A------C-AT--AGT--A---C--A--T--C-AG--G--------A 1310
MUS-2.CM.01.CM1246          TGT-ATCA--G-CAGCA----------A---------T-GA...........CAGC-AGAC--AG-AAAT-----GA-T--A----G--A------C---T----CC--C-G-G-TGT--C-A------C-AT--AGTA-A---CT----T--T-A------C--T--A 1393
MUS-2.CM.01.CM2500          TAT-CTCC--A-CAGCA----------AG--------C-GG...........CAAC-AGAC--AG-GAAT--------C--A----GG-A------T---------C--GC-A-G-TGT--C-A--------A----GTA-A---CT-A-----C-A---G--C--C-GA 1355
MUS-3.GA.09.09GabOI81       TCT-CTCA--A-CAGCA--G--G----A------A--CCGC...........CAA---GA------AAAT---TC---C-------GG-A------C-----A--GC-GC-A-GGTGT--G-A------C-A----GTG-A------A-----------G-----T---  891
MUS-3.GA.11.11GabPts02      TGT-ATCC--G-CAGCA--G--G----AG----------GA...........CAG---GAT--A--TAAT---TC----------G--A------C---G----C----C-G-GGTGT--G-A------C-A---AGTA-A----T-A-----T--------G--C---  903
OLC.CI.97.97CI12            -C-------ATT-G---C-G---C-G-CT---G-A-ATGCA...........GGA..........--AGAT-------C-C-C-----G------G-T---A-A--G--GC-G--GT-G--G-AG------TCG--AGTT-CAG-G--A-----C----------G-T--A 1521
RCM.CM.00.SIVagi_00CM312    -C--C--C-ATA--GCT-----G----CC----CA--C-GA...........GCA---AAC--A--AG---------A-C-------G--AT------G-A--A--GC-TC-A--GTGT----A---------A---AGTA-A---CT----T----AG--G--------A  876
RCM.CM.02.SIVrcm_02CM8081   -A--A--C-GTG--GCA-----G----CT---TCA----GA...........GCT--CAAC--A--AAT------GA-C--------G--AT-----T--T-----G--GC-A---TGT-----G--------A---AGTA-ATG-------T--T-A---G--------A  876
RCM.GA.x.SIVRCMGAB1         -A-----CT-AA-AGCA-----G----CT----CC--C-GA...........GCA--CAAC--C--TAAT--G--CA-TC-G--C-G--AT---------AG-A--G---C-A---TGG----A------C-AT--AGTG-A---C--A--T----AG-----------A 1087
RCM.NG.x.NG411              -A--A--C-ATA--GCT----------CT----CC--T-GA...........GCC--CAAC--A--TG----G---A-C--------GG-AT-----TG-G--T--G---C-A---TGT--G-A--------A---AGTA-A---CT-A-----C-AG--G--------A 1087
SAB.SN.x.SAB1               -------C--AA-A--A--G-------A-----CC--C-GA...........GCAC-GAATG--G--AAT------A-T---------G--------C---------TC-CC-G---TGT--C-A------C-A---AGTA-A----T-A-----T-A------G--T--A 1928
SMM.CI.79.SIVsmCI2          -A--A-----AG-GG-T--G-----CCA-------TAC-GG...........CAAC-AAAC---G-A-----T--GA-C--C----GG--------TCAA-----C---C-G---TGT--C--G-------AT--A---A---C--A--TG-G-A------C--C--- 1024
SMM.LR.89.SIVsmLIB1         -C--------AG-GG-T--G------CAG------TAC-GG...........CAAC-AAAC-----A-----T---A-C-------GG--------CCAA--AA-GC-GC-A---TGT--C----------AT--A--T-A----------TG-G-AG-----G------ 1013
SMM.SL.92.SIVsmSL92A        -A-----C--AG-AG-A----------CAG------TAC-GG...........GCTGCCAAC--C-----G--G--CA-C--C--CCGGC-------CCAAT-A--T---C-A---TGT--T--------C-AT--AGT--A---C-----T--C-AG-----G--C--A 1021
SMM.SL.92.SIVsmSL92B        -A-----C-A--CAA-T-----G----A-------TAT-GG...........GCAC-AAAC--AG-T-----G-----C--C----G---G-----CCAG------C-CC-A---TGC--C--G-------A---A--A-A---C-----TG-T-A------T------  997
SMM.SL.92.SL92B             -A-----C----CAAGT-----G----A-------TAT-GG...........GCAC-AAAC--AG-------G-----C--C----G---G-----CCAG------C-CC-A---TGC--C--G-------A---A--A-A---C-----TG-T-A------T------ 1368
SMM.US.04.G078              -C--C-----AG-AG-T--G------CAG------TAC-GA...........CAGC-AAAC--C--A--G--T---A-C-------G---------CCAAT-----C-GC-G---TGT--C--G-------AT--A-TG-A---C--A---G-G-A------G-----A 1194
SMM.US.05.D215              -C--C-----AG-GG-T--G-----CCA-------TAC-GG...........CAGC-GAAC--CG-A---------A-C-----C-G---G-----CCAAT----G---C-A---TGT--T----------A---A--T-A------A--T----AG-----G--T--A 1188
SMM.US.06.FTq               -------C--AG-AG----G--G----A-------TAT-GG...........CAGC-AAAT-----A--------TA-C--C--C-G----------CCAG--T--GC--C-A---TGT--T----------AT--A--A-A------A--TG---AG--G--------A 1204
SMM.US.86.CFU212            -C--C--C--AG-GG-A--------TCAG------TAC-GG...........CAGC-GAAT--G--A--G------A-C--C--C-GG--------TCAA--T--G--GC-A---TGT--G--G-------AT--A--A-A------A--TG-G-A---G-----C--A 1180
SMM.US.x.H9                 -------C--AS-AG-C--------TCA-------TAY-GG...........CAGC-AAAC-----A--------CA-C--C--C-GR--G-----CCAAT-A--GC-GC-G---TGC-------------AT--A--A-A----T----TG-G-A------------A 1410
SMM.US.x.PGM53              ----------AG-GG-A--G-----TCAG------TAT-GG...........CAGC-GAAC--C--A-----G---A-C-----C-GG--------CCAA--T--G---C-A--GTGT--C--------C-AT--A--A-A----------G---A------G--G--A 1854
SMM.US.x.pE660.CG7G         ----------AG-AG-T--------CCA-------TAC-GG...........CAAC-AAAC--C--A--------CA-C-------G---G-----CCAAT-A--G--GC-G---TGT--------------A---A--A-A----T-A--TG-G-A------------A 1926
STM.US.89.STM_37_16         -------CT-A-CAG----------TCA-------TACCGG...........CAGC-AAAT-----A--G------A-C--C--C-GG--------TCAA--A--GC--C-A---TGT--C--------C-AT--AGTT-A----T-A---G---AG-----G-----A 1587
SUN.GA.98.L14               TA--ATCA--AG-AG-A--G--G---CAG-----CTATGGC......CTTAATGGAGCAA--AGAG-TG-----CA---C--C------G-C---G-C--AGA---CA-GG-A-G-G----T---C-CC--CAGA-----TCAG-AG-A--A--T-----G--------A 1902
SYK.KE.x.KE51               -A--C------A-C--A-----------AG-----A--T-GA...........CAG--CAAC--A--T-AG--T---C-G--C--C-GGCAG---------T--A------C-A--GTGT--GCA-G-C--C-AT---GT--AT---T-------C--G--------T--A 1370
SYK.KE.x.SYK173_COMGNM      -A--A--CT-AA--GCT--G-------A-------T--T-GG...........CAG--CAAC--AG-A-A------------A----G---------C--T--A--GC-GC-G--GTGT--GCAGG-C--C-A----GTTTCT--CT-A-------A---G--C-----A 1742
TAL.CM.00.266               -A--A--C--T-CCGCT-----GG-G-A-----CC--C-GA...........GCAC-GAAC--CG-AGAT--G----C---A--C---------G-C---A----C---C-GCGCTGT----A------C-AT--AGTT-A---A-----TG-G-A------C--C-G- 1547
TAL.CM.01.8023              -A--C--C----CAGCA------G-G-AG----CC--C-GG...........CAGC-GAAT--AG-AGAT--G--GTCC--A--C--G--G---G-C---A-A--C---C-GCGCTGT----A------C-AT--AGTG-A---C-----TG-G-A------C--C-GA 1045
TAN.UG.x.SIVagmTAN1         -A--C--CT-TG--GCT--G-------AG----CTTTT---............GCT------AGAG-AGAT-----CGT--C--C-G-G-----G-T---------C-CC-A---TGT----A------C-AT----T-TCAG-C---------C--------GG-C--- 1798
VER.DE.x.AGM3               -C--C--C--AG-G--A-----GC-G-A-------ATAC-CA...........GCC-----CAGAG-AGAT--G--T-CC--C----G---G-----T--------G--GC-A---TGT----A------C-A---AGTGTCTG-CT-A-----------G--G--C--A 1339
VER.KE.x.9063               -C--------AG-G--A------T---A-------TTAC-CA...........GCC--C--CAGAG-AGAT--T--G-CT--C----G---------T-----A--GC-GC-G--GTGT----A------C-A---AGT-TCAG-AT-A-----C-----G-----T--A 1838
VER.KE.x.AGM155             -C--------AG-G--A--G--GC---AG-----CTAT-C-...........GCT--C---AGGG-AGAT-----G-CC--C---CG---------C-----A--G---C-A---TGT----A------C-AT--AGTGTCTG--T-A--T--C---------G--C--A 1827
VER.KE.x.TYO1_patent        -C--C--CT-AG-A--A-----GT---A-------CTAT-C-...........GCT--C--C-GGG-AGAT-----T-CC--C--CCGG--------T--T--A---C-TC-A--GTGT--C-A------C-A---AGTATCAG-C--A-----T--G--G-----T--A 1835
WRC.CI.97.97CI14            TA--AGCA-GTG-GG----TG--C---CC---T--G-TC-A...............--------AG-AGAT-----GTC---A------------G-CC-TAGA---C--C-G---G----TCAG--C---GTT--AGTA-----------------A------G--T--- 1863
WRC.CI.98.98CI04            TA--AGC-T-AG-AG----TG--T-G-CC---T--G-TC-A...............--------A--AGAT-------C-G-G--C-----G---G-GC-TAG---GC--C-A--GG-----CA---T---GTG--AGTT------T-A--T----AG--G--------- 1881
WRC.GM.05.Pbt_05GM_X02      TA--ATC---TA-GG-----G--T---C----C-AG-TC-G...............C-A-----A--AGATA------C---A-----------------TAGA---C--C-G--GG----GCA---T---GTA--A--ATCT--CT-A--T--C-A------G-----A  825
```

Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

```
                                                                                                                                Gag-Pol fusion TF protein start
                                            Gag p24 Capsid end Gag p2 start                                     Gag p2 end Gag p7 nucleocapsid start
H1B.FR.83.HXB2             GACAGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGGCAAGAGTTTTGGCTGAAGCAATGAGCCAAGTAACAAATTCAGCTACC..............ATAATGATGCAGAGAGGCAATTTTAGGAACCAAAGA................................ 1950
Gag                         T  A  C  Q  G  V  G  G  P  G  H  K  A  R  V  L  A  E  A  M  S  Q  V  T  N  S  A  T  .  .  .  .  .  I  M  M  Q  R  G  N  F  R  N  Q  R  .  .  .  .  .  .  .  .  .  .  .
H1O.CM.91.MVP5180          -GT---C-----A----------G--AACT--C-------A-A-AC-A--A---------GCTTCT-CCCAGC-AGATTTA-AAGGAGGATACACAGCAG--T-C-----A-----GC-GAA-CCA--TAG--A-........................ 1988
H1N.CM.95.YBF30            ---G--C--C--------G--G-----A-CA-------------GC-A--A--G--T---TCA--G--GCAGC-GC--A-A--TAGT...........G-CT-TGCA--A--G--A--C----AAGG-AT---G.......................... 1517
H1P.CM.06.U14788           --AT--C-----A--------G-----AACA--------C--G--C-----A-----T---GCAGC--CT-ATC-AG-TAGCCAAGAATTAAAA...GGAGGGTAT-CAACAGTTTTT-TGCAG--TGGA--G---....................... 1427
CPZ.CD.06.BF1167           ---T--C--------------G-----A-CA----------G--CC-A--A-----C---CAG---AC-CAGC-GAAT-T--ATATGGTACAGGGG......-GAGCAGC-CCTCAAGGAAAA--AGGGGC-.......................... 2043
CPZ.TZ.00.TAN1             C-----C-----A-----G---------TCT----------AGA--C-------------GC-TC--C----GCAGGG-GA..............GTA-AT---C-------..--AGGAAAA--ACCA-CCTT-AAA......................... 1591
GOR.CM.07.SIVgorCP2135con  --AT--T-----------G--------TACT--C--A--------C--R--A--G------GCAGC--CCCACTCAGATYTA-AA...GGGGGATATACAGC-G-CT-TAT-CA-AAGGGGGGA-RC--T---G-GATT.......................A 1502
                                           Gag p24 Capsid end Gag p2 Spacer start                             Gag p2 Spacer end Gag p8 Nucleocapsid start

MAC.US.x.239               ---G--T-----A--------G--G--G--A--G-----T---T-AA----A-----CC---AAG-G-CCCTCGCAC---TGC-AATCCCT......TTTGC-GCAGCC--ACAGA-GGGACCA............................. 2213
Gag                         T  A  C  Q  G  V  G  G  P  G  Q  K  A  R  L  M  A  E  A  L  K  E  A  L  A  P  V  P  I  P  .  .  F  A  A  A  Q  Q  R  G  P  .  .  .  .  .  .  .  .  .  .  .  .  .
H2A.SN.86.ST_JSP4_27       A--C--C-------G-----T-----A-----G--A--C---T-AA----A-----CT-A-AGG-G-CC-TGGCAC----CC--ATCCCA.....TTTGC-GCAGCC--ACAG.................................... 1702
H2B.CI.x.EHO               A-----C--C-----GA-------C--A--G--G---------GC-AA-------G--TT-A-AAG---CCTTG-CAC-TT-C---AATCCG......TTTGCCGCCGCT--ACC-A-AGCAGGG.......................... 2228
H2G.CI.92.Abt96            ------C-----A--GR----------A--G--A--A--C--GC--A-----------T-A-AAG---CCTT---CC-TA-AG--CTACCA......TTTGC-GCTGCT--ACA-AAA-CAGGAG-A.........................A 1633
H2U.CI.07.07IC_TNP3        ------T--C--A-----G-----T--A--G--A-----T---C-CA----A-----GC-A-AAG-G-CCCT-GC-C-C-GAC--CTGCCC......TTCGCTGCTG-C---CA-A-A-CAGGAG-A.........................A 1738
H2U.FR.96.12034            A-----C-----A---A----------A--A--A--A--T--GC-GA----A-----CT-A-AAG-G-CCTT-C-AC-TATGC-TATACCC.....TTTGC-GCAG-C--ACAGA-AGGCGGCCCA........................ 1748
ASC.UG.10.RT03             T--------C--A--C--------G--ACA-T-----A-C-AGC-G--A--A-----C---GCATC--CTTTC...CA-AAG--AGGAAATTACAACATGG-GCA-GGC-CAC-----CCGAAAG--GCT-GC--GGGGCCAATG........................A 1309
ASC.UG.10.RT08             T--------C--A--C--------G--ACA-T-----A-C-AGC-G--A--A-----T---GCATC--CTTTC...CA-AGG--AGGGAATTACAACATGGCGCAAGGC-CACA----CCGAAAG--GCT-G---GGGGCCCATG........................A 1324
ASC.UG.10.RT11             T--------C--A--C--------G--ACA-T-----A-C-AGC-G--A--A-----C---GCATC--CTTTC...CA-AAG--AGGGAATTACAACATGG-GCAAGGC-CAC-----CCGAAAG--GCT-GC--GGGACCAATG........................A 1309
COL.CM.x.CGU1              AG-------C--------G--------A-A---C-------AG--GC-A--A--G--T---CAG--GT-CCAGC-AGAGAGA...............ACC-AC-----TG-A............................................ 1713
COL.UG.10.BWC01            -G----------A--------------A-A-T---------A---C-----A---------CAG--GT-T-GGC-AGA-........................-C-----ACAT--T--AA.................................. 1488
COL.UG.10.BWC07            --GT-----C--A--G--G--------A-A-T----A--C-AG--CC-A--A-----C---CAAGGGA-GCA---GAT--GCCTGGTGGAA............................................................. 1725
DEB.CM.04.SIVdeb04CMPF3061 -CAT-----C--A--G-----------CAA--G--A-----GC-CA----G--G-----GT-TC--CC-TG-GGGAG-GAGGAGGC.........CTT-CC---G-A--ACAGA-G...GGACCAGGAGGGCC-..............................AGAC 1896
DEB.CM.99.CM40             ACAT-----C--A--G--G-----C--ACAG--A-----C---C--A-----G--G--T---GC-TC--CCTT---AGAG--AGGGTCC.........CTAGGC---G-A--ACA-A-G-GAGGACCCCCTGGTTCT............................ 1735
DEB.CM.99.CM5              -CAT-----C-------------T---CAA--G--A--CC-TC-GA----G--G--T---GC-TCG-CCTT---AGA---AGG-TCC.........CTGGGG---G-----CA-A-AGGAGGACAAGGAGG-G-T............................C 1727
DEN.CD.x.CD1_CMO580407     -CA---C--C--------------G--A--A--A-----CC-TCCCAC---A--G---T-AGCTTC--C-TTTC-GAAGCAAGGGTTATGCCATGGTTCA-GG-AAGG-GCCCA-CC-CCAG-AGAA-GA............................ 1732
DRL.DE.11.D3               -TT---T--C--A--T------------A-------------G--CA----A-----T-----AG--CA-CAG-CAG-T...............................------------CA-AATCCCCCC---GGT-C-CCG...............CGGGGACCCAGGGGACCCC 1380
DRL.x.x.FAO                -TT---T--C-----T------------A-------------G--CA----A-----T-----AG--CAGCAGGCAG-T...............................-----------CAGAATCCCCCC--AGGT-C-CCG...............CGGGGACCCAGGGGACCCC 1473
GRV.ET.x.GRI_677_gri_1     A-T---T--C--------------G--ACAA--------T-AGC-AA---TA---ATG-----TA-T-G-CAG...........................-AT---G-C--AGTG--ACC-CAG-AA--GGGCCCC.................................. 2054
GSN.CM.99.CN166            T--------C--A--G-----G-----AACT--C--AT-T--GC-C--A--A-----C---GCAAC--CCCTC--AGGCA---GTTCA.........TAC-AC---G------A----CCCCC-G-A-GG---............................... 1630
GSN.CM.99.CN71             T--------C--A--G-----G-----AACT-----AT-T---C-C-----A--G--C---GCAAC--C-CTC--AGGC...-G-TCA.........TAC-AT---G----A------CCCCCAG-A-GG---............................... 1633
LST.CD.88.SIVlhoest447     --GG--C-----A--------------A-C------A-G---GA-------A---------GCAAC--CC-TGC-AAGTCAA-TGAGA.........CAG-AT---G-A--AGTGACACCACCA-----TGC-CAG...GGAAGGTTTGTAAGAACAGGAGGAGGAGGCC 1159
LST.CD.88.SIVlhoest485     --AG--C--------------------A-C------A-G--AGA-------A--G------GCAAC--CC-TGC-AAGTCAA-TGAGA.........CAG-AT---G-A--AGTGACACCACCA-----TG-GCAG...GGAAGGTTTGTAAGAACAGGAGGAGGAGGCC 1159
LST.CD.88.SIVlhoest524     --A------------------------T-C------AAG--A-A-C--A--A--G------GCAAC--CC-TGC-GGGGCAG-TGAAA.........CATC-C---G-----CA-ACACCACC---A--TGC-CA-...GGGAGATTTGTAAGAACGGGGGGAGGTGGCC 1153
LST.KE.x.lho7              --A---T--------------------T-CT-----A-G--A-A-A----A----------GCTTC--CC-T-C-GCA-CAAGTAAGA.........CAA-AC---G-----GTCACTCCGC----A--TGCGCA-...GGCAGGTTTGTGAGAACAGGAGGAGGAGGTC 2243
MAC.US.x.251_1A11          ---G--T-----A--------G--G--A--A--G-----T---T-AA----A-----CC---AAG-G-CCCTCGCAC---TGC-AATCCCT......TTTGC-GCAGCC--ACAGA-GGGACCA............................. 2212
MND-1.GA.x.MNDGB1          -GA-----------------T--G--ACAA-----A--C---T-AA----A---ATG-----AAC---GGT-GGACA-......TCA.........CAA-ATT-TG-----CAGA-AGGGCC-CAA-GA....................GGACCAGTTAGACAACCTA 1593
MND-2.CM.98.CM16           -CT---C--C--A--G--------C--TAAAT----A---CA-A-GA----A--G--T----AGG---CCCAGTCA......................GCTG------------ATTCGGGAGGGCCACCG-GGG-C......................CCCCCGAGACAAC 1913
MND-2.GA.x.M14             ACT---------A-----G--G-----AAAAT--------CA-A-GA----A--G--C---CAGA-GA-GCAGTCA......................GAAG-C----------ACTCGGGGGGACCACCG-GGG-T......................CCTCCGAGACAAC 1837
MND-2.x.x.5440             -CTG--C--C--A--G-----------TAAAT-C------CA-A-GA----A--G--C---CAG----C-CAGGCAG-T.....................G-----------ACTCGGGGGGACC-CCG-GGG-T......................CCCCCGAGACAAC 1470
MNE.US.x.MNE027            ---G--T--C--A--------------A--A--A-----A--A-T-AA----A-----CC---AAG-G-CCCTTGCAC---GGC-ACTCCCT......TTTGC-GCAGCC--ACAGAAGGGACCA............................. 1693
MON.CM.99.L1_99CML1        ACAG--------------G-----T--AATG--C--AT-T--GA--C----A-----T---GCAGGG-C--T-GCCAAC...........ATGCCCATG-AT---G-A---GCCA-AGGACCACCCC-A..........................A 1612
MON.NG.x.NG1               CCA---------A--G--------C----CA------AGC--GT-A--A--A--G--T---GC-AC--C--TT---AGCAAC.........ATGCCAATG-AC---G-A--AG-CA-AGGAGG-R-TC-G-C-C-C......................CGACAAGGCRWGC  199
MUS-1.CM.01.CM1239         -CAG-----C--A--G--------G--A-CA------T-C--GC-AA----A--G--T---GC-TCG-CCCTT...AG-CAA-A-AGT...CAGCTG...-AC---G----AG--.........GCC-GAGGG-A-..................GGCAGCCAGGGAGGTC 1613
MUS-1.CM.01.SIVmus_01CM1085 -CAG-----C--------G---------ACT-----AT-C--GC--A----A--G--TC--GC-TCG-CCTTC...AA-CAGG-AGGC...ACACTG...-AC---G----AG---C---GGGAGCC-GGGG---G...............GGCAACGGCAGCCAGGGAA 1623
MUS-2.CM.01.CM1246         ACAG--C--C-----GA-T-AG-----ATG--C---T----GC-GA----A-----C----CTA-T-CTTTG...AGGCAA-A-ACCATTAACACCATC-AC---G-A--A--------.......CAG--TCC-AGG..................GGAGTAATGGGAAAAAAGA 1703
MUS-2.CM.01.CM2500         CCAG--T-----A-----T-----G--TATA--C--AT-C-A-T-AA----A-----C---GCATC--CTCT-...AG-CAG-A-ACCATTGAGACCATC-AC---G-T---G-G......CAAG-TCC-AG-..................GGAGGGGCCGGAGGAGGGC 1665
MUS-3.GA.09.09GabOI81      ACAG-----C--A--------------A-CAAGC-----T---T-A-----A-----C---GTTG-C-CGTTT--CAA-A-AGGAGCAACAGCC.............---G-C--------CAGCACCA-TACCTAG-...............................GGACCAA 1189
MUS-3.GA.11.11GabPts02     -CA---------A--------G-----A-CA--A-----T--GT-A--A--A-----C---GT-AGT-C-TTT---AA-T-AGGAGCCAGCACAGCCATG--GCA--G-GG-CCGCCTCCAGGA-------CC..............................C 1204
OLC.CI.97.97CI12          ------C-----A--------G-----A---T----A-------C--A--A-----GC----A-C---GCAG.............................---G----ACAGA-ACCAGCACCA.............................GGAGGAGGGCCTC 1798
RCM.CM.00.SIVagi_00CM312   A-----C-----A--------------ACAGT----A--C--G--CC-A--A-----C---CAA----CG-AT-GCG-T......................--C---GCT--A-TGA-AGG-GGGCAA-GA................TCAGGCCCGCCAAGAGGACAAGGAG 1180
RCM.CM.02.SIVrcm_02CM8081 -CT------C--A-----T---------CAA--C-----T-----A-----A-----C---CAG--T-CT-AT-GCATG......................--G---GCT--A-TGA-AGGGCAAG-A-GG................TCAGGGCCACCAAGAAGAGGTGGAG 1180
RCM.GA.x.SIVRCMGAB1        A-----C--------G-----------CAG-----A--TC-G--AC-A--A--G--C---CAAATGA-GCA--G-AAT......................--C---GCT---CA-TCAGCAAAC.....................AGGGGGCCTCCAAGAAGATCAGGAG 1385
RCM.NG.x.NG411             A--------C--A--G-C------T--ACAG--C-----T--G--GC-A-------------CAGATG--GCA-.....................AGTAAT--C---GCC--ACAG--AGGACC---A--AGG-CCC..................CCTAAACTAGGGGGAG 1388
SAB.SN.x.SAB1              T--T-----C--A--CA-T------G-ACAA---------C-TC-AA---G-----T----CAGCG-CCTTTC-GCA-CAA--AGTT.....GGGAAC--CT-TG-----CA---GGCAAGACCC-GGGGTCCC............TTAGGGGGAAGAGGTAGACCTC 2250
SMM.CI.79.SIVsmCI2         A-----T--C------A-------G--A--A--G--------C-GA----A-----CT----GG---CCTTTC-GC-T-GAC-GCTCCCT......TTTGCGGCAGCC---CA-A-AGGAGG-G-AGTGAGG................................ 1321
SMM.LR.89.SIVsmLIB1        ---------C--A--G--------T--A--G--G--A--C---T-AA----A-----TT-A-AGG---CTTTCC-AC-----T-TTTGCCT......TTTGC-GCAG-C--ACA-A-AGGAGGA-CA--G...................................... 1307
SMM.SL.92.SIVsmSL92A       ------C-----A-----G--------T--G--G---------T-AA----A-----CC-T-AGG-G-CCCTTGGGC--A-CCAGCTACCA......TTCGCGGC-GCC--ACAGAAGGG-.......CC-.................................... 1309
SMM.SL.92.SIVsmSL92B       A-----C-----A--G--------C--A--A--G-----CC-CC-CA----A-----C----AAG-C-CCCTC-CGGGGC--TTAGTG...............GCAGCA---TTTA-AGGAGCAGCA--AGGCCA-..........................GGGA 1292
SMM.SL.92.SL92B            A-----C-----A--G--------C--A--A--G-----CC-CC-CA----A--G--C----AAG-C-CCCTC-CGGGGT--TTAGTG...............GCAGCA---TTTA-GGGAGCAGC---AGGCCA-..........................GGGA 1663
SMM.US.04.G078             ---------C--------G-----G--A--A--A-----C---T-AA----A--G--CT-A-AAG-G-CCCTGGCAC---GCC-ACTTCCA......TTTGC-GCAGCC--ACAGA-AGGAGGACCA-GA................................ 1488
SMM.US.05.D215             A--------C--A--G--C--------A--A--A-----C--GC-AA----A--G--TT---AAG-C-CTTTGC-AC---GGCATCTGCCA......TTTGC-GCAGCC---CAGAAGGGAGGACCT-GA................................ 1482
SMM.US.06.FTq              -T----C-----A-----G--T-----A--A--A-----T--GT-AA----A--G--CT-A-AGAGT-CTTT-GCAC---GAGTACTTCCA......TTTGC-GC-GCC--ACAGAAAGGAGGACT--GA................................ 1498
SMM.US.86.CFU212           A--R--C-----A--G-----G-----A--A--A-----T--GC-AA----A-----CC---AAG---CCCT--CAC---GGC-ACTCCCA......TTTGC-GCAGCC---TT-A-AGGAGGACCA--A................................ 1474
SMM.US.x.H9                ------C--------------R--G--A--A--A--AS-T---C-AA----A--G---T---AAG-T-C-CTG-CGCA--GGC-ACTCCCA......TTTGCRGCAG-C--ACAGAAAGGACAA...................................... 1698
SMM.US.x.PGM53             ------C-----------------G--A--A--A--A--T--GT-AA----A--G---T---AAG-G-C-CTG-CGC---GGCAACTCCCA......TTTGC-GCAG-C--ACAGAAAGGACCA...................................... 2142
SMM.US.x.pE660.CG7G        ------C---------A-------G--A--A--A--A--T---T-AA----A------T---AAG-G-C-CTG-GAC---ACCAACTCCCA......TTTGC-GCAG-C--ACAG...--AGGACAA-GG................................ 2217
STM.US.89.STM_37_16        A-----------A--G--T--G-----A--A--G--A--T---C-GA----A-----CT---AAG-G--CTTCC-AC---ACC-ACTGCCC......TTCGC-GCAGCC--ACAGCAGGGA........................................ 1872
SUN.GA.98.L14              --G---C--C--A--G--T--G--G--GA-------A-G--AGA-CC----A----------GC-TCT-CTTTT-GGCAGCAAGTAGGA......AGACAAGCT---G-----CA-AA-TTACCACCC-GAA-C--CCAGGGGAGGTTTGTAAGAATAGGAGGAGGAGGAC 2236
SYK.KE.x.KE51              TCA---T--C--A--G------------ACAAT-C--A----A-C-A--A--A----------GTAATGAC-CA-C-CAGCTTG.................GGC-----A--AG-GCCA-GACAAG-A-G-A-TCCC...............................A 1656
SYK.KE.x.SYK173_COMGNM     -CA---T-----A-----C----------TTA--C-------A-C-CA----A--G--C---GTAATG-CCCAGC-G----TA...................-AT---G-A--AG-GCCATCAAAAG--................................. 2018
TAL.CM.00.266              CCAG--C-----A--G-------------CAG------AGC--GC--A----A-----T---GCAG---CCTT---A...CAA-A-ACCCAACAAGTCATAGCC---G-A--ACAG--GC-AGGAG-CCCA............AGGGGGAGAGGGGGACCTAGAAGGAACC 1872
TAL.CM.01.8023             -CAG--------A--G--C--G-----TCAG------AGC---C-CA----A-----C---GCGG-G-CCTTG--A...CAG-A-ACCCAGCAAGTCATAGCC---G-A---....--AC-GGGAG-TCCG............AAAGGGAGAGGGGGACCAAGAAGAACTC 1367
TAN.UG.x.SIVagmTAN1        ------T--------------------A--A-----A-----AGC-CA---TA---------CAA---A-GCA-GGA....................GTT-AC---G-----AG---CACCGCGAG-AGGTAGGG-C............................AGAG 2087
VER.DE.x.AGM3              A--T--C-----A-----G--------AA-TT-C--A----A---GA----A---ATG---CAAA-TA-GCA--GCCAG.....................-AC---------ACAG---GG-CAG--AGGAAG-CC-............................. 1624
VER.KE.x.9063              ---T--C-----A--------G--C--AA--T-C--A---------A----A---ATG---CAAA-CA-GCAG-C-CA-.....................-AC---G-----CA---GGCAGCAG-AGGAGT---G.........................AGAC 2127
VER.KE.x.AGM155            ------C-----A--G--G-----C--AA-TT-C--A--C-A---CA----A---ATG---CAGA-CC-GCAG-G-CAG.....................-AC---G-A---CAG--AGG-GGA---GGAAG-CC-............................. 2112
VER.KE.x.TYO1_patent       A--G--T--------G--------C--AA--T-C--A----A--AA----A---ATG---CAGACCA-GCA----CA-.....................-AC---G-----CAG--AGG-CCA-AA-GA---.............................. 2117
WRC.CI.97.97CI14           A-----------A-----------G--A--AT----AAGC-A-A-AA----AC----TT-ACAAG---CTTTG...CAT...T-ACAG............GCC---G-A--ACAG-CAGGAGCA-AACCT................................. 2142
WRC.CI.98.98CI04           ---T--------A-----------G--A--AT----AAGC-AGA-AA----AC----CT-ACAGG-G-CTTTG...AGC...T-TCAG............GC----G---ACA-CAGCCAG-G-AACCT................................. 2160
WRC.GM.05.Pbt_05GM_X02     -----------------------T--A---T----AAGT-A-A-AA----AC-----T--CAGG---CCCTG...--CCAAT-ACAG............GC----G-A--ACAGCAGGCAG-C-A-.................................. 1104
```

```
H1B.FR.83.HXB2             .....AAGATTGTTAAGTGTTTCAATTGTGGCAAAGAAGGGCACACAGCCAGAAATTGCAGGGCCCCTAGGAAAAAG................................................................................GGCTGTTGGAAA 2034
Gag                        . .  K  I  V  K  C  F  N  C  G  K  E  G  H  T  A  R  N  C  R  A  P  R  K  K  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  G  C  W  K
H1O.CM.91.MVP5180          .....GG-CCCA-A--A--C----------A-----G--A--T-T---A-A---C--TC-A--A-----A----G-................................................................................--T--C------ 2072
H1N.CM.95.YBF30            .....--ACCCA----A------------------G--C--TTTG--A-----C--T-A---------A-G-GGA................................................................................-----------G 1601
H1P.CM.06.U14788           .....---CCA--------C--T--C-----A----T---A----T---A-AG--C----A---A-----A-G--G-................................................................................--R--------G 1511
CPZ.CD.06.BF1167           .....GGA-ACA-A--A-----T---------A--GAT------T-----A-----------A-----A--A-G----................................................................................--A--------G 2127
CPZ.TZ.00.TAN1             ..AAGGGTCAGC-GC-------T--C-----G----T---C--T-----A--------T-----T--A--A--G--A................................................................................--T--C----GG 1678
GOR.CM.07.SIVgorCP2135con  AGAAGGGACCA--A--A--C-----------G----T---A--T-T---A-A------T-----A--C--A-G----................................................................................--A--C------ 1591

MAC.US.x.239               ..AGA---CCAA---------GG--------G-----G--A---T-T--A--GC-A-----A-----A--A-G-C--................................................................................--A--C------ 2300
Gag                        . .  R  K  P  I  K  C  W  N  C  G  K  E  G  H  S  A  R  Q  C  R  A  P  R  R  Q  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  G  C  W  K
H2A.SN.86.ST_JSP4_27       ..AGA-G--CAA-------C-GG-----C--A--G---------T-G--A---C-A---C-A--A-----A-G-C-A................................................................................-----C------ 1789
H2B.CI.x.EHO               ..AAG-G--CA--G-CA--C-GG--C--C-----G-CG--A--T--------GC-G----A---------A-GGC--................................................................................--A--C------ 2315
H2G.CI.92.Abt96            AAAGG-GC-CAA-A--A----GG-----------G--G--A--T----TT---C-G-----A-----C--A-G-C--................................................................................--G--C-----G 1722
H2U.CI.07.07IC_TNP3        AGAAA-GC--GA-A--A--C-GG--C--------G-----A---T----A--GC-G---C-C--A--C--A-G-C--................................................................................--G--------G 1827
H2U.FR.96.12034            ..AGG-G-CCA--C-GA----GG--C--------G-----A-----T--A--GC-G--T-AA---------A-G-C--................................................................................--A--C------ 1835
ASC.UG.10.RT03             AAAACCCT-A-CC--G---------------AC-GTTC--A----TGT---A-C-G---CCACA---G-AA--G-TA................................................................................AAG--C-TT--G 1398
ASC.UG.10.RT08             AAAACCCT-A-CC--G---------------ACG-TTT--A----TGT-----C-A---CCACAG--A-AA--G-T-................................................................................AAG--C-TT--G 1413
ASC.UG.10.RT11             AAAACCCT-ACCC--G---------------GC-GTTC--A----TGT-----C-G---CCACAG--A--A--G-TA................................................................................AAG--C-TT--G 1398
COL.CM.x.CGU1              ..GTA----CA-CA--A--C---------CAGGG-AT-------TT--------TG--TCCCAAAAGACCA-T-GGAGGAGCCGGTCGA..................GGAAGAGGCCGAGGACGGGGAGGCTTTAGAGGAGCTCCCAGAAGGCCAGTGA-G----TT-CC 1863
COL.UG.10.BWC01            ..GTA--A-CA-CA-----C--TC-G-----AG--AT---A---TTGCAG-AGC-G---CCAAAAAGAG-AGC-GGAGGATTTAGAGGAGGAGGAATGAGAGGAAGAGGAAGAGGAGGAAGAGGAGCCTGGGGAAGAGGAAGAAACCCAGGACCTCTGAAA--C-TTGG- 1656
COL.UG.10.BWC07            ..GTG-GC-CA-C--G---C-----------GC-G-T---C--TTTG--A--G--C---CCTAAG--C--AC-GTTTCCAACTAGGGGA.........AGAGGAAGCTCAAGAGGAAGAGGAATAATGAGAGGAGGAATGAGAAGAGGAGGACAAGCCAAG--C-TTGG- 1884
DEB.CM.04.SIVdeb04CMPF3061 CTGGG--T-GAA---GA--C-AT---------AC--ATT--C--TGTTCAA-AGG-C--T---AAG--C--A-C-GTT................................................................................A-A--C-TC--- 1985
DEB.CM.99.CM40             ..AGA-GACAGA---GA--C------------AC--ATT--A--TCTCCAA-AGG-C----AAAGA-----A--G-CC................................................................................AAA----TT--- 1822
DEB.CM.99.CM5              CTAGA-GACAGC---GA--C--T---------AC-GATT-----TGTGCAA--GG----T-AAAAG-----A---GTT................................................................................AAA--C-TC--- 1816
DEN.CD.x.CD1_CM0580407     .....-G-GGAC---GA---------------C--ATT--A---CT-----AGG----T--AAAG--C-A--G-GT-ACT..................................................................CAACCAGGA-CG--C-TT--- 1828
DRL.DE.11.D3               CTCCA-GA-ACCC--GA---CCA--C--C--GC-GT-T--A--T--CCTA---C-A--T-CAAAG-----A------................................................................................--A--C-TT-G- 1469
DRL.x.x.FAO                CTCCA--A-ACCC--GA---CCA--C--C--GC-GTTT--A--T--CCTG---C-A--T-CAAAG-----A------................................................................................--A--C-TT-GG 1562
GRV.ET.x.GRI_677_gri_1     ..CGAGG-CCGC-A--A--C--T------------TTT--A--T-TGCAA--GG-A----A---A--A--AC-G-TC................................................................................AAA--C-TT--G 2141
GSN.CM.99.CN166            ..GGA--A-CCCCA--A----AT---------AC--TTC--A---CT---A---G----TCCTAAA--C-AAG---GA................................................................................AAG--C-TT--- 1717
GSN.CM.99.CN71             ..GGA--A-CCCCA--A----A----------A--GTTT--A--T-TC--A---G-C--TCCCAAA--C-A-G---GA................................................................................AAG----TT--- 1720
LST.CD.88.SIVlhoest447     CTAGA---CC-T-A-CT--C--T---------A--GCC---T--T-----A----TG--T--ACAG-----AC-GG--................................................................................--A--C-----C 1248
LST.CD.88.SIVlhoest485     CTAGA---CC-T-A-CT--C--T---------A--GCC---T--T-----A----TG--T--ACAG-----AC-GG--................................................................................--A--C-----C 1248
LST.CD.88.SIVlhoest524     CAAGA-GACC-C-C-CT---------------G--GCC---A--TT----A-A--TG------CAG-----AC-GG-A................................................................................-----------C 1242
LST.KE.x.lho7              CCAGG--ACCCC-C-CT--C------------A--GCC---T---------A----TG-----ACAG--C--AC-GG-A................................................................................--T--------C 2332
MAC.US.x.251_1A11          ..AGA---CCAA---------GG--------G-----G--A---T-T--A--GC-A-----A-----A--A-G-C--................................................................................--A--C------ 2299
MND-1.GA.x.MNDGB1          CTGGA-G--AACC--TC--C-----C---AAT-----------TGT---A--GTTC-T--A---------A-GG--A................................................................................--G--C-----T 1682
MND-2.CM.98.CM16           CTCCT-GA-ACA-C-GA--CCCT--C-----A---TTT--T--TGGGCTG--GG-C--T-TAAG----C-A------................................................................................--A--C-TC--G 2002
MND-2.GA.x.M14             CCCCT-GA-ACCCC-GA--CC---C-----A-GGT-T--A--TGTTCTG---G-C--T--ACT---C--A--G---................................................................................--A--C-TT--G 1926
MND-2.x.x.5440             CCCCT-GA-ACCC--GA--CCC------C--A--GTTT--A--TGT-CTG--GG-------A-----A--A--G-GA................................................................................-------TT--G 1559
MNE.US.x.MNE027            ..AGA---CCAA---------GG--------G-----G--A---T-T--A--GC-A-----A-----A--A-G-C--................................................................................-----C------ 1780
MON.CM.99.L1_99CML1        GGAAGGGACAACCC--A--C-----------A---TTT-------TG--G-A---------AA-----CCAA-GG-G-................................................................................AAA----AC--T 1701
MON.NG.x.NG1               AAATC-G-TG-TA---T---GGG--A-T---AC-T-TG-CAA-G-ACTGT-CTGC-CC---AAAGA-AG-TTGCTTC................................................................................C----CG-C--G  288
MUS-1.CM.01.CM1239         CCAGGGGA-ACCC--G-----A---C-----GC-GTTT--A----TG--A----GC---CCCAAG--A-AA-CT-G-................................................................................AAA--C-TC--- 1702
MUS-1.CM.01.SIVmus_01CM1085 ACCGAGGA-A-CCC-GA--C-A---C-----GC--T-T--C---GT------GG-----CCCAAA--C-----C--A................................................................................AAG----TC--- 1712
MUS-2.CM.01.CM1246         GGGAA--C-GCACACGT--C-A---C-----AC--TTT--A--TTTG--A---G-----CCCAAG----A-TCT-CC................................................................................A-A----TT--G 1792
MUS-2.CM.01.CM2500         CTAGA-G--CCCCACG---C-----C-----GC--TTC--A--T-T-----A-G-----CCCAAA--A--AGTT-G-................................................................................AAA--C-TT--G 1754
MUS-3.GA.09.09GabOI81      ACAAC-GACAGCC--G---C--T---------AC-GTTT--A--TCTG--A-AG--C--TCCCAAA--C--AGT-GTA................................................................................AAA----TT--G 1278
MUS-3.GA.11.11GabPts02     AAAAG-GAGCACC--GA--C------------AC--TTT--A--TCTG--A---TC---TCCAAAG--C-AAGT-GTA................................................................................AAA--C-TC--G 1293
OLC.CI.97.97CI12           AGCAAGG--GGC-G-G-----AT---------T-G-TTT-----T--TAAA-AT-----TTCACAAAAGG-TCC-................................................................................ACA--C-AC--G 1884
RCM.CM.00.SIVagi_00CM312   GAAAG-CAT-CT-G--A----A---C-----A--GTTG-------TTT-A---T----T-AA-----C--A-G---A................................................................................--A--C-----G 1269
RCM.CM.02.SIVrcm_02CM8081  AAAGACCAG-ACCA--A----A---------A---AC---A--T-TTT----GTT---T-A---------A-G---A................................................................................--A--C-----G 1269
RCM.GA.x.SIVRCMGAB1        GAAATCCA-A-T-A-GA----A---------T--GCC---A----TTT-T---T----T-AA---------A-GG---................................................................................--A--C------ 1474
RCM.NG.x.NG411             GACCT-G-T-CC-C--A--C-AT------------AC---T--T--T--T---T------AAA-AAG---A--G--A................................................................................--G--C----G- 1477
SAB.SN.x.SAB1              TGAACCCA-A-A----A----A---------A--GCCT--T---CTG--T---TT---T-A-----------G-C-A................................................................................-----C------ 2339
SMM.CI.79.SIVsmCI2         .....-G-CCAA-C--A----GG--------G--G-----A----GT--T--GC-A--T-A---A-----A-GGC--................................................................................------------ 1405
SMM.LR.89.SIVsmLIB1        .....-G--CA----G-----GG---C----A--G-----A--TT-T--A--GC-A--T-A---A--C----G-C--................................................................................--A--C-----G 1391
SMM.SL.92.SIVsmSL92A       .....-GC-AA--C-GA--C-GG--C------------------T----G--GC-G--T-AA--A--C--A-G-C--................................................................................-----C------ 1393
SMM.SL.92.SIVsmSL92B       ATAAGCCT--CA-C-G------T--C-----A--GAC---A---T-T--G---C-G-----A--T--C--A-GG---................................................................................--A--C------ 1381
SMM.SL.92.SL92B            ATAAGCCT--CA-C-G------T--C-----A--GAC---A---T-T--G---C-G-----A--T--C--A-GG---................................................................................--A--C------ 1752
SMM.US.04.G078             ....----CA--G--A--C-GG--C-----G-----------T--T--A---C-G--T-A------A--A-GGC--................................................................................--T--C------ 1572
SMM.US.05.D215             ....-----GA----A--C-GG--C--------G-----A---T-T--A---C-G------------A--A-G-C--................................................................................--A--C------ 1566
SMM.US.06.FTq              .....C-T-CA--A-----C-GG--------A--G--G--A---T-T--A-A-C-A-----A-----------G-C-A................................................................................-----C------ 1582
SMM.US.86.CFU212           .....-GA-CA--C--A----GG--------A-----T--A-----T--A---C-G--T-AA--T--A--A-G-C-R................................................................................-----C------ 1558
SMM.US.x.H9                ..RGA-----AA--W------RG--C---R-A--G-----A---T-T--A---C-A-T---A--------A-G-C-R................................................................................-----C-----G 1785
SMM.US.x.PGM53             ..AGA----CAA-------C-GG--------A--G--G--C---T-T--A-A-C-A-----A--------A-G-C--................................................................................-----C------ 2229
SMM.US.x.pE660.CG7G        .....-G--CAA-C-------GG--------A--G--G--A---T-T--A---C-A--------------A-G-C--................................................................................-----C---GGG 2301
STM.US.89.STM_37_16        ..CGA-G--CA--A--A--C-GG--------G----C---------T--A-A-C-G--T-AA-G---G--A-G-C-A................................................................................--T--------- 1959
SUN.GA.98.L14              CTAGG---CCCA-G-CT-----T--------T---CC---A--TCT---A-ATC-G-----A-AG--A-A----GGA..................................................................CCACCCGGCTC---C------ 2334
SYK.KE.x.KE51              GGAGAGG-CCAACA-GA--C--T---------GC-GTT---A--TCTTCAA-A-G-C---CCTAGG--A-AA--GCTT................................................................................AAG--C-TT--C 1745
SYK.KE.x.SYK173_COMGNM     ..AGATCA--GA----A--C-A---------GC--ATT-------TGCAA-A-G-C----AAAAG--ACTT---GCT................................................................................AAA--C-TT--T 2105
TAL.CM.00.266              CCCCAGG-CAGA---G-----AT--C-----A---TTT--C--T-T---A-AG--C---CC------GCC--G-C--AAGGTG...............................................................CCGCCAGGCACT--C-AT--G 1976
TAL.CM.01.8023             CCCCAGGACAAA-C-GA----AT-----------GTTT-----T-TT--A-A---C--TCC---T---CA---GCCACGGGGG...............................................................CCTCCGGGGA------TC--G 1471
TAN.UG.x.SIVagmTAN1        GCAGAGGCCCACCCCGC--C-----A-----AC--ATT-----T-T-CAA-A-G----TCCAAGAG-AG-ACC---C................................................................................AAA--CCTA--- 2176
VER.DE.x.AGM3              ..AGACCCCCA--A--------A---------A---TTT--C--T-TGCAA---C-A---CCT-AA--A--A--G-T-................................................................................A-A--C-T---- 1711
VER.KE.x.9063              AGAGACCCCCAC-A--A----A------C--A---TTC--A----TGCAA--GC-A----CA-AA--A--A--G-TA................................................................................AAA--CCTT--- 2216
VER.KE.x.AGM155            ..AGACCCCCGCCA-------A---C-----A---TTT--C----TGCAG--GC-G--TCCT-AG--A--A----TA................................................................................AAA---CTT--- 2199
VER.KE.x.TYO1_patent       ..AGACCCCCAC-A-GA----AT---------A---TTT--C--T-TGCAA---C-A--TCC--AA--A-------CA................................................................................AAA---CTA--G 2204
WRC.CI.97.97CI14           ..AAAGG-CCAA---GA-----T---------AC--ATT-----T-TG--A-A-G-A--TCCCAAG--A-A----CCAGTCCCTATGAGA......................................................AGAGGACAGGGACCCACA--C---GG- 2256
WRC.CI.98.98CI04           ..AAAGGACCAA---GA--------------AC-GATT-----T-TG--A-A-G-C--TCCCAAG--A-A----CCAGTCCCCTTGAGA......................................................AGAGGACAGGGACCCACT--C---GG- 2274
WRC.GM.05.Pbt_05GM_X02     ..AGAGGACC-C-A-GA-----T---------AC--ATT-----T-TG--A-A-G----T-A-AAA--A-AA-GGTCTCCTCAGCAAAGG..................................................................GGACCTCCCAC---C---GG- 1212
```

```
                                                                      Gag-Pol -1 ribosomal slip site
                                                                               Pol start
                                                        Gag p7 nucleocapsid end Gag p1 start                                   Gag p1 end Gag p6 start
H1B.FR.83.HXB2             TGTGGAAAGGAAGGACACCAAATGAAAGATTGTACT.........................GAGAGACAGGCTAATTTTTTAGGGAAGATCTGGCCTTCCTACAAGGGAAGGCC.........AGGGAATTTTCTTCAGAGCAGACC.........AGAGCCAACAGC 2159
Gag                         C  G  K  E  G  H  Q  M  K  D  C  T  .  .  .  .  .  .  .  .  .  E  R  Q  A  N  F  L  G  K  I  W  P  S  Y  K  G  R  P  .  .  .  G  N  F  L  Q  S  R  P  .  .  .  E  P  T  A
Pol                                                                                               F  F  R  E  D  L  A  F  L  Q  G  K  A  .  .  .  R  E  F  S  S  E  Q  T  .  .  .  R  A  N  S
H1O.CM.91.MVP5180          ------C-------T----------------C-AAAAT........................-GA--------A-------------TA-------C-GGGGGGCACG-----............---C----A-G-G---A.........ACAAGT......GT-CC- 2194
H1N.CM.95.YBF30            -----GC-A--------T----------------AAAAT....................GAA-GA-------------------------G----T--C---T---A--G--A--............---A--C--C-CC----CA-C-A-AAGGAA...------C----- 1738
H1P.CM.06.U14788           ------C-------T--T-----------C--C-AATCA........................-GA--------A---------------------C-GGGGGGCAAG-----............---C--C-A-G-A----AACA-GT.........-C-A--G----- 1639
CPZ.CD.06.BF1167           ------C-------T--T--GC----G--C--C--A...................CAGTCAA-A--TGGA-TA-------------ACG-ACC--C-GG-GGGG-T-C--A--............---A------G-G----ATCAGAGTCAGACTCAGA-T-AGGA-CA 2267
CPZ.TZ.00.TAN1             ------C-A--G-----T--------G--C--C--CACCAGAAAC...............AACAGC-CTGG--TA-------------ACG-ACC--C-TG-GGGG-T-C--A--...............------C---G-G----A--CC--.............AGAG-A--G 1812
GOR.CM.07.SIVgorCP2135con  -----GC-A--G--G-------------C--C-GATCA........................-GC-------G-------------ATA-------C-AGGGGGC-AG-----............---C--C-A-G-G----AGCA-GT.........GC-A-------- 1719
                                                                        Gag-Pol -1 ribosomal slip site
                                                                                 Pol start
                                                        Gag p8 Nucleocapsid end Gag p1 Spacer start                    Gag p1 spacer end Gag p6 start
MAC.US.x.239               --------AATG-AC--TGTT---GCCA-A--CC-A..........................--C--------GGG---------CCTTGGTCCATGGGG......AAAG-A---.........CC-C-----C-CCAT-GCTCA.......................--T 2407
Pol                                                                                                  F  F  R  P  W  S  M  .  .  .  G  K  E  .  .  .  A  P  Q  F  P  H  G  .  .  .  .  .  .  .  .  S
Gag                         C  G  K  M  D  H  V  M  A  K  C  P  .  .  .  .  .  .  .  .  .  .  .  .  D  R  Q  A  G  F  L  G  L  G  P  W  G  .  .  .  K  K  P  .  .  .  R  N  F  P  M  A  Q  .  .  .  .  .  .  .  .  .  V
H2A.SN.86.ST_JSP4_27       -----C----C-------ATC---GC-A-A--CC-A..........................--A--------GGG---------TT-GG-CCATGGGG.....AAAG-A---.........CC-C-----C-C-GT-GC-CA......................-AT 1896
H2B.CI.x.EHO               ------C--C--------ATC---TC-A-A--CC-A..........................--A--------GGG---------TTCGG-CCCTGGGG.....AAAG-A---.........TC-C-----C-CCGTCCAGGC.......................-C- 2422
H2G.CI.92.Abt96            -----G---CC---C---ATT---GCCA-A---C-A..........................------------AGG---------TTTGGACCTTGGGG.....AAAG-A---.........TC-C--C--C-CCAT--C-CA......................--T 1829
H2U.CI.07.07IC_TNP3        -----C---CC-------ATC---GCGA-A--CC-A..........................--A--------GGG---------TTTGG-CCATGGGG.....AAAG-A---.........TC-C--C--C-CCAT--C-CA......................GAT 1934
H2U.FR.96.12034            ----A----CC-------A-C---GCTAGC--CC-A..........................--A-CCA---TGGG---------CTTGG-CCTTGGGA.....AAA--A---.........TC-C--C--C-CCAT-GC-CA......................--T 1942
ASC.UG.10.RT03             -----C-GA--G--C--TATGGCA-G-C-G---CGAGCCTCA...........GAAGGTGCAC--G--GGA--A--------------GCT-TTTTAGA.....CAAGCAT-TGGGGCCCAGGAA-GGCCCCAGGA-CTT.......................... 1523
ASC.UG.10.RT08             -----C-GA--G--C--TATGGCA-G-C-G---CGAGCCTCT...........GACAGTGCAC--G--GGA--A--------------GCT-TTTTAGA.....CAAGCAT-TGGGGCCCAGGAA-GGCCCCAGGA-CTT.......................... 1538
ASC.UG.10.RT11             -----C-GA--G--C--TATGGCA-G-C-G---CGAGCCTCT...........GAAGGTGCAC--G--GGA--A--------------GCT-TTTTAGA.....CAAGCAC-TGGGGCCCAGGAA-GGCCCCAGGA-CTT.......................... 1523
COL.CM.x.CGU1              --CAATC-----------ATGCAA-G-------C-C........................A-C-AG----A-----------GCA-C-A--GGGGT......--AGCTA-AGAC...............AGCCATCTT-CCC--...................... 1964
COL.UG.10.BWC01            -----CC-ACT---G---ATTCAACG---C--CCAA........................A-C-AG----A-----------GCAGG-A--GGGAT.....A-ACCT--A.........GACCGCAA-CT---CCCCAGC.....................-AT 1763
COL.UG.10.BWC07            ------C-ACTT------ATCCAAGC-ACC--CC--........................A-T-AG---TG------------CACCAG--GGGAT......--ACCTA-A.........GACAGC-G-CT-C-CCCCA-A.....................GAT 1991
DEB.CM.04.SIVdeb04CMPF3061 ---CA-C--AC---G---ATGGCA---A-C--CC-A...............AAAGCAGCTCCTC--TC-A--AA--------------AC-TCCCTGGGG......-A-G--AG-CCC...CAGGAATTTCA-AGAG---CAACAC-A...GACATGG-GCAT-GGGAA 2128
DEB.CM.99.CM40             -----GC-------C---ATTGCA---A-----GGA.....................CAGACTCCGAG---A-------------T-CTCCCTGGG-......CTCC---AGACC...CAGGAACTT-C-AGAAGGA-TTG-......GAAGTA...CATCC-G-A 1950
DEB.CM.99.CM5              -----T-----G--C---AT-GCA--GA-C---GGA.....................CAA-TACC-AG---A-------------TGCTCCCTGGG-......CTCCC--AGACC...CAGGAATTT---GGAAGG--TTG-......GAAATA...CATCC---A 1944
DEN.CD.x.CD1_CMO580407     ---------CT---T--T---GCA--GA------GA...........AGTAACACTAATGATC--G------TA-------------TGGTCTTGGGAG.....-AAG-AT--TCC...CAGGAACTT-G-A-AGT-CTA-CCT-A...GATGAATC--GATTTGC- 1974
DRL.DE.11.D3               ------GCTTT--AT--TATGT-A-G-A----C...........................CCA-A-----TG-------------T-CTCCCTGGGG.....CA-CG--AAACC...CAGGAATTT-CC--AGGCCTC.......................... 1573
DRL.x.x.FAO                ------GCTTT--AT--TATGT-A-G-A----C...........................CCA-A-----TG-------------T-CTCCCTGGGG.....CA-CG--AAACC...CAGGAATTT-CC--AGGCCTC.......................... 1666
GRV.ET.x.GRI_677_gri_1     --C--C--AATT--C--TATGGCA-----C--C-AG........................A-TG-------A-----------T-TGG-CATTGGGG.....A---GC-AAACC...AAG-AATTT-G-G-AAT-C--AG--GACACAGT...T-GT-TGGA-C- 2272
GSN.CM.99.CN166            -----C-GA-C---G---TTT-GC---CCA----GA..............ACAGGGACACCAATC-AGG--CT-T---------ACCTTGG-C--GGGCG.....CC-CC-AG-CCC......CA----C----CAGTA-CTTTC-TGCTGCC...-ACAG-CC-G-- 1857
GSN.CM.99.CN71             -----C----C-------TT-GCT-GGC-A--C-AAACA..............GGAACACCA-CC-AGG--CT----------ACCTTGG-C--GGGCG........-C-AG-CCC......CA----C----CAGTA-CTTCC-TGCTGCC...-TCAG-CC-G-- 1857
LST.CD.88.SIVlhoest447     -----T-GCA-G-A---TAGGT-TGCTC-A---C-AAAGCCT......................A--G--A---TA----------CT-TGGACCCTGGAG......-A-TG-T--TCC..........---A--C----C-TT--TGG-GGGGAA......T-CAGG-GT... 1376
LST.CD.88.SIVlhoest485     -----T-GCA---A---TAGGT-TGCTC-A---C-AAAGCCT......................A--G--A---TC----------CT-TGGACCCTGGAG.....-AATG-T--TCC..........---A--C----C-TT--TGG-GGGGAA......T-CAGG-GT... 1376
LST.CD.88.SIVlhoest524     -----TGCTA---AC---AG----GCCC-G--CC-AAAACCA......................A--G-GA---TG-----------T-TGGTCCCTGGAG.....-AATG-A--TTC..........---C--C----CAATC-TGG--GGGGC......T-CAGG-GT... 1370
LST.KE.x.lho7              -----C-GTA---AG--TAG-T-TGC-C-G--CC-A...................AAACCAA-AG--A---TG----------AT-TGGTCCCTGGAG....--GCAA-CCT--AGG......CAATT--CC---A-T-G-AG--G-.........T-CAGGG-G... 2460
MAC.US.x.251_1A11          --------AATG-AC--TGTT---GCCA-A--CC-A..........................--C--------GGG---------CCTTGGTCCATGGGG.....AAAG-A---.........CC-C-----C-CCAT-GCTCA.......................--T 2406
MND-1.GA.x.MNDGB1          ------GCAATG-AT--T--G-AAGCTC-A--CC--...............AAGCCAGCTCAGC--CAGAG--T-------------T-TGG-CCTTGGGG.....TCCCTCCAAACC......G-------A--C.................................. 1792
MND-2.CM.98.CM16           ------G-TCT-------AT-----G-A---CC-A...........................-AGAT--TG-------------T-C-CCCTGGGG.....CA-TG-CAAACC......CA----C--C-C..................G-C-ATGC--TT 2109
MND-2.GA.x.M14             -----GG-CAC------TATG----G-A-C--CC-A...........................-A-AT--TG-------------TGCTCCCTGGGG.....CA-TG-CAAGCC......CA----C----C..................T-CCATGC--CT 2033
MND-2.x.x.5440             ------G-TCC------T-T-----GGA-C--C...........................CCA-AGAT--TG-------------T-C-CCCTGGGG.....CA-TG-CAAGCC......CA----C--C-C..................T-CCATGC--TT 1666
MNE.US.x.MNE027            ------C-AATG--C--TGTT---GCCA-A--CC-A..........................--C--------AGG---------CTTTGG-CCATGGGG.....AAAG-A---.........CC-C-----C-CCAT-GC-CA......................-AT 1887
MON.CM.99.L1_99CML1        -----CC-ACCG-------TCGCC-----C--CC-GCAGCCCCCC...........AAACAAA-C-A-GG--TG-------------TCCT-TCGG-C-.....CAAG-A-GGGCC......-A----C--C-C--TC-C-TCTGTGCAGCC...CAGCG--C-GA- 1844
MON.NG.x.NG1               GAA--GC-TRT-TCCA-AA-CTGCCC-A-CG-AGGACAAAACCAA.............CCCAGA-ACAG--TG-------------TCCT-TTGGCC-...-GGCAAG---GGGCC......-A----C----CAGTA-C-TCCTTGCAGCC...-ACAG-CC----  431
MUS-1.CM.01.CM1239         -------G---G------TTGGCA---C-G--C-GG..................TCTGAAGGT-CC-A-TC---------------G-TCCTCTTGGCCT...-C--A-----AGGCC......-A----C--C-CC----C-TCG-TTTTACC...-AC-G-CC-GC- 1842
MUS-1.CM.01.SIVmus_01CM1085 -----G-GA---------TTGGCA-G-C-A----GG..................TCTGATAATACC-A-GC--TG------------G-TCC--TCGGCCT...-C--AA----AGGCC......GA----C--C-CC----C-TCT-TCTTGCC...-ACTG-TC-GC- 1852
MUS-2.CM.01.CM1246         -----T------------TTGGCC-G-C-G--CCGA...............ACCGACACAGGCA-A--CGCT-TG---------------GATCTCGGC-T...-CG-TCTC--AAGCC......GA----C--C-CAGTT-CATCG-TCATGCC...TAC-G--C-CC- 1935
MUS-2.CM.01.CM2500         -----C-G------C----T-GCA---C-G--C-GA...............ACCAACTCAGACA-AG-CGCT-T--------------AGGT-TCGGC-T...-C-CA-...-AAGCC......GA----C--C-C-GTT-CATCG-TCATGCC...-TCAG-TC-CC- 1894
MUS-3.GA.09.09GabOI81      --C--C-GA--G-----TATGGC---GC-A--C-AC....................ACATCAACTG-CAG--TG-------------TCCT-TTGG-C-...-C--AA----GGACC......-A----C--C-CAGTC-C-CAGTTGGTGCC.........G--G-- 1409
MUS-3.GA.11.11GabPts02     -----C-GA-----T---ATGGCC---C-A--C--C......................TCCC-AG-CA---TG-------------TCCT-TCGG-C-...-C--AAG---GGACC......-A-------C-C-GTC-C-CAG-TGGTACCCAC--CC---C-G-A 1430
OLC.CI.97.97CI12           -----G--ACC---C---AT-GCA---A----C-GG.........GTGGGAGGACAACAACAAC----CA---TG-------------ATGGA--GGAGG..G--A-C-GA---TCCCCTCAACAT-C-ACAGAGGACAGAG-TGGACATCTT...TCCCT-TG--C- 2039
RCM.CM.00.SIVagi_00CM312   ------G----G-AT---ATCC-T--G-----CC--..............................-AG----TA-------------TGCTCCCTGGAA......-CAGG--AAACC......GA---------C-TTA-C-TCCGTGACACCCTCT-CT---GGGAT 1397
RCM.CM.02.SIVrcm_02CM8081  ------G-T--G-AT--T-TT-----G-------C..............................-AG----T-----------A-GTGCGCCCTGGAG.....TC-GG--AAACC......GA----C----CGTT--CTTCCTTGACTCCCTCT-CT---GGG-T 1397
RCM.GA.x.SIVRCMGAB1        ------TCCCC--AC--T-TCT-------C--C--A..............................-AG--AATA------------GAC--CCCTGGGG.....TCAGG--AAGCC......GA----C----C-TT--CTTCCTTGACTCCCTCT-CT---GGGAT 1602
RCM.NG.x.NG411             ------G-------T--T-T------G-----CC-A..............................-A-G---T---------------TGCTCCCTGGAA......-CAGG--AAACC......GA----CC---CCTT--CTTCCTTGAGACCTAC............ 1593
SAB.SN.x.SAB1              -----C-GCCC--AC--T--G-----------CAA..............................-A-----T-------------TTTGG-CCCTGGGG.....CA--G--AAACC...CAGGAATTTCCC---GACCTC--TCAGACCAAC...--CT--CC--AT 2467
SMM.CI.79.SIVsmCI2         -----C--ATC---C--TGTT---GCTA-A---C-A..........................--A--------GG---------ATTTGG-CCATGGGG.....AAAG-A---.........TC-C--C--C-CCAT--C-CA......................-CT 1512
SMM.LR.89.SIVsmLIB1        -----C--AAC---C---ATC---GC-A-A--CC-A..........................------------GG---------CTTTGGTCCTTGGGG.....AAAG-A---.........TC-A-----C-CCAT-GC-CA......................-AT 1498
SMM.SL.92.SIVsmSL92A       -----T--AATG------ATC---GCCA-A--CC-A..........................--A--------GG---------ATTTGGTCCTTGGGG.....AAAG-A---.........TC-C--C--C-CCAT-GC-CA......................-AT 1500
SMM.SL.92.SIVsmSL92B       -----GG-A--------TATTCA-GCCA-C--CC-A........................A-CCAGA---GGG---------CCTTGG-CCATGGGG.....AAAG-A---.........TC-C--C--C-CCAT-CAG-C......................... 1485
SMM.SL.92.SL92B            -----GG-A-------GTATTCA-GCCA-C--CC-A........................A-CCAGA---GGG---------CCTTGG-CCATGGGG.....AAAG-A---.........TC-C--C--C-CCAT-CAG-C......................... 1856
SMM.US.04.G078             --C-----AAC---C--TGTC---GCCA-A--CC--..........................--C--------GGG----------TTTGG-CCATGGGG.....AAAG-A---.........TC-C-----C-CCAT-GC-CA......................--T 1679
SMM.US.05.D215             --------AATG--T---ACT---GC-A-C--CC--..........................--C--------GGG---------ATTTGG-CCTTGGGG.....AAAG-A---.........CC-C-----C-CCAT-GC-CA......................--T 1673
SMM.US.06.FTq              --------AATG--C--TATC---GCTA-A---C-A..........................-----------TGGG---------CTTTGGTCCATGGGG.....AAAG-A---.........TC-C--C--C-CCAT--C-CA......................GAT 1689
SMM.US.86.CFU212           --------AATG--C---GTC---GCTA-A--C--A..........................--A--------AGG---------RTTTGG-CCATGGGG.....AAAG-A---.........CC-C-----C-CCAT-GC-CA......................--T 1665
SMM.US.x.H9                --------A-C---C--TGTT---GCTA-A--CC-C..........................--A--------GGG----------CTTGG-CCATGGGG.....AAAG-A---.........CC-C-----C-CCAT-GC-CA......................GAT 1892
SMM.US.x.PGM53             --------AAC---C--TGTC---GCCA-A--CC-C..........................--A--------GGG----------CTTGG-CCATGGGG.....AAAG-A---.........CC-C-----C-CCAT-GC-CA......................GAT 2336
SMM.US.x.pE660.CG7G        --C-----AACG--T--TGTT---GCCA-A--CC--..........................--A--------GGG----------CTTGG-CCATGGGG.....AAAG-A---.........CC-C--C--C-CCAT-GC-CA......................GAT 2408
STM.US.89.STM_37_16        ------C-AC-G--C--T--G---GCCA-A--CC-A..........................--A-------TGGG---------CTTTGG-CCATGGGG.....AAAG-A---.........CCAC--C--C-CCAT-GC-CA......................-AT 2066
SUN.GA.98.L14              ---------ATG-----TA--CAAGCTC-A--CC-A.....................CAA-GA-A-----TG------------T-TGG-CCCTGGAA.....CA-GG----GCC............------C----CAGTC-TGCCTT-TGCTCC...GCC.......... 2456
SYK.KE.x.KE51              ------GGAACG------AT-GCA-GGC-G----GACAACCAAGAAAAGGACAAGGCAACCCTCCTCC------C--------------AGGG---GGGAA.....CA-G-AA--CCC........--CA-------C-GT--TG--GGAGACACC...-ACAG-CC--C- 1900
SYK.KE.x.SYK173_COMGNM     --------AAC--------T-GCA-GG-CC----GGCAACCCAAGAGAAATCAAGGCCCTCCA-TTGC---------------------AGGA---GGGGT...-TCCA-G-----CCC........--CC--C--C-CAGT---G--CGAGCCATC...T-CT--CC-TCT 2263
TAL.CM.00.266              -----C--ACC---C---AT-GCA---C-A--CT-C..........................TCAGC------C---------------TCT-TTTTAGG.....-AA--TC--ACC......G-TC-GAGGA-CGAG--ATTTT--....................... 2086
TAL.CM.01.8023             --C--GC--ATG-------GTGCCGC-C-G--C-GA..........................C--CC------A---------------GCT-TTTTAGG.....-AAG-TC--CCC......G-TC--GGGA-CGAG--A-TTC--AGTGGT...GC-A--C--C-- 1599
TAN.UG.x.SIVagmTAN1        -----G---CC-------TT-GCA--------CCGC..........................AGTG-------C-------------GA--GCCAA-C-G..........-----CAAAGCC...GCGCAATTT---AGA...-CAAG-CGGAGCCGT.........TC-GA- 2298
VER.DE.x.AGM3              -----G--ACC---G--TTT-GCA--------C-GA.............................G-----TA-----------T-TGG-C--TGGAT..........----GC-AAACC...CAG-AATTT-CCCGCCGCT-CTCTTGGGGTGGA...GCCAA-TG-GC- 1839
VER.KE.x.9063              -----G---CT-------TGGCA-----C--C-GG.............................G-----TG-----------T-TGGAC--TGGA-..........---G-CAAAACC...GAG-AACTT-C-CGC-GCC-CTCATGGAGTGGA...-CC-AGCG-GC- 2344
VER.KE.x.AGM155            ---------CC---G---TT-GCA-----C--C-GG.............................G-----TG-----------T-TGG-C--TGGAT..........---G-CAAAACC...AAG-AATTT-CCCGCAGCC-CTCTTGGGGCGGA...-CCAAGTG-GC- 2327
VER.KE.x.TYO1_patent       --------ATTG-------T-GCA-----C--C-GG.............................G-----TG-----------T-TGGAC--TGGAT..........---GGCAAAACC...GAG-AATTT-CCCGCCGCT-CTCTTGGAGCGGA...-CC-AGTG-GC- 2332
WRC.CI.97.97CI14           ------G-A-T------TGT-CA--G-A-C---C-A...................AAAATTAGA-------TA-------------TGG-A-AGGGCA......-ATGCCCT-GAACAATTATCCAGT-CAGAAG-T---G--CTTGAACAA...TTTTGGGC--A- 2396
WRC.CI.98.98CI04           ------G-AATT-----TGTGCAA-G-A-C--CC-A...................AAAAGTAGA--G---AT--------------TGGGA-AGGGCA......-ATGCCCT-AAACAATTACCCAGT-CAAAAGGT---G--CTTGAACAA...TTTTGGGC-G-- 2414
WRC.GM.05.Pbt_05GM_X02     ---------ACC--C---ATCCA----A-C--CC-A...................GAAACAAGA--G----TA-------------TGG-C-AGG..........-CAG-----.........CTCC--CAA-TA--CAGTTCA.......................... 1319
```

Draft of 30 September 2015, not for public release
plvdna_main.tex r8876–10546


```
H1B.FR.83.HXB2          CCCACCAGAAGAGAGCTTCAGGTCTGGGGTAGAGACAAC.....................................................................AACTCCCCCTCAGAAGCAGGA...GCCGATAGACA 2230
Gag                     _P_P_E_E_S_F_R_S_G_V_E_T_T_____________________________________________________T_P_P_Q_K_Q_E__P_I_D_
Pol                     _P_T_R_R_E_L_Q_V_W_G_R_D_N_____________________________________________N_S_P_S_E_A_G___A_D_R_Q
H1O.CM.91.MVP5180       AT--G-CCC-CCA-TGGAGGA-G-A-T-AAG-----------------------------------------------------ACAAGAGAATCA..GAGTCAGAAGGGGGATCA.....................AGA.....-----2263
H1N.CM.95.YBF30         ---G---CT------T-ATG---T-CA--AG--.....................................................GAAGAGCACACAGGGGAAGGAGATG-AGG----C-----GAG.....C---G- 1821
H1P.CM.06.U14788        ------TAT.....................................................................GGAGGAGGAGGAGATGACTCAGAA-AAG----R-G----AAA-GA-GAC--G- 1701
CPZ.CD.06.BF1167        GAGT-AGAC-C--GAGAAACTAGTG-A-AA--C.....................AGAACCCTCAGCTCCTCC.....................................AAT....GA-GAG-TG- 2332
CPZ.TZ.00.TAN1          GAA.....................................................................GGCTCAGGAGCAGGAGAC-G-A-AGA-A-CAGT-GT-CCAACTG-CCC-CCAC 1868
GOR.CM.07.SIVgorCP2135con -------AT.....................................................................GGAGGAGGGAGAGAAAG-A-T-AG-A-ACCAGA--AG--GG 1769

MAC.US.x.239            G-AT-AG-GGCT--TGCCA-CTG--CCCCC.....................................AGAGGACCCAGCTGTGGATCTGCTAAAGAACTACATGCAGTTGGGCAAGCAGCAG-GAGAAAAG----GAG-AAGCAGAGA--AGCCTT 2526
Pol                     _S_A_S_G_A_D_A_N_C_S_____________________________P_R_G_P_S_C_G_S_A_K_E_L_H_A_V_G_Q_A_A_E_R_K_A_E_R_K_Q_R_E_A
Gag                     _H_Q_G_L_M_P_T_A_P_P_____________________E_D_P_A_V_D_L_L_K_N_Y_M_Q_L_G_K_Q_A_R_E_K_Q_R_E_S_R_E_K_P_
H2A.SN.86.ST_JSP4_27    ---G-AG-GGCT--CACCA-CAG-ACCCCCGATAGA-----CCCAGTAGAGGACCTACTAGAGAAGTACATGCAGCAGGAAAAAGGCAGAGAGAGCAGAGAGAGAGGCCATACAAAGAAGTGACAGAGGACTTCCTG-AG-TCG-------A-ACAAGA--C-CCAT 2054
H2B.CI.x.EHO            --A...G-GGATAGTGCCATCTG-GCCCCCGAT--ACC---AGCATTCGGCATGACACCTCAGGGTGCGATTCCATCTGCACCCCCTGCAGATCCAGCAGAGGAGATGCTAAAGAACTACATGCAACTAGGGAAG-AG-AGAAGG-------A-CAGAGA--G-CC-T 2580
H2G.CI.92.Abt96          ---T-AG-GTCT--CTCCATCAG-ACCACC.....................................GATGGACCCAGCAGTAGACCTACTGAAGAATTACATGCARCTGGGGAGGAAGCA--AGGA.............C-GAGAAAC-A-CC-T 1936
H2U.CI.07.07IC_TNP3     ---T-AG-GGTT--CACCA-CAG-GCC.....................................CCCAGCAGACCCAGCAGTGGATCTGCTGAAGAACTACATGCAGATGGGCAGGAAACA--AGGA.............-AGCAG-GAA-G-CCAT 2041
H2U.FR.96.12034         ---G-AG-G-CT--C-CCA-CAG-CCC..............TCCAGCAGAACCAGCAGTGGATCTGCTGAAGAACTACATGCAACAGGGCAAG-AA-AAAAGG-------A-CAGAGA--G-CCAT 2091
ASC.UG.10.RT03          ..........................................CCCAGTAGTACAGCCATCAGCTCCACCTTTGACTCTGCAGGCAGACACTTTTCAGGAGGAGAACTTAGGCCCAGCTGTAGGGAT-CAAAAGATC-C......A--AGG.........GG 1627
ASC.UG.10.RT08          ..........................................CCCAGTAGTACAGCCATCAGCTCCACCTTTGACTCTGCAGGCAGACACTTTTCAGGAGGAGAACTTAGGCCCAGCTGTAGGGAT-CAAAGA-C-C......A--AGACAAC-AT-GGG 1651
ASC.UG.10.RT11          ..........................................CCCAGTAGTACAGCCATCAGCTCCACCTCTGACTCTGCAGGCAGACACTTTTCAGGAGGAGAACTTAGGCCCAGCTGTAGGGAC-CAGAAGA-C-C......A--AGG-AAC-AT-GGG 1636
COL.CM.x.CGU1           GAAGATGAGCA-AGA-C-GCCAC-CA-.....................................................................G- 1993
COL.UG.10.BWC01         GAGCAAGATGC---AGCAGGA-GGAC-.....................................................................G- 1792
COL.UG.10.BWC07         GAGT-AGAGGA--GAACC-TCCAAG--.....................................................................GG 2020
DEB.CM.04.SIVdeb04CMPF3061 GTACAT-CCG---G-AGGGGACGAGAT.....................................................CTTCTCCTCTTCCAGGGTACCGGA-CAGAAA--------C---C-GCA-GA-GAGACG- 2208
DEB.CM.99.CM40          GAGCA---C---TCAGG-ATAC--GAC.....................................................TACCAGGGTACCAGTACAGGGGCAGGAGAA-AAC-CCC--AG---GGAAA-A-CGAGG- 2036
DEB.CM.99.CM5           GAGCA---C---TCAGG-TTAT--AAC.....................................................AACCAGGGTACCTGTGCAAGGGGA--GGAATG---CCC--G-----CACAA---CCAGG- 2030
DEN.CD.x.CD1_CM0580407  ATGGAGCCC--T---GG-.....................................CCAGCAGCAACCACAGTACCCACAAAATTTTCAAGAGGGGGGAAACAGACAATCTTGTCAGGAAGACAAGAGGCAGAAC-AG-AGT---------CC-TCA-GA-GAC--GG 2096
DRL.DE.11.D3            .....................................GCTGACCCCATCAGCTCCTCCATTGCCAGGATACCTCCAGGAGGATCCGGCCGAGAGGATGCTCCAAAAGTACATGGAGCAGGGAGCTCAGCAG-AGAGA-AA----G...AC-ACAA-A-CAGA-G- 1698
DRL.x.x.FAO             .....................................GCTGACCCCATCAGCTCCTCCATTGCCAGGATACCTCCAGGAGGATCCGGCCGAGAGGATGCTCCAAAAGTACATGGAGCAGGGAGCTCAGCAG-AGAGA-AA----G...AC-ACAA-A-CAGA-G- 1791
GRV.ET.x.GRI_677_gri_1  AA--G-CCCCCC.....................................AATGGAAACAGCTTACGATCCAGCAAAGAAGCTCCTCCAGCAGTATGCAGAGAAGGGACAGCGCCTGAGAGAGGAGAG.....A--ACA-A-A-GGA-AC 2379
GSN.CM.99.CN166         GGA...G--G--AGAACC-CA-CA.....................................................GCTCAATCTCTCGAC-C-AGTAGGGG-ACCC---AGCAA-GGACAGA-GG 1928
GSN.CM.99.CN71          GGA...G--G-T-GAACC-CA-CA.....................................................GCTCAACCTGTCTCCGC-GGTGGGGG--CCC---A-CAA-GA-CAGA-GG 1928
LST.CD.88.SIVlhoest447  AGT----TC--CTCC-CC.....................................AATGGAGAGGAGCCCAACACAAGCAGAGAGAGCCTTAGAAACATACAGGAACCTCGGACAACAGCTG-GGAG.............GCAA-A-CAGAT-C 1480
LST.CD.88.SIVlhoest485  AGT----TC--CTCC-CC.....................................AATGGAAAGGAACTCAACAAAAGCAGAGAGAGCCTTAGAAACATACAGGAACCTCGGACAACAACTG-GGAA.............GCAA-A-CAG-TTC 1480
LST.CD.88.SIVlhoest524  GATT--TTC--C-CCACC.....................................CATGGAGAAAGTGCCTCCAACAAGAGCAGAGAGGGCAATAGAGACATACAGGAACCTAGGACAACAACTG-GGAA.............GCAA-AACA--TTC 1477
LST.KE.x.lho7           AATT---TC--CTCC-CC.....................................AATGGAGAGTGCTCCAACAAAAGCAGAGAGAGCACTGGAGACTTACAGGACTTTAGGACAACAGCTGAAAAG.........C-GCA--A-CA--T-C 2567
MAC.US.x.251_1A11       G-AT-AG-GGCT--CGCCA-CTG--CCCCC.....................................AGAGGACCCAGCTGTGGATCTGCTAAAGAACTACATGCAGTTGGGCAAGCAGCAG-GAGAAAAA----GAG-AAGCAGAGA--AGCCTT 2525
MND-1.GA.x.MNDGB1       .....................................GGCACAAGAGGTGACTCCAACAGCTCCACCATTAGAGGAGAAACCTCTGCAGAAAACTCTGAGCACTTATCAGAAATTAGGGAGAGGGCT.............CAGGCA-AA---GA-GG 1899
MND-2.CM.98.CM16        GA-C---AC...GC-CC.....................................TCCAATGCCAGGGATGGAAGACCCAGCAGAGAGAATGCTGTTAGATTACATGAAGAAAGGGCAACAGCAGAAAGCGGAAAGC-A.............AC-GGAAAA--A---G- 2225
MND-2.GA.x.M14          GA-C---TC--CACC-CC.....................................GATGCCAGGGATGGAGGATCCAGCAGAGAAGATGTTATTAGACTATATGAGAAAGGGGCAACAACAGAGGGCAGCAAAGGAGA.............CA-GCA-GA--AG---- 2152
MND-2.x.x.5440          GA-C---TC--CACC-CC.....................................GATGCCAGGGTTGGAAGACCCGGCAGAAAAGATGCTGTTAGACTATATGAAGAAAGGGCAACAACAGAGGGCAGCAGC-GGAG.............CC-GAAAGA-GA-A-G- 1785
MNE.US.x.MNE027        G-AT-AG-GGCT--CACCA-CTG--CCCCC.....................................AGAGGACCCAGCTGTGGATCTGCTAAAAAACTACATGCAGTTGGGCAAACAGCAG-GAGA.............AAGCAAAAG--AGCCTT 1994
MON.CM.99.L1_99CML1     T-TG----C.....................................AGGGACTGCGGAGCCAGTCAATCTCAACCAAGAGATTGG-T-------------A-GAC-GG--CCCCA- 1915
MON.NG.x.NG1            GGA----ACCC-RG-GGCYTACAAG-T-AACTTCT--C-.....................................TCCAGTGGGACCACC-GAGA-----------A-CGCC-A-AGCCCCG- 511
MUS-1.CM.01.CM1239      AGAG-AG--G-GAGA-C-GGTCAA-CTCTCCCC-C--GT.....................................--GTCCCAA-GAT-GAG 1898
MUS-1.CM.01.SIVmus_01CM1085 AGA--AG--G-G-GAA--GGTCAA-CTCTCTCCAG--GT.....................................--GTCCCGG-GAG-GG- 1908
MUS-2.CM.01.CM1246      AGATA---C----CCGG-G-AC-TCTCCCC-CTAGT-GGGCCCAAGGA.....................................A--GCCA--C-AG---- 2000
MUS-2.CM.01.CM2500      AGATA---C-----TAG---AC-TCTCCCC---.....................................GGTGGGTCCACGG-AGGAGAA.........CC-AACC--CCA--C-C 1965
MUS-3.GA.09.09GabOI81   --------GGACTG-GGAGCCAGTCAACT-TTC-C-.....................................GAAGGT-GGGGA.............CCGCA-AA-GAGA-GG 1474
MUS-3.GA.11.11GabPts02  TTTTGGCA-GCCAGT-AA-TTTG-CCC-AA--TAGG.....................................ACCTCC---CA-.............CTCGGG-GAAGAG-GAG 1495
OLC.CI.97.97CI12        TA-C-TG-----.....................................ACAGATTTTGGAGAAGGGAACAAG.....------AAA-AGAGAGATG- 2089
RCM.CM.00.SIVagi_00CM312 GGAT-AGT-C--CCC.....................TGCGGAAGCCATGTTAAAGAGATACTTGGAGCAGGGGAAAAAGCAGAAGCAGCAAAAGAGGCAGGA.....---GAAAAGCCAGC-GG 1495
RCM.CM.02.SIVrcm_02CM8081 GGA--AG--CA-TCA.....................TGCCTCCCAGTACGATCCAGCAGAAGCCATGCTAAGGAAATACCTGGAACAGGGCAGAATGCAGAA-CAG-AA-AGAG----G-A--GAA-G-TCAGA-GG 1513
RCM.GA.x.SIVRCMGAB1     GGAGAGCA-TT-CGA-CC.....................TGCAGAAGAGATGCTAAAGAATTATCTGAGGAGGGCAGGGGAACAAAAGAGACAACAGAGGCAGGA.....A--GAGCAA--AGAGAG 1703
RCM.NG.x.NG411          ...TG-TCC--G.....................GAGAGAGGAAGGATGGAATCCCCAGTACGATCCAGCAGAAGAGATGCTCAGGAAATACCTAGCATTGGGGAGACAGCACAAACAGGAGCAG-GGAAAGA.....A-CAA-GA--A--TGG 1718
SAB.SN.x.SAB1           GGA--GG--TT-C---AGGCC-GAG-AAAATTG-TATG-AGACAGACCTCCGACCAGAGGACCAGGGCCAGACGATCCAGCAAACAGCCCTGTTAAAGCAGTATGCTGTTCAGGGGAAACGGCAGAAACAGCAGTGGCA.....AA-CCACT--CC-C-AC 2622
SMM.CI.79.SIVsmCI2      --AG-AG-GGCT--CACCA-C-G-GCCCCC-AT-GATC-AGC.....................TGTGGCGGCGCCCCCAATGGATCCAGCTGTGGACATGTTGAGGAACTATCTGAAACTAGGTCAGCAACAG-GAGA.....-AGGAAAGAA-GGCCTT 1646
SMM.LR.89.SIVsmLIB1     G-AT-AG-GGCT--CACCA-CTG--CCCCC.....................................AGCAGACCCAGCTGTGGATCTGCTGAAGAATTACATGCAGCTAATGGGCAAGAAGCAG-GGGA.....-A-CAGAGAA-G-CC-T 1605
SMM.SL.92.SIVsmSL92A    G--T-AG-GCCT--CACCG-CAG-CCCCCC.....................................AGCGGACCCAGCAGTGGATCTGCTGAAGGACTACATGAAGTTGGGAAGGAAACAG-AAGA.....-AGCAGAGA--G-CCAT 1607
SMM.SL.92.SIVsmSL92B    ....A--TC-TT--CACCGTCAG-ACCCCC.....................................AGATCCAGCAGCCCAGATGGTCAA...GGAGTATCTACAGAAAGCACAGAGGGA--AGA-.....AAGGAG-AGC-GGCCTT 1586
SMM.SL.92.SL92B         ...-A--TC-TT--CACCGTCAG-ACCCCC.....................................AGATCCAGCAGCCCGGATAGTCAA...GGAGTATCTAGAGAAAGCACAGAGGGA--AGA-.....AAGGAG-AGC-GGCCTT 1957
SMM.US.04.G078          A--T-AGAGGCT-GTGCCA-C-G--CCCCC.....................................AGAGGATCCGGCAGTGGATCTGCTAAAGAATTACATGCAGCTAGGGAGGAAGCAG-GGGA.....-A-CAGAGA--G-CC-T 1786
SMM.US.05.D215          A-AT-AG-GGCT--CGCCG-CTG-CCCTCC.....................................GGCGGATCCAGCGGTAGACCTGCTGAAGAATTACATGCAGCTAGGCAGGAAGCAG-GGGA.....-A-CAGAGA--G-CCTT 1780
SMM.US.06.FTq           G--T-AG-GGCTT-CGCCA-C-G-CCCTCC---AGACCTAGC.....................................AAGGGATCCGGCGGAGGATCTCTTGAAGAGGTACATGCAGCAGGGCAGGAAGCAG-GGGA.....-AGCAGAGA--G-CCAT 1808
SMM.US.86.CFU212        A-AT-AG-GGCT--CGCCA-CAG-ACCCCC.....................................AGCAGATCCAGCTGTAGAYCTGTTAAAGGAATACATGAAACAGGGCAGGAAACAG-GAGA.....-A-CAGAGA--G-CCTT 1772
SMM.US.x.H9             G--T-AG-GGCT--CACCA-CTG--CCYCC.....................................AGAGGATCCAGCTGTGGATCTACTGAAGAATTACATGAAGGTGGGCAGAAGGCAG-GGGA.....-A-CAGAGA--G-CCTT 1999
SMM.US.x.PGM53          G--T-AG-GGCT--CGCCA-CTG--CCCCC.....................................AGAGGATCCAGCTGTGGATCTACTGAGGAATTACATGAAGATGGGCAGAAAGCAG-GGGA.....-A-CAGAGA--G-CCTT 2443
SMM.US.x.pE660.CG7G     G--T-AG-GGCT--TACCA-CTG--CCCCC.....................................AGAGGATCCAGCTGTGGATCTACTGAAGAATTACATGAAGATGGGCAGAAAGCAG-GGGA.....-A-CAGAGA--G-CCTT 2515
STM.US.89.STM_37_16     A--T-AG-GGCT--CGCCA-CTG-CCCTCC.....................................AGAAGACCCAGCTGCGGATCTGCTGAGAAGTTACATGCAGCTGGGCAAGAAGCAG-GAGA.....-AGCAG-AA--C-CC-T 2173
SUN.GA.98.L14           .....................TCTAGAGGATCTGACATTGGGGAACAGAATGACACCACCCCAGTCAAAGGCAGAGAGAGCTCTGGAGACCTACAGACTGTTAGGACAGGGCCTCAG.....A-CACA--A--AGAGG- 2569
SYK.KE.x.KE51           AG-GGAG--TTG-CCA-GGCA-CAACA-AC-TG-GG--A....TTATGCTTCACCCCAACAGAAGCACATAGCAATGGCCACCCCGCTACAGCCACAGTCCTCTCCAAAGACAGTCAACGGACCTCTTGGTCCCTGGG-AA--TT-.....A--GAA-AAAGA---G- 2058
SYK.KE.x.SYK173_COMGNM  GGAGGACAT.....................C--AGATGGGCCATGGCTCACATGGTCAGCACAGATGAGCCAACAAGCACAAGCGAAGGCACAGAACTCTCCCTCGAAGAAACCCCCCCACCAACAGGGAAGTGCTCT-G-.....-A-GGA-AGC-GT-GG- 2400
TAL.CM.00.266           ....GTG-TGC-ACCA-CGGCAC-GCCAA-.....................................AGAGCCTCTCACAGCATCATGGCAGGAGGGGACTACCACGTGGACCCCACCAGCC-AGAA.....-GCCC-AAGA-ACA- 2187
TAL.CM.01.8023          A--G--GAT.....................................AGAGCCTCTCACAGCATCATGGCAGGAGGGGACTACCACATGGACCCCTCCGACC-AGGA.....T-CCGCAG-ACCCA-GC 1685
TAN.UG.x.SIVagmTAN1     GG-T--CCCGAT-CCTGCGCACGGATTCCC.....................................CACAGGTTCGCCAGCTGCGGGAGCATACGACCCACGCCAGGAAACTGCTGGGAGCAGTATG-CAAAAAGGG.....A--TCAATT--G-A-GC 2414
VER.DE.x.AGM3           ---T---CCGAGTCC.....................................ATACGACCCTGCAAAGAAGCTCCTGCAGCAATATGCAGACAAGGGGAAGCAGTTGAGGGA-CAAAGGAA.....AA-ACCA--AGC--TG- 1940
VER.KE.x.9063           ---T---CCGACACC.....................................ATACGACCCAGCAAAGAAGCTCCTGCAGCAGTATGCAGAGAAGGGAAAACAGTTGAGGGA-CAAGGGAA.....-AGAACA--ACC-ACA- 2445
VER.KE.x.AGM155         ---T---CCGA-C-A.....................................CTCTACACCTTACGACCCAGCAAAGAAGCTCCTGCAGCAGTATGCAGAGAAAGGGAAACAAATGAGAAATCAGAA-AG.....AA-CCCC--AGCGA--- 2437
VER.KE.x.TYO1_patent    T--T---CCGAGCG--AC--CCC-.....................................ATACGACCCAGCAAAGAAGCTCCTGCAGCAATATGCAGAGAAAGGGAAACAACTGAGGAGCAAAAGAG.....A-TCCA---GC-ATG- 2442
WRC.CI.97.97CI14        AA--A--ATGCCTTCTGCTCCTC-AAT-AG---GATC-AAACAT.....................TCAGGAGCTGGAAAAATATTTATTGAAGGGAGCACAATTGAGGGAACAACAGGAACACAAGAA.....-ACAGA-GGACAGA-G- 2521
WRC.CI.98.98CI04        ....A--AT-CC-TCTGCTCCTC-AAT-AGG---GATC-AAACAT.....................ACAAGAATTAGAGAAATACTCCTGAAAGGAGCTCAATTGAAGGAGCAGGCACAGGCAAA.....AA-GGAAGA--AGA-G- 2533
WRC.GM.05.Pbt_05GM_X02  .....A--AT--ATGCAG--GTGAATACCTTTGGGC-AGTGGT.....................GCCAACAGCACCAGTGGATCCTACCATGCAAGAACTAGAGCAATATCTTCTGAGAGGAGCTCAGATGAAGAGGGA.....-A-AGA-AAAGA-ATGC 1450
```

```
                                          Gag-Pol TF end  Pol protease start          Gag end
                                                                                      Gag p6 end
H1B.FR.83.HXB2             AGGA......ACTGTATCCTTTAACTTCCCTCAGGTCACTCTTTGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTG...........CCAGGAAGATGGAAAC 2382
Gag                        K  E  .  .  L  Y  P  L  T  S  L  R  S  L  F  G  N  D  P  S  S  Q  *
Pol                         G  .  .  T  V  S  F  N  F  P  Q  V  T  L  W  Q  R  P  L  V  T  I  K  I  G  G  Q  L  K  E  A  L  L  D  T  G  A  D  D  T  V  L  E  E  M  S  L  .  .  .  .  P  G  R  W  K
H1O.CM.91.MVP5180          ..........G-----C--A--TG-C-------AA--C--------G-CA----AA-A------GC----G-T-----T--T---TGT--G---T--C-G--------G-----------------A-TA-C--ACAA--A...........GA----------C-- 2411
H1N.CM.95.YBF30            -CTCTCTGTA...CCCA-----------------A-------------------G--A-----G----A------AAAG--G---GA--------T-----------------T-----------------A-------GC-ACAA--A...........GAG----A-------- 1976
H1P.CM.06.U14788           -A--GTT.......---C------G-C-------AA-----------G-CA----AG-AA-AC--G----A-----------TA-TTGT--G-----------------G--T-----C------G----TA-CT-ACC---A...........GA-------------- 1853
CPZ.CD.06.BF1167           -TC-TTGGAACAGCA-GGA-GGCTT------G-AA--CA---------G-----AA-GA-GGA-G---CAG-GCA---A--GG-CTGCAC---C--T-----C-----G--T---------------T-C---CC-T-AG--A...........AA----TC---C--C- 2490
CPZ.TZ.00.TAN1             T---GATGACGA--A-AGGCGGGTTC-------A----A---------G-----AA-GA-G---G----AG-CCA---A---G-CTGTC-------T--------T-----------C-GT--T--TTGTA-C--C-AA--A...........AAG---CAG----C-- 2026
GOR.CM.07.SIVgorCP2135con  -ACT.........T------C------------AA-----------G-CA----AA-AA-AC-TG-G-----------A--T---TGT--G---T-G-----C-----G--T-------------G-----C-CT-ATC---A...........GA---T--G------- 1918

                                                                            Gag end
                                              Pol Protease start        Gag p6 end
MAC.US.x.239               -CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T--------AGGA----AG-A-----TGCTC-T--T-AA--A--G-CTGTA----TAT--C-G--------G--T------T-TA-TG--AC--G---AGAG--A...........GGTCC-CAT-AT-CC- 2684
Pol                         L  Q  G  G  D  R  G  F  A  A  P  Q  F  S  L  W  R  R  P  V  V  T  A  H  I  E  G  Q  P  V  E  V  L  L  D  T  G  A  D  D  S  I  V  T  G  I  E  .  .  .  .  L  G  P  H  Y
Gag                        Y  K  E  V  T  E  D  L  L  H  L  N  S  L  F  G  G  D  Q  *
H2A.SN.86.ST_JSP4_27       GCAGAGAGACGACAG-GGAC--GCTGCA-----AT--T--------A--A----AG-A------GC-C-TG-T-A---C--G-C-GTA----T-T-GC-----C-----G--T--C--CT--A--G---C--GCG-AGAG--A...........GGGA-C-AT-AT-GT- 2212
H2B.CI.x.EHO               -CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T--------AG-A----AG-A----A-GC--CT--T-A---T---TC-GTA----TAT--C-----C---------T-----CT--A--G---C--GG--AGAA--A...........GGCA-C-AT-AC-CC- 2738
H2G.CI.92.Abt96            -CA-GGAGGTGACAG-RGAH--GCTGCA------C--T--------AG------AG-A-----TGC-T-C----AAA-T--G-C-GTA----TA--------C-----G--T-----CT-TA--G---C--G---AGAA--A...........GG-A-T-AT-AT-GT- 2094
H2U.CI.07.07IC_TNP3        -CA-GGAGGTGAC-G-GGA---GCTGCA-----AT--T--------AG------AG-A-----TGC-C-C----A---T--G-CGGTA----TAT-------C-----G--T--C---T-TA--G-G-C--G---AGAA--A...........GG--T--AT-ATC-C- 2199
H2U.FR.96.12034            -CA-GGAGGTGAC-G-GGA---GCTGCA-----AT--T--------A-------AA-A-----TGC-T-C----A--------C-GTAA---TA---C-GA----C--G--T------T--A--G---C--GG--T-AA--A...........GA-CC-TAT-AT-CT- 2249
ASC.UG.10.RT03             --A-GACAAGGGA--C--T--ATC-C-----GTCA--C--------GG------AA-A---T-CTGCTTAG-G-A-A-----GA-GTT--CCTA-----G--C--G--T-----------A--A-CC-T--TCA-GACA-AGAACTA......GG--AT-AC---C-C- 1791
ASC.UG.10.RT08             -AA-AACAAGGGA--CC-T--ATC-C-----GTCA--C--------GG------AA-A---T-CTGTTTAG-G-A-------GA-GTT--CCTA-----G--C--G--T-----------A--A-TC-T--TCA-GACA-AGAATTG......GG--AT-AC---C-C- 1815
ASC.UG.10.RT11             --A-GACAAGGGA--CC-T--ATC-C-----GTCA--C--------GG------AA-A---T-CTGTTTAG---A---A---GA-GTT--CCTA-----G--C--G--T-----------A--A-TC-T--TCA-GACA-AGACTTA......GG--AT-AC---C-C- 1800
COL.CM.x.CGU1              G---GAAGGGGGA-AG--TC-ATC-C---T-G-A---T-----C--AG------AA-A--T-T-G--G-AG---AA--A---AAGTGCC----G--G-------------------------T--TA-----GAAGA-G-TCAA......TTAA----CGCT----C-G 2157
COL.UG.10.BWC01            --AGAGACGAGAACAC--TG-ATC-C---T-G-----T-----C--AG-A----AA----G-T-GC-G-A----AA------AA-TGTC-G-TC--C-------------------C------A--A-----TCAGA-A-AGAT......ATAAA---G-CC----C-G 1956
COL.UG.10.BWC07            --ATGGACAGGGAAGGAATC-ATC-C---T-G-AA--C--------AG-A----TT-T---TGGGC-G-AT---AA--A---AG-TTCC-G-TG---G----------G----------C------C--A-----GAAGAGA-TCAA......TTAA----C-CC----C-G 2184
DEB.CM.04.SIVdeb04CMPF3061 G-AGAGGGAA-TACCCAGA----T-GGA-T-G-A---C--------AG-A----AG-AAC-------G------A---AAG-AGGGT-C-G-T---------------G-----------TA--A-TC-T--G-AAGA-A-T......CAACTTA-CTC-CCC-TTGTC- 2372
DEB.CM.99.CM40             GA-GGAAGTA...CCCGGA-----TGGGG--G-AA--T--------AG-A----AA-GAC----G--C-C----AA--AAG-AA-GTCC-G-TG---A----------G-----------CA-TA-TTC-----AAGA-A-A......GATTTAGAG-C-CCT----CC- 2197
DEB.CM.99.CM5              GA-GGAAGTA...CCCAGA-----TGGGG--G-AA--T--------AG-A----AA-GAC-G--G--T-TG-G-AA--AA--AGGGTCC-G-TG---A-----C-----T-----------A-TA-TTC---G-AAGA---A......GATTTAGAT-C-CCT----CC- 2191
DEN.CD.x.CD1_CMO580407     --ACAACGAGGGA-AG--T--ATC----GT-G-A---T--------AG------AA-A-GGCAG--CTG-G-TCA---A---AGGGTA--G-TC--TA----------G--------C--CA-TA---C----CAAGA-A-A...AATTTGGGAA---AGG------TC- 2263
DRL.DE.11.D3               ---GCCCCTACAA-G-GG-C-AC-AC--------T--T-----C---TCA----AACTACA-T-G--G-A----AA------AAGGTA--G---T--C----C-------------------------A-TC---G-T--GAA--A...........A------A------G- 1856
DRL.x.x.FAO                G--GCCCCTACGA-G-GG---AC-AC--------T--T-----C---TCA----AACTACA-T-G--G--G---AA------AAGGTA--G---T--C----C-------------------------A-TC---G-T--GAA--A...........A------A------G- 1949
GRV.ET.x.GRI_677_gri_1     --A-GGAGAA-GAAGTGGAGGATGT----T-G--C--C--------AGGA----AA-GAAACG-G-C-TC----AA--AACG-C-GT-C----CT-G-------------------C--TA--A-TC------A-GAC---CACTTTCCCCCA-ATAA-CC----CGTT 2549
GSN.CM.99.CN166            ----GAGCAGGAACAC--TC-ATC-C-----GTCA--T--------GG-A----AA-GAAGGA-G-CTCC--------A---G-GGT-TCCATGT-------C-----G--G--C--C--CA-CG-CC----CTCAGCAA-T......GAATTAGATAC-CC-----C-- 2092
GSN.CM.99.CN71             CA--AAAGAGGAACAC--T--ATC-C-----GTCA--T--------GG-A----AA-GAGAGA-G-CTCC--G-----A---G-GGT-TCCATGT-------C-----G--G--C-----CA-TG-CC-----TCAGCAA-T......GAATTAGACAC-CC-----C-- 2092
LST.CD.88.SIVlhoest447     CAC-GAAATGTG--G-CGAGCC-TGCCTGAGTTTC-TCT-----CCAG------AA-A-AA-AC--TCT-G-CAAT-----G-C-GT-TCT-----TC-T--C--T--G-----------TA-C--TTC-----ATTCAG-T......AGAATAGAG---CCT-AT-C-- 1644
LST.CD.88.SIVlhoest485     CAC-GAAATGTG--G-CGAGCC-TGCCTGAGTTTC-TCT-----CCAG------AA-A-AA-AC--TCT-G-CAATA----G-C-GT-TCT-----TC-T--C--TA-G-----------CA-C--TTC-----ATTCAG-T......AAAATAGAG---CCT-AT-C-- 1644
LST.CD.88.SIVlhoest524     CAC-GAAGTGTG--G-CGAGCCTTGCCTGAATTTC-TCT-----CCAG------AA-A-AG-A-G--CT-G-CAAT-----GA-GTAACT--C-----G--C--T------------A-C--TTCT----ATTCAG-T......AGAATAGA---CCT-AT-CT- 1641
LST.KE.x.lho7              CAC-GAAATGTG--G-CGAGCCCTGCCTGAATTCACTCT-...-CCAG------AA-A-AA-GT---TT-G-CAAT-------C-GT-TC---A--G--G--C--T--G--T-----C---A-T--TTC-----ATTCAG-A......AGAATAGA----CCC-AT-CT- 2728
MAC.US.x.251_1A11          -CA-GGAGATGACAG-GGA---GCTGCA-----AT--T--------AGGA----AG-A-----TGCTC-T--T-AA--A--G-CTGTA----TAT--C-G--------G--T------T-TA-TG--AC--G---AGAG--A...........GGTCC-CAT-AT-CC- 2683
MND-1.GA.x.MNDGB1          ----GAAGAG-GA-G-GGA---TCA------G--TA-T------CAAG-A----AA-A-AAGAGG-CTCAG-G-AT--TGTCACT-TAAG-------C-----------T---------CA-C--TA-T----GA-A-A-A......AAATTAAA-----AT---C-G- 2063
MND-2.CM.98.CM16           G--GTCCATACGA-GCAG---AC-AC--------T--T--------A-CA----AACTACAGT-G--G------AA------AA-GT---G--CT--C-G--C--------T-------------A-CA----TC-AGA---A...........AA---T-AT------- 2383
MND-2.GA.x.M14             ---GTCCCTACGA-GCAG---AC-AC--------T--T--------A-CA----AACTACAGT-G--G-A----A---A---AAGGTA-----C---C-G--C--T-------------------A-CA----TG-AGA---A...........A------AT------- 2310
MND-2.x.x.5440             ---GTCCCTACGA-GCAG---AC-AC--------T--T--------A-CA----AACTACAGT-G--G-A----A-------AAGGTA--G--A--TC-------T--------C----------A-TA----TT-AGAG--A...........A------AT-----G- 1943
MNE.US.x.MNE027            -CA-GGAGGTGGCAG-GGA---GCTGCA-----AT--T--------AG-A----AG-A-----TGCTC-T--T-A---A--G-CTGCA----TA-----G--------G--T------T-TA-TG---C--G---AGAG--A...........GGTCC-CAT-AT-CC- 2152
MON.CM.99.L1_99CML1        ----GACGAGG-G-G---T--ATC-C---T-AGCA--T--------GG-A----AA-AAGAT-TG-GCT---------A---AAGGT-TCCATG-----G-------T--C--C--C---A-TA--C----TCA-CACG-A......AATTTAGGG-AG-AT----C-- 2079
MON.NG.x.NG1               ----GACGAGGAA-G---T--ATC-C---T-AGCA--C--------GG------AG--AGA---G--G-A--C-CT-----GTCCGT-TCCATGT----G--C--G--R--G--C-----CA-CG-CC----RCAAGACA-A......AATCTAGG--A--AT----C--  675
MUS-1.CM.01.CM1239         C-AGCGACAGGAA-AG--T--ATC-------GTCA--T------CAGG-------A-TAGAGA-G---T-----AA--A---G-GGT-C-CATG--GC-C--C-----G--------C--CA--G-CC-G--TTCC-ACA-T......CAGATAGATAA-CC------T- 2062
MUS-1.CM.01.SIVmus_01CM1085 -AC-GGACAGGAA-AG--TC-ATC-------GTCA--C------CAGG------AG--AGAC--G--G------AA-----GG-GGT-TC-ATG--GC-C--C--C-----------C--CA--A-CC-G--TGCCCA-A-A......CACCTAGA----CC------T- 2072
MUS-2.CM.01.CM1246         G-A-AGACAG-AA-AC--TC-ATC-------GTCA--T--------GG------AA--AAA---G--G-C--TAAT-----G-C-GT-TCCATG--CC-G--C--C-----G-----C--CA-CA-CC----CTCAGA-A-A......AAAATTAGT--GGA--T--CT- 2164
MUS-2.CM.01.CM2500         CACCGGACAGGAA-A---T--ATC-------GTCA--C--------GG-------A--AGA---G---T-G-CCAT--T--G-CGGT-TCCATG--CC-G--C-----G--------C--CA-TA-TC-G--GTCTGAAA-A......GAAATAGATAA-GA-----CT- 2129
MUS-3.GA.09.09GabOI81      -A--AACCAGGAAAAC--T--ATC-------GTCA--C--------GG------AA--AGA---G--G------A-------AGGGT-TC--TC---C-----C-----G--T-----C--CA--A-------GTCTCACA-A......GTCTTAAA---GGA-----CT- 1638
MUS-3.GA.11.11GabPts02     --ACGACAAGGAAC-C--T--ATC-C-----GTC---C--------GG-------T--AGA---G-G-GC----AA------GAGGT-TCG-TC-----G--C--G--G--C--C-----CA-CA----G---GGA---G-C......CAAATTAAG--G-------CC- 1659
OLC.CI.97.97CI12           ----GGAGAAGAAACC--T--ATC-C-----AGCA--T--------AG-A----AA-GA-AGAGC-GT-T----AA--A--G-A-GTCACT--A--G-------------G--------C------A-TA-G-GG-A-GA-A-C......TACATGGA---T-ATG---TT- 2253
RCM.CM.00.SIVagi_00CM312   G-AGAGCATACGAAG-GG-A--GT-C-------AT--T--------A-------AG-AAC-TGGG-C-CC----A---T---AA-GTA-----C--T-----C-----T--T-----------CA-CAGT--C--AGAA--A...........GG--AT-AT-----G- 1653
RCM.CM.02.SIVrcm_02CM8081  G-A-AGCCTACGA-G-AG-AC-TG-C-------AT--T--------A-G-----AG-AACT-TGG-C--T----AA------AAGGTA-----C--C-----------G--------C------A-TA-T-GG--AGAA--A...........GG--A--AT-----G- 1671
RCM.GA.x.SIVRCMGAB1        ---GAGCATAT-A-G-AG-C-----C-------AT--G--------A-G-----AACTACA-T-GCTC-A----AA-----GAA-GT---G-TC--------C---------------C------A-T----G---AGAA--A...........GG-AATGAT----CC- 1861
RCM.NG.x.NG411             G-AGAGCCTACGA-G--G-G--GT-C-------AT--T--------A-G-----TG-AACATGGG--G-A----A-------AAGGT---G--CT---------------------G-----T---G------TA--C--A-TT-AGA---A...........GG--A--AT-----G- 1876
SAB.SN.x.SAB1              --AGCCCCTACGA-G-GG---AC-GC----------T--------AG-A----AA-GAAA--TG-GT-C----A---A--GAA-GTCACT-----------C-----------CT----A--C-G-GT--AGAA--A...........GG--AC-AT------- 2780
SMM.CI.79.SIVsmCI2         -CA-AGAGGT-AC-G-AGAC--GCTGCA-----AT--T--------AG------AG-A-----TGC-T-C----A---T--G-C-GTA----TA-----G----------T-----CT-TA--G-G-C--G---AGAA--A...........GGGCT-G-T-AT-CC- 1804
SMM.LR.89.SIVsmLIB1        -CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T--------AG------AG-A------GCCC-C--T-A---T--G-CTGT-----TAT-------C-----G--T--C---T--A-TG---C--G---AGAA--A...........GG-CC--AT-AT-CC- 1763
SMM.SL.92.SIVsmSL92A       -CA-GGAGGTGAC-G-GGAC--GCTGCA-----AT--T--------AG-A----AG-A------GCTC-T----AA--A--G-CTGT---G-TAT-------C---------T-----CT--A--G---C--G---AGAA--A...........GG-TT--AT-AT---- 1765
SMM.SL.92.SIVsmSL92B       -CA-GGAGGTGAC-G-GGA---GCTGCA-----AT--T-----C--AG-A----AG-A----G-GC-CTA----A---T--G-CCTTA--G-T----C-G--C-----G-----C--CT-CA--G---C--GG--AGAG--A...........GGTCC-G-G-AC-C-- 1744
SMM.SL.92.SL92B            -CA-GGAGGTGAC-G-GGA---GCTGCA-----AT--T-----C--AG-A----AG-A----G-GC-CTA----A---T--G-CCCTA--G-T----C-G--C-----G-----C--CT-CA--G---C--GG--AGAG--A...........GGTCC-G-G-AC-C-- 2115
SMM.US.04.G078             -CA-GGAGGTGAC-G-GGA---GCTGCA------T--T--------AG-A----AG-A------GCTT-C--T-A---T--G-CTGT-----TAT--C-T--C-----G--T--C--CT--A-TG-G-C--GG--AGAG--A...........GGGCC-CA--AT-CC- 1944
SMM.US.05.D215             -CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T--------AG-A----AA-A------GCCT-T--T-AA--T--G-CTGT-----TG---C--------------T--C--CT--A--G---C--GC--AGAG--A...........GG-CC--AT-AC-CT- 1938
SMM.US.06.FTq              -CA-GGAGGTGAG-G-GGA---GCTGCA------T--T--------AG-A----AG-A------GCCC-T--T-AA--T--G-C-GTA----TAT-------C-----G--T--C--CT--A-TG---C--G---AGAA--A...........GG-TT--AT-AT-CC- 1966
SMM.US.86.CFU212           -CA-GGAGGTGAC-G-GGA---GCTGCA-----AT--T--------AG------AG-A--G---GCGC--G-C-A---T--G-CTGT-----TAT--C-G--C-----G--T-----CT-TA--G---C--G-G-AGAA--A...........GG-AC--AT-AC-CC- 1930
SMM.US.x.H9                -CA-GGAGGT-AC-G-GGA---GCTGCA-----AT--T--------AG-A----AG-A-----TGCCT-C--T-AA-AA--G-CC-TA----TAT------------G--T------T--A-TG---C--GG--AGAA---...........GGTCC--AT-AT-CC- 2157
SMM.US.x.PGM53             -CA-GGAGGTGAC-G-GGA---GCTGCA-----AT--T--------AGGA----AG-A---G-TGCCT-C--T-AA-AA--G-CCGTA----TAT------------G--T------T--A-TG-G-C--GG--AGAAC-A...........GGTCC--AT-AT-CC- 2601
SMM.US.x.pE660.CG7G        -CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T--------AG-A----AG-A-----TGCCT-C--T-AA-AA--G-CCGTA----TAT------------G--T--C---T--A-TG---C--GG--AGAA---...........GGTCC--AT-AT-CC- 2673
STM.US.89.STM_37_16        -CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T--------AG-A----AG-A------GCCC-T--T-A---T--G-CTGTA----TA-----G--C-----G--T--C--CT--A--G-G-CG-GG--AGAGC-A...........GGGCT-CA--AC-CC- 2331
SUN.GA.98.L14              ----AAGGGGGGAA-G--AGGAGC-A-G-T-G-ATATG------CCAG-A-----T-A-AG-AGG-CTTC--TAAT-------CTGT-ACT---T----G--C--G-----------C---A--G-GTCT---GATTC-G--......TCCATAGA---TCAC-----G- 2733
SYK.KE.x.KE51              T-A-GGAGAAGGGACC--T--ATC-C-----TCAA--T--------AG------AA-GA-GGA------TG---AA---ACCAAGGTAAC-ATGT-GC-G--C-----G--------C--CA--A--AG----CAAGACA-C......CAGTTA-ATCA-CCT----CC- 2222
SYK.KE.x.SYK173_COMGNM     ----GGAGACCAA-AG--TC-ATC-C--T-----C--T--------AG-A----AA-GA-AGA-G--G-TG---AA--TG-C--GGT-C--ATGT--G----------G----------TA--A-TAG----GAAGA-A-T......CAATTA-ACCA-CC----TC-- 2564
TAL.CM.00.266              CCACAACAAGGGA-AG--TC-ATC-C--TT-GGCCAGT--------AG------AA-A----A-G-GCT-----AA--A---GAGGT---G-TA--CC----C--------------C---A-T--A-TA--C---T-GGAGAAA-A......GAATTGGA----GT-CCCC-C- 2351
TAL.CM.01.8023             -AC-GACGAGGGA-AG--T--ATC-C--GT-AGCCAGT--------AG------AA-A--A-AGG-GCTC----AA--AAGGGAGGT---G-TA--CC---------G--------C---A-CA--C-----GG-GACA-A......GAATTAGA----GTGCCCC-C- 1849
TAN.UG.x.SIVagmTAN1        --A-GGAGAAGGAACTGGAGGACTA----T-G--T--C--------AGGA----AA-GAGT-AG--CT-C---CA---C--G-C-GTAC---T---G-------------T--------CA--A--A-G---GAAGA-A-C......CATTTAGA----C-C---CGG- 2578
VER.DE.x.AGM3              -TCCCGATTGGACAG-GGGA-ATT---T...G-AC--C--------AG-A----AA-AAAA---G-TT-C----AA---GTC-CC-TCAG---AT----------G--G-------------CA-TA--A-----GCAGA---A......CAATTAT------C-------- 2101
VER.KE.x.9063              -TCCAGACTGGGC-G-AGGA-ATT---T...G-AC--C--------AG-A----AA-AAAA---G-C-TT--T-AA--AGTC-CTGTAA---TC--G-----------G--------C--CA-CA--A------CAGACA-A......CAACTAA-----CCT----G-- 2606
VER.KE.x.AGM155            -TCCAGATTGGAACG-GGGA-ATT---T...G-AC--C--------AG-A----AA-AAGG--CTGT-TA-----A--AACTGCCGTTA-G--AT-------------G--------C--TA--A--A-G--T-CAGA---A......CAATTAAGG---TC-----G-- 2598
VER.KE.x.TYO1_patent       -TCCGGATTGGACCG-GGGA-ATT---T...G-AC--C--------AG-A----AA-AAAG---G-GT-T----AA---GTC-CC-TTA-G--A--GC----C-----G--------C--CA--A-TA------ATGA---A......CAATTAT----TCC-----G-- 2603
WRC.CI.97.97CI14           G---AGTGGGGG--CCGGAA-GGC-C-------AA--T-----C-AACCA----AA-GA-ACA-TGC-TA----AA--A--TAGTGTCTC---AT--C---------------------------------A----T-T--ATTTA...............AATAAG-AT---C-G- 2676
WRC.CI.98.98CI04           GA--AACGGGG---CCAGAG-GGC---------AA--T-----C-AA-CA----AA-GA-ACA-TGT-TA----AA--AA-TAGTGTCTC---CT-G---------------------------------A----TATCCATTTAAAT............AATAAG-AT---C--- 2691
WRC.GM.05.Pbt_05GM_X02     --A-GGAGGGGAAAGTAGAA-ATC-C-------AA--T-----C-AACCA----AA-GA-ACA-TGT-TC--T--TC-AG-GAGTGTCTCT--GT--------------------C--T---G----T-TTCAATTA...............GA-----AT---C--- 1605
```

Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

Pol Protease end Pol p66 and p51 RT start

```
H1B.FR.83.HXB2            ACTTTAAATTTTCCCATTAGCCCTATTGAGACTGTA.........CCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAAAAGCATTAGTAGAAATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAAT 2698
Pol                        T  L  N  F  P  I  S  P  I  E  T  V  .  .  .  P  V  K  L  K  P  G  M  D  G  P  K  V  K  Q  W  P  L  T  E  E  K  I  K  A  L  V  E  I  C  T  E  M  E  K  E  G  K  I  S  K  I
H1O.CM.91.MVP5180         --AC-------C--T--A--T--C--A-CCC-A--G.........--------GC---A-------------A-------A----------CC-AT-TAG---G-----G-----C--ACT-C---A---CA---A-----C-------A-----C----G--- 2727
H1N.CM.95.YBF30           --------------A--A--T--------A-------.........-------------A-----------------G--A----------T-----AC---------G-G-----AG--------------A-------------A--------T-G--- 2292
H1P.CM.06.U14788          -----G-G---C-----C-----C--A--A-TA---.........-------------AG------------G--T-GG--G--G-------C---T--A-G---------G----T---AG--CC--C---CA---------C-------A-----AA------- 2169
CPZ.CD.06.BF1167          ---C-------C--T-----TA-AG-A--A-TA---.........AA---G--GC-C---GA------------G--T--G--A-----------C-AT-CAG--G-----TG----CC--AC----------GA-A-TT-AAG-GCA-----------CATTG--G- 2806
CPZ.TZ.00.TAN1            --C-----------T--------A--A---GTA---.........AA----C------AGA------------G-------A--G--G----CC-CT-CA-G--G-----TG-G------AC------A---AGACAT------------A--------TGC-G- 2342
GOR.CM.07.SIVgorCP2135con --C--------------A---T-C--A--TGTA---.........--C---TCT-----A-----G--------A--C-----C--G--------T---T--A-----------G----C---AC--CT------CAG--A------C-------C--G--A-------- 2234
```

Pol Protease end Pol p66 and p51 RT start

```
MAC.US.x.239              T--C-------------AGCTAAAG-A---C-----.........AA---CGCC------------A-----A-----T-G--G--G-------AT--A-------G---GTT------AG------C---GA-A------------T--TC-GT-GGAGG--GC 3000
Pol                        M  S  L  N  F  P  I  A  K  V  E  P  .  .  .  V  K  V  A  L  K  P  G  K  D  G  P  K  L  K  Q  W  P  L  S  K  E  K  I  V  A  L  R  E  I  C  E  K  M  E  K  D  G  Q  L  E  E
H2A.SN.86.ST_JSP4_27      T-A------C-A--AG-CGC-AAG--A--AC-AA--.........AA-A---TGC-G----------A-----A-----C-G-G---------C--A---A-----------G-G---C--AA---G--C---GAGA-A-------GA--G--CC-GC-AGAGG-GGC 2528
H2B.CI.x.EHO              ---C-------C--AG-AGCAAGG--A--AC-A---.........AA---CC-G--------T-A--AA-----G-----A-C-G---------CC-AT-CA----G------CT---CC-CAA------C---GA-A-A--------A--G--AC-GT-AGA-G-GGC 3054
H2G.CI.92.Abt96           T--C-----------A---GCAAAGG-A---C-A---.........AA----GGC----A--T--C--------G-----A-C-GG--------CC-AT-TA-G---------C-----C-CAGG---------GACA-A------C---------C--T-GGA-G--GC 2410
H2U.CI.07.07IC_TNP3       T--C-------C--AG-GGCAAAGG-A--A--A---.........AA----C-C----A--------------A-----A--G--G-----RC-AT-TA---G------G-----C-TAA----------GA-A-------GC----G--AC--T-AGAGG--GC 2515
H2U.FR.96.12034           T--C----C-----A--AG-AAAAGAA-CACTA--C.........ATT-GGGG-C----A-----C-A------A-----T-A-G---G-----T--AT-TA-G------T--G--------AA---G--A---GAGA-A-----------G--AC--T-GGA-G--GC 2565
ASC.UG.10.RT03            --------CA-GGTAG-AG--AAAT--A--C-GAC-.........GA----CGGC---GAGA---G-AA-----A--C-----C--------------A-GT-C---G-------C-G---C--AA---G--AGTACA---T----T-G-TA----AC--T-GGA------ 2107
ASC.UG.10.RT08            ---C-----A-GGTAG-AG--AAAT--A--C-GAC-.........GA----CGG-----AGA---G-AA--C--A--T-G---C--------------A-GT-C----------C-G---C--AA---G--AGTGCA---T----T---TA----AC--T-GGA------ 2131
ASC.UG.10.RT11            ---C-----A-GGTAG-AG--AAAT--A--C-AAC-.........GAG---CGG------GA---G-AA-----A--C-----C--------------A-GT-C----------C-G---C--AA---G--AGTACA---T----T-G-CA----AC--T-GGA----G- 2116
COL.CM.x.CGU1             GTAG--GC-CAATTATCAGATAGG--CCCAGTCACC.........AAG--G-CG-----AGAG--C---------G---------A--G--G-----C--A-GCA------------G---GT---CA-A-G--A---GA-AGAT-----GCA--G--A-----AGA----GC 2473
COL.UG.10.BWC01           GT----GC-CAATTGTCAGA-AGA--CCC--TAACT.........AA------GC-C---GA------------G----G--A--G---------C--A-GCA-------G---G---G-C--CAGA----A---GA-AG-C-A-----A---------G--AGA----GC 2272
COL.UG.10.BWC07           GTC---GCACAG-TGTCAGC-AAC---CCAGTGAC-.........AA---GTC-C-G-----T--GC----C--A--C-GG--G--G---------TC-T-GTA-----------TG---GGC--CA-A-G--A---GATAG-C----GGCA-----A-----AGA---GGC 2500
DEB.CM.04.SIVdeb04CMPF3061 T-GA----CA-G-TAGAG--T-AG---C-ATA-ACC.........-----G---------AGA---G-A-TCA--A---------C---------G-C-CT-GCC-G-G----GAT-G--GC-CTCAG-G--A-A--GGGA-AGAT-GTGGA-G-A--CC---GG-TCCTGC 2688
DEB.CM.99.CM40            T--A----------AG-ATTAGA-TA-ACC...............--------------AGAG--GGCA-----A-----GA--------------C-TT--AGG---------G-G--C---AA---T--C-TA-AGCCC---AT-G---C---AC-G--AGTCCCGGC 2507
DEB.CM.99.CM5             T-CA----------AG-ATTGGACT--ACA...............---------G--G---GAG---GCT---------GGA-------G-----CC-TT--AG----------G-----C---AA---G--CATA-GGCCT----T-G-A-C---AC-G--AGTCCCC-C 2501
DEN.CD.x.CD1_CMO580407    --AC----CA-GG-AG-ACAG-AA--GCG-C--CA-...GTCACC-------------A------T-T--------T-G-A-A-------------T--A--TAC---G---T-------T---CA-C-G--AGTG---C-------------G--A-GGT-GGA---GGC 2585
DRL.DE.11.D3              ---C-T--C-----TG----T-AAG-A--AC-A---.........AA-----C-C----A--------------A-----GA-A----------CC-CT-CA-G--G-----TG----C--GAA------A---GA-C-A-----G---------TC----AAGT-G--- 2172
DRL.x.x.FAO               ---C-T---------TG----T-AAG-A--AC-A---.........AA-----C-C----A--------------A-----GA-A----------CC-CT-CA-G--G-----TG----C--GAA------A---GAGC-A-----G---------TC----AAGT-G--- 2265
GRV.ET.x.GRI_677_gri_1    -AA---GT-A-GGGAG--CTAT---G-C---T--AG...GAAACAAA----C--C----AGA---G-AA-----A--T---T-G-----------C--AT--AG---------TG----T---AC------A--C-A-C-A------G----G--A---T-A--T-GG-- 2871
GSN.CM.99.CN166           --------CA-GGTGCAGCAGAAG--A--AC-CA--.........GA---GC-CC-----G------AA--G--A-----GA-C-----G-----CC--T-TAGG-----G---G-G--C---ACCC----AACACA-------CTT-GATT---TC--T-GGA------ 2408
GSN.CM.99.CN71            --C-----CA-GGTGCAGCAGAAAC-A---C-C---.........GA---GC-CC----AT------AA--A--A-----GA---------------C---T-TA-G------G---G-G--C---ACCC----AACACA---A---CTC---TT---CC--T-GGA---G-- 2408
LST.CD.88.SIVlhoest447    -GA---GTAC-AT-ACAATTAAG-GACA-A-TACCC...ATCACAAAG--C-----G--------CTGT-----T--C-G---A-----G-----T---T-TA-------G---G---GCC--CA--C-------GATAGAT-A-----A-------------A------GC 1966
LST.CD.88.SIVlhoest485    -AA---GTAC-AG-ACAATTAAG-GACA-A-TACCT...ATCACAAAG--C-G---G--------CTGT-----T--C-G---A-----G-----T---T--A-------G---G-G-GCC--CA--C-------GATAGAT-A-----A-------------A------GC 1966
LST.CD.88.SIVlhoest524    -GA---GTG--AG-ACAATTAAG-GAGA-A-TACCT...ATAACAAAG--GC-T--G--------GTGT-----G----G---A-----G------C-AT-CA-------------G---G----CA--C-------GATAGAT-A-------------------A--CCC-G- 1963
LST.KE.x.lho7             -AA--G-T---AG-TCAGTTAAG-GA-A-A-T-CCT...ATTACAAAG---TCT-----A--T--GTGT-----A----G---A--G--G-----TC-AT--A----G------G---GCC--CA--CT------GATAG-T-A-----A-----A-----A--TCC-G- 3050
MAC.US.x.251_1A11         T--C-----C-------AGCTAAAG-A---C-----.........AA---CGCC--------------AA-T---A------T-G--G--G---------AT--A-------G---GTT------AG------C---GA-A-------------T--TC-GT-GGAGG--GC 2999
MND-1.GA.x.MNDGB1         -AA----TA--GG-ACAATTGT--GA-A-A-TACC-...ATAACAAA---G--------A------G-A-----A--C-G-A-A-----------T--A-GTA----G------GTT-GTC-TCAGA----A---GATAGAT-A--GG--------A------AGT-GGG- 2385
MND-2.CM.98.CM16          --AC----------AG----TAAAG-A--A--A---.........AAG----C-C----A-------CT--C--A--T---A-A-----------------T-TA-G-----G--TTT---C---CA------A----GTC-A-----G--------CC-G--C--T--G-- 2699
MND-2.GA.x.M14            ---C-G---------T-----TAAG--A-----A---.........AA-----C-C----A---------------A----G-A--------G------C--T-TA-------G--TTT---C---CAGAC---A---GAT--A--------------C-----C--T-G--- 2626
MND-2.x.x.5440            ---------------T-----TAAG--A-----A---.........AA-----C-C----A--------CA------A--G-G-A-A----G------C--T-TA-------G--CCT---C---CA------A--C-ATC-A---------A-----C--G--C--T-G-G- 2259
MNE.US.x.MNE027           T--C--------C-----AGCTAAGG-A---C------.........AA---C-CC--------------AA-----A-------T-G-GG--G--------AT--A-------G---GTT------AG------C---GA-A-------------T--TC-GT-GGAGG--GC 2468
MON.CM.99.L1_99CML1       --CC-C--CA-GG-TGC-G-GGGAGACCTCCAACCCCCATCTTTAGAG--C-CT-----AG-------AA--G--A-----GT-G-----G--------AT-CAG---------G---G-G--TC-TAC-C-----AC-CA-------CTT--ATT---AC--T-GGA-CC--C 2404
MON.NG.x.NG1              --CC-G--CA-GY--CAGG-AATCGCCC-AT-CCC-...TCCTTGA---G-CCC-T--A-------AA------G------T-G--------CC-CT-TA--------G---G-G--TC-CACCACC---RCCCGG------AT-G-ATT---CC--T-GGA-G--GC  997
MUS-1.CM.01.CM1239        --C--G--CA-GATACAGGAAAAA--A--AC-CA--.........AA---GGCC------GA---GGCAA-A--A----TG--C-----G------C-----TC----------G--T--G--C-----A--GC--AGT-CAGC-A---TT---ATT--ATC-G--AGA-G-G-- 2378
MUS-1.CM.01.SIVmus_01CM1085 ---C----CA-GATACAGG-AAAAT-A--AC-AA--.........AA---G-CCC----AGA----GCAA-A--G----TG--C--G-------T-----CC-G-----G--C-GG--C--G-AG--C--AACCCAGC-----CT---A-T---CC----AGAGG---- 2388
MUS-2.CM.01.CM1246        --AC----CA-GGTA--GGCAGGACAGC-ATTGAA-...CCCACTAA---T-GCC-T--AGA-A---AAA-A--G----TG--A-----------C-CT-TA-G--G--G---G-G---C--AA-------AC-CAGA-A---AT---ATT---TC--C-GGAGG--GC 2486
MUS-2.CM.01.CM2500        --C--G--CA-GGTAGGAG---AAC-G--AC--ACG.........AA---TGG-C-C--AG----G-AA--A--G----TG--A-----------------T-CA-G---------G----TC---A-------AC-CAGAGA---TT----CTG--AC--T-AGA-G---C 2445
MUS-3.GA.09.09GabOI81     --A--G----G-GT-CAG-C-T--T-A--AC--C-C.........AA---T-GGC----AGA-----AA--A--G--T--G--A-----------C-A---C----------G-T------A--A-C--AACACAG------CTT---TT----C--T-AGAGG-GGC 1954
MUS-3.GA.11.11GabPts02    --AC-------C--GG-A-CAGGG---A-A--AA-C.........--G--G---C----AGA-----AA-----A--------C-----G-----CC-T---C---------C-----CC-G-AGC-G--AACCCAG-------T---ATT---AC--T-GGA-G---- 1975
OLC.CI.97.97CI12          CG-C-GGTC-GCG-ACAGCTTT-AGA-A---TCCCT...ATTACTAAG-----------AGAG---CACA-GCCT---C-GA----G---------T--A-GTA-G-----G---G----T----AGA----A---GACAGAT-----G-A-----T--G--ACAG---GC 2572
RCM.CM.00.SIVagi_00CM312  T-A-----------A--A--TAAAGCA--A-TAA-T.........AA----G-G------GA---TCA------A----G---A--G--G-----CC-C-GTA------------G----C---AC------A--C-AT-CA-----------G--A------AGT-G--- 1969
RCM.CM.02.SIVrcm_02CM8081 T-------------A------AAGGCA--A--AA-T.........GA---G-----G--A------CAA-----G--T-----A------------T--A-GCA-------G---G----G---AC-----------AT-CT--------A-----A-----A--C----- 1987
RCM.GA.x.SIVRCMGAB1       --AC----------A--A----AA--A--A---A---.........AAG---G--------T-T---CAA-----A---CG---A-----------------T--A-----------G----T---AC----------C-AT-CA-----G--------A-------------- 2177
RCM.NG.x.NG411            T--C----------A--A--TAAAGCA--A--AA---.........AAG---G---------------CAA-----G--T-G---A-----------T----GTA----G-----TG----T---AC---G--------AT-CC-------------G--A---------T-G--- 2192
SAB.SN.x.SAB1             ---C-------GA-ACAA--AGAG--A--AC--A---.........AA---CC-C-------------CAA-----G----GGA-A-GG--------T---T--A------------G-G--C---AAG-CC------GA---CT-A-----AC----AC-TT-GGA--G--- 3096
SMM.CI.79.SIVsmCI2        T-A-----C-GC--AT-GGCAAAGG-A--AC-A---.........AAG----CT--G-----------AA--A--A------T-A-----------GC--T--A-------G---G-G---C-CAGG-----C---GA-A-A---------A-----AC--T-GGA-G--GC 2120
SMM.LR.89.SIVsmLIB1       T--C----------AG-AGCTAAAG-A--AC-A---.........AAG------C----A-----G-AA-----A------T-G-G-----------C-TT--A----G--G---G-G--C---AAG--G--C---GA-A-----------------TC-GT-AGA-G--GC 2079
SMM.SL.92.SIVsmSL92A      T--C------A---AG-GGCAAAAG-AC-AC-A---.........AAG----C-C----A-------A------A------CAA-G---G-----CC--T-TA-------G---G------C--AA---G-------GA-A-A-----------------C--T-GGA-G-GGC 2081
SMM.SL.92.SIVsmSL92B      T---------AC--AG-GGCAAAGG-A--ATA-AC-.........AA------G--G---GAG-----------G----G-T-A-----------C--AT-TA-G----G---C---------AC---G--C--CGAGA-A-----------G--AC--T-GGA-CG-GC 2060
SMM.SL.92.SL92B           T---------AC--AG-GGCAAAGG-A--ATA-AC-.........AA------G--G---GAG-----------G----G-T-A-----------C--AT-TA-G----G---C---------AC---G--C--CGAGA-A-----------G--AC--T-GGA-CG-GC 2431
SMM.US.04.G078            T--C-----C-A--AG-GGCTAAGG-A---C-A---.........AAG--G--G-----A-----G-AA--A--A-----GT-A-----G-----GC-AT-TA-----------G-G---C--AA----------GA-A-A-----------G--CC--T-AGA-G-GGC 2260
SMM.US.05.D215            T-A-----------AG-AGCTAAGG-A---C-A---.........AAG--G-C---G-A-------A-------A------C-A-G------------AT--A-------G---G-------AAG-----A---GA-A----------A--T--CC-GT-AGA-G-GGC 2254
SMM.US.06.FTq             T-------------AG-AGCTAGGG-A--AC-A---.........AA---G-C-C----A-----GCA-------A------T-A-G--G-------AT-TA-------------G-G-----GAA---G--C---GA-A-A-----G-----G--CC-GT-GGAGG--GC 2282
SMM.US.86.CFU212          T-C-----C--C--A--GGC--AAG-G---C-G---.........AA-----C-C----A--------AA------A--------A-GG---------C-AT--A-------G---C-G------AA---G--C---GA-C-------------T--TC--T-AGA-G--GC 2246
SMM.US.x.H9               T--C-----C-C-----AGCTAAGG-G--RC--A--.........AA-----C-C----A--------AA------A------T-A-G--G-----GC-AT-GA-------G---TT------AG------C---GA-A-A-----G--A--T--CC-GT-AGAGGR-GC 2473
SMM.US.x.PGM53            T-C------C-C---G-AGCTAAGG-G---C--A--.........AA-----C-C----A--------AA--C--A------T-G-G--G-----GC-AT--A------------GTT------AG------C---GA-A-A----------A--T--CC-GT-AGAGG--GC 2917
SMM.US.x.pE660.CG7G       T--------C-C-----AGCTAAGG-G---C--A--.........AA-----C-C----A-----G-AA--A--A------T-G-G--G-----GC-AT--A-------G----TT------AG------C---GA-A-A----------A--T--CC-GT-AGAGG--GC 2989
STM.US.89.STM_37_16       T-A-----------AG-AGCTAAAG-A--AC-A---.........AAG--G-C------A--------AA----A------A-A---------------AT--A-------G---G-G---C--AA---G--C---GA-A-A-----G-----T---C-GC-AGA-G--GC 2647
SUN.GA.98.L14             -GA---GT---AG---C-CTTAG-GAAA---TCCC-...ATTACAAA---T--G-----AGA---GGCC-C---A---------G------------CC--T--AGG--G------G---GCC--CAGA---------GATAGAT-----GCA-----A--G---AGC-GGGC 3055
SYK.KE.x.KE51             -AA---GTAA-GATAGCACAAAAAT-A--AC-AA-G.........AA---CC--C-----------G-A---GCTA--T---T-A-----G-----TC-----AG-------C-TG-G---C-GCG--C---AGT-T----T---CTTG-AA-G---C--T-GGA---GGC 2553
SYK.KE.x.SYK173_COMGNM    -AA---GTAA-GGTTCAA-CAG-AT-A---C-A--C.........AA----GTC-G--A----AC-AA--ATTG--CCGGC-A-------------AT-G-T---G---T--G-G--CC-CAAG-C---AGTAGA---C---CTC----C---AC--T-GGAG---GC 2895
TAL.CM.00.266             --AC-T--CA-GGT--GCTCAGAC--ACC--TAACC.........AAG--T---------AGAG--T-A---ACC---C---A-C--G----------C--A--CAG-----------G---GGC--CA-ATC--CATAGAC--C----TG----C---AC-GT-AGAGG-GG- 2667
TAL.CM.01.8023            --AC-T---A-GGT--GC-CAGAA--ACCAGTAA--.........AA------G--G---GA-----A---ACC------GA-C--G---------C-----CA-G---------G---GGC-GCAG-CC--CATAGAC--C----TG----C---AC-GT-GGAGG-G-- 2165
TAN.UG.x.SIVagmTAN1       -GA----TAA-GGGACAGCTAAA-GAGA-A-TACCC...ATAACCAAG---C----G--A--T--G---TCA--A--G---T-A-----------CC-AT-----------------CC-GAC------AAC-CAG-CTC-A-----------A--------T--GG- 2900
VER.DE.x.AGM3             -AA---GTAA-GGGTCAACTGT-AGAAC-A-T-CCC...ATTACC--T--G---------AGA---GGCTAGA--A--TTTCT-A-----------CC-CT-CA-----------------C---CAG-----A---GACC-AT-A--G--A-----A------AGC--G-- 2423
VER.KE.x.9063             -GAC--GTGA-GGGACAATTGT-AGAAC-A-T-CCC...ATCACC-----C---C-G--AGAT---GCAAGG--A---C-TA-A-----------TC-AT--A-G--G---------G--C---AC---G--A---GA-C--T-A-----A-----------CAGC--G-- 2928
VER.KE.x.AGM155           -AA---GTGA-GGGGCAATTGT-GCAGACA-TACC-...ATCACC--G---CGC------GA---GGCCAGA--A---CG-T-G--G-----------C-CT-TA-------------T---CC-GCA---------C-A-ACAT-A--GG-A-----A---T-AAGC-GGG- 2920
VER.KE.x.TYO1_patent      CGG---GTAA-GGGACAATTAT-AGAAA-A-T-CCT...GTCACA--T--C-----G---GA---GGCTCGG--A--CTGT--A-G---------TC-CT-TA---G--G--TG----T---CAG-----A---T-CC-AT-A--GC-------A-----CAGT-G-G- 2925
WRC.CI.97.97CI14          CA----GCC---G-ACAATTAT-AGACA---T-CCT...ATAACAAA---G-GC--G--A------TGCTCA--A--T-G---A-----G-----C---T--A--A--------G-G-GGC--CAGA--------GATAGAC-A---GCA-----A------AGC-G-GC 2998
WRC.CI.98.98CI04          CA---GGCC---G-ACAATTAT-AGA-A---TACCC...ATAACAAA-----GT--G---------TGCACA--A----G---G--G--G-----T--AT--A----G------G---GGC--CAGA----C---GATAGAT-A---GCA-----A------AGT---GC 3013
WRC.GM.05.Pbt_05GM_X02    CA-C--GCA---G-ACAATTAT-AGA-A-A-T-CC-...ATAACAAA---CTC------A--T---TGTC-A--A----G-A-A--G--------TC-CT--A----G-----TG-G-GGC--CA-A-G--A---GATAGAT-A---GCA-----A-----AAGT-GGGC 1927
```

```
H1B.FR.83.HXB2              TGGGCCTGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGAAAATTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCACATCCCGCAGGGTTA..................A 2850
Pol                          G  P  E  N  P  Y  N  T  P  V  F  A  I  K  K  K  D  S  T  K  W  R  K  L  V  D  F  R  E  L  N  K  R  T  Q  D  F  W  E  V  Q  L  G  I  P  H  P  A  G  L  .  .  .  .  .  .  .
H1O.CM.91.MVP5180           A--A-----------T--T-----A--TA-T-----T-----A--G-----T--C-----G-----------G-----C---------T-A----A----A----T--------G--G--------T--T--------A-GG--T---..................- 2879
H1N.CM.95.YBF30             A--------G--------T--C------A-T-----T-----A--G-----T--C----------------------------G---T-A----A--G--C-----T--T--------G--GC-------T--------A-----A---..................- 2444
H1P.CM.06.U14788            ------------------T--C--C--CA-T-----A-----A--G-----TG-C-GC--G-----G-----------C-----------A-----A-----A-----T-----------C-----G--R--C--T-----A-G---TC-T..................C 2321
CPZ.CD.06.BF1167            A--A-----C-----T-----------CA-C-----T--T--A--G-----------A----------------------T-----G--C-----A--G--A-----T--T-----G--A------------------C--A----------..................- 2958
CPZ.TZ.00.TAN1              ---A--A-----C-----T--C--A---A-T--------T--------G--T-CCT-------------------------------G-----A-----------T--T-----GT-A--GC----------C-----G----------..................- 2494
GOR.CM.07.SIVgorCP2135con   ------------------T-----A--TA-T--C--------A--G-----------K------------G--------T-------T-A----A--G--A-----T--T-----------G-----G--T--T--C--A-G----C-T..................- 2386

MAC.US.x.239                -CCC--GACC-----------C--C--CAC------T-----------G--T-AG-AC-----------TGC-GA-------T--G-----A----G-GTC-----G-----TAC------C----------------C--T-----AC--..................G 3152
Pol                          A  P  P  T  N  P  Y  N  T  P  T  F  A  I  K  K  K  D  K  N  K  W  R  M  L  I  D  F  R  E  L  N  R  V  T  Q  D  F  T  E  V  Q  L  G  I  P  H  P  A  G  .  .  .  .  .  .  .
H2A.SN.86.ST_JSP4_27        ACCT--AACT-----T--T-----C--CAC------A--C--------G-----AA-AC----------TGC--A-------T---------A--C---GT------------ACA---A-C--G--------T-----C--A-----AC--..................G 2680
H2B.CI.x.EHO                GCCT---ACT------------T-G--CACC--C--------A--G------AA-AC--------G-TGC--A----------------A--C---GT---C-----A--TACA--G--C--GC-G--T--T--T--C--A-----AC-G..................G 3206
H2G.CI.92.Abt96             ACCC--CACC-----G--T--C--A--CACC---A-T--C-----G--G-----AA-AC----------TG---A----C---------------C-G-GT---C--G-----TACT-----C--G--G--C-----T--C--T-----T--G..................G 2562
H2U.CI.07.07IC_TNP3         GCCC---ACC-----T--T-----A--TACC-----T--------G-GG----AA-A------------TG---A-------------------A--C-GAGT------G--T---ACA--G--G--------T-----T--C--T-----AC--..................G 2667
H2U.FR.96.12034             -CCC--CACT-----------C-----CACC---------------G--G-----AA-AC----------TG---A----C--T-----GT-A--C-GAGT---------T--TA-T-----C--G-----T-----C--C--A-----AC--..................G 2717
ASC.UG.10.RT03              A-----A-----C---------T-----------TGT--T-G---G--G-----AA--A--------G-TGC--A-T---------GA----A---G--CAC--A--G-----AGCA-----A--G--G--G--C-----C--A-----T---..................C 2259
ASC.UG.10.RT08              A-----A-----C---------T-----------TGT--T-----G--G-----AA--A--------G-TGC--A-T---------GA----A---G--CAC--A--G-----AGCA-----A--G--G--G--C-----C--A-----TC--..................C 2283
ASC.UG.10.RT11              A-----A-----C---------T----------CTGT--T-G---G--G-----AA--A--------G-TGC--A-T---------GA----A---G--CAC--A--G-----AGCA-----A--G--G--G--C-----C--A-----T---..................C 2268
COL.CM.x.CGU1               A-AATTA-GG--------T-----A---A-----TG---T-G----------T-AA-A-G-------G------A----------G--G--A--C--------A---------AT------C--------G--------C--A-G---TC-T..................- 2625
COL.UG.10.BWC01             --CA-A--G-----T--T-----C---A-----TGT-----G------G--T-AA-A-G---------------A--------T------T-A-----------A-----T----TT--G--A---CA---G--T---------T-G-------..................G 2424
COL.UG.10.BWC07             A-C-TTA-G---------T-----C---A-T---TG-------------G--T-AA-A-G--------------GA----C--T--G--GT-A-----A--G-------------------A-A--GC---------T-----C--A-G---A--G..................C 2652
DEB.CM.04.SIVdeb04CMPF3061  CTCA--CACC---------T--CT-A--T--G---------T--A--G--------AG-AC-----------TG---A-----C-T------T-G--C---GC---A---CCAG-T-----G--C--GAC---G--C--G--C--A-----TC-G..................C 2840
DEB.CM.99.CM40              A-C---CACC--C-----T--CT--------G----TA--------G--G-----AA-G---G---C---TGC--A----C--A----CCT-A--C---GCC--A--G--GG-G-----G--G--GAC---T---------C--AT-T-CAC--..................C 2659
DEB.CM.99.CM5               A-A---CACT--C-----T-GCT----T--G----TT--------G--G-----AG-A-----------TGC-TA----CC-G----CC----------GCT--C--G--TG-----------G---AC---G--C--G--C---T-G-----G..................C 2653
DEN.CD.x.CD1_CMO580407      GCCA--CACC--C--T---------A--T--G----TA--TCCA--G--------AA--AC---AT--G-TG---A-----------GA--T-A---G--GCC--C--G----------------A---------T-----------T-----A---..................C 2737
DRL.DE.11.D3                A--T--A--------------C--A--G--G---TGT--T--A--G-----TG----C-------------C------------GC--T-A--C--AGC---A--G-------TT--G--G--G-----G--------C--A-G---A---..................- 2324
DRL.x.x.FAO                 A--T--A--------------C--A--G--G---TG---T-----G-----TG----C-----------------------------GC--T-A-----GC---A--G-------TT--G--G--G--G--G--------C--G-G---C---..................- 2417
GRV.ET.x.GRI_677_gri_1      A--AGGA--------T--T-----A-----G-----------------G--T-AA--AC---------TGC-T-----------G-----A--C--AGCT-----------T-TT--G-----GC-G-----T--T--C--A--G--CC-T..................C 3023
GSN.CM.99.CN166             A--A--A--------------CT-A-----C-----------A--G--G----AAT-AC---------TGC--A-------T---A--T-A---G--GCT--G--------GGCT--G--C--G--G-----T--G--C--A-----AA-TCCGCACCCTGCAGGACTAG 2578
GSN.CM.99.CN71              A-----A--------------CT-G-----C-----------A--G-------AAT-AC---------TG---A---------GA--T-G---G--GCT--G--------GGCT--G--C--G--G--G--T--G--C--A-----AC--..................G 2560
LST.CD.88.SIVlhoest447      A-ATG-A-GG--C-----T--C--A---A-----------------G-------AA-A-G-G------------A-T-----T---C----A---GCTTT--A--------TCAT-----A--------C--T-----C--T-----AA--..................- 2118
LST.CD.88.SIVlhoest485      A-ATG-A-G---C-----T--C--A---A-C---------------G--G-----AA-A-G-G---------C-GA-T-----T---C----A---GCTTT--A------C-TCAT-----A--------C--T--------T-----AA--..................- 2118
LST.CD.88.SIVlhoest524      --AT--A-G-----T-----C--C---A-C-----T-----------G-----AA-A-G----------G--GA-------------A--T-A---G--TT---A--------TCAT--GC-A--G--------------C--A-----AA--..................- 2115
LST.KE.x.lho7               G-AT--A-GG--C-----------------A-------T--C-----G-------AA-A-G----------G---A-T----------A-G--C---G-ATT-------------TCAT--GC-A--------T--T-----C--A-----TA-C..................- 3202
MAC.US.x.251_1A11           -CCC--GACC-----------C--C--CAC------T------------T-AG-AC-----------TGC-GA-------T--G-----A----G-GTC-----G-----TACA-----C-----------------C--T-----AC--..................G 3151
MND-1.GA.x.MNDGB1           A-AT--A-G----AAT--------A--TA-C---------------G--G--T-AA-A-G--------------A----C--T------T-A--C---TT---A--G--T--TCAT---T-A--G-----T-----T--C--A-----AA--..................- 2537
MND-2.CM.98.CM16            A--T--A-------T-----C--A--G--G---TG---C--A--G-----TG-A--C-GC-----------G--------T---C-GT-A--C--AGT-----G-------TT--G--G--G--G--G-----C-----C--T-G---TC-C..................- 2851
MND-2.GA.x.M14              A--T--A----------T-----A--------TGT-----A--G-----TG-A--C-GC--------------------T--GC--T-A----GAGTG-----G--T----TT--G--G--GC-------C--------G-G---C---..................- 2778
MND-2.x.x.5440              A--T--A--------------C--C--------CTGT-----A--G-----TG-A--C-GC--------------------T---C----G--C--AGTC-----G--T----TT--G--G--G--------CA-------A-G---CC--..................- 2411
MNE.US.x.MNE027             -CCC--GACC-----------C--C--CAC------T------------G----AG-AC-----------TGC-GA-------T--G-----A------GTC-----G-----TACA-----C-----------------C--T-----AC--..................G 2620
MON.CM.99.L1_99CML1         C-A---CA-T--C-----T--CT-A--CTGT------------------G----AG-GC-----------TGC-CA-------------------C---GC------G-------TT--G--G--G--G-----C-----C--A--G--AC-T..................G 2556
MON.NG.x.NG1                A-A---AACC--------T---T---TTGC--C--G---G-----G--T-AA--C--------G-TGC-CA-----------G---T-A--C--AGC---A-----T----TT--G--A-----------T-----C--G--T--AC-C..................C 1149
MUS-1.CM.01.CM1239          A-----A--C-----T-----CT-----TGT-----A--C-G---G-----T-AAT------------TG---A-------T---C-G--A---G--GCC--C-----G---ACA--G--A--GC----G--C--------A-----A---..................G 2530
MUS-1.CM.01.SIVmus_01CM1085 A--A--A--G-----T--T-----T-----TGC-----A--T-G------G----AAT----G------TGC--A----------GC--T-A---G--GCT--C-----AC--GCA--------C--------C--------A-----A---..................- 2540
MUS-2.CM.01.CM1246          A------A-T--C-----T---T-A---------C-------GA--G------AG--A-----C---TG---A----------A-G------G--GTG--A-----A--TGCA-----C--GAC---G--C---------T----------..................G 2638
MUS-2.CM.01.CM2500          A------C---------------GC----C------T----G------G-----AAT-C----------TGC--A----C------A-G--C---G--GCC--A-----------CA-----C--GAC---G--C--G--C--AT-----C-T..................T 2597
MUS-3.GA.09.09GabOI81       A-----A--C--C--T-------T-A---TGT-----T--T--A--G--G-----AAT-A----------TG---A-------T---A--T-G---G----C--C-----------C-----A--GC----C-----C-----A-----AC-G..................T 2106
MUS-3.GA.11.11GabPts02      A-----A----------------T-A---TGT-----A----GA--G--G-----AA--A--------G-TGC-CA------------A--------G-A-AT--A------C---TT--G-----G-----G--C--T-----GT----CC-T..................G 2127
OLC.CI.97.97CI12            A-A-G--TC-----A--GG-----C---A-----TG------A--G--G----AA-A-G-G-AT-----G--GA-------T--G-CT--A---GCAGCC--A-----T--------G--A--------------T-----A-----------..................- 2724
RCM.CM.00.SIVagi_00CM312    A--A--A--G--------T---------A-----TGT--T-GA--G-----TTCG--C-----------GC--------------------A--------G--A--G--T----TT--G--A--GC-G--G--T--------A-G---A--G..................- 2121
RCM.CM.02.SIVrcm_02CM8081   A-----A--------T--------G--TA-T---TG---T-G------G--TTC---A-------------G----------------T-A---------------G-------TT--G--A--GC-G---------------T-GG--T--G..................- 2139
RCM.GA.x.SIVRCMGAB1         A-----A-----C--C-----C--A---A----CTGT--T--A--G-----TCC--A-------------G--------T--------A-----------A--G-----T-TT--G--G--GC-G---------------A-G---A---..................- 2329
RCM.NG.x.NG411              A--A--A--------T-------------A-----TGT--C--A--G-----TTCC--A-----------GC-----------------T-G-----A-----A-----T----TT--G--A--GC-G-----C--G-----A-G---A---..................- 2344
SAB.SN.x.SAB1               A-----A-----------T--C--C-----C-----A----G---G-----T-AA---C-G-------T-C-CA-G--C-----GC-GT-A-----A--T-----------TCA------C--G-----G-----C--C--A-----C---..................C 3248
SMM.CI.79.SIVsmCI2          ACCT--CACT-------------A--TACC-----A-----A--G------AA-AC--------G-TGC--A------C---G-----A-----AGT---C--G-----TAAT--G-----GC-------------C--A-----A--G..................G 2272
SMM.LR.89.SIVsmLIB1         -CCT--AACC--------T--C--A--CACT--C--------------G----AG-AC--------G-TGC--A-------T---------A------GTG--A-----T---ACA---A----G--G-----C-----C--T-----AC--..................G 2231
SMM.SL.92.SIVsmSL92A        CCCA--AACT--------T-----C--TACC--C--T-----A--G--G----AA-AC--G-------TG---A-------T-----T-G--C---GT---A--G-----ACA---A----G--------C--T--C--A-------C--..................C 2233
SMM.SL.92.SIVsmSL92B        CCCT---AC------T--T-----A--GACC-----T--T--------GA-T-AGGAC--G-------TGC-CA---------------T-A--C--A-TG--A-----T---ACA-----A--GC-G--T--C--------G-----CC--..................- 2212
SMM.SL.92.SL92B             CCCT---AC------T--T-----A--GACC-----T--T--------GA-T-AGGAC--G-------TGC-CA---------------T-A--C--A-TG--A-----A---ACA-----A--GC-G--T--C--------G-----CC--..................- 2583
SMM.US.04.G078              CCCT--AACC-------------A--CACC-----T-----A--G--G--T-AA-GC--------G-TG---A-------T------T-G----G-GTC--A-----T--TACA---A----A---G--------T--G--C--------A---..................G 2412
SMM.US.05.D215              CCCT--AACT-----G--T--------CACT-----T--C--------G-----AA-AC--------G-TGC--A--------------GT-A-----AGT---A---------TACA-----G--G--G-----T-----C--A-----AC--..................G 2406
SMM.US.06.FTq               -CCT---ACT--------------G--CACC------------------G-----AA-A---G-------TGC--A----C-----------A----G-GTT--A--G--T---ACA---------C--------------C--T-----AC--..................G 2434
SMM.US.86.CFU212            -CCA---ACC--------------A--TACT-----T-----------G-------AA-AC--G-------TGC--A-----------G-----A-----AGTC-----G--T---ACA-----G--G--G-----T-----C--A-----A--G..................G 2398
SMM.US.x.H9                 -CCT--AACC--------T--C--C--CAC------T------R-GR-----G-A-A---------G-TGC--A-------T------T-A--C---GTC--------T--TACA-----A--G-----G--------C--TR----AC--..................G 2625
SMM.US.x.PGM53              CCCT--AACC--------T--C--C--CAC------T-----A--G-------AA-A---------G-TGC--A-------T------T-G----G-GTT--------T--TACA-----A--G-----G--------C--T-----AC--..................G 3069
SMM.US.x.pE660.CG7G         CCCT--AACC-----G--T--C--C--CACT-----T-----------G-------AA-A---------G-TGC--A-------T------T-A------GTC--------------TACA-----A--G-----------------C--T-----AC--..................G 3141
STM.US.89.STM_37_16         -CCT--AAC-------T---------C--CACT-----T-----------G-----AA-AC--G-----G-TGC-GA-----------------A-----AGTG--A-----T--TACA---A----GC----T--------C--T-----AC--..................G 2799
SUN.GA.98.L14               A-AC--A-GG--C--T--T--C--A---A-T-----T-----T--------G-------AA-A-G-----------G---A-------T----T---A---GGT-TG-----------TCAT--GT-A--GT-A--G-----G--T--T-----A-----A---..................- 3207
SYK.KE.x.KE51               CTCC---AC------T--T-----C-----T----TA---C-A--G--G----AATTA---------G-TGCCTA-G-----T--GA----C---G-AGC---A--G-----T--------A--G-----------T-----A-G---C-C..................- 2705
SYK.KE.x.SYK173_COMGNM      -A-T---ACC-----T--T-----C----------TA--T-G------G----AA-AG--------G-TGC-CA-T-----T---A--T-A---G--GCC--A-----T--T-TC--G--A---TC--G--A-----G--------C--C--A-G--AC--..................- 3047
TAL.CM.00.266               ---A--G-----------T---T------G--C--T--C-GA--G--G----AA---C---------TGC-CA---C---C-C-CC--A---C-AGC---C---------GC------A------A---CAG--G-----T--C---T----CC-G..................C 2819
TAL.CM.01.8023              ------G-----C--------CT-------G-----T--C-GA--G--G----AG--AC---------TG--GA--------C---CC--G---C--GC-----G-----TGCA--G--C---CA--G--C---CAG--G--------C---T-T--AC--..................C 2317
TAN.UG.x.SIVagmTAN1         ---AGGA--G-----T-----------C--G---TGT--CCG-----------AA--AC-G-----G-TG---A-------T----CT--C-----AGC--A--G------TT--G--A---TT--G--A--------T--T--C-----A-----A--G..................- 3052
VER.DE.x.AGM3               A--AGGA--G---G-------C---------G---TG----------------AGT-AC--------TG----------T--G-----A--C--AGC--A-----T--T-TC-----A-TC-----A--G-----C-----T-----AT-------C..................G 2575
VER.KE.x.9063               A--AGGA------G-T--T-----C---A-C---TGT--C-GA--G--G----AA---C-------G-TG---------C---------T-A-----AGT---A---------TT-----------------G--------T--C--G-------G..................C 3080
VER.KE.x.AGM155             A---GGA--C---G----------A--------CTGT----G-----------AAT-AC-G-------TGC-G-----------G-----C--C--AGCT--A---------TT------------C--------T-----C-----A------------..................- 3072
VER.KE.x.TYO1_patent        A--AGGA------G----------C---A-----TG---------G--G----AAT-CC-G-----G-TGC-------C--T-----GT-A------GC---C-----T----TT------G--------G-----C--C--A-----A---..................- 3077
WRC.CI.97.97CI14            A-AT--A-G---------T--------TA-------T--T--A--G------AA-A-G-------------GA-----------A--T-A---G-AGTT--A--G-----T-TT-----A--G-----------T--C--T-G---AC-G..................- 3150
WRC.CI.98.98CI04            -TCC--A-G---------T--------TA-------A-----------G----AA-A-G-------G--GC-CA-------T---A--T-A---G-AGT---A--G-------TT-----A--G--G--C--T--T--C--A-G---AC--..................- 3165
WRC.GM.05.Pbt_05GM_X02      A-AAG-AA-T-----T--------A---A-C-----A----GA--G-----T-AA-A-G-------------A-------T----G---A---GCAGT---A----------TT--G--A--------C--T--T-----A-G-------..................C 2079
```

```
H1B.FR.83.HXB2            AAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCA 3020
Pol                       K__K__K__K__S__V__T__V__L__D__V__G__D__A__Y__F__S__V__P__L__D__E__D__F__R__K__Y__T__A__F__T__I__P__S__I__N__N__E__T__P__G__I__R__Y__Q__Y__N__V__L__P__Q__G__W__K__G__S__P__
H1O.CM.91.MVP5180         -GC-A-GGC----T--T-----CT-A-----A--A-----T-----C---TGC--T------CC------T--A--A--C-----C--C--T--T------G-G-----------C----AG-A-----C-----------C--C--G--A--G--------T------ 3049
H1N.CM.95.YBF30           -GC-------------G-----TT-------A--A-----T---------TG------G--CA----T--T--A--------A--T--------------------------------T--T-------C-----T--------G-----A--C--------G------ 2614
H1P.CM.06.U14788          -G--A-GR--------G--CA--T-A-----A--R-----T-----C--CTGY--T------CC---T--T--A--A-----A--T--C--T---------G----T-----C--------AC----R---GT---Y-----T--G--T-----------G--------- 2491
CPZ.CD.06.BF1167          ------G----GTCA-T--T--GT--------T--G-----T---------A-C--A-----CCC---A--TCAA---------A--------------A---G-C--T-----A-GG--T--T-AA------TGT----CA-----A--------------G--G--T--- 3128
CPZ.TZ.00.TAN1            G---A-G---TATG--G--------------A--G-----C--C-----CA-----C-G---CC---------A--------A--T-----------C---C-C--T---A-C----------AA----T------T--C---T-A--T--A--T-----G-----T--- 2664
GOR.CM.07.SIVgorCP2135con ---------G--------------------T--G--C--T---------TTG---------CCC---T--T--A--A-----A--T--C--A--T------G-G--T--------------G-A--G--------------A--A--C------------------T--- 2556

MAC.US.x.239              C---A-GG---AG-A-T------------A-A--------------C--CA-A--TC----------A--T---C----C-----C-----TT----ATCAG----T----CAGAG-----A-AAC----CATT--T--G--T--G--T----------G--G------ 3322
Pol                       L__A__K__R__K__R__I__T__V__L__D__I__G__D__A__Y__F__S__I__P__L__D__E__E__F__R__Q__Y__T__A__F__T__L__P__S__V__N__N__A__E__P__G__K__R__Y__I__Y__K__V__L__P__Q__G__W__K__G__S__
H2A.SN.86.ST_JSP4_27      CC-----G---CG-A-T--T--C--A-----A--G-----T--C-----CA-A--AC--C----G--T--T--AC----------------TC----ATCA----------CTGA------A-AA-----CATA--T--A--CTCA-------------G--------- 2850
H2B.CI.x.EHO              C-TCA--G---AG-A-----------A-----A--A-----C--C--CAGT--C--AC-----CC----------AC-A-----A--------TT-G--AGCAG----T----CAGA------A-AG-------TT-----A--C--A--------------G-----C--- 3376
H2G.CI.92.Abt96           C-G-A----GGAG-A--------------A-A--G-----C--C--CAGCA----TC------T---T------C----C--A--T--C--TT----ATCAG---------CAGA------A-AA-----CATC-----G--C--A--G--A--------G--C------ 2732
H2U.CI.07.07IC_TNP3       C-G------G-AG-A-------------C--A--A-----C-----CAGCA----TC--------C-A------CCA--C--A--T--C--TC----ATCAG---------CAGA------A-AG-----CATC-----A--T--A--G--A------------G------ 2837
H2U.FR.96.12034           CTG-------GAGGA-------T--A-----A--G-----C--C--CAGCA-C--GC-GTG-------------C-A---------T-----TT----ATCC--T--T----CAGAG-----A-AA-----CATC-----G---T-G--T--A--C--------------- 2887
ASC.UG.10.RT03            CC-----GGC-CAG--TT-CA--G-A--CA-CAAA-----C--C-A-G--A-C--GC--C----G-----TCAA--A--C--A-----C---------TCAG-G-----CAT-GG------AGAA----------TT--G--A--G--------G-----G-C---T--- 2429
ASC.UG.10.RT08            CC-----GGC-CAG--TT-CA--G----CA-TAAA-----C--C-A-G----C--GC--C-C--G-----TCAA--A----A-----C--G------TCAG-------CAT-GG------AGAA----------TT--G-----G--------G-----G-C---T--- 2453
ASC.UG.10.RT11            CC------GC-CAG--TT-CA--G-A--CA-CAAG-----C--C-A-G----C--GC--C----G-----TCAA-GA--C--A-----C--A-----CTCAG-G-----CAT-GG------AGAA----------TT--G--A--G--------G-----G-C---T--- 2438
COL.CM.x.CGU1             T-G-A--G-CCCATA--------T-A---A-A--A-----C-----C---A-------G---CCG--T-ATCAA-----C--A--T-----A--C--ATCAG-G------TT-G--------C-------GT------A-----A--G-----T--------G--T--- 2795
COL.UG.10.BWC01           T-G-------GAATA-T------T-A---A-T--G-----C-----C--TA-A---C----CCC---T-ATCAAC-------A--C-----A------TCGG-G--T---TTAT-------C-C---G---ATA-----A---T-A--C-----C--------G-----C 2594
COL.UG.10.BWC07           -TG-------GAACA-T-----T--A--CA-A--G-----C-----C--T-----AC-G---CC---A-ATCA------C--A--C-----AG-C--ATCAG----T--CAT-T-T-----ACAA-------GT------G--C--G--G--A--C--------------- 2822
DEB.CM.04.SIVdeb04CMPF3061 C-C------GGCAGA-------TT----CA-A-CA---------C-A---TA-A--A------CC-TC---TGCTC--------G--------T-----ATCAG-T--T--CACAG-C--T--AGAG--G---G---T--GG-----G-----A-------GT-C-------- 3010
DEB.CM.99.CM40            C-C-A-TGG--CAGA-C--T--CT-A---C-T-CA-----C----A----A-AG-AC----CCC-ACA--TGCA--A--C--A-----C--AG----ATC-G-C--T---ATACA------AGAA-------G---T--GA--A--G--T----------GC-CT------ 2829
DEB.CM.99.CM5             C-C-A-TGG--CAGA-------C--A--CT---CA-----C----A----A-A-----G---CC-C-G--TGCACGT-----A-----C--A------TCGG-G------ACACA------AGAG-------G---T--GA--A--A--T--A--C----GT-C------- 2823
DEN.CD.x.CD1_CM0580407    --C----GG--CACA-C--T--GT-A--CT--AAA-----------A-AGT--G--T--G---CC---T--TGCAC-A--C--A---------AG-G--GTCC--T---------AGG------ACAA--G--CATT-T---A--CT-A------A------GT------G--T 2907
DRL.DE.11.D3              --GCATGCG--CAGA-C--C-----------A---------C--C--C---TGC--TC-G---CCG-----T--A-----C---------C-----T--ATCGG-G--T---C--GGC-----A--C--G---------T-----C-------------------C-----T 2494
DRL.x.x.FAO               --GCATGCG--CAGA-C--T-----------A---------C-----C---TGC--TC-G---CCG-----T--A-----C---------C-----T--ATCGG-G--T---C--GGC-----A--C--G---------T-----C--C----------------C-----T 2587
GRV.ET.x.GRI_677_gri_1    -G--A--G--GCA-A-C------A-A--CA-A--G-----C----A----A-A-A---TGCA-G--A-----A--A-----A-----------C--CTCAG----T---ACAGGG--------A--G-------T----TGT--G--T-----------------T--T 3193
GSN.CM.99.CN166           -GC-A--GG-GCAT-----TA--A-A---A--AAA-----T----A-AGTA-A--A--GT-----G--T-----------C--C--T---T-AG----TCAG----T---C--G-C----CTGAA--G-------T---A--CT-A--G-----T-------CG--G--- 2748
GSN.CM.99.CN71            --C-A---G-GCAC-----TA--A-A---A--AAA--C--T----ACAGT--A--A--GT-------A-----A-----C---------T-AG----ATCAG-------C-----C----CTGAG--G-------T---A--TT-A--T-----T-----G-CG--T--- 2730
LST.CD.88.SIVlhoest447    -G--ATGC--GCAG--------GG-A---A-A--A-----C------AGCA-----C-G---CC-----A---AC---C--A-----C--AG----ATC---T--T--CC--GA--------GA------------T-----A--A--C--A--------G--GAGT--- 2288
LST.CD.88.SIVlhoest485    -G--ATGC--GCA-A-------GG-A---A-A--A-----C------AGTA-----C-G---CC-----AT--AC---C--A-----C--AG----ATC---T--T--CC--GA--------GA------------T-----A--A--C--A--------G--GAGT--- 2288
LST.CD.88.SIVlhoest524    -G--ATGC---CA-A-T------G-A---A-A--A-----C-----CAGCA-A--T------CCT--T-A---AC-A---------G--C--T-----GTCAC--------C--GA------A-AA------------T------T-G-----A--------G--GAGT--T 2285
LST.KE.x.lho7             -G--ATGC---CAGA-T------G-A--CA-A--A-----C-----CAGTA-C--T------CC-A-T-A---A---------A--------T------TC-C-C--T---C-AGAG-----A-AA-----------------------G-----A--G-----G--GAGC--T 3372
MAC.US.x.251_1A11         C---A-GG---AGGA-T-------------A-A-----C--------C--CA-A--TC----------A--T---C----C-----C-----TT----ATCAG----T----CAGAG-----A-AAC----CATT--T--G--T--G--T--A--------G--G------ 3321
MND-1.GA.x.MNDGB1         -----TGT---AG-A-------C--A---A-A--G-----C------AGTA-A--TC-G---CC---T-AT--ACCC-------C-----GG----ATCAG-T--T---C-AG-------A-AA-----CAT---T-----T-----T--A--G-----G---AGT--- 2707
MND-2.CM.98.CM16          --C-ATGTG-GCA-A-T-----------CA-A--G-----C-----C---TG---T--G-----G-----T--A-----C-----------C-----T--ATCGG-G--T---C--GGC-----A--C--G--C-----T-----C--A-------------G--C--C--- 3021
MND-2.GA.x.M14            --C-ATGTG--CA-A-C---A--T-----A-A--A-----C---------TG---T--G----GAG---T--A-----C---------------T--ATCGG----T---C-AGGC-----A--C--G--C-----T-----T-------------------C--T--G 2948
MND-2.x.x.5440            -GC-ATGTG-GCA-A-T------T----CA-A--A-----C---------TG------G---TG-----T--A-----------G--C-----T--ATCGG-G-----CC--GGC-----A--C-----C-----T-----C--A-----------------C--T--- 2581
MNE.US.x.MNE027           C---A-GG--GAGGA-C--------------A-----C---------C--CA-A--TC----------A--T---C----C-----T-----TT----ATCAG----T----CAGA------A-AAC----CATT--T--G--C--G--T-----G-----G--G------ 2790
MON.CM.99.L1_99CML1       -G----TGG-CCAT--G---A--G-A--CA--AAG--C--C-TC-ACAGCA-------GTGG---CCG-----A--A--C--G--G--CT-A-----CTC-T---------CTGA---T-CT-AA--G-------T---G--C--C--G-----C--------G--C--- 2726
MON.NG.x.NG1              -G--C-TGGC-AATA----TA--G----CA--AAA-----C-T--ACAGCA-C--TC-G-------A-T--T-----A-------------T-A------TCAG----T----CAGA-----CCG-G----------T--GA---T-G--C-----R-----------C--- 1319
MUS-1.CM.01.CM1239        CGG-AC-TG-GCAT--C---A--G-A---A-CAAG--C--C-T--ACAGC--C--A------CC--CA-----A--A--C--A--C--CT-GC-G--CCAGG----T--CC--GG-----CA-GA--G-------T---A-----G--------G-----------C--C 2700
MUS-1.CM.01.SIVmus_01CM1085 T-G--C-TG-GCAC--T---A--G-----A-CAAG--C--C-T--ACAGT--C---C-----CC-----T--------C--A--C---T-GC-G---CAGG-------CTT-GG-----CA-GA---------A-T---A-----A-------C--------G--T--T 2710
MUS-2.CM.01.CM1246        CCC-A--GGC-CAT--C--TA--G-A---A--CAG-----C-T--A-AGT--C--A-----CA----A---C-ACCC--C--G--C--CT-GG-C---GCAG-G-----CAT-GG-----CT-AA----TC--A-TT--G---T-G--------G-----G--------C 2808
MUS-2.CM.01.CM2500        CT--A--GGCTCAC--C--CA--G-A--CA--CAA-----C-T--ACAGTA-C--AC-CT----G--G--T---CCT--C--A--C--CT-AG-T------G----T--CCTAGGG--C-C--AG--G-----A-T---G--C--G--C--A--G-----G--G--C--C 2767
MUS-3.GA.09.09GabOI81     C-CGA---CG-CAT--G---A--G-A---A--AAA-----C----A-AGT--A--A----CCC---------------C-----C---T----T--AGCAG-T--T---C---TG----CTCGA-----C--A-T---A--A-----G-----G-----G--C--T--C 2276
MUS-3.GA.11.11GabPts02    C-G--C-TG--CAT--C--G--C--A--CA--AAG--C--C----A-AGCA-C--T-----C--G----------A--C--A--C---T-A-----ATCG-----------ACA-----CA-AG--G-------T---G--C--C--T--A--G-------C---C--T 2297
OLC.CI.97.97CI12          --G-A-GGG--CAGA-T------T-A---A-A--A-----T-----C--TA-A--GC-----CC---A-AT--AG-A-----A--------A--T--ATCAG----T-----T-G------A-GA--G---ATA--T------T-A--T--A--------G--G------ 2894
RCM.CM.00.SIVagi_00CM312  -G--ATGC--GCAGA-T--T--G-----CA-A--G--C------------CTGC---C-CT----GCC------A---------A--T-----T-----AGCA-----T---C-AGGG-----T--C--------A-------------T-----------G--C--C--- 2291
RCM.CM.02.SIVrcm_02CM8081 ----ATGT---CAGA-------TT----CA-A--G-----T---------TG---TC--T-----CCT--T--A--A------------C--A-----AGCA--C--T---C--GG------AG-----G-----A------------C--------------G-----T--- 2309
RCM.GA.x.SIVRCMGAB1       -GC-ATGTG-GAG-A----G---T-A--CA-C--G---------------TG--TTC-GT----GCCT--T-----A---------------A-----AGCAG----T---C-AGG------AG-G--G-----A--T---------G--T-----------G-----T--C 2499
RCM.NG.x.NG411            --C-ATGC-GGCA-A-T-----CT-A---A-T--G-----G---------TGC---C--T----GCCA--T--A-----C-----T--C--G---TAGC---T--T---C-AGG---T--C---------------A--T-------T-G--C-----------G-----T--- 2514
SAB.SN.x.SAB1             -GC-A-GGG-GCAGA-T-----GT----CA-A--A-----C------AGCTGC--T------CC------TCA---A-----A-----C-----C--ATCAG-C--T---AG-GAG-----C--A------------T--G-----C--------G--------C-----C 3418
SMM.CI.79.SIVsmCI2        C-G-A--G-G-AG-A--------T-A--CA-A--A-----C------AGTA----A---TGG-----G--T--AC-A-----A-----CG-TC-C---TCAG-C-------CAGA------T-AA--G---ATT-----G--C--A-----A--------------T--T 2442
SMM.LR.89.SIVsmLIB1       C-----GG-G-AG-A-T-----GT-A-----A--A-----------C--TA-------G--------A--T---C-----------C------T-G--ATCGG---------CAGA------A-AG------GT------G--C--A----------------------- 2401
SMM.SL.92.SIVsmSL92A      -GC-A-GG-GGAG-A-C------T-A---------A-----T---------CA-C--A--G---CCG-----T---C-A-----A--T-----TT-G---TC-G----T-G--CTGA------A-AA-----CATC--T--A--C--A-----------------C--T--- 2403
SMM.SL.92.SIVsmSL92B      -GG-A-TGG--AGGA-------G--A---A-A--G-----C-------------A---C----CCC--A--T--AC-A--C--A-----C--T-----ATCAG-G-----CC-AGA------A-AG--G---ATT-----G--C--G-----A--------G-----T--T 2382
SMM.SL.92.SL92B           -GG-A-TGG--AGGA-------G--A---A-A--G-----C-------------A---C----CCCG--A--T--AC-A--C--A-----C--T-----ATCAG-G-----CC-AGA------A-AG--G---ATT-----G--C--G-----A--------G-----T--T 2753
SMM.US.04.G078            C---A-G--G-AG-A-C------T-A-----A--G--------------CA-A--A--G-----G--A-----AC----C-----C------C---ATCAG-T-----C-CAGA------A-AG--G---GT------A--T--A-----A--------G--------- 2582
SMM.US.05.D215            C-----G--G-AGGA-T------T-A----------A--------C--C--TA-A--A--G-----GA----T---C----------T------C-G--ATCAG---------CAGA------A-AA-----CATT--T--A--T--A--G----------------------- 2576
SMM.US.06.FTq             C---A-G--G-AG-A-T--------------A--G--------C-----CA-A--AC-G--CCC---A-----AC-A--------C------C----ATCAG-T-------CAGAG-----A-AG-----CATC-----G--CT-A-----A-----------G------ 2604
SMM.US.86.CFU212          C-----G----AGGA-T-----C---------A--G-----------C--TA-A-----------G----------C----------T--C---C----ATCGG---T----CAGAG-----A-AA------ATC-----G--T--A-----A--G--------------- 2568
SMM.US.x.H9               C-----G--G-AGGA-C-----GT-------A---------------C--CA-A--T-----------A------C-A--C-----C------T----ATCGG----------CAGA------A-AA-----CATC--T--G--AT-A--T--A--G--------------- 2795
SMM.US.x.PGM53            C-----GG-G-AG-A-C------T-------A---------------C--CA-A--TC----------A------C-A--C-----T-----TT----ATCAG----------CAGA------A-AA-----CATC--T--G--AT-A--T--A--G--------G------ 3239
SMM.US.x.pE660.CG7G       C-----G--GGAGGA-C------T-------A---------------C--CA-A--TC----------A------C----C-----C-----TT----ATCAG----T----CAGA------A-AA-----CATC--T--G--AT-A--T--A--G-----G--G------ 3311
STM.US.89.STM_37_16       C-----GG-G-AG-A-C------T-------A--G--------------CA-A---C-------G-G---T---C----------T--C---T-----ATCAG----------CAGA------A-AA-----CATC-----G--T--G--------------G--------- 2969
SUN.GA.98.L14             -G---TGT---CAGA-T------G-A---------G-----C--C---A-A--A--G---CC---T-AT--A---------A-----C--A-----GTCAG----T---C--G-------A-AA------GT---T-----C--A---------C---------AGT--T 3377
SYK.KE.x.KE51             GT-CA----G-...A-G------T-A--CT--AAA--C--C----ACA----A--TC-----------G-----AC--------A-----------AG-G--C-------------TCTT-T-----AGAG-----C----T---G---T-G-----A--------G-----G--C 2872
SYK.KE.x.SYK173_COMGNM    --CG-C------...C-----A-TA-A---T-AAAA-----------AC-----C---C-G--CA----G--T---CCA--C--G--------AG-G--A--C-----T-----CCT-C--G--AGAA--------A-TT-CA---T-A-------------------T--C 3214
TAL.CM.00.266             CTC---TG--GCA-A----G--C--A--CA--AAG-----C----ACAGCA-C---------CCC--------TCAC----C--A--C--C-----C--GTC-C-G---------T-GG-----A-GA--------A-T---G-----A--C--A--------G--T--G--C 2989
TAL.CM.01.8023            C-C-------GCAGA-------CT-A---A--AAG-----C----A-AGCA-C--A-------CC--------GCCCGT--C-----T--C-----C--GTC-C-G--T-----C-G--------GGC-G--C----T---G--A--C--C--A--G-------------C 2487
TAN.UG.x.SIVagmTAN1       ---GA--G---CAGA-C--------A---A-A--A-----C----A-AGTA-A--A--G---CC---A------CCC-----A--T---------C--ACAAG----T----TAGGG--T--A-----G-----A-T----TGT--A--T--A-----------------C 3222
VER.DE.x.AGM3             ------TG-CGGA-A--------T-A--CA-A--G-----C----A-----A-A-A-----CCC---G--T--A---------C--T---------T--ATCAG----T---C-AGGG-----T-C-------------T----CTGT---------A--------G-----C--- 2745
VER.KE.x.9063             -G--A-TG---CAGA--------T-A---A-A--G-----T----A---CA-A--A------CC---G--T--A--A--C--A---------A--C---TCAG----T---C-AGG------TG-----G-----A-T----TGC--------A--C--------G--T--- 3250
VER.KE.x.AGM155           -G--A-TG--GCA-A----CA-TA-A--------G-----------A-AGCA-A--AC-G---CCT--G--T--A--A--C--A--T--C-----C----CGG------------GG------C--A--------A-TT---TGC--A--G-----C-----G-----C--G 3242
VER.KE.x.TYO1_patent      G-----TG-G-CAGA-------TT-A-----A--A--C--C----A---CA-A--A--G---CC-A-T--T-----A--------T-----T--T--C-CAG-G--T---C--GG---C--------G-----A-T---CTGT--C--G--A--G-----------T--T 3247
WRC.CI.97.97CI14          G------T--GAACA-TT-----G-A---A-A--G-----------C--TA-A--A------CC---A-AT---C-A--------T-----AG----ATCAG----T------T-T-------AG------GTA--T-----A--G--T------------------T--T 3320
WRC.CI.98.98CI04          G------T---AATA-TT-----G-T---A-A--G-----C--------CA-C--AC----CCC-----AT---C---------A--T-----TG----ATCAG----T-----AT-T--T----AG--G---GTC--T-----A-----T--A--G-----G--T--C--T 3335
WRC.GM.05.Pbt_05GM_X02    -T-------G-AATA--------T-A---A-A--A-------------TTA-C--A------CC---A-AT--AC----C--A--------AG-----TC------T-------A------T-CC------GTA--T-----A--C--T--A-----------G--T--T 2249
```

```
H1B.FR.83.HXB2          GCAATATTCCAAAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTGGGGACT 3190
Pol                      A  I  F  Q  S  S  M  T  K  I  L  E  P  F  R  K  Q  N  P  D  I  V  I  Y  Q  Y  M  D  D  L  Y  V  G  S  D  L  E  I  G  Q  H  R  T  K  I  E  E  L  R  Q  H  L  L  R  W  G  L

H1O.CM.91.MVP5180       --C----T--G---TCA--------G--TC----T--A--A-------AGC--C-----AG---AA--T---G------A-----C--A---------A--T---CC-T-G-CAG-A------AG-GGG-C--ATT---T--GG-----T-A-ATCA------T- 3219
H1N.CM.95.YBF30         -----T--T--G----CA--------G--TC----T--A--A--C---G-GA--C-----G---A-C--T--C--G---------CC-C---G---------A--T---CC-T--CA--A------GAGGC-G----A--C--T---G-T-----T----A-----CT- 2784
H1P.CM.06.U14788        --C--T--Y--GCA-TCA-----T--G--------A--C-----R--GAGT--------AG---AA----Y-------------Y-A------G--A--TY-GCCTT-GTCTG-A------CAGCGGG----A-A--Y--GG-R--C--C-ATGTA-----GT- 2661
CPZ.CD.06.BF1167        -----C------TAC--T-----------T-----T--A--C--G--GA-GC---T--A----A-G-T--C-------------CC-------------T--GAC-GCT-AAA-A----A-GA-GTGG-TC-----T-A---A-C--CT--AAAGCA-----CT- 3298
CPZ.TZ.00.TAN1          -----T--T--G--C--T----------C----T------C-----G--C-C-----TG-G-AC--T--C-----T--------C-T--CA----T--A--TC-TA-TGA--A-G-A-----G-A-CTG---A--A----------G-------A-CA------T- 2834
GOR.CM.07.SIVgorCP2135con -----T-----GCACTC----------AC-T--T-----C-----GA----C----A----AA----------G---------C----------G--A--T---CCC---ACAG-A-----GCAG---G----AA-A--C--GG-----T-A-ATGCA------T- 2726

MAC.US.x.239            --C--C------TAC-CT---G-C-TG-GC----A--C--C--G--GGC---------TG-GACCT-AGTC--G--T--------CA-C-TAA---CTAG----AGGAC-GACCT-G-A---GAC-GGGT-G-TTTAC--TCA-AGG---TCT--AAT--CATA--GT- 3492
Pol                      P  A  I  F  Q  Y  T  M  R  H  V  L  E  P  F  R  K  A  N  P  D  V  T  L  V  Q  Y  M  D  D  I  L  I  A  S  D  R  T  D  L  E  H  D  R  V  V  L  Q  S  K  E  L  L  N  S  I  G
H2A.SN.86.ST_JSP4_27    -----T--T--GTAC-CA----GGC-GG-------A--A--C------GC---C--G--T--CA--C--AT---G----------A-C-TGA---CCAGC---AGGAC-GATTTAG-A---GAC-G-GTGG-TCT-C----A-AGG---T---AAATG-CCT----T- 3020
H2B.CI.x.EHO            -----T-----GTAC-C---G----GG-AC----C-----C------GCC--CAAT--TG-CAC---AATC--G----------CA-TCTC--G-C-AG----AGGAGCGATCT-G-----GAC-GGGT-G-GTCTC-A--A-A-G-G-TAT-AAAT-ACAT----T- 3546
H2G.CI.92.Abt96          --T--C-----CA--CA----GG---GG-AC----C--C-----C------GCC---GAG--TR--A----YATC----T----------CA-TCT---G-CCAG----AG-ACTGATTTAG-A---GATCA-GT-G-TTT-C---A-AGG---TCT-AAAT-AAACR--GT- 2902
H2U.CI.07.07IC_TNP3     --C--C--T--GTA--CT----G----G-GC----C--C-----G---GCC--CGAG--TG--A-C---ATC--G--T--------A-CCT-----CCAG----AG-AGTGACCTAG-----GAC-AGGTGG-CTTAC--T-A-AGG-T-T----AAATCA-CT---GT- 3007
H2U.FR.96.12034         A----T--T--GTA--CA----GG--CC-AC-G-----A-------GGC----GA---TGCT--C-----AT----G--T--------CA-TCT-A----CCAG----TAG-A-TGATTTAG-----GATCAGTT-G-TTCAC----C-AGG-G-TA---AAT-A-ACA---T- 3057
ASC.UG.10.RT03          A-T-----T--G-AC-CA--AGG-G-CC-G--GC-A-AGA-C--G---A--T-C---C-A--CC----GAT---G----------CC-CCTCA----G--A--TGA-AC-G--ACAG-A--C--G-A--T-G-----C-AA-T--GA-CATC--TC-A-AACAA--GT- 2599
ASC.UG.10.RT08          A-T-----T--G-AC-CAG-AGG-G-CC-G---C-A-AGA-C--G---A--T-C---C-A--CC-C--GAT---G----------CC-CCTCA----G--A--TGAGAC-G--ACAG-A------A--T-G-----C-AA-T--GA-CATC--TC-A-AAC-A--GT- 2623
ASC.UG.10.RT11          A-C-----T--G-AC-CA--AGG-G-CC-G---C-A-AGA-C--G---A--T-C---C-AG-CC-C--GAT---G----------CC-CCTCA----G--A---GAGAC----ACAG----C--G-A--T-G-----C-AA-T--GA-CATC--TC-A-AACAA--G-- 2608
COL.CM.x.CGU1           A----C-------G-A-CAG-AG---G----C----A--G---C--GAG-GGA-A-TTG...CAA--AGC---G----------A---C-------A--TC-GCCTT---A-G---------GGGC-G--C-------C---G-TGCA--C---TAC-AT--GT- 2962
COL.UG.10.BWC01         A-------T--GGCA-CAG-AG----G--T-----A--AA-------GA--GGGAA-TTG...CAA--TGCA--G--T--------C-T--------G--A--T--GCC--A--AAG-A-------A-G-----C-----T-A---G-TGCAT-AC-CTAT-AT--GT- 2761
COL.UG.10.BWC07         A-C-----T---G-A-CAG-AG----G--T----------G----G-GGG--GGAAAGCTA...CAC--AGCC--G--T--------C-A--A-----AG---T-----C---AAAG-------GCA-G--G--C-AA--G-T--GACT-TCT--C-ACAT-AT--GT- 2989
DEB.CM.04.SIVdeb04CMPF3061 --T---------CAC-CTG-TCATCGGG-A--G--AGGA--------GA-T-C-----TG--C-C--AGTC--G----------CC-A-TG-----G--A---AG-AC-GA--AA--A--C---CT-GT-G-CA-AACA--A---G-TT----CC-AGA-A----GT- 3180
DEB.CM.99.CM40          --C--T--T--GCCTCAG-AGGG-G-CAG---C--ATC-----GG--A-G---------GC---A-A--AGTC--G--T----------ACTA-----G--A--T---AGG-A--AT--A-CCT--AG--GG----CAC-AT-A-----GTTC--CC-AGA-A----T- 2999
DEB.CM.99.CM5           --C--------GGCA-CAG--GGC-GGCAGC-GA-A-T-----AGG--A--T-C------T--A----AGTC-----T-------CC-A-TG--------AA-T------T-AG-TCTCA---TT-GA-C-GG--A-AC----T--GA--TT----CC-AGAAAA---GT- 2993
DEN.CD.x.CD1_CM0580407  A-C--T--T--GGC-CAGCAGAT----CG-TAGC-AA-----G--GGCCC----T--TG-GA--T-AATC--G--T--------C-C-TG-----C---A--TAG-TCCC---AA-----C--G-A--TGG-TA----A--A---G-CTTCT--GCAGTA-----CT- 3077
DRL.DE.11.D3            -----C------GC--CAGCTGAC--G-----C----A--C---G-GA-G--------TG-GA----A-----G----------C---T------GAG---TAG-CC--AGCATTTA---GAGCAG-TG--TA----AT-A---AC---CT-AAA-TTC-----G-C 2664
DRL.x.x.FAO             -----C------GC--CAGCTGAC--G-----C----A--C---G-GA-G--------TG-GA----A-----G----------C---T------AG---TAG-CC--A-CATTTA---GAGCAG-TG--TA-A--AT-A---AC---T-AAA-TTC-----G-- 2757
GRV.ET.x.GRI_677_gri_1  A----T------G-A--CGGCAG-----C--T-----GAGA-C-A--GG---C-C---T-GGT---AA--TGTC-----------C-----GGT-G-CG--A---CAT--TGAGACTAGA----ATCA-C-GG----CATAG-A---A-GATG----C-AGAAAAA--T- 3363
GSN.CM.99.CN166         A-C--C-T---AC-TTG-AG-TTCCT-GC--CGA-AGA-----G-T--GG-G-----TG-G--AT-AATA--G----------ACTGA-------A--TAGGA--C-G--AG-A--C--GCA-GT-G-GC-CA-AA-A---A-TTTA--AACATCC-ATAACA- 2918
GSN.CM.99.CN71          A-C--T--T--GGCC-CTG-AG-TTCCT-G---T-T-AGA-----G-C--GG-G-----TG-G--A--AATA--G----------C----TGA-------A--TAGGA--T-G-CAG-------GCA-GT-G-GC-CA-AA-A---A-TTTGT--AC-TCC-ATAACA- 2900
LST.CD.88.SIVlhoest447  TGC--T--T--GC--CAG-AG--GG-C--C--TCAGAA---------TT---C--T------A-CT-G----------------A-TCA-------A--TAGG--C-AGAAAGG---GATCAGGC-G--A-A--AT-----G-T-TCT-AG---CA---AATT- 2458
LST.CD.88.SIVlhoest485  TGC--T--T--GC--CAG-AG--GG-C--C--TCAGAA---------TT---C--T------A-CC-G----------------ACTCA----G--A--TAGG--C-AGAAAGG---GATCAGGC-G--A-A--AT-A---G-G-TCT-AA---CA---AATT- 2458
LST.CD.88.SIVlhoest524  TGT--T--T---GAA-CAG-T-GTGG-C--C-T-CAGAG--C--G---TT----------C-CC----C--G----------TCA----T--A---AGG--G-A--AAGGA-----TCAGGC-G--A-A--AT-A---G---T-T-AA-A-CT---AACT- 2455
LST.KE.x.lho7           TGT-----T--GG-A-CAG-AG--GG-C-TC-CTCAGAG-----G--TT----------GA-C--T-C-----T----------A-TCA-------A--TAG---G-G-AAAGGA---GATCAGGC-G--A-A--A---C---G---T---TA---CA---AACT- 3542
MAC.US.x.251_1A11       --C--C------TAC-CT----G-C-TG-GC----A--C--C--G--GGC---------TG-GACCT-AGTC--G--T--------CA-C-TAA---CTAG----AGGAC--ACCT-G-A---GAC-GGGT-G-TTTAC--T-A-AGG---TCT-AAAT--CATA--GT- 3491
MND-1.GA.x.MNDGB1       TGT--C--T--G-G-CAG-AG--TC-C-GC-G---GTA-------GA-CC----ACAG--CAGT-A----------------------T-----G--A----ATAC-GC--AAG-----GAG-A-GCT----TA--AT-A--GGCTTTAT-AA---CA---AACT- 2877
MND-2.CM.98.CM16        --C--T-----GGCA-CAGCAGAT---------C----C--C---G-GAG-C-------TG----G----------T-------C-C-T---T--GAG---TAG--TTGCCCCAG-A--C--C-G--TG--TC-A---T-A---G-C--C--C---TTT-----G- 3191
MND-2.GA.x.M14          --------T--GCA-CGGCAGAC-----TC--A-AA-A--C-A-G-GG--T-C-----AG--T-A--T----------C---T---G---AG---TC-GAGTGCGTCAG----C-AC-A--TGG--A-TAG-T----GG-----T-AAG-TTC-----G-- 3118
MND-2.x.x.5440          -----T-----GGCA-CAGCAGAC-----TC--A--G------AGG--A-GC------AG-CT-A--------G--T--------C---T---G--GAG---TC--AGTGCCTCTG-A--C--T-A--TG----TA--AT-A--GG-----T--CA-TTC-----G- 2751
MNE.US.x.MNE027         --C--C------CAC-CT----G---TG-GC-G--A--C--C--G--GGC---------TG-GACCC-AGTC--G--T--------CA-C-TA---CTAG----AGGAC-GACCT-G-A---GAC-GGGT-G-TTTAC--T-A-AGG---TCT-AAAT--CATA--GT- 2960
MON.CM.99.L1_99CML1     A-TG-G--T---CA--CAGC-G--G-GC-ACC-C--GAGA-C--G---A--T-C----GG--AC-T-AATA--G----------CC-ACTGA------AA-TCACTCTC-GCCAG-A---C--G-A--T-G---G--TCA-A---GCGACC--AC-A---AAA--GA- 2896
MON.NG.x.NG1            A-----T--T--GTACTCAGCAT--G-GC-GC-RA-AGRAA--C--G-G-A--T-------GG-CACCC--ATC--G----------C-ACTCA----T--A---CATTCCC--CCTG-A--C--G-AG-T-G---ACA-CCA-T---ACG-TR--A-CAGAAAA---GA- 1489
MUS-1.CM.01.CM1239      A-C--C------TAC-CGGCAG-----T-A---C--GAAA-C--GG--GC-------T--CACCC--ATC--G----------C-CTAA------CA--AG---G---TC-GGA-----G-G-GT-G---CAC--A-T--GA-CATG--C----AACT---GA- 2870
MUS-1.CM.01.SIVmus_01CM1085 A-C--C-----GTTC-CAGCAG-C---C-------AGAGA---AG---GC---C-----TG-CAC-C--ATC--G----------CC-CCTGA------CA--AG----T--TCAAGC--C--G-A-GT-G---CAC--A-T---A-CATG--C--A-A-CTT--GA- 2880
MUS-2.CM.01.CM1246      A-C--T--T---GCC-CA-CAG-C--GC-TC------AAG---G--GG-C----AG---GCCT-A--AGTC----------CC--CT-A---G-A---AG---------G-A-----G-GGCTGG-CA-AAGAG-T---G-C-TC--AACA--CAAA--GA- 2978
MUS-2.CM.01.CM2500      A----T--T---GCA-CA-CTG-----T-A------AAA-C--G---G------AG--TGC-T-A--AGTA-----T-------CC-ACTCA---G--A---AG---------G-A-----G-GGCTGG-CA-AAGAG-T---G-C-TC--AACA--CAAA--GA- 2937
MUS-3.GA.09.09GabOI81   A-T--C--T---CAC-CAG--T----CC-A--GA--GACA----GG--A--T--AAG------A-AT-GAT---G--T--------C-A-TAA---C-------AGG--CC-T---G-T-------A-GT-G----A-CAA-A--GA-TATG-----A-A--ACAACA- 2446
MUS-3.GA.11.11GabPts02  A-T--T--T--G-AC-CAGCAG-TTCTC-GC-GA---AGA-A--GG-TA-GT--AAG--TG-GACC---AT-------------CA--CTA-----G--A---AG---G-AG---G-A-----G-AG-T-G-G--A--AA-A--G--G-TC---G----AA-AT--GA- 2467
OLC.CI.97.97CI12        TGT--C-----G----CAG--G-TC-TT-GA-G-------GG--G--GG-------T--G----A-T-A--C--G--T--------C-------G----A-----GCC--G--CAG-----TT-GAG---G--A-AC-------GG-G-TCT--AG---T-ATAAC-- 3064
RCM.CM.00.SIVagi_00CM312 --C-----T----CC-CTGCT-AC-----T-----A--AG-C--G-----------TT----C--------G--T--------CC---T--------A---AG-AC-GA-TTAG-A--C--TCA--TG---A--A-AT----GG-----T--C-CGCC------T- 2461
RCM.CM.02.SIVrcm_02CM8081 --T-----T----CA-CTGCA-AT-G-------------A--C-----G---------TG----A-----------------A-T-----------TAGGAC-GA-TTAG-A----ATCAG-TG---A-AAGA--C---G-----T------TC-----GT- 2479
RCM.GA.x.SIVRCMGAB1     --C--C--T--GGCATCAGCC-AT--G-----C----A-----GG--G-G--------TG-CA-C--T--C--G----------C-C-T------C--A--TAG-AC--AGTT-G-A---GAC-A--TG--CA-AC-AT-A---G-T-----AC--TTC-----CT- 2669
RCM.NG.x.NG411          --T-----T----CA-CAGCT-AC-----T-----A--A-----G---G-G--------TG----C--------G--T--------T------G--A--TAGGAC-GC-CATG-A-----TCAG-TG--CA-AAG---C---G-G-----C---CC-----TT- 2684
SAB.SN.x.SAB1           A----T--T--G-CA-CAGCC-AC-----TC-GC--GAA-----GC--A-G--C-----TG---A---A----------------A---TAA---CCAG---TAGGCC--AG-CAG-A----CT-GT--TGG--C--C--T-A---G-CT---AGAA-CT----GT- 3588
SMM.CI.79.SIVsmCI2      --C-----T---GA--C-----GG--GG-G-----A--C--------GCT---GAG--TGCTAC---ATC--G--T--------A-CCTCA---C-AG---TAGGA-TGATTT-G-----GACCA-CTG-G-CAC----C-AGG-GTTGT-AAATG-A-T----T- 2612
SMM.LR.89.SIVsmLIB1     -----T--T--GTT--CA----A---GG-AC----A--------------ACC-------TG-GACCC-AAT---G--T--------A-CCTAA---CTAG----TAG-AC-GACCTAG-----GAC-AGGT-G-CTTAC--T-A-A-G----T-AAATG-C-TA---T- 2571
SMM.SL.92.SIVsmSL92A    -----T--T--GCA--C-----G-C--G-----C-----G---GC----T--TG-ACCC-TGT---G--T--------CA-TCTGA---CTAG----AGGTC-GATTTAG-A---GAT-AGGT-G--ACAC----A-A-G---TTGT-AAATGAT-T---GT- 2573
SMM.SL.92.SIVsmSL92B    ---G-T-----GGCA-C-----GGC-GG-A-----A--C--C------GCT--C-----TG-GC--C--GTC--G--T--------ACTAA-T--CAG-A--AG--GCC--ACAG-----GAC-A--TGG--ACAC--T-A---G-CATG--CAAC-ATCT---GT- 2552
SMM.SL.92.SL92B         -----T-----GGCA-C-----GGC-GG-A-----A--C--C------GCT--C-----TG-GC--C--GTC--G--T--------ACTAA-T--CAG-A--AG--GCC--ACAG-----GAC-A--TGG--ACAC--T-A---G-CATG--CAAC-ATCT---GT- 2923
SMM.US.04.G078          -----T--T--GCA--CA----G-C-GG-A-G--C--C--C------GCC-------TG--AC-T-GAT-------------A-C-TAA---CCAG----AGGAC-GACTTAG-A---GAC-A-GTGG-CCT-C-A--A-AGG---T---AAATG--CTA---T- 2752
SMM.US.05.D215          -----T-----GTAC-CA----G---TG-AC-------A-----G---GC---C----TG-GAC-C-AATC-------------CA-C-TAA---CCAG----AGGAC-GATCTAG-----GAT-AGGT-G-CTTAC-AT---A-G-G-TCT-AAATG-CATA---T- 2746
SMM.US.06.FTq           --G--T------CAC-CA----G---TG-A--G-------------GGCC-------T---ACCT-AGT--G----------CA-C-TGA---CTAG---TAGGAC-GACTT-G-A---GAC-GGGTGG-CCT-C--T-A-A-G---TAT--AATG-CCTA--GT- 2774
SMM.US.86.CFU212        -----T--T---CA---CA----G-GGTG-A--G-----A--------GC--------TG--ACCC-AAT---G----------CA-CCTGA---CCAG----AG-AC-GACTTAG-A---GAC-G-GT-G-GCT-C--T-A-A-G---TCT-AAACG-AAT---CT- 2738
SMM.US.x.H9             --T--T--T--RCA--CTR---G---TG-A-----A--------------GC--------TG-GACCC-GATC-------------CA-C-TAA---CTAG---TAG-AC-GATTTAG-A---GAC-GGGT-G-TTTAC-AT-A-AGG---T----AAC--TATA---T- 2965
SMM.US.x.PGM53          --T--T--T--GCA--CT----G---TG------GA-----C-----GC--------TG-GACCC-GATC-------------CA-CTAA---CTAG----AG-AC-GATTTAG-----GAC-GGGT-G-TTTAC--T-A-AGG---T----AAC--CATA---T- 3409
SMM.US.x.pE660.CG7G     --T--T--T--GTA--CT----G---TG-A-----A--------GC--------TG-GACCC-GATC-------------CA-C-TAA---CTAG---TAG-AC-GATTTAG-----GAC-GGGT-G-TTTAC--T-A-AGG---T----AAACG-CATA---T- 3481
STM.US.89.STM_37_16     -----T--T---TA--CA----G---T--A--------A--C--G-G-GCC--C-----TG--AC-C-GATC-------------A-C-TGA---CCAG----AG-AC-GATCTAG-----GAC-GGGT-G-TTTAC--T---AGG---TC--AAAT-AT-TA---T- 3139
SUN.GA.98.L14           TGC--C--T--GG-A-CGG-AG-G-GCC---------AAA-------G---GT-------AG-TCAAT-G--C-------------ACTCA---------T-AT--G-A-AA-A-----GA-GA--T-G-GA-AC-AT----G--G-TA--CA--GAA---AATT- 3547
SYK.KE.x.KE51           A-T--C--T--GGCA-CAG-A-ATC-G--AC-CC-A---A-A--GG--A-G---T----T---A-A--AGTA--G----------CC-T-TG-----G---A--AG-TC-GA-AA-G-A---G-TCA--TTG--AGT--AA-TGT-A--AG---A--AGCAGTA---T- 3042
SYK.KE.x.SYK173_COMGNM  A-T--C--------C-CA--A-ATC-G--AC-C--A--------GA-GT--T---TG-CAC-C-AATA--G----------C--A-TGA-T--CA-----TAG-TC-GAGAA-GCA---CAGGAG-T-G---C-AC-AA-AGTGAC-GCA--A---A-AAGT----T- 3384
TAL.CM.00.266           A-C--C-----CA--C-TC-GATC--T-A--GAG-AAA---GG-CA-GT-C--TC-GG---CAC---T-------------CC-A-TG------AGCA-TGAGAGCT-G-A-G-A------GG-G-T---GA---A--A---GCT-TCT-AAAC-A-GTA--GA- 3159
TAL.CM.01.8023          A-C-----T---TAC-CTTCAGATC-GC-G---AG-AAG-----GG---A--T-C---C-GG---CGC---TC--G----------CC-ACTAA-----AG---TGA-AGCT---AA-------G-G-T-G-G-GA-----T---GCC-TC--TAAC-A-GTA--GA- 2657
TAN.UG.x.SIVagmTAN1     A-C-----T--GCAC-CAGCACAGCG-G--C--A-AGAAA-AG-----A-GGT--A---TT----AA--AAT---G----------CA--GGA-----A--T--GCC-GAGAAAG-A---GATGA---GG----A--T--GTCC-A-ATTAA-A--CT- 3392
VER.DE.x.AGM3           A-T--T--T--G-AC-CAGCAG-TTCC--TC----AGAAA-A-A---GG-GTTAAA-CC-C---ACC--TGTG----------CC-A-GG-----G---C-GGA---TGA-TACAC----GATCGGTTGG----AC-A-A---ATGA-AT-AAGTGCC------T- 2915
VER.KE.x.9063           A-T--T--T--G-A--CAGCAG------T------AAAA-A-A-G-GG--TTAAA-CCAT-GACA--AGTG-G----------------GG-----G---C-AGA---TGA-TACACC---GAC-AGTT-G----AC-----T---GTCA-G--AAAA-CA------T- 3420
VER.KE.x.AGM155         A----T-------AC-CAGCAT-------TC----AGAAA-A-AG---G--TTAAA-C-GC-GACG--TGTC--G----------CC-C-GG-------AC-AGA---GGGTCCAA-----GATCAGCT-G--C-AACA--T--GA-TAGAT--CAAGAA------T- 3412
VER.KE.x.TYO1_patent    A----C-------A--CAGCAG--TCC--T--G---GAGA-A-A--G-A-CTTG----CAC---ACC--TGTA----------------A-GG-----T---C-AGA-A-TGA-CACACC---GAC-A-TT-G----AC-A-T--A---AC-A-AT-ACAAGCC-----CT- 3417
WRC.CI.97.97CI14        TGT--T------TCA-CTG-AG--TCTC-A-----AATA------GG-A-G------A-T---CAAG-TCT-----T--------C-A---A-T-----A-----G--C-AGACAG-A---CT--AC--GG----A--AT-A---A-G-TA-------CA---AATT- 3490
WRC.CI.98.98CI04        TGC--T------GCA-CAG-AG----GTT-G-----TGTA--C-AG-----G-----TACT--GCAG--TCTC--G----------------A---A------AG------T---A-CCAAG----TT-GAG---G-----A---T-A---------TG-----CT---AATT- 3505
WRC.GM.05.Pbt_05GM_X02  TGC-----T---GCA-CAG--G-TTCTT-GC-G---ATC--------GA--C----TA-AG-GCAG--AGTC-----T--------C-------------A--T-AT----AGAAAG-A---TTGCA--TG--TC-----T-----G-T-T---T--A-CT---AAC-- 2419
```


Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

```
H1B.FR.83.HXB2              AGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAAACTCCTTAGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGAGAGATTCTAAAAGAACCAGTACATGGAGTGTATTATGACCCATCAAAAGACTTA 3524
Pol                         Q  I  Y  P  G  I  K  V  R  Q  L  C  K  L  L  R  G  T  K  A  L  T  E  V  I  P  L  T  E  E  A  E  L  E  L  A  E  N  R  E  I  L  K  E  P  V  H  G  V  Y  Y  D  P  S  K  D  L
H1O.CM.91.MVP5180           -A--C--T-A---A----G-----AAG----G--C--GT-AA-C------------T--T-G-----G---G----TT---GTA----G-----A------T-A-A------------A-AG--------G--------------A-----CC-G--TGAC--------G 3553
H1N.CM.95.YBF30             ----C--T-----A--C-G-----AA--GC-C-------T-AA-C---------------TT-G----------G-CAACT-T---------------AT-------A------------G-----AT----------CC-G--------C------------GG------A--- 3118
H1P.CM.06.U14788            -A-----T-----A----GG----AAG-G-----Y-----AA----------T--RT-CT-----------G-TG-YT-T--TAG----------AT-------A-AG-----T-A---A---T--------------G--------T-----CC-A---GA---G--A--- 2995
CPZ.CD.06.BF1167            -A-----T---------C-G-AC---AG--C-C--C---T-AA----G---GT-----G---------T---G---AGT-G--T-C---------AT-----T-A-AG-----T-A-C-A---T--------A-------G--GACC--C---------AAGG--CCTC-T 3632
CPZ.TZ.00.TAN1              -A-----T-AT--A--A---ACT-AAG--C---C---T-GA--------GTA---G-T----T--GCC-G--GA-A-G--CAGG--------AT-G--GT-A-A------T-AGC--------------AAG--T--A----CA--C-----T--TAA-TT-CCTC-G 3168
GOR.CM.07.SIVgorCP2135con   -A--A--T-----A--A-G-----AAG--C-C--C---T-AA----G--T--T---T----T---------G-GA-C--T-GTA----G-----AT--------A-A---G-----G--A--AT----------C--G--------------CC-G--TGA---G--A--- 3060

MAC.US.x.239                -A-----T-----T--A---ACC-AA--TC-C----GGT-AA----------AA-TGA-T--------G-A-G-T-AGTGG--T--GATG-----AGC----TAT-AG-----T-A-ATA-----C-GTC-GGA-CA-G-A---TGT-----CC-AGA-GGC--GCCA--- 3826
Pol                         A  Q  I  Y  P  G  I  K  T  K  H  L  C  R  L  I  R  G  K  M  T  L  T  E  E  V  Q  W  T  E  M  A  E  A  E  Y  E  E  N  K  I  I  L  S  Q  E  Q  E  G  C  Y  Y  Q  E  G  K  P
H2A.SN.86.ST_JSP4_27        -A--C-----------A--GACC--AA-C-------GGT-AA-C-------AA-TGA----C-------AGG---AGTGG------TT------AGCG----A-A--------A-ATC--CT---GCC-GGA-CA-G-A---TGC-----CC-AGAGGA---G--GC-- 3354
H2B.CI.x.EHO                -AC---T---G---------GACC-----TA-----------AA-A--G---AG-TGA-C------------AGG---A-TGG--T---TTG------GC----T-CCAG-------A-ATC--C---G--C--GAGCAGG-A---TCC-----CA-AGA-GGGGT-CCT--- 3880
H2G.CI.92.Abt96             --C-A--T-----A--C---AC--AAA-CC-G--Y---T-AA-------T-AR-TGA-CT-G--G----A-G---AGTGG------TTG-----AGC----T--CA--------AGATT--C---R--C-GGARCAGG-A--C-CC-----CA-AGA-GR-G-GCCT--R 3236
H2U.CI.07.07IC_TNP3         -A--A--T-----A-----GAC--AAA-C-----C-----GA-C-------AA-TGA-C---------G-A-G---AGTGG------TTG------GC---G---CAG--G--T-A-ATC--CT--G--C--GA-CAGG-A--CACA--C--CAGAGA-GGGG--CCT--- 3341
H2U.FR.96.12034             -AG----T-----A--A--GAC--AA--TC-T-------T-AA---AG--G-AA-TGA----C--------AGG---AGTGG------ATG-----AGC----T-T------G--T-AGATT---T-G-GTC-GGA-CA-G-G--G-CT-----CAGGGA-GA-G--A-TC-- 3391
ASC.UG.10.RT03              -A--A--T---------A---AC--A-GCTC-G------T-AA-----------A----GCT---------A-G---AGG-C---------G-----AAC---G--T----------CA--CC---T--G-CC--GA--CTA-A--G-GA--C-----TG--GAG--GCCA--- 2933
ASC.UG.10.RT08              -A--A------------A---AC--A-GCTC--------T-AA-------G--A----GCT-------G-A-G-G-AGG-C---------G-----AAC---G--T-T---G---CA--CC---T--G-CC--GA--CTA-A--G-GA--C-----TG--GAG--GCCA--- 2957
ASC.UG.10.RT11              -A--A------------A--GAC--A-GCTC-G------T-AA----G-----A----GT-----------A-G---AGG-C---------G-----AAC---G--T---------TCA--CC---T--G-CC--GA--CTA-A--G-GA--C-----TG--GAG--GCCA--- 2942
COL.CM.x.CGU1               -AT-A---------A--C-GGAC--A-G-GA--A-----TG-AC--A---T-AG--GCATT-GTTG--G-A-G-G-AG--T-GT-C---------AGAG--G--ACA---T----A-C----AA-C-------GA----T--G---------C----ATA---CATG-G---A-T 3296
COL.UG.10.BWC01             -AG----T-----A-----G-AC--AAG--A-CA----GTGTACA-A---G-AG--GCAT---TT------A-G---A-----GT-C---------AGA------A-A-------T-AGA----AG-------GGA-CA-A-A-----T-----CAGA---GAGG-----A-T 3095
COL.UG.10.BWC07             ----C-----------A-GGACT-AAG-----A-C--GTG-ACA-A---G-AG----GG--TCT---T-A-G---AG--CT---G-C-------AGA---GT-A-AGC-G--T--G-----CT-G------GAGCA-G-A-----A-------A-A---GA-G-G--GA-T 3323
DEB.CM.04.SIVdeb04CMPF3061  -A--A--TT--------A---AC--AA--CC-G--C--GAGTA--------GGTTCC--GGT--------CCCTG--AAGT----------G---C--GCC--G----A----GC-----C----CT---------A---GTCA--GACC-----CA-AGA-CAGG-G---C-C 3514
DEB.CM.99.CM40              -A-----TT--------A---AC--AA--C--------GGGCAA-C------GGT-CCG-G-T--------CCCGG-TA-C--CT-T--------TC-AGC------TAG---G----AG--A--C--G--GC--GA---GTCA------T--C---A-AGAGCAGG--CCA--- 3333
DEB.CM.99.CM5               ----A---T----A--A--GACC-AA--C--G--C--GGCAA-C------GGTTCC--GG--------TAA-G--AGC--TT----G------C-AGC-------AGG--G--T-A---A-C--CC---C-GGA---CTCA--G--A--C--CA-AGA-GATG-GCC-C-- 3327
DEN.CD.x.CD1_CMO580407      --T-G---T----A--A--GAC--AAG-----------GAG-A---------GTA----ACT--GG----ACC--TT----C--G--G-G-------GC---G--TAA----GCACA---A--AT--------AGTA-TTCA---AGT------AGG---GAG---CCT--G 3417
DRL.DE.11.D3                -A-----TT-C--A------AC--A-G-GC-C-------T-AA----G---G-----C-C---GAT----A-G--GA-TGG--CAGG--------A------TAT-A---------A-CT---AT---------A-A-G-----G--A------C-G---GA----CC-C-- 2998
DRL.x.x.FAO                 -A-----TT-C--A------AC--A-G-GC-C-------T-AA----G---G-----C-C---GAT----A-G--GA-TGG--TAGG--------A------TAT-A-------T-A-CTA--AA-------G-A-A-G-----G--A------C-G---GA----CC-C-- 3091
GRV.ET.x.GRI_677_gri_1      -A-----T-----A------ACC-AA--TACC---GCCA-GT-G-----G-AA--GAAC--CCT-----A----GT-TGG---------G------GC----TATAAGA-C--TCA--G---AG-GC-G---A--CA-G-A---ACA--C---------TCTC-----A--- 3700
GSN.CM.99.CN166             -A--AC----T--T--A-GGAC--AA--G--G--C---TGTA----G---G-T----GG--------CCAGG--GAG--C--T--G--G---C--A-----T-A--------T--G-----CT-----C--T-TTC-G-G--T-GC----------G--GA---GCCGC-C 3252
GSN.CM.99.CN71              -A--AC-T--T--C--A-GGAC--AA--G-----C-G-TGTA--------G------GG--------CAGG-GGAG--C--T--G-----C-AA----GT-A-----------------------C--T-CTC-G-G--CAGC--C--C---G--GA---GCCAC-T 3234
LST.CD.88.SIVlhoest447      -AC----TGAG--A----G-ACT-AAG--C-T-------T-AA----G---GTA-----T---GAT-----TG--A-CTGG-----T-C----TTAGA---TAT-A-C-----CA-C----A---------AA-C-G--G----CT---------T--TA-T-----A--G 2798
LST.CD.88.SIVlhoest485      -AC----TGAG--A----G-ACT-AAG--C-T-------T-AA----G--GGTA-----T---GA------TG--A-CTGG-----T-C----TTAGA---TAT-A-C-----CA-C----A---------AA-C-G--G----CT-----------TA-T-----A--G 2798
LST.CD.88.SIVlhoest524      -AC-G--TGAT--A--C-G-AC--AAG-GC-C-----GT-GA---------GTA--GC-CT-GGA-C--C--G-GA-TTGG-----G------CTAGA---TAT-AGACC--T-AGACA-------------AA------G--G-CT--C-----T---AGT-----AC-T 2795
LST.KE.x.lho7               -AT-G--TGAT--T----GGAC--AAG--C-C--C---T-AA-A--G---GTA---C-CT-GGA----A-C---AACTGG-----G-----CTT-GA---TAT-GGC-G----A-----G-A---T------AAGA-G--G----CC-------------GA---G--AC-T 3882
MAC.US.x.251_1A11           -A-----T-----T--A---ACC-AA--TC-C----GG--AA----------AA-TGA-T--------G-A-G-T-AGTGG--T--GATG-----AGC---TAT-AG-----T-A-ATA-----C-GTC-GGA-CA-G-A---TGT-----CC-AGA-GGC--GCCA--- 3825
MND-1.GA.x.MNDGB1           --T-A--T--T-----C---AC--AA---C-G--C--G--AA-A------GGACT-AACA-------GAA-G-CA--A-G------------AGA--G---TAT-A-C----T-A------CT-GGCT---GA-CA-G-A--GTCT---------T--TAAT--G--A--- 3214
MND-2.CM.98.CM16            -------TT-C--A------AC--AAG-GC-C----G---AA-------GG-A---C-C---GAT---AA-G--GA-TGG--TAG----------AT-----TAT-A---------A-CTC--AG-GC-G--GGA---G---------A--C---C-A---GAG--GCCT--- 3525
MND-2.GA.x.M14              -A--C--TT-C--A------AC--AAG-GC-C--C---T-GA-C-----GG-A---C-T---GAT-------G--GA-TGG--TAGG--------AT-----TAT-A-------T-AGTTA--AG-GC-G--GGA------------T-----C-G---GG----CCTC-G 3452
MND-2.x.x.5440              -A--C--TT-C--A------AC--AAG-GC-C-------T-GA-A--G--GG-A-G-C-C---GATA--A--G-GGA-TGG--TAG----------A------TAT-A---------AGCTA--AG-GC-G--GGA----------------C--CC-G---GA----CCT--- 3085
MNE.US.x.MNE027             -A-----T-----T--A---ACC-AA--TC-C----GGT-AA----------AA-TGA-T--------G-A-G-T-AGTGG--T--GATG------GC----TAT-AG-----T-A-ATA-----C-GTC-GGA-CA-G-A---TGT-----CC-AGA-GGC--GCCA--- 3294
MON.CM.99.L1_99CML1         -A--C--TT----A--A---ACC-AAA-CC-G--C---AGTA-C-----GG-ACGGAGC--G-----TAC-G-GGTC--CT----TTT---TC-AGCT--G--A--G-----T--G---------C--G--GGA-TCCACC--G-GA-----C-----CAACG-GCCA--- 3230
MON.NG.x.NG1                ----C--------A--C--GAC--A---TC----C-R-TGTA-CC-R--GRTA-----C-----G--CA--G-GGAG---T----CTT---CC-AGC----T-A---------T--G--A-------GTC---AGTCGAC---G-GA-----C-----CCAGCTC-----G 1823
MUS-1.CM.01.CM1239          -AGCC--------AG-----ACC-AA--GC----C---TGTA-A--G---GTT---T-C--G--------A-G-T-A---------G-----TC-AT-----T-----------CAG--A-----T--GC-GG----C--G--G-CA--C-T---T---GA-C--CC-C-G 3204
MUS-1.CM.01.SIVmus_01CM1085 -AGCA--------T--C--GAC--A---------C---TG-A-C--G---GTA---T-T--G--T---AGG-C-AGT-G-----G------C--GC----T----------TCAG----C--TC--C-GA----C--G--G-CA--C---------GA-C-GCCA--- 3214
MUS-2.CM.01.CM1246          -AG-C--TT--------C---ACT-AA--T-----C-G-TGTA-C--G----TA--G-GG--------GC--G-GGAG--C-GT--GC----CC-AT---------A---------T--G--A--C--G--GC--GAG-CCGG---G-CC-------------GA----CC-C-- 3312
MUS-2.CM.01.CM2500          ----A---T----A--C--GACC-AA--TC-G--C---TGTA-C------GTA-----C--T------AC-G--GA---T--C--GC-G--TC--T----GT--------------G-----CA-G---C--GAG--CA-A----CA--C-----T---GA----CCA--- 3271
MUS-3.GA.09.09GabOI81       -AG-A--T-----A----G-AC--AA--TC----CC-TTGTA--------G-A--------------CAG-G-GGTG---T-----TT------A--G---A-------------G--A--AA-G---C-GGA---G--A--T-AC--C-----T---GA-------C-- 2780
MUS-3.GA.11.11GabPts02      -A-----T--T--A--C-GGAC--A----C-G---GCCTG-A----G---GTA--G-----------CA--G---A-T-------G-C------AGC----T----------T--G-----AA----GC--GA---G--G--CAG---C-----T---GG------AC-T 2801
OLC.CI.97.97CI12            -A--C--TT----A--A---ACG-AAGCT------------GA---AG---C-T----GGT-GCT---TAA-G-GGAGTGG---------G---A--GA----T-A-A---G-----------T----G---GA-CA-G-A------ATG---A-AGAGGAGG-GCCTC-C 3398
RCM.CM.00.SIVagi_00CM312    ----C--TG----A--C---ACT-AAG----G--C---T-AA----------TA---AG-T--GAT--TAAGG--GA-T-T---A----------AT----GTAT-A------T-A-CT-C-AT---------AA-T-G-----G--A-------------GA----CCA--G 2795
RCM.CM.02.SIVrcm_02CM8081   -A--C--TA----A--A--GAC--AAG------------T-AA--------GTA---AGG---GAT---AAGG--GA-T-C--CAG----------A------TAT-AG-------AGCTAT-G--G--------A-A-G---------C--C-----T---GA----CCTC-G 2813
RCM.GA.x.SIVRCMGAB1         -A------AGT---C-A-GGACT-A-G----G--C--GT-GA-A--G---G-A------T--GAT---AA-G--GA-A-G---A-----------AA-----TAT-A---------AGAT--------------AAGT-G-----G-----------TGA-AAG---CC---- 3003
RCM.NG.x.NG411              ----A--TT----A--A--GAC--A-G-G-----C-----AA-A--G---GTT--G---T-GGAT--CAAGG--GA-T-T---AG----------AT-----TAT-A---------A--CTG-C--G--G--GAA-A--------G--T------A-----GA----CCA--- 3018
SAB.SN.x.SAB1               -A------A--------A--GACT-AA--T--G----G---AA--------G-T-G-C-TT-G-----GA--G-C-A-TGG---------G-----A------T-A-A------T---C----CT---G-C--AAGCAG--G---CAA-----------TG-TCTTCCT--- 3922
SMM.CI.79.SIVsmCI2          --T-A--T--T--A--A--GAC--A---T-----C--G--GA-C-A---G-AA-TGA----C-----G-A-G-G-AGTGG------TTG-----AGC----T---------------ATC--C--GG-GC--GAGCAGG-A----CC-T---CAGGGA-GA-G--A-T--- 2946
SMM.LR.89.SIVsmLIB1         G------T-----A--A---ACC-A---TC-T--C---T-AA-------G-AA--GA-TT-------G-AGG-C-A-TGG--T--GATG-----AGC----TAT-A---G-GT-A-ATC-----T-GTC-GGAGCAGG-A--CTGC-----CC-AGA-GAT--GCCT--- 2905
SMM.SL.92.SIVsmSL92A        -A-----T-----A--A--GAC--A---TC-G-------T-AA-A-------AAGTGA-------------AGG---A-TGG-----GTT--------GC---GT-A-----G---AGATT--A--C-GTC--GA-CAGG-A--CTCT------C-GGAGGG-G-GCCA--- 2907
SMM.SL.92.SIVsmSL92B        ----A--T------------AC--AAA-CC-G--C---A-GA-C------G-AA-TGA-T--G---------A-G-T-AGTGG------TT------AGC----T-------G--T---ATT---------TC--GAGCAGG-A---AGA-----CAGAGA-GATG--CCGC-- 2886
SMM.SL.92.SL92B             ----A--T------------AC--AAA-CC-G--C---A-GA-C------G-AA-TGA-T--G---------A-G-T-AGTGG------TT------AGC----T-------G--T---ATT---------TC--GAGCAGG-A---AGA-----CAGAGA-GATG--CCGC-- 3257
SMM.US.04.G078              ---------------A---AC--A---TC-T----G---GA----------AA-TGA--T--------A-G-T-A-TGG-----GATG--T---GC----TAT-A------T-AGATT--C--C-GTC-GGA-CAGG-A--GTGC-----CC-AGA-GAT---CC-C-- 3086
SMM.US.05.D215              -A-----T-----A--A--GACT-AA--TC-G-------T-AA----G---AAGTGAGT-----------A-G-G-AGTGG--T--GATG-----AGCT--GTAT-A-----------ATT---T---GTC--GA-CA-G-A--GTGT------C-GGA-GAC--GC-A--- 3080
SMM.US.06.FTq               -A-----T---------A--GAC--AA--CC-C--C--G--AA---------AG-TG---T-G--------A-G-C-A-TGG------ATG------GC---GT-T-AG-----T--ATC-----T--TC--GA-CA-G-A--GTGT--C--CC-AGAGGAT--GCC---- 3108
SMM.US.86.CFU212            -AG-C--T-----A--C---ACC-AA--TC-T----------GA-C-------AA-TGA-T--G-----G-A-G-C-A-TGG--T---TTG-----AGC---GTAT-AG-------AGATC-----C-GTC-GGA-CA-G-A---TGT--C--CC-AGAGGG---GCCT--- 3072
SMM.US.x.H9                 -A-----T-----A--A--GACT-AA--TC-T--C-G-T-AA-C-------AA-TGA-TT---------AGG-T-AGTGG--T--GATG------GC----TAT-A--------AGATA----CC-GTC--GA-CC-G-G---TGT--C---C-GGAGGG---GCCAC-- 3299
SMM.US.x.PGM53              -A-----T-----A--A--GACT-AA--TC-T--C---T-AA-C-------AA-TGA-TT---------AGG-T-AGTGG--T--GATG-----AGC----TAT-A--------AGATA-----C-GTC--GA-CAGG-G---TGT--C--CC-AGAGGG---GCCA--- 3743
SMM.US.x.pE660.CG7G         -A-----T-----A--A--GACT-AA--TC-T--C-G---AA-C-------AA-TGA-TT---------AGG-T-AGTGG--T--GATG------GC----TAT-A--------AGATA-----C-GTC--GA-CA-G-A---TGT--C--CC-AGAGGG----CCAA-- 3815
STM.US.89.STM_37_16         -A-----T-----A------ACT-AA--TC-G--C---T-AA-C-----T-AG-TG-----G--T--G-AGG---AGTGG------ATG------GC----TAT---------T-A-ATT-----C-GTC-GGAGCAGG-A--GCAA--C--CCGAGA-GAC--GCCTC-- 3473
SUN.GA.98.L14               -AT-A--TGA----C-C-G-AC--AAG-------------AA-------G-AT---C--T--G----GAA-G--AAGA-G-----------TAG-GA---TACCAGAGT--TCAG---G-AT--C-G--GT-----TCA--CAGT--C-----A---GAT-----A--- 3884
SYK.KE.x.KE51               -A--A--T--T--T--C---TGT-AAGC---GA----GTG-A-A-----GG-A---------T--C--CA-TG--GA---G---------G-----AGC----A----T-----T---------A--C----T-GA-CA-ACA--GTCT------C-G---GA----CCA--- 3379
SYK.KE.x.SYK173_COMGNM      -------TT----A-----GAC--AAG-GC-C-----GTGTA-C---------T---CA-T------------C-CA----T-------C---T---GC----T-A-A---G--T---C-A--CT----G---GA-CAGGCA--GTCT------------TAA----CCT--- 3721
TAL.CM.00.266               -A--A--TT-T--T------AC--A-GCCC-C--C-GGTA---C--G---GTA----G---TCT-----A-G---AGT--T----G----------GC----T-A------------G--A--CT--GC----GA-ACT--C--TACA-----CC-GGA-GG-G-GCC-C-G 3493
TAL.CM.01.8023              -A--A--TT--------C---AC--AAGCCC----C---TG------G-TGTA---GCT-GTT-----AGG-TA--T-G-----------------AGC----T-----------------CA------GC---GGAGACC-----TTCA-----CA-GGA-GG-G--CCA--- 2991
TAN.UG.x.SIVagmTAN1         -AT-A--T-----A-----GAC--AAGCCC----C---T-A--AC-----G-TC---ATT-----CT---G-G-A-TGG--G--GC-G-----G----GT-CAG-C----TCA------CT--------GAGCA---G----CT--C----------AGT---TCT--G 3726
VER.DE.x.AGM3               --C-----------AC-C-G-ACT-AAA-C--G------T-AA-C-------AA--GAACT--CT----AC-G--A-CTGG-----G---------AGC---TAT-----------A------CT-----ACGGA-CAGG-A--GACC--C--CA-A---GG--G-CC-A-C 3249
VER.KE.x.9063               --C--------C---T-A-G-AC--AA--GA----C-GGT-AA-C--G----AG---AAT---TT---G-A-G-G-A-TGG--T-----G--G--AGC---GTAT-----G----A---A---T-G---AC-GA-CA-G-A---ACT-------CA---GG---GCCTC-T 3754
VER.KE.x.AGM155             -AT-A--T-----AT-G-G-AC--AAA-TA-C-----G--A------------AG---AATT--TT---C--GG--GA-TGG--CCC---G-----AGC---GTAC-A---------AG-----C------AC-GAGCA-G-A--TACT-------CA---GA----CC-C-T 3746
VER.KE.x.TYO1_patent        -AT-G--T-----TC---GGACC-A-A-TA----C--GT-AA----------AG---AAT--GTT---GC--G-GA-TTGG---CCT--G-----AGCT---TAT---------TGC---------T---AC-GA-CAGG-A---ACC-----CA-A---GG--T-CCTA-T 3751
WRC.CI.97.97CI14            --C-A--TGA---A--AC-GAC--AA---------C-G-T-GA-------G-AT---C--T-G--------A-G--GA-TGG-----------TTTAGA---GT-A-A-ATC--T-AGC----AT---------GA-CA-TCA----CC--C---A-AGG-AA-G-CTTAA-- 3821
WRC.CI.98.98CI04            -------TGAG--A---C--AC---A---------C---T-AA----------A----C-C--C--------A-G--A--TGG-----G------TTAGA---GT-A-A-AT---T-A-AGC------------GGAGCA-ATG--G-CT-------ACTGGGAA-G--TTAG-G 3836
WRC.GM.05.Pbt_05GM_X02      -A--A--TGAG--T--A---AC--AAG----G------T-AA----G----AT---C-T-------C--A-G--GA-TGG--CATG--------AGC------A-A-ATC--T----CT--------G---T--CA---A---AGT-------GAGG-GG-G-CTTAA-- 2750
```

```
H1B.FR.83.HXB2             AATATGGGGAAAGACTCCT...AAATTTAAACTGCCCATACAAAAGGAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCTGAGTGGGAGTTTGTTAATACCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCCATAGTAGGAG 3858
Pol                         _I__W__G__K__T__P_.__K__F__K__L__P__I__Q__K__E__T__W__E__T__W__W__T__E__Y__W__Q__A__T__W__I__P__E__W__E__F__V__N__T__P__P__L__V__K__L__W__Y__Q__L__E__K__E__P__I__V__G__
H1O.CM.91.MVP5180          T--------G--ATTA---..------C-GG-----AG-TACT-GA-----T--------T------G----A-------G-----------------A-----A-----C-GC--A--C--A--GA-C-------------------A-C------T--T-----G- 3887
H1N.CM.95.YBF30            ---T---------------..--------G-T-A--AG-----------GTG-----GG----------C--TC----------A--T-----------------A-----C--C--T------C-T--A-----------T---------A-C---G--A--CAGT--G- 3452
H1P.CM.06.U14788           G-----------ACTA--C..--G----G---A--AG--A-T-GAA-TGTG-----G--T------T----C--------------------A-----------A-------GC--A--C--TC-TA-T--GC-C-----T---C----A-----C------CC----A 3329
CPZ.CD.06.BF1167           -----------GAGTA---..GT---CC----T--AG-TAC-------GTG---C--CA-------T-T--C--------G-T--G---------A--A---------------C-----C--A---A-A--G---------A-TC-GTTA-C------T--GCC--A-- 3966
CPZ.TZ.00.TAN1             ---T-----T-TTGTG---..------TT-T-A--TG--TCC--A--G------AGCCAG--------T--T--C-------TT-----G-A-----------A---A---C-----A--AC--A-C-GGC--------A-TC-GTT-TCT--C-----CCC--A-- 3502
GOR.CM.07.SIVgorCP2135con  T--C--------ATTG---..--G--C-GG--A------TGTCGAAGT--C-----CG-T-TT-----T--C--------G--TT-T-----------------A-------G---A--C---C-CA-T--GC-C-----T--A--G--A------------CC----- 3394

MAC.US.x.239               G--C------C--GTC--A..------CC-CT-A--AG-TG-G-----TGT-------CAG-----------C--------G-TA--------A--G--A-----T---A-CTCA--A--A--GC----A-G----GTC-T-A-TC---T---G--C--T----AG---- 4157
Pol                         _V__I__W__G__Q__V_.__P__K__F__H__L__P__V__E__K__D__V__W__E__Q__W__W__T__D__Y__W__Q__V__T__W__I__P__E__W__D__F__I__S__T__P__P__L__V__R__L__V__F__N__L__V__K__D__P__I__E__
H2A.SN.86.ST_JSP4_27       C--T-----CGA-TA--A..------C-C--A--AG--G---GA--T--C------CAG------GATA-C--C-------TG--A-----C--A--C-----C--CA-ATC------G--AC-G--C-G----GTA-TTA-CC-G-T------T------C----C- 3685
H2B.CI.x.EHO               C--C------G---TA--A..-TG--CC-T--A--AG--G---GA--G--------TCAG-----------T--C-------TA--------C--A--A-----T-----CTCA-----A--A---A-A-GG---GCC--TA-CC-G-TC-----C---C---A----- 4211
H2G.CI.92.Abt96            ------------ATTA--C..TTC---C-T-----AG--G---G---C--C------CA-------------C----------TA--T-----C--A--T-----C-----GTC---A--G-ATC----A-GG---GTC---A-CC---TC---------TC---AGCAGA 3567
H2U.CI.07.07IC_TNP3        ---C-----G--AGTG--A..-TC--CC-C-----AG-GG-G-GA------------CAG-----------C----------TG-----------A--------T-----ATC---A--A--T--G--A-GG---GCA--TA-CC-T-TC-----G--TC---A-C--- 3672
H2U.FR.96.12034            ------------A-T---A..-TG--CC-T-----AG--G---GA--GGTC------CA----------T--C--------G-T-------G-C-TA--C-----T-----ATCC--A--A---C-G----G-C-GGCA--TA-TC---TA---------AT--C--...- 3719
ASC.UG.10.RT03             -TAC-----G-GA-TA--A..C-----C--TG-C--AG-TA-C--A--GCAG-----CG---------AT--AC-------G-TGT----------G--TATAA-ACCA--GC-C--T--A--T--GA-AC--A-G---------A-----TAC-G--T--A--TTC--A-- 3264
ASC.UG.10.RT08             -TAT-------GA-TA--A..C--------TG-C--AG-TA-T--A---CAG-----CG-G-------GT--AC--------TGT----------G--TATAA-ACCA--GC-C-----A--T--GA-AC--A-G---------A-----TAC----T--G--CCC----- 3288
ASC.UG.10.RT11             -TAT-----G-GA-TA--A..C--------TG-C--AG-TA-C--A--GCAG-----CG-G-------GT--AC-------G-TGT----------G--CATAA-ACCA--GC-C-----A--T--GA-AC--A-G---------A-----TAC-G-----A--CCC--A-- 3273
COL.CM.x.CGU1              C-CC---------GTG--C..----A-G---G-C--AGC-A-G-GA--GGAT------CAG-------GT--C----------AGTG-T------C--A---ATA--------CTCAT--T--TATG-----A-------GTA-GGA-C--G-TTTT---G--TC-GAAG-AG- 3627
COL.UG.10.BWC01            C--C---------GTC---..-G-A-G--GG----AGC-A-G-G---GGAC------CAG-------GT--T--C-----GAGT--A-----A--A--AATA--A------TCC-----ATATG----C-------GTA-GGA-TC--AT-TT------AG--AG-A-T- 3426
COL.UG.10.BWC07            ---C---------GT----..CTTA-G--GA-A--AGC-A-G-GA--GGAC------CAG-------GT--T---------AGT-----------A---ATA--A-----CTCC--AT-A--AG-G--A---C-GGT--GGA-T-----T-TTGAGC--TC---A-AA-- 3654
DEB.CM.04.SIVdeb04CMPF3061 --C------G-G-TTA--C..-C-----GGT-A--TG-C--G---C-GGAC-----TGG--------CAT---C----------------T-----C--A--A------GCG--AC-C--T-------C----C-G----------TTCC--G-TTTC-------T---AA--A-- 3845
DEB.CM.99.CM40             TT-T-------GA-TA--A..C-G--C-GG--C--AG-GGT---A---GA------C-AT------CATA-TC-C----------G-------C--A--C------GCAA-AC---------A-ATC-G--C--GC-C-----TG--C-G-T-TC-----------CC--AT- 3664
DEB.CM.99.CM5              CTAT---------AGT---C..-------GGT----AG--GT--------GA------T-AC------CATA-CC-C-----G-----T-----C-----C------GCG--GC--------A-AAC----C-GGC-------TG--C-T-TATC---G--A---CC--AT- 3658
DEN.CD.x.CD1_CM0580407     ---C-----C-GACT-----..-CT--C-G---C--AG--AGC----T-GAT-----TG---------T----A----------TT-AT-----A-----C-----AA-G--ATCA--T--A---C-TC-A-G-AAT-TT--TA-T--G-TATC---G--------ACAA-- 3748
DRL.DE.11.D3               G--C-------GATTG--C..-CT--CT-CT-A--AG--G---G----GTT------CA-------CAG-----C-------G-TA--A-----C--G--T-----A--C--ATCA--A--A-ATC-CA-C--G---------TA-T--GCT--GT--C--A---CC----- 3332
DRL.x.x.FAO                ---C--------G-TTG--C..-TT--CT-CT-A--AG--G---G----GTT------CA-------CAG-----C-------G-TA--A-----C-----T-----A---A-ATCA--A--C-ATC-CA-C--G---------TA-C--GCT--GT--C--A---CC---G- 3425
GRV.ET.x.GRI_677_gri_1     C--T-------GACT---A..CG-G-AC-----C--AG--G-C---A----------T-TG------CAG--C-----------TAT--------A--A--A---------------GC--A--A-T-C----A---C-G-----TTCC----TA---------A--CAA----- 4031
GSN.CM.99.CN166            T-CC---------GTA--AGAC-------GGA-C--AG--GTC-------CA------TG---------TGA-T--------G--A--A------A--CACCATA--AGCAA-AC-C--A--A-A-C--C---G-CA-------ACCC---TCTCG------C---A--AT- 3586
GSN.CM.99.CN71             T-CT--------AGTA---GAC-----C-GGA-C--AG--GTC------GCA------T--G-------TGA-C--------G--A--A--------CACCATA--AGCA--AC-C--A--A-AT--GT---G-CAG------ACCC---TCTCG------C---A-AAT- 3568
LST.CD.88.SIVlhoest447     C--T-----TC-TGTG--C..--GA--C-GG-A--GG--AC--GA---GTG------TCA--------GT--TC--------G-T----------A-----A-T---A---A-CTCA-----A-AG----AAC-GGA------TAGC-----AGC-------T-----A---G- 3129
LST.CD.88.SIVlhoest485     C--T-----TC-TGTG---..--GA-CC-GG-A--AG--AC--GA---GTG------TCA--------GT--TC--------G-T----------A-----A-T-------A-CTCA-----A-AG----AAC-GGA------TAGC-----AGC------------A---G- 3129
LST.CD.88.SIVlhoest524     C--T-----TTTTGTA--A..--GA-AC--G-A--TG-CACT-GA---------------AT--------T-A-C-T--------TA--T-----A--A--C-TA--A--CA-CTCC--A----AA----AAC-GGA------TAGC--G--AGC------------A---TA 3126
LST.KE.x.lho7              C--C-----CTTTGTG--A..---A--C--G-C--AG-GACT-GA---GTC------GCAC-------GC--CC-C------G-TT--A-----------A--C-TA--A--CA--TC-------G-AA----AAC--GAG------ATT-GG---GC-----------A----G- 4213
MAC.US.x.251_1A11          G--C------C--GTC--A..------CC-CT-A--AG-TG-G-----TGT-------CAG-----------C--------G-TA--------A--A--A--A-T---A-CTCA-A--A--A--AC------G----GTC-T-A-TC---T---G--C--T----AG---- 4156
MND-1.GA.x.MNDGB1          ---C-----TTTTGTA--A..---A-GC-GA-A----CTAC--G---G-T-------GAT------CAT----------GTGT--A-----A--A--AGTA--A---A-C-GC--A---ATG----AA-GGGA------TAGC--GTCCCC------TC---AG--G- 3545
MND-2.CM.98.CM16           G--C--------ACTC---..--G---T-T-----A---G---GA---GTC-----CCAG------C----A--------------A-------C--A-----A----ATCC--T--A-A-C-TA-CGGG--G-----TA-CC-GTT--G------TG-GCC----- 3859
MND-2.GA.x.M14             G--T------G-ATTG--A..------CT-T--C---C-GG-G-G----GT------TCAG------CAT--T--------G--A--A-----C--A--A-----A--CA-CTCA--A--A--AC--A-A-GG--G-----TA-CC-GTTA--------A---CC---G- 3786
MND-2.x.x.5440             ---C---------CTA--G..------CT-T--C--GT--G-G-G---GGT------CCA-------CAT--C----------TA--A-----C--A--------C--CA--TCA--A--A--A---A-A-GG---------A-CC-GCTA-----C--A---CC---G- 3419
MNE.US.x.MNE027            G--C------C--GTC--A..------CC-CT-A--AG-CG-G-----TGT-------CAG-----------C--------G-TA--------A--A--A-----T---A-CTCA--G--G--AC----A-G----GTC-T-A-TC---T---G--C--T----AG---- 3625
MON.CM.99.L1_99CML1        C-C-------CGACT---C..-TT---CGC--A--AG-GGTG--A--GCA------CG---------G----TA---------AT--AT---G-A--A--CATA---GCG--GT--T----A-A-C-GC-C-G-CAG-----TACTC-G-TAC----C--------A---G- 3561
MON.NG.x.NG1               C--T---------CTA--A..-T-----G-T-A--AG--A-C--A---CAG-----C---------G-C--CA----------T---T-----A--A---ATA---CCA--AC-CT-T--A-A-C-GC-T-G-CA-------ACC----T-C--R-T-----T-A-CAC- 2154
MUS-1.CM.01.CM1239         C--T-----G---C-C--CGAA--GG-C-GGA-T--AG-GATC-----GCAG-----TCA--------------AC-------G-TAT-A-----A-----CATA---GCA--AC-----A-C-A-C-GC--CG-CAG----T-ACA--G-T-CC---G--AT-GTC--A-- 3538
MUS-1.CM.01.SIVmus_01CM1085 C--T-----G---C-C---GAC--GG-C-G-A-C--AG-GATC--A---CAG-----CCAG-------T--CAC-C------G-T-T-G-----A--ACCATA--AGCA--GC-----A-C-A-C-GT-A---CAG----T-ACA--G-TACC------AT--AA--A-- 3548
MUS-2.CM.01.CM1246         CTGT-----G-----C--AGAC----GC-GGA-A--GG--GTC------CAG-----C-AT------G----C-CA------A-A--T--------C--CATA--AGCA--AC-C-----ATA-C-TC---G-CAG----TTACC--G-TCCC---G--------A-AC-- 3646
MUS-2.CM.01.CM2500         CTGT-----------C---GAC----G--G-A-A--GG-GGTC-----GCAG-----CCAC------G----A-CA-----GA-T--A---G----A---ATA---GCG--GC-C--A-----A-C-GT---G-CA-----T-ACTC-G-T-CC---G--TC---AG-CG- 3605
MUS-3.GA.09.09GabOI81      -TAT--------AGTA--AGAC-------GGA-T--AG-GATT--A--GCA---------A-------G----CC----------TA--A-----C-----TGT---AGCA--AC-CT-G--A-ATC-GT---G-CA--AC---ACC--G-TAC--------T----A-AA-- 3114
MUS-3.GA.11.11GabPts02     -TAT--------A-TA--GGATGT---C-G-A-A--GG---AT-------CA------T-AT------G-T--A--------G-TA--A---G-A--AA-TGT----GCAA-AC----T----A---GC-T-GGCA--AT--TACCC---TAC-G--G-----T-A----- 3135
OLC.CI.97.97CI12           G--T---------TTA--G..-TT---C-GG-A--T-CTAC--GAA--GAT------G-----GCATTT--TG----------TACAT-----A--A--CATA---------ATC--ATT--GAAAATCAA-GGAC--TC-GGG------T-CC--C------G...-AG- 3723
RCM.CM.00.SIVagi_00CM312   T--T--------ATTG---..------G--T----TG--G---GA--GGT--------CA-------G-T--T-----------T-T-A---------A--A-----A-----G-GC--G--G--T---A-T-GGC-G-----TA-CC---TA--G--T------CC----- 3129
RCM.CM.02.SIVrcm_02CM8081  C--T--------ATTG--C..------G-G-----A---G-G-GA--GGT--------CA-------G-T--T--------G-TAT-A-------------------A--C--G-GC--G--C--A---A-T-GG--------TA-TC-G-TA-----T------CC---G- 3147
RCM.GA.x.SIVRCMGAB1        T--T-------G-CTG--A..-C---C-G------A---G-G-GA--GGT------CTA----C--T-CC-A--C-----G-TT--------A--A--T-----A-------GC--A--A--TC-TA-T-G-C-CG-----A-CC---TT--G--T------CC----- 3337
RCM.NG.x.NG411             ---T---------CT---A..------G-G-----A---G---GA---GT-------CAG------T-G--T--C-----G-TTT--------------A-----A-------G---T--A--AC--A-A-GGC--------A-TC---TA--G--T--T---CC----- 3352
SAB.SN.x.SAB1              C--C------C---TA---..-TTA-GG-------AG--G---GA--GCTC------CA-------T-T--C--------G-T---------C--G--A-----AA-G--C-G---T----AA--GA-T-G---G------A-A----TA-----T-----CCC----- 4256
SMM.CI.79.SIVsmCI2         T--C-----------TA--G..TTC--CC-T-----AG--G---GA---GT------GCA-------------C-----------TA--------C--A--T-----T-----ATCC--A--A--A--G--C-G----GTA-T-A-TC-G-TC-----T------CC---G- 3277
SMM.LR.89.SIVsmLIB1        ---C-AA--GC---T---A..--G--CC-CT-A--AG--G---GA----TC-A----CA-----------T--------G-TG--A-----A--A--A-A--T---A-ATCC--A--C--T--G--T-GG---GTC-T-A-CCCG-TA--G--G---C---AG---- 3236
SMM.SL.92.SIVsmSL92A       ---T---------GTG--A..------CC-CT----AG--G---GA--GCT------GCA-------------C-----------TA--T--G-------T-----C-----ATCC-----C--AC--A-A-GGA-GGCT-TTA-T--G-TA--G--T------CC----- 3238
SMM.SL.92.SIVsmSL92B       ---------------A--A..TTT---C-C--T--AG--G---G---GGTT-----TCA-------------T-----------T--T--------A-------C--CA-CTCC--A--G--G--A-C-GG---GT--TTA-T----TA-----T--A----A-AAG- 3217
SMM.SL.92.SL92B            G--------------A--A..TTT---C-C--T--AG--G---G---GGTT-----TCA-------------T-----------T--T--------A-------C--CA-CTCC--A--A--G---A-C-GG---GT--TTA-T----TA-----T--A----A-AAG- 3588
SMM.US.04.G078             ---T------C-AGTG---..------C-TA----AG--G-G-GA-----T------CA-------------T--C-------TG--A-----A--G--A-----C--C--GTCG--------TC----T-G----GT--T-A-T--G-TA--G--T--T--T-A----- 3417
SMM.US.05.D215             ---T------C-AGTG--A..------CC-T-----GG--G-G-GA---GTT------CAG-----------C--C------G--A--T-----A--A---------C-----GTCC--A--C--AC----A-G----GTC-T-A-CC---TT-----C--T----A---G- 3411
SMM.US.06.FTq              ---T---------AGTG---..------C-TA----AG--G---GA----TT-----GCAG-----------C--------G-TA--A-----A--A--A-----T--CA-ATCA--A--C--T-----T-GGC--GT--TTA-TC-G-TA-----C--T----A---G- 3439
SMM.US.86.CFU212           G--C------CG--TC--A..------CC-T--A--AG--G-G-GA----TC-----GCA------------C--C------G-TG---------A--A--C-----C-----ATCA--A--C--T--G--C-G----GTT-T-A-CC---TA--G--T--TC-T-A----- 3403
SMM.US.x.H9                ---T------C--GTG--A..-G----C-TT----AG--G-G-RA----TT------CA-------------T-----------TA--------A--A--A-----C-----GTCA--A-----------C-G-C--GTT-T-A-CC---TA-----G--T---CAG---- 3630
SMM.US.x.PGM53             ---T------C--GTG--A..------C-TT----AG--G-G-G-----TT------CA-------------T-----------TA--------A--A--A-----C-----GTCA--A-----------C-G---GTC-T-A-CC---TA--------T---CAG---- 4074
SMM.US.x.pE660.CG7G        ---T------G--GTG--A..------CC-TT----AG--G---GA----TT------CA-------------T-----------TG--------A--A--T-----C-----GTCA--A-----------C-G---GTC-T-A-CC---TA-----G--T---CA---G- 4146
STM.US.89.STM_37_16        G--C------C---TC--A..------CC-CT----AG--G-G-GA--GGTC------CA-------G-----T--------G-TA--A-----A--A--A-----T-----ATCC--A-----T-----A-G----GTC-TTA-T----TA---------T-G-A----A 3804
SUN.GA.98.L14              T---------TTTGTG--A..--GA-AC-GG-A----------GA--G-T----AGTCAG------G-----C--C-----GTGT--A-----C--------TA--A-----C-GC--T---AAG--G-A-C--GAG-----TACTC---CC-C-------AG-GCC----. 4214
SYK.KE.x.KE51              G--T---------A-TC---..G----CC-C--C--AG--A---G----GA---------A-------------TC-C---------A--A---T-G--A--AGTAA-A-G---GC-C-----CATA-----T-G--GG-AT---A-T--G-TTTC---G--AG--...-A-- 3713
SYK.KE.x.SYK173_COMGNM     GTAT---------GTA--A..-C---CC-T--C--AG--A---GA--GGA-------A-------------C---------------A---G-A--A--AGTCA-A--CA-A-G------G--A--G--T-GG-GG-AC--TA-C--G-TACC-----------CC--A-- 4058
TAL.CM.00.266              CT-T-----GC--GTA--A..GT---C-GGA-C--AG-GGTC-----GGA-------CAG-----------AC-C-------TG-T-A-----A--CACCATC-T-CCCA--C-C--A--A--A-----T-GGC-------TA-CC---T-C-G--G---------A----- 3824
TAL.CM.01.8023             CT-T------C--GTG--C..-T-----GGG-A--AG-GGTC-----GGAG-----GCA-------T---------------TG-T-A---C-G--CACCATT-TCCCCA-AC-----------------C-GGC-------TA-CC---TATC---G--------A----- 3322
TAN.UG.x.SIVagmTAN1        CT-T--------ATTA--C..-TCA-GG--G-A--AG--G---GA---GT-------CA--------GT--C-----------T---A---C-A--A---ATA--A--CA-CTCC--A--G-------A-C-----------TA-C--G-TA-----T------CC-AA-- 4057
VER.DE.x.AGM3              T---------G--TTG--A..GTCC--G----C--A---G-G-G-G---GT-------CA-------G-T--T--C-----G-TA-GT--------A--C-----A-----C-G-------A---C-----A-----------TACCC-GACA-------------CC-AAG- 3580
VER.KE.x.9063              C--T------G-ACTA--A..-TT---G----C--AG--G---G-----GT-------CA-------G----C--C-------TG-G--------------A-----A-----G-G-GTA--A--TC-----------G------ACTC-GACC--G-----T---CC----- 4085
VER.KE.x.AGM155            ---T------C-ATTA--C..-CT---G----C--AG-GG-G-G---C---------CA-------G----C----------T--GT-----A--C--A-----C-----C-G-GTT--G--------A-CT--G-----TACAC-GACT--G---------CCCG---- 4077
VER.KE.x.TYO1_patent       T--T-----G-TATTA--A..GTTC-AG----C--G---G---GA--GGT-------CA-------G-G--T--C-----G-TA-G---------C--A-----T-----C-GC-----A--T--GC-C---C--------ACA---ACA-------------CCCAAG- 4082
WRC.CI.97.97CI14           ---C-------G-GTA--C..------C-G-A-A------G-------T-T------TT-T------T----C-----------TGT-A-----C--A--CACA--ACA---GTC---G---TT-C-T-CT-GGGA------TA------T-CT------TC--CC--A-- 4155
WRC.CI.98.98CI04           ---C---------GTA--C..--------GA-A------G-------T-TC------CT---------T----TC-C-----G-T-T-A---------A--TACA--ACA---ATCA------TA---G-CA-GGGA--------A-AA---TACTG--G--T--------AG- 4170
WRC.GM.05.Pbt_05GM_X02     ---T-----G--ACT---A..-------G-A-C--AG--G-------T-T------C-TG------T----A--C-----G-----A-----C--A---GTA--ACA---GTC---T---TTTC--AAA-GGGA-------ACA----TATC------T----A---G- 3084
```

p51 end and p66 RT continue_Pol p15 RNAse H start

```
H1B.FR.83.HXB2          CAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCTAACTGACACAACAAATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAGAAGTA 4028
Pol                     A  E  T  F  Y  V  D  G  A  A  N  R  E  T  K  L  G  K  A  G  Y  V  T  N  R  G  R  Q  K  V  V  T  L  T  D  T  T  N  Q  K  T  E  L  Q  A  I  Y  L  A  L  Q  D  S  G  L  E  V
H1O.CM.91.MVP5180       ---------T----------A---------T---A-T--A---C------G--G----------AG-ACA-----A---G---CA-AA-A-AGT--GAA--G-----C-----A---G----A---ATG--TG-A-TAA-----C-----------CAAGGAGC----- 4057
H1N.CM.95.YBF30         ---------T----------A---------T-----A--A-----G----------T-T----G--AG-----A---G-----G---T-TA-TG-A-----C--C-----A---G----A---ATG----T--CCT-A-G--C--A--A--G--A---CGG--T--- 3622
H1P.CM.06.U14788        ----------T----------A-----A-----C--G----Y-----T--------A--AG----G--T--G--G--GA-AA---AA---GAA--R----T-----A--RG-A--A---G----G-G-TGT----C--A--A--A---GA-C-G-CT 3499
CPZ.CD.06.BF1167        ----G----AT----------A-----A---------CT---------G--------C-----G--AG---------A-AGC-GG--AA-AGAAT----A--A--T--T------C--G-A--A-----G--T--ACTG--G--A--AA-A---G-AAA-GA-A-G--C 4136
CPZ.TZ.00.TAN1          ----------T----------A-----A--C-G----AAG----G-----------A--A--C-----C--T-C-GGTCAAAGGA-T--GAGA----C--T-----AC-AG-A--A---TGG---G-AG-------C--AA-A--C--A---GC-C-G--- 3672
GOR.CM.07.SIVgorCP2135con ----------T-----C-----A---------T--A--A-GC---CA---------------C--A--AG---G--T-----G---A-AAGG-AA---GAG-G------T--------AG-A--AC--G-G---G-CCTCT-G--C---G-A--A--A---GAA-A-T--T 3564
```

Pol RT end_Pol p15 RNAse H start

```
MAC.US.x.239            A--------AT---AC------AT--TG---T-AAC--T-A---GA---G-----------A-C--AG----G-C-A-G-C-----AAAAGTGT--GAAC-G--T--T-----AC-AG-A--A--GG-----T--CTCA-G--A---ACA--C--A--GCC-A-G-C- 4327
Pol                     G  E  E  T  Y  Y  T  D  G  S  C  N  K  Q  S  K  E  G  K  A  G  Y  I  T  D  R  G  K  D  K  V  K  V  L  E  Q  T  T  N  Q  Q  A  E  L  E  A  F  L  M  A  L  T  D  S  G  P  K
H2A.SN.86.ST_JSP4_27    -------------CAC------AT-CTGC--T-A-C-AT-A-G-GA--------------CA-A--AG---------G-C--G-GAGGCTAT--GAGC-A--C--C------C-AG-A--A---G----CT--GCGA-G--AG-AACA--C--A--TCC-A-G-CC 3855
H2B.CI.x.EHO            T------T-A---CAC------AT-CTG-----AA-CCT-A---GA---G-------------C--AG-C--G----AGG-T------AAAC-AT--GAAC-A-----------C-AG-A---C-TG-----T--GCA-----AC-A-----C--A---CC-C-G--C 4381
H2G.CI.92.Abt96         ----G----AT---AC---C---T-CTGC--T-AAACCT-C--GGAR----------C--C---C--AG-C-------A-G-C--G--AAARGT-T--GAGC-G-------C----C--G-R--AC-TG-G---T--GCAA-G--C---------C--A--TCC-A----C 3737
H2U.CI.07.07IC_TNP3     TG-----T-A---CAC------AT-CTG------AA-CT----GGAG--G--G----------C--AG-C-------AGG-T--G--CAAGGTGT-GGAGC-A---------C----C-AG-A---C-TG-G---T--CTGA-G--C---------C--A--T-C-C---CT 3842
H2U.FR.96.12034         A---G--A-AT---AC------CT-TTG--------CTT-A---CAG--G---------C----G--AG---------A-G-C-G---AAAAC-TT--GAAC-G--C--C--C--AC-AG-A--A---G----CT--GC-A-G--C--AG-A-----A----C--A-----C 3889
ASC.UG.10.RT03          -------T-GG---A-T-----A-----A--T-----AT---------G-------------G--AG-------A-G-G---A--AG-AT--GAAC-T--C--C-------A--AG-A-----------T-ACT-------C-A--A---GG----AC-A--CAG 3434
ASC.UG.10.RT08          -------T-GG---A-T-----A-----A--T-----AT---------G-------------G--AG-C-----G-A-G-------A--AG-AT--GAAC----C--C-------AG-A-A---------T-ACT-------C-A--A---GG----AC-A--CAG 3458
ASC.UG.10.RT11          -------T-GG---A-T-----A-----A--T-----AT---------G-------------G--AG-------A-G-G--G--G--AG-AT--GAAC-----C--C------A---G-A-----------T-GCTC------C-A--A---GG----AC-A--CAG 3443
COL.CM.x.CGU1           --CC---T-AT-GGAC------A-G-TGCTC----A-ATTAGG-GCG----G-------GGA-A-A-TCA-A----GA-G--G----A----AT--TCA---AGGGT--------C-AG-A---C-GAC--G-G---TG------C-AA-AC--GG-CC-AA-AG-A-G 3797
COL.UG.10.BWC01         --CCC--T-AT-GGAC------TGGTACTCT--AA--GA---T-GG----------G----GGA-A-A-TCATCT---GA-G-GG-G--AA----T----AA--AGG-T-T------C--G----A---AGG---G---TAA-G--CC-AG--C--GGACCTAAGAG-A-G 3596
COL.UG.10.BWC07         A--TC----A--GGAC------A-GGTACTCA--AA-AGAAGG-A----G---------GG--A-ACTCA-A----GA-G--G--A-A-AAGTT--CCA---AGG-T-T------C-AG-A--A---A---G---CCTGA-G--AC-CA-AC--GGACC-AA-AG-A-G 3824
DEB.CM.04.SIVdeb04CMPF3061 ----T--A-AT----G-----A-----------A--AT-C--GGA---G-------G--CA--T-AG-G-------A-AGT-G-A--AGGGAAT--GAAA----C--T--C--AC-AG-A--A---AC----G--A-GA-G--A---G----CAGT---GA-A-T--- 4015
DEB.CM.99.CM40          ----C----A-----------------------T-C-------G-----T--C------C--AG-GT-G----A-----GT--AAAATG-T--GAGA-T--C--T--------AG-A--A---G----C--A-TGT-G--C---G-A--AGGCCC--C-A--A-G 3834
DEB.CM.99.CM5           -T--C----AT----------A-------T----AT-A---A-G----------C-CA-AT-AG-AT-G----A---GGG---AAAA-GT--GGAAA-T--T--C-----A--G-A--A---G--C--CCTA--G--C---GGGACCC-CCA-GA-G 3828
DEN.CD.x.CD1_CMO580407  -GCCC----AT---AC------A-----A--T--AC--------CA-----G-----------GG-C-CC-AT--C-AG-----G--A--A---T-GGAA--A-----C-----A---G-A--A---AT--G-G-C-TAA-G--C--A--A---AGTCCCAACA-G--C 3918
DRL.DE.11.D3            A----GTT-AT-----G-----A--C--------A-TT-G---GA---------T--C-------CTCC--G-G-A-G-G--G-GA-AG--T--GAA--A--C--C---------G----A-----G-C--CCTGT----C--AA----C--C--TCC-A----C 3502
DRL.x.x.FAO             A----GTT-AT-----G-----A----------A-TT-G---GA---------T--C-------CTCC--G---A-G------AA-AG--A--GAA--A--C--C---------G----A-----G-C---CTGT----C---A----C--T--TCC-A----C 3595
GRV.ET.x.GRI_677_gri_1  A---TGTT-AT-----G----------ATC--AA-T---C--------T--G---------C-GT-AG-G-------A-AGT-G-A--AGGGAAT--GAAA----C--T--C--AC-AG-A--A---AC----G--A-GA-G--A---G----CAGT---GA-A-T--- 4201
GSN.CM.99.CN166         ---TT--T-AT----------A-----A--T---ACCT-C---C----------T--T-----G--C----C-----A---T-GG-CCA-AGAAT--GAA--A-T---------C-AG-A--AC-C--T---G--CT-T-G--A--AA-A--AGG-CC-CC-A--A-G 3756
GSN.CM.99.CN71          ---TC----AT----------C-----A--T---ACCT-----------T--C-----G--C----C-----A-T-T-G--CCA-AGAA---GAA-GA--T---------AC-AG-A--AC-C--G---G--CT-T-G--A--AA-A--AGGACC-CC-AG-A-G 3738
LST.CD.88.SIVlhoest447  T---T----AT----------A--T--AG-A-AA-TAGGG---AC------G---T-----C--AC-ATC-----AGG-G------AAAGAAT----A-----C--C-----AC-AG-A--A---G----TG-C-TAT-G--A--AAGA--CAGTAAGAGCAGG--C 3299
LST.CD.88.SIVlhoest485  T---T----AT----------A--T--AG-A-AA-TAGGG---AC----------T-----C--AC-ATC-----AGG-G------AAAGAAT----A--T--C--C-----AC-AG-A--A---G-G--TG-C-TAT-G--A--AA-A--CAGTAA-AGCAGG--C 3299
LST.CD.88.SIVlhoest524  ----T----A-----G-----A--T--AG-A--A-TAGG---AC----------G---A-A--C-ATC-----A-G-G--A-AAAAGAAT--CAA--T--T--T------C--G-A--A---G-G---G-A-TGT------A----AGCAA--C-AC---T 3296
LST.KE.x.lho7           T---C----A-----------A-----AG-A-A--TAGGA---AC----------------A-A--AC-ATC--G-A-G-G-----AAAGGAGT----------C--------C-AG-A--AC--G-G---G--CTAA-G--A--A--A---AGCAATAGTA----- 4383
MAC.US.x.251_1A11       A--------AT---AC------AT--TG---T-AAC--T-A--GA---G-------------A-C--AG----G-C-A-G-C-----AAAAGTGT--GAAC-G--T--T-----AC--G-A--A---GG-----T-CCTCA-G--A---AACA--C--A--GCC-A-G-C- 4326
MND-1.GA.x.MNDGB1       T------A-AT-----T----A----------C-G----A-G----------A----AG------TTTT----GG--A-AAGAATATCTAA-T--C--C-----C----A--A-----T--G-AA-A------C-AG-A---AGT---AGTT-T--T 3715
MND-2.CM.98.CM16        A---TGTG-AT----G-----A----------AA-CT-C---GA---G-------------A-AGC----AAT-A-TCC-G---AA-AG-----GAGA-T--C----C--A--AG-A--A---GG-G-----AA-AA-G--A--A--A---C--A---CC-A----- 4029
MND-2.GA.x.M14          AG--TGT--AT-----G----------A-T--AA-CT---GGA---G-------------CTA---AGC---G---A-AGT-----AA-AG---T-GAAA------T--------AG-A--A---G-G-----ACTGT----C---A-A-----A--TCCTAGG-C- 3956
MND-2.x.x.5440          A---TGTA-AT----------------A-----AA-CT-C--GGA---C--G-----C--TA---CTC------T-A-AGC--G--A--AG--T-GGAA--A-G--C-----A--G-A--A-GAC---T--AA-GA-G--------A-----A---CCCAGG--G 3589
MNE.US.x.MNE027         A--------AT-----------T--TG---T-AAC--T-A---GA---G-----------A-C--AG----G-C-A-G-C-----GAGAGT-T-GAAC-G--T--T-----AC-AG-A--A--GG-----T--CTCA-G--A---ACA--C--A--GCC-A-G-C- 3795
MON.CM.99.L1_99CML1     -TCC---A-A---C--G--------C--C--T---ACAT-A--GGA---G-------T-----A--C--CTTT--GGC-AG---G-CC-----AT--GAAC-G--C--C--C--A---G-A---C--G-G--CG-G-TA-----CC-CAGA---GGACC-CCTC-GA-G 3731
MON.NG.x.NG1            -TCCC----A---------------G-------A-CAT-A---GA---G--G----------C--C-G--C----CCA-AG---G--AA-TC-AT--GAGC-G-T--C--C------G----AC--G-----G-GCT-T----A-----AA--GG-CCCGCTAGTA-G 2324
MUS-1.CM.01.CM1239      -TCCC--A-AT----------A-----AC-T-AA-TAT---G-----G--G----------C-C-----A-G-----G--G-T-G-T--GAAA-T--T--C-----A---G-A--AC-G---TG-C-TGT-G--C---A-A--AGGCCCTCCCAGCA-G 3708
MUS-1.CM.01.SIVmus_01CM1085 -TCCC--A-AT-----G-----A-----A--T-----TAT-C-----G----------C--C--A--C-C--G-A-G-----G--AA-T-G---A-GAAAGT--T--C-----A---G-A--A--GG-G---G-C-TAT-G--C---AA-A--AGG-CCTCCTAGGA-G 3718
MUS-2.CM.01.CM1246      --CCC----A-----G-----A-----A--T---A-CT----C-----G----T--------A--CG----G--CAGG-G-G--C-A-A-ATT-GGAAA----C--C----C-AG-A--A--C-C-----C---CTCT-G---C-AAT----GG-CC-CC----A-G 3816
MUS-2.CM.01.CM2500      --CCC----AT----T-------C-----T--AA-CT-C------T-GG----G---C--AG---------CAGG---G-CCA-AGA----GAAA-T-----C--C--A---G-A--A--G--G-T--A-TAA-G---CC-ACA---GGACC-CC-A--A-G 3775
MUS-3.GA.09.09GabOI81   -CCCC--A-AT-----T-----A--G--------AACCT-C---GC-----GG---------C--C----C-----G--G----CAA-AGATT-GGAAA----C--T-----C-AG-A--A-----G-----A-TG-----C--A-----GGCCC-CC-A--A-G 3284
MUS-3.GA.11.11GabPts02  --CCT---AT---AC----------A-----AA-CT-C---------------G-----G--C--C-CG--CCA-AGG----G--TGAGT--CAA-GG--T--C--C------G-A--A--GG-----G-GCTAT-G--AC-TA-A--CAGCCC--A-CA-GC-G 3305
OLC.CI.97.97CI12        T---------------G----A---TA--T---A-A-GC--GGA------G---------C-G-TCAG-G---AG--T---A-A--ACA-A-CAG--G-GT-GT--------AG-A--AA--AG---T---ATAGA---CC-TAG---C--AAA-AGT--G--C 3893
RCM.CM.00.SIVagi_00CM312 AG--GGTG-AT----G----------A-----AA-T-G---------------CA-A--AG---------A-G------AAAAGAAT--GAA--T--C--C--------G-A--A---G-G--TG-ACT-T-------A--A-----A---AA-AG---- 3299
RCM.CM.02.SIVrcm_02CM8081 A----GT--AT-------C--------------AA-T-GC---CAG---------CA----AG-C--------G----G--AAAAGAAT--GAA--T--C--T-----A---G-A--A--GG-G---G-C-TAT------A--AA-A-----A--GAA-A-G--- 3317
RCM.GA.x.SIVRCMGAB1     AG----GT--A-----G---------------AA-T-G---A------G-G----------A----------A-G----G--AAAAGAAT--GAA--A--T--T-----A--AG-A--A---G-----G-A-TA--G--A--AA-A-----A--GCCTA----- 3507
RCM.NG.x.NG411          A----GT--AT-----------------A--T---A-T-G---G--G-----G---------A--AG----G---A-G----G--AAAAGAAT-GAAA-T--T--C-----A--AG-A------GG----G-A-TA--G--C--AA-A--C--A---AA-AGG--- 3522
SAB.SN.x.SAB1           A--C-GT--AT----------------T---AA-TT-C--GA---------------CT-A--AG-------GGAT-----G--G--AG-AT--GAGA-T--T--C--C-----AG-A------G-G--C---CTGT----C--AAGA--C--T---AGCA----- 4426
SMM.CI.79.SIVsmCI2      A------G--T--CAC------AT--TGC-----A-TCT----GGA---G-------G-----A--AG-------C--GG-------AAAAC-T---GAAC-G--C--T-----A-GAG-A--A---G-G--T---GC--C----C-TAGA---G-A---CC-A-G-C- 3447
SMM.LR.89.SIVsmLIB1     T-----------------A-AT--TGC--T--A--AT---GA-------G-----G--CA-A--A--C--------G-C-----AAAGG-TT--GAAC-A--T--C--------AG-A--A---C--G-----T--CT-A-G--AA----A--C--A--TCC-----C- 3406
SMM.SL.92.SIVsmSL92A    A---G--A-A-----G-----AT-CTGC--T-AAC--T-A---GC---------------C---G-C-------A-G-T--G--AAAGGT-T-GAAC-G--T--C--C----AG-A--AC--G-G--CT--GCAA-G--CC-AGCA--C--A--GCCGA----CT 3408
SMM.SL.92.SIVsmSL92B    AG---GTA-AT--CA-T-----CT-CTGC--T--AA-C-G----GA---G-------------C--AG-C-------A-G-T--G--CC-GC-AT--GAAC-G--T--C------C-AG-A---C-G---G--CC-G-TAT-G--C---A-A-----TCC-AGTA-G--C 3387
SMM.SL.92.SL92B         AG---GTA-AT--CA-T-----CT-CTGC--T--AA-C-G----GA---G-------------C--AG-C-------A-G-T--G--CC-GC-AT--GAAC-GG-T--C------C-AG-A---C-G---G--CC-G-TAT-G--C---A-A-----TCC-AGTA-G--C 3758
SMM.US.04.G078          ------A--T----G-----CT-TTGC--T-AAC--T-A---GA---------------CA-A--AG-C-----G--GA-C---A-AAAGT--T-GAAC-G--T--C------C-AG-A--A---G-G----CT--CT-A-G--AG-AACA-----A--AC-----C- 3587
SMM.US.05.D215          A------A--T----------AT--TGC--T--C-T-AT-A---GA---G-------C--CA-A--AG---G------G-C-GG--AAAA---T--GAAC-G--T--C--C---AG-A---C--G-----CT-CCT-A-G---G-AA-A-----A---CCTA---CC 3581
SMM.US.06.FTq           -------A--------T----TT-TTG---T---C----G--GG---------------CA-A--AG-C-------A-G-T--G--AAAACT-T-GAAC-G--T--C-----AC-AG-A--A---G----TT--CTCA-G--AG-AACA-----A---CC-----C- 3609
SMM.US.86.CFU212        A------A-----CAC------AT--TGC-----C--T----GA---G-------T-----A--AG---G-----A-G-C-GG--AAAGGTGT-GAAC-G--C--T------C-AG-A---C--G-G--CT--CTCA-G--A--AACA--C--A--C--AAAGTGT- 4244
SMM.US.x.H9             -------A--T--------A-AT-CTG---T---C--T-A-G-GA---------C------G--AG-------C---GGC--GACAAAACTG--GGAAC-G--T--C-----AC-AG-A-----GG----CT----------C--AGCA-----A---CC-A---C- 3800
SMM.US.x.PGM53          ----G--A--T----------AT-CTG---T---C-AT-A-G-GA-----G----C------A--AG-------C---G-C--G-CAAAAC-TT-GGAAC-G--T--C-----AC-AG-A-----GG----CT-C--------C--AGCA-----A---CC-A-G-CC 4244
SMM.US.x.pE660.CG7G     -------A--------------AT-CTG---T---C--T-A-G-GA---------C------G--GG---G-C---G-C----CAAAACTTT-GAAC-G--T--C--C--AC-AG-A-----GG----CT-C--------C--AGCA--------CC-A---C- 4316
STM.US.89.STM_37_16     ------A--T-----G-----CT--TG-------C-AT-----GA---G-------C--CA-A--AG-C-----G--GA-C---A-A----AAAAG-TT--GAAC-G--T--T-----AC-AG-A--A---G----T--GCCA-G--C---CA--TCC-A---C- 3974
SUN.GA.98.L14           ..--C----A-----G-----A------G-A-A-CTAGAA--GAG---G----------A-A--CC-GG--------TCC-G----AAG-AAT-GGAAA----C--C------C-AG-A--AC--G-G--T--AA-AA-G--C---G-----AGTA-GAGCTC---T 4382
SYK.KE.x.KE51           -------A--T----C--------C-----T---A-T-GC---GA----C-----T--CA-A--AG----G--T-A---G-C-T-AAAG-GGT-GGAAA-------C-----A---G-------G----G-A-TAA-G--A--A---C--A---CC-----T 3883
SYK.KE.x.SYK173_COMGNM  ---T--------C-------A--------T--A--C-GC--GAC------T----------AG-A-G-GAT--G-C---G-G---ACAGTAT--GGAGC-A--------C--AC--G-A---A---G-G-GCT-GCTCA-G--A--------C--AAA-GACA----- 4228
TAL.CM.00.266           ----C--G-AT---C----C--T-----CC--C-A----T-A---GA---C--G-------C--G---GCA-C-----A-G--C-T--CA---AGT--GAAA-A--T--C--C--A---G-A--A---GG------G-ACTAT-G--AC-CA-A-----A--GCCCA-G--- 3994
TAL.CM.01.8023          ------A-A-----G-----T--T---CC----A--AT-A--GGA---C--------C-----G---C-CGTC-----A-G--C-T--A---AGGA--T-GCA-T--T--T--------G-A--A---G-----G-GCTAT-G--A--AA-AC-C--T--TCC-A-G--- 3492
TAN.UG.x.SIVagmTAN1     --A-GGTG-AT---TGC--------TGC--T--AA-TT-C-----------G-----C-TCC-A-------C----GA-TC-C-G--CA-AGAAT--GAG--T--C--G--C--AC-AG-A--A---------G-CCT------A--AA-AC--AGTAAGAGCA----T 4227
VER.DE.x.AGM3           A---TGT--A-----G-----A--TTG---T--AA-TT-A-GGGA-----------A-C--AC-ATAT--G-A----GG--G-AA-AAT--GAAA-T--------C--C-AG-A--A---ATG--C--AA-AA-G--AC-AG-A---AGT---GCCTA-T--- 3750
VER.KE.x.9063           AG--TGTA-AT----------CTG---T--AA-CT-A---GA-----------CA----AC-ATAT---A---G-G---A-AG-AAT--GAGC-G--T-----C---C-AG-A--A---GGT---C---AA-AA-G---C---G-----AGT---GA-A----- 4255
VER.KE.x.AGM155         AG--TGT-A-----------A--CTG---T--AC--T-G---GAG----------C--CA-A--CC-ACA---C-A----G---ACAACAG---GAAA----------AC-AG-----AC-GAC---C--AA-AA-G---C---G-----AGC---CCCTA----C 4247
VER.KE.x.TYO1_patent    AG--CGTT-A-----------A---TGC-----AA-TT-A---GA----------CA-CT-AC-ATAC----A---G-G---A-AA--AT--GAAA----T--C------C-AG-A--A---AC---T--AA-AA-G------G-A--CAGT--GCCTA-T--G 4252
WRC.CI.97.97CI14        T------T-AT--A-------AT-T--A-----A--C-G---GA---G-----------...TGG-CC-A-TC--A-G----------A-GTG-GCAAA-T-----T------C-A--A---A--GAT----T-AA-A-----ACA-A--A-G 4322
WRC.CI.98.98CI04        TG-----A-AT--CA------AT-T--A--T--A--C-G---GGA---G--------C-TC...TGG-CA-ACTCT-A-G--------AA-T-ATT--TA-A----C--T------C-AG-A--A---GCG--TT-GA-GT----CC-AG-A---GG---T-C-A-GA-C 4337
WRC.GM.05.Pbt_05GM_X02  T----GAAA-G---A-------T-T-----T-AA----GC---GA---G----------...TGG-C--AGACT--GA-C--GA-ACAG-AAA-C--AA-T--T--------C--------C-GATT---T-AAGAA-G--A--AG-A--CAGTTC-GATA-G--- 3251
```

Draft of 30 September 2015, not for public release
plvdna_main.tex r8876–10546

Pol p66 RT, Pol p15 Rnase H end Pol p31 Integrase start

```
H1B.FR.83.HXB2              ATTAGTCAGTGCTGGAATCAGG......AAAGTACTATTTTTAGATGGAATAGATAAGGCCCAAGATGAACATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAGTGATTTTAACCTGCCACCTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGTCAGC 4362
Pol                          L  V  S  A  G  I  R  .  .  K  V  L  F  L  D  G  I  D  K  A  Q  D  E  H  E  K  Y  H  S  N  W  R  A  M  A  S  D  F  N  L  P  P  V  V  A  K  E  I  V  A  S  C  D  K  C  Q
H1O.CM.91.MVP5180           ------A--CAAA-AC--T--A......-G---C--G--CC-G--A---------C----A-----A--T-----A--------T-----------------T-A--------C---GGAT-A-----AA-------C--G-----CA-T--T--T---CC------C--TA 4391
H1N.CM.95.YBF30             ------T--CT-A-----T--A.........--A--T----CC-------T-----A--A--------A--T-----C-G---------C-------A---------C------------T-A--C--CA----G---------------------------C-----C---- 3956
H1P.CM.06.U14788            ---------YAAR-AC--T--A......-G---GT----C--G--A--C-----CC-A--A--G--G--T-----A------------------------CC-A--------R---GG---T-----AA----G--------G--TA-YAA--AT--CCC---------TG 3833
CPZ.CD.06.BF1167            ------A--CAAAAAC--T--A..........----T--T--CC----A---------C-A--A-----A--T-----A-G--TC--T----------A-ATGT-A---GA---A-A---TT-A--C--CA-------T--G-----CA----TCAA--CA----------CA 4470
CPZ.TZ.00.TAN1              ---G--T---AAAAAT--T--A.........--A--T----CC-G--------TA--G-A--A--G--A--C-----T------------------A---TT-A---GA---A-A---T-----C--A--T--G--T--------TA-T--TCAG---CCA-----C--TA 4006
GOR.CM.07.SIVgorCP2135con   ----------AAA-AT-----A......-G---CT----CC----A--C------C-A--A--G--A--T-----A------------------------C-A--C-----C---GGT-----G--A--G-----C--------CA-CAAT-----CCC---G--C--TG 3898
```

Pol RNAse H end Pol p31 Integrase start

```
MAC.US.x.239                CC----T---CAA--G--T--A......C----T--C--C--G--AAAG-----GCCA--A-----A--------T-----C--T------GTA-A--A-T---TATTCA-A---GGAT-A--CAGAA----G--C-G-C-G-------A--C---------------T- 4661
Pol                          H  L  V  S  Q  G  I  .  .  R  Q  V  L  F  L  E  K  I  E  P  A  Q  E  E  H  D  K  Y  H  S  N  V  K  E  L  V  F  K  F  G  L  P  R  I  V  A  R  Q  I  V  D  T  C  D  K  C
H2A.SN.86.ST_JSP4_27        C-----A---CAG--C-----A......C-----T----CC----GAA-------ACCC--T--G--G--------A---------T--C---GTA-A--A-C-AT-CCA-A-A---GGA-----CAAAT----G----G-C---------AA--CA--CACCC---------- 4189
H2B.CI.x.EHO                CC----G--CCAA-----T--A......C--A-C------C----AAA------ACCA--T-----A--------A---------T-A----GTA-A--A-C-A-TCCA-A-A---GGGA-T----AAT----G----G-C--------AATTC--------------C--A- 4715
H2G.CI.92.Abt96             TC----A--CCAG-----T--A......C-----------C--GR-AAA------ACCT--T-----A--------A--R-T---T------GTA-A--AGT-AA--CA-A-G---GG-A-T--T-AAT------------C-G------AATTCA-ACC-C-GT--C--A- 4071
H2U.CI.07.07IC_TNP3         T--G--T---CAG---------A......C-G--TT----CC----AAA------ACCA--T--G--A--------C----T---T------GTA-A--AGC-A-TCTTCA-A---GG-A-A---AGAT----G------C--------G-AT-CT--------------C- 4176
H2U.FR.96.12034             T--G--G---CAG-----T---.......C--------G---C----AAA------ACCA--T--G--A-----------------C--T--C---GTT-A--AGC-ACAGCA-A-A---GGTA-A----AAT-------C---C--------AAT-CT---C--GTT--C---- 4223
ASC.UG.10.RT03              -C----A--CAAA---------A......-GG---T-G--CC-G--ACA----CCAC-A-----G--A--------A-GG-----T-AC------CA--ACC-TAGAGACAGG--CC-GA-A---G-AT--A----------G--CT--AAAGCA--CCCC--G-----AG 3768
ASC.UG.10.RT08              ------A--CAAA---------A......--G---T-G--CC-G--ACA----CCAC-A-----G--A--------A-GG-----T-AC------CA--ACC-TAGAGACAGG--CCGAA-A---G-AT--A----------G--TT--AAAGCA--CCCC--G-----AG 3792
ASC.UG.10.RT11              ------A--CAAA---------A......-GG---T-G--CC-G--ACA----CCAC-A-----G--A--------A-GG-----T-AC------CA--ACC-TAGAGACAGG--CC-AA-A---G-AT--A----------G--CT-GAAAGCA--CCCC--G-----AG 3777
COL.CM.x.CGU1               -AAG--A.........---G-A--A......C-G---A-G-GGA-----AA------AGCA--AG----A--C---C-A--G-T---T------GT-CAGTATC-AAAAGAAC-A---GGGT-A--CA---------------------TTGG-AA--A---TCAG-G-------A 4125
COL.UG.10.BWC01             -AAG--A.......--TG----A......C------A-G-GGA-G--AAAG-----AG----AG--CTA--T---C-----------T------GTTCA-TAC---CAAGAGC-G---GGAA-A--T-AGT-G------T--G------TTGG-AG--G---CC-C---------A 3924
COL.UG.10.BWC07             -AAG--T.......---G-A--A.........----C-A-G-GGA----GAA------AG----AG-G--A--T--CAGC-----TC------C---GTTCAGCATT-AAGGGAGACA-A-GGGT-A---G-AA--------T----------TGG-A---A--CTCAC---------A 4152
DEB.CM.04.SIVdeb04CMPF3061  ------T---AAG-----T--A.........-GA-C--T---C----AA--G----AG----T--G--A--T--C--T------C---------------A-CAGC-AAGGGAG--G-----T--C--CA-CT----T-----GC------A-TAATCAA---CCC---------AA 4349
DEB.CM.99.CM40              ----------AAA-----A--.......-GGA-C--C--C--G--AA-GG----GG-A---------G--C-----A-----C--TG-G-------A-CA-T-AAGGGA---A-ACC-A--A--TA-AC-CA----T--GC-----TA--A-ACAA--CCCC-----------A 4168
DEB.CM.99.CM5               ---G--A---AGA-----A--.......-G-A-C------C--G--AA--G----AG----T-----G--T---------------TG--------A-CA-T--AGAGA---A----C-GT-A--TA-CT--A-----C--GC-G--TA-T--GCAG---CCA--G--C---A 4162
DEN.CD.x.CD1_CMO580407      GC----T---AAA--G------A......C--A-C-----C-----ACA---TCCAC-A--T-----A---------T-----C--T--C-------A--AG--CAGAGA-A-G-A-GG---A-----A--A-G------T--G------CA-TAATGT-----CCA----------A 4252
DRL.DE.11.D3                ------T---AGA-----A--A......C----GT-----C-----AAT--G--A--A--A--G--A---------TTT---C--T-A-----------T-CT-A--ACAA--A---GGT--A---GGGA----G-------------------G--ACAA---CC----G------AA 3836
DRL.x.x.FAO                 -C----T---AGA-----A--A......C----GT----CC-----AAC--G--A--A--A--G--A---------TTT---C--T-A-----------T-CT-A--ACAA--A---GGTT-A---GGGA----G-------------------G--ACAG---CC---------AA 3929
GRV.ET.x.GRI_677_gri_1      ------A--CAAA-----A--A......C-GA-C--C--C-------A-------AG-A--A---------C-----CA--G--C--T-AC----------AGT----TACAG--A---GGAT-A--TAA-A----------------G---------GGCA---CCC------C--AA 4535
GSN.CM.99.CN166             ----------CAAG-------C-C......C-----G--T--CC-G--AAAC-----GCCT--AGT---A--C-----A--G------T-----------AGTACC-AAGAGA-C-A-AC--AA-C--CG-AC--T-G--T-------------GAA--AA---TCC--G-------G 4090
GSN.CM.99.CN71              ----------CAGA--G---C-C......C---------C--CA-G--GAAC-----GCCC--AGT---G--C-----A---------T-----------AGTATT-AAGGGA-C-A-AC--AA-C--GG-AC--T-G--T---------------AA--AG--CTCC-----------A 4072
LST.CD.88.SIVlhoest447      CC---------AAA------T--ACAAAAAC--------G---C---GAAAG-----GCCC--A--G--A--GT--AGC----T---T-A----GCT-A--ATT-A-AAGAAA-A---C--A-A--C---A-G------T--GC---------GAAT--T---A------------A 3639
LST.CD.88.SIVlhoest485      CC---------AAA------T--ACAAATAC--------G---C----AAA------GCCA--A--G--A--G---AGC----T---T-A----GCT-A--ATC-A-AAGAAA-A---C----A--C---A-G------T--GC---------GAA---T---A------------A 3639
LST.CD.88.SIVlhoest524      ------T-T-AAA------T---ACAAAAC-G--CA-G---C-G--GAAG-----ACCA--TGT---A--G----GA--G-T------A----GCA-A--ACT-A-AAGAAA-G-A-GGGA-A--C--CA-G------T--GC---------AATGAG----C---G------- 3636
LST.KE.x.lho7               CC----A---AAA--T--A--ACAGAAAC-G--CA-G---C-----AAAG-----ACCT--AGT---A------AGC-----TC--T-AC---GCA-AG-ATC-A-AAGAAA-A---------A--C--CA-G------C-----C-----T--CAATGA-----CA--C------AA 4723
MAC.US.x.251_1A11           CC----T---CAG--G--T--A......C----T--C--C--G--AAAG-----GCCA--A-----A--G-----T------C--T------GTA-A--A-T---TATTCA-A---GGAT-A--CAGAA----G--C-G-C-G-------A--C---------------T- 4660
MND-1.GA.x.MNDGB1           GC----A---T-A---------A.........-----CT----CC--C-AAAT-----ACCA--A--G--A-------------------T--C---GAAGC-CA-T-AAGAGAGA-A--CC--T-A---G--C--------C---C-G--T--GCAA--T--CAG---G--CTGT- 4049
MND-2.CM.98.CM16            -C----G--CAGA---G-A---......C--------G---------A--C----------A--A-----A--G-----T-----------A---------G---T-A--ACAA--C--CTGTA-A---AACA----G------------------CAG---CCA--G-----AA 4363
MND-2.GA.x.M14              ------A--CAAA---------A......C-----------CC-G--A------------A--T--G--G--------C------C--T-A----------T--T-AAG-CAG--A----G-A-A--C---A----G--------------T--ACAG--CCCA--G------G 4290
MND-2.x.x.5440              ------A---AAG-----A--A......C-G---------------G-------------A-----G--A--------C-----------A-------------GT--T--CAG--C---------A--C--CA----G--------G-----T--ACAA---CCA----------A 3923
MNE.US.x.MNE027             CC----T---CAA--G--T--A......C----C--C--C--G--AAAG-----GCCA--A-----A--------T-----C--T------GTA-A--A-T---TATTCA-A---GGAT-A--CAGAC----G--C---C-G-------A--CA-------------T- 4129
MON.CM.99.L1_99CML1         CC----GG-CCAG-----T---.......C----C--G---A-G--AAATT-G--ACCC--AAGG--A--T-----A--G-----T----------A-TACT-AAGAGA-ACC-ACC--A-T--TA--C-GT-G--C-----------CAA-CAT---C-C---------AA 4065
MON.NG.x.NG1                ------A--YCAA--G-----A......C-G--C--C--CA-G--GAAC--T--ACCR----G--G--T-----A-----------C-------A-TATC-CAGGGA-AC--A-GG-A-T---A-AC-CC----C-----G-----CAA-CAT--CCCC--G-----CA 2658
MUS-1.CM.01.CM1239          -C-G--A--CCA------T---.......C-G--G-----CA-G--GCA------GCCT--AA----G--C-----A--G--C--T--C-------AGTATC-CAGGGA-A-G-AC---A-C---G--T--C-G--T--G----------AA-CTT---CCC-----------A 4042
MUS-1.CM.01.SIVmus_01CM1085 ------A-A-AG-------------......C----C-----CA-G--ACAG-----ACCT--AA-G--A--T-----A---------------------AGTACC-CAGGGA-A-A-ACC-AA-C---G--T-GT----C------------AA-CAT---CCA-----C---A 4052
MUS-2.CM.01.CM1246          CC----T--CAGG------C-C......C-G----------A-G--AAAT-----CCC---AGTT--A--C---------G--C--TTC---------AGTATC-CAGGGA-A-A-AC---A-A--CA-CA-CT----C--G----------AA--AA--CTCAGC-------AA 4150
MUS-2.CM.01.CM2500          CC----T---AAA--G---C-C......C-G--T--G---A-G--AAAC--T--GCCA---AT---A--C-----T----------TCC--------A-TATT--AGGGACA-A-A---TA-A--CA-AC-TC-T--C--G----------AA---A--CTCAGC---C--AA 4109
MUS-3.GA.09.09GabOI81       ------T--CCAA--T--A---.......C-----T-G---C-G--AAATG----ACCA--AA-G--G--T-----C---------T-----------A-TACT-AAGGGA--T--ACC-GA-A---G-AC--T-G----------------AA-CA---CAGC-----C--AA 3618
MUS-3.GA.11.11GabPts02      G-----T---CAG---G-A---......C-GA---------A-G--GAAC-----GCCA--AGT---A--T---------G--C--T--C-------A-TACC-AAGGGACACA-AC--GA-A--CA-AT--C-G--T------------TAA--AA----G---G-----AA 3639
OLC.CI.97.97CI12            ----------AAA-----A---ACTAAGC----GA-G-----G-CA-AC-----GCCA--A-----GAGC--CA-T--G-----TCAA---GCACA-AGTT--CAGCAA--G--CC-AA-A----TAA----G------AGG-------ATGAA-----AG-----A-AG 4233
RCM.CM.00.SIVagi_00CM312    ------G--CAGA-----T--A......C-----T----CC-T------C-----A--A--A-----A---------T---------T--C------------T-----AGAA-----C-GAA-A--T-AA---------T--G---------G---CAA---CC---G--C---G 3633
RCM.CM.02.SIVrcm_02CM8081   ------A---AAA-----T--A......C----C--C-------------------A-----T-----A--G-----A---------T-A-------------------AGAA-----C-GA-A----AA-----------G--G----------CAA--CCC-----------G 3651
RCM.GA.x.SIVRCMGAB1         ------T---AAA---------A......C-G-----G--CC-------G-----A--A--A-----A---------A--G-----T-A-------------T-----AGAA-----C-AA-T----AAA----G----------------ACAG---CCA---------AG 3841
RCM.NG.x.NG411              ------G---CAG-----T--A......C-G---------CC----C--G----GA--A--A--G--A---------A---------T--C-------------------AGAA--A---C-AA-A--T-AAA-----------------------CAA--CCCC--G-------G 3856
SAB.SN.x.SAB1               -C----T---CAA-----A--A......C-----------CC-G--CA----T--AG-A--A-----A--G-----T-----C--TGCC---------GAGT---CAACAA--A---GG-T-A--TG--A----------------------GGCA---CC------C---A 4760
SMM.CI.79.SIVsmCI2          TC----G---CAA--G--A--A......C-G--C--G---C-T--ACA------ACCA--A--G--A--G-----A---------T--C---GTC-AGCAGT-A-TGCACA-A---GG-A-A--C-AAT----G--T---C-------GAA--C---CAC-C-G-------- 3781
SMM.LR.89.SIVsmLIB1         T--G--T---CAAA----T--A......C----T---.-C-----GAA------ACCA----------A-----------G-----T--C---ATA-A--A-T---TGTT-A-A-A-GG-A-A---AGAT-G---------C-G-------A--C------C-----C--T- 3739
SMM.SL.92.SIVsmSL92A        TC---------CAA---------A......C-G--------C-----AAAC-----ACCA--A-----G--------T-----------------ATA-A--A-C---TCTA-A-A---GGTA-C--CA.AT-------T---C---------AAT-C------C-----C---- 3741
SMM.SL.92.SIVsmSL92B        T-----T---CAA-----A--A......C-G--C--G--CC----AA-T-----GCCA--A-----A--C-----C---------T--C---GTA-AG-AGT-A---CAGA-A-A----A-C----AAT-------T---C-G------AATGCT---A-C-----C--A- 3721
SMM.SL.92.SL92B             T-----T---CAA-----A--A......C-G--C--G--CC----AA-T-----GCCA--A-----A--C-----C---------T--C---GTA-AG-AGT-A---CAGA-A-A----A-C----AAT-------T---C-G------AATGCT---A-C-----C--A- 4092
SMM.US.04.G078              TC-G------CAA--G--T---......C----CT-G--CC----AAAG-----GCCA--T-----A--G-----A-------------C---GTA-A--A-T---TCTT-A-A---GGGA-A---AGAT-G---------C-G-----G-AT-C--------------C---- 3921
SMM.US.05.D215              C--G------CAA---------A......C----T------C----AAA------ACCA--T--G--G--T-----A-----C---------GTA-AG-AGC---TGTT-A-A-A-GG-A-A---AGAC----------GC---------A--CA-----C-----C--C- 3915
SMM.US.06.FTq               T-----T---CAA---------A......C----C--G---C----GAA------GCCA--A-----A-------------G--C--T------GTA-A--A-C---TGTTCA-A----GAA-A--CAGAT-----------C-G---A-T-AT-CT-----C-----C--C- 3943
SMM.US.86.CFU212            CC---------CAG---------A......C-----T-G---C----AAAG-----ACCA--T-----A--G-----A-----C-----C---GTA-A--AGT---TATTCA-A-A-GG-A-A---AGAT-G---------C----------AT-C-------C-----C--C- 3907
SMM.US.x.H9                 TC----T---CAA-A---T--A......C----C-----C----AAA------ACCA--G--G--A--G-----A-----C--T--C---GTA-A-AA-T---TATTCA-A---RGTT-A--TAGGC-----------C---------A--CA-------------C--C- 4134
SMM.US.x.PGM53              CC-G--T---CAA-----T--A......C----C-----C----AAA------ACCA--A-----A--G-----A--G--C--T--C---GTA-A--A-T---TATTCA-A---GGTA-A--TAGGC-----------C-G-------A--CA---C--------C--C- 4578
SMM.US.x.pE660.CG7G         CC-G--T--CCAG-----T--A......C----C-----C----AAA------ACCA--A-----A--G-----A--G--C--T------GTA-A--A-T---TATTCA-A----GTT-A--TAGGC-----------C-G-------A--CA-------------C--C- 4650
STM.US.89.STM_37_16         T--G--A---CAG-----T--A......C-G--C-----CC----AAA------ACCA--T-----A--G-----A---------T--C---GT--A--AGC-A-TCTTCA-G---GG-A-A---AGGC----------GC-------G-AT-CA-----C--G--C--C- 4308
SUN.GA.98.L14               -C----T--CCAG---------AAAGAGAC-G---A-G---A----AAAG-----ACCA--TGT---A-------G---G-TC--T-AC---GCAGCCT-TT--CAGGAGATG---G--A-A--C-TA--G---------GC-G-----CAATGAA----CAC--------A- 4722
SYK.KE.x.KE51               ------A---AAG-----T---.........---A----G---C-G---AAC--CCCA--A--A--G--A---------------------C----GTAGA-TAC--AAGGCAA--A---C-T--A--CAGACAG-----C----CT-------AA-TG---CCC-----C--AA 4217
SYK.KE.x.SYK173_COMGNM      -C--------AAA-----A--A.........---A-C-----CC----AA----TCCCC-A--A--G--A--T-----A-GG--C---TCA---AT-GAGTATC-TAGGCAG--G---C-T--A--C-GACAG-----------C---TA--CAACAA---CCA---------AA 4562
TAL.CM.00.266               G-----A--CCAA--------------......--GA-C--T---G----ACAG--CCCAG-A--A-----A-----------G---C--T-A----------G-ATC-TAAGGCAAGA-----GT-A--CA-CA-------C--G-CT--TA---AAGCA--CCCC---------AG 4328
TAL.CM.01.8023              GC-G--G---CAA--------------......-G-A-T--C---G----ACA---CCCAG-A--A--G--A--G-----A-GG--C--T-AC-------AG-ACC-AAAGGCCAGG-----A--A--CA-CA-------C----CC--CA---AGGCA--CCCC---------G 3826
TAN.UG.x.SIVagmTAN1         GC----A--CCAG--G--T---......C----G-----C-----CA----G--AG-A-----G--ATC-------T---------T-C--------CA-TTT--AAGAGA--CA--CGGTA-A---G-AC-G-------------------TGCT--CCCA---------AA 4561
VER.DE.x.AGM3               -C----A---CAA--G--G---.........--A-TT-----------AAA------AG-A-----G--G---------A-GG--C--T-A----------GAACT-A--AGACAC--A-GGG--A----AAA-T--G------------------TG---CCA----------A 4084
VER.KE.x.9063               ------A---CAA-----A--A......-G---T------A---GAA-------AG-A--T-----A---------C-G-------T-----------A-AATT-A--GGACACA---GGAT-A----AGA----T--C-----G---------A-TG---CCA--TGCCA-AA 4589
VER.KE.x.AGM155             ------A--CAAG---G-T--A......-G-A--T-G--CA-T-GCA-G-----AG-A--A-----A---------T-GG-----------C------AATC-A--AGACACA---GGAT------AAA--------T---------T------A-TG--CCCA---------AG 4581
VER.KE.x.TYO1_patent        ------G---AAA--C--T--A......-G---TT----C-----AAA------AG-A--T-----A--G-----A-G------T-A---------A-AACC-A--AGA-ACA-A-GGG--T----AAA--------------G-----G----TG---CCA----------A 4586
WRC.CI.97.97CI14            ------G-----A--G--T--ATCA...---A-TA-G---C----AAAT---ACACCA--AAT---AAGT---ACT-----C--TCAAGG-----AG--CT-ACAAGAA--A---C-T--T--T-AAT-GA--------AG----A---AGGAA---AG--CC--C--AG 4659
WRC.CI.98.98CI04            ------A-----A-----T--ATCA...-------A-G--C-----AAAT---AC-CCA--GAT---AAGC---ACT--------TCAGGGC----A----T-ACAAGAA--A---C-TA-T----AAT--A-T------AG--------AAGAA---AG--CC-----AG 4674
WRC.GM.05.Pbt_05GM_X02      ------A-----A--C--T--ATCT...-G----A-G---C-----AAAT---ACCCCA--AAT---GAGT---ACA--------TCAAGGG---G-G---T-A-AGGAA--A---C-T-----CA-CA-TC-T------AG--------AAGAA----C--CC------G 3588
```

```
H1B.FR.83.HXB2                CTTTTAAAATTAGCAGGAAGATGGCCAGTAAAAACAATACATACTGACAATGGCAGCAATTTCACCGGTGCTACGGTTAGGGCCGCCTGTTGGTGGGCGGGAATCAAGCAGGAATTTGGAATTCCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAAGA 4699
Pol                           _L__L__K__L__A__G__R__W__P__V__K__T__I__H__T__D__N__G__S__N__F__T__G__A__T__V__R__A__A__C__W__W__A__G__I__K__Q__E__F__G__I__P__Y__N__P__Q__S__Q__G__V__V__E__S__M__N__K__E_
H1O.CM.91.MVP5180             --G-------------C---------T--C---GT---------A--------ACCT-----T--AA----AG-CA-G-AA--T--A---------A-A--C--AC-A--T--G-----G--A--A--T-----A---------------------G-C----------- 4728
H1N.CM.95.YBF30               A-------G-----T------------------GTT-----C-----T-----ATC---------TA----C--T--A-AA--A--------------AAAT-----A-----------G--A-----------T--------G----C-------G-C----------- 4293
H1P.CM.06.U14788              G-GC-C-----G---TC---------T------GT------C--A--------GCCT--C--T--TA-C---G-A--A-AA--A------------CTTAAT--A-CT--T--G-----G--A--------------------G--------G-----A----------- 4170
CPZ.CD.06.BF1167              A--------C------------------------AG--------A--T-----AGC------T--TA----G-A--A-A--A--A--C--------AAAC---C-A----------------------------A--------------G--------C---------C- 4807
CPZ.TZ.00.TAN1                A-CC-------G--------------T-----G-A---------A--T-----ACCA-----T--TA--A-AG-A--G-A--A------C--------ACA--TC-A--T---------G-----A--T-----T---------------------------------C- 4343
GOR.CM.07.SIVgorCP2135con     --GC-C--G-------CT----------------T---------A-----C--GCCT-----T--TA----C--A--C-A---T------------CTCAATG-A-CA--T--G--------------T-----------------------G-----------C----- 4235

MAC.US.x.239                  --G--------G-----C--------TA-T-C-CATC----C--A--T-----TGCT--C--TG-TTCGCAAGAA--A-A-ATG-TTGCA--------A--G--AG----CACC-----GG-A--A--------A--G-----G--------G---G-A------C-CC- 4998
Pol                           _F__L__L__K__L__A__G__R__W__P__I__T__H__L__H__T__D__N__G__A__N__F__A__S__Q__E__V__K__M__V__A__W__W__A__G__I__E__H__T__F__G__V__P__Y__N__P__Q__S__Q__G__V__V__E__A__M__N__H_
H2A.SN.86.ST_JSP4_27          --AC-----C-G--CA-T--G------A---C-CATT-G--C--A--------TGC---C-----TTCACAGGAA--A-A-ATG-TGGCA------ATA--T--AG-A--ATCC--C---G-A--T--------A-----C-----------G---G-A------C-CC- 4526
H2B.CI.x.EHO                  --G-----GC-G--CA-C--------TA-C-C-CACC-G--C--A-----C--TGC---C-----TTCACAAGAT--G-AAATG--AGCC------ATA--G--AG-A--AAC---C---G-G-----T-----AG-------G------------G-A-----CC-TC- 5052
H2G.CI.92.Abt96               --G--------------G--------TA-C-C-CACC-------A--------TGC---C------TCGCAAGA---G-A-ATG-TAGCC-------TA--C---G----A-C-------G-A--A-----C-----G--------G-----------A------C-TC- 4408
H2U.CI.07.07IC_TNP3           --G--------G---ACT--------TA-CTC-CATT-G-----A--T-----TGC---C------TC-CAG-AA--G-A-ATG-TAGC--------TA--C---G----ATC------GG-G-----T--Y--A--G--------------G---G-C------C-CC- 4513
H2U.FR.96.12034               --AC----GC----------G-----TA-C-CTCATT-G--C--A--T-----TGC---C-----TTCACAGGAA--A-A-ATG-TAGCA-------TA--T--AG----TACC------G-A-----T--C--G-----C--------G------G-C------C-CC- 4560
ASC.UG.10.RT03                --G---C--C--TGT-CT-----------C-C-CAG-----C--A--T-----ACCA---------A-CCAGGAAT-AGCA--A--AGCA--------TAATG-AC-A--CAGCACA--TG-G--T--------A--G-----G------------AAC-A---C--GC- 4105
ASC.UG.10.RT08                --G---C--C--TGT-CT---------A-C-C-CAG-----C--A--T-----ACCA---------A-CCAAGA-T-GGCA--A--AGCA--------TAAT--AC-A--CAGCACA--TG-G--T--------A--------G------------AA--A---C-GGC- 4129
ASC.UG.10.RT11                --G---C--C--TGT-CT-----------C-C-CAG-----C--A--------ACCA---------A-CCAGGAAC-AGCA--A--AGCA--------TAATG-AC-A--TAGTACA--TG-G--T--------A--G-----G-----------GAAC-A---C-GGC- 4114
COL.CM.x.CGU1                 --CA-------G--TA-TCAG---GAG--T-G-CA-G-T--C--A--T-----GCCA----TGTAA--CAGCACT---A--A--AGTG--------CTA-----AGCA--CACCACA---CAC-----T-----T---TCA--------T-----GCAG-GA--C--G-- 4462
COL.UG.10.BWC01               --CA-----C----TA-TCA----GAG-----GCAGG-G-----A--------TCCA-----T--TA--CAAG-AT---A--A--TGTA-------CTA-----------TAC-ACA--------A-----C--A---TCA-----CA--------CAA-GA-----GAT 4261
COL.UG.10.BWC07               ---A-T-----G--TA-C-TG---GA------GCAGG-G-----A--T-----A-CA--------A--ACAGG-CT-C-A---A--GGTC------TTA--G--T-CA--CACCACA--G-C---T--------T---TCC--G---A-------GCAA-G------GAT 4489
DEB.CM.04.SIVdeb04CMPF3061    --G---C------GG-CT--G-----CA--C--CA-G-T--C--A--------ACC---C--T-TTA--CAGG-AT--GCT---------------TTA--G--AG----CACCACA--GG----A-----C--A--G------------------AAC-AA----G-C- 4686
DEB.CM.99.CM40                --G---C-G----G--C--------------CTCAG-----C--A--------ACC---C--T-T-A--CAAC-A--C-A---A--T--C------CTA--T--TG----CACCACA--GG-A--A--T--C--A-----C--G-----G------AG--A---C-GGAT 4505
DEB.CM.99.CM5                 T-A---C-G--G-GG-C---------CA-T---CAG-----C--A--------ACCA-------TAA--CAGC-A--C-A--A--A----------TT------AG-T--TACTACA--GG-A--A--T--C--A-----------G---------AGC-AA----GG-T 4499
DEN.CD.x.CD1_CM0580407        T-AC-GC-GA----T-CT---------A-C-CC-A-T----C--A--------ACCA-----T-T-A--CAAG-CA-GCA---A--A-----------AAAT-----A--CACCACA---G----A--T-----A--------G--TA-T------AA--A---C--GC- 4589
DRL.DE.11.D3                  --GC----------C--C-----------GTC-CACT-G--C--A--T-----ACCA---------A--AA-A--AGCAA-A-TA--C-----------TAA--AG----TACCACA---G-------T--C-----G-----------G--------A-----C--TC- 4173
DRL.x.x.FAO                   --GC----------T--C------------TC-CACT-G--C--A--T-----ACCA---------A--AA-A--AGCAA-A-TA--C-----------TAAG-AG----TACCACA---G-------T--C-----G-----------G--------A-----C--TC- 4266
GRV.ET.x.GRI_677_gri_1        --GC-G---C-GTT--C------------G---CATC----C-----T------CCA--C--T---TC-CAG-AT--GGCA--G-TG--C-------GTAAT--AG----CACCAC------A--T--T--C--A--GTCA--G--TAGT------AGC-----C-G-C- 4872
GSN.CM.99.CN166               --C---G---------C--C------T----CCCACC-------A--T-----CC-------GT-TCCAAGGAAT-GGAA--A------C--------A-AC--AC-A--CACCACA--GG-G--A-----C--G-----------G---------AAC-----C---C- 4427
GSN.CM.99.CN71                --A---G---------C--C-----------CTCACC-------A--T-----CC-------T-TC-AAGGAAC-GGA--A--T--------------AAAC--AC-A--CAGTACA--GG-G--A-----C--G-----------G---------AAC-----C---C- 4409
LST.CD.88.SIVlhoest447        --G------C-GTGCA-C------------C-GCA-G-C-----A--------ACCA--C--TGT-A-CAAGGA---ACAA----TAAT-------CTA--T--TG-A--TAC-ACA--------T--T--C--T--G-----G--------T---G-C-AA-------C 3976
LST.CD.88.SIVlhoest485        --G------C-GTGCA-C------------C-GCA-G-C-----A--------ACCA--C--TGT-A-CAAGGA---ACAA--A-TAACC------CTA--T--TG-A--CTC-ACA--G-----T-----------G-----G--------T---G-C-AA-------T 3976
LST.CD.88.SIVlhoest524        T-G-----G---TGTA-C--------T------CA---T--C--A--T-----TCCA----TGTGA--AAAGAT--ACA---A-TGACC-------TTAA-T--AG----CAC-ACA--G--A--A--T--C--T--------------------GG-A-AA-------T 3973
LST.KE.x.lho7                 --G---------TGCA-T--------T------CA---------A--------TCCA----TGTAA--AAGGA---ACA---A-TAACC-------ATA--G--AG-A--TAC-AC---G--A--T--T--C--T--G--------G--T------G-A-AA--C----T 5060
MAC.US.x.251_1A11             --G--------G-----C--------TA-T-C-CATC----C--A--T-----TGCT--C--TG-TTCGCAAGAA--A-A-ATG-TTGCA--------A--G--AG----CACC-----GG-A--A--------A--G-----G--------G---G-A------C-CC- 4997
MND-1.GA.x.MNDGB1             --G-----G-------C-CA------TA-T-GT-A-C----C--A--T-----GCCT--C--T--TA--CAGGAA--AGAAA--ATG---------TTA--G--AG-A--CAC---------A-----T--C--A--------G--G-----G--AA--AA------GT- 4386
MND-2.CM.98.CM16              --G--------------T---------------CATC----C--A--T------CCT---------A--AA-AA--AGCCA-A-T--------------TCA--AG----CACTACA-----------------T--G-------------A---GG-A-AA--CC-TC- 4700
MND-2.GA.x.M14                --GC-----C-------------------G---CATT----------------TCCA--C--TGT-A--AG-A---AGCCA-A-TG------------TCAG--AG----CACCACA-----A-----T--C-----------------T--G---G-A-AA---C-TC- 4627
MND-2.x.x.5440                --G-----GC-------T--G------------CATC----C-----------TCCT--C--TGT-A---AA-A---GGCCA-A-T---C---------TCA--AG----CACCACA---G-A--T-----C--A--G------------------G-A-AA---C-TC- 4260
MNE.US.x.MNE027               --G--------G---A-C--------TA-T-CGCATC----C--A--T-----TGC---C--T--TTCGCAAGAA--A-A-ATG-TTGCA--------A--G--AG----CACC-----GG-A--A--------A--G-----G--------G---G-A------C-CC- 4466
MON.CM.99.L1_99CML1           --G--GG-GC-G--C-CT-TG-----G--C-CTCAG--------A--T-----GCCA---------A-CCAAGAAT-CGAA--A--TGCA--------AAAT--TC-T--CACCACA--GG-A--------C--A-----C-----G--T-----GAA--------G-C- 4402
MON.NG.x.NG1                  --G--GG--C-T----CG-C-------A--C-TCATG----C-----------ACCA---------A--CAGGAAT--GCA--Y--TGCA--------AAAC--AG-A--CAGTAC----G-G--------C--A-----------T---------AAC--------GC- 2995
MUS-1.CM.01.CM1239            --G---G---------CTCTT-----T----C-CA------C--A--T-----AGC---C----TTA---AAGAAT-CG-A--T--A-----------TA-C--AG-T--CACCACA--GG-A--A-----C--T--------G------------AAC-AA--C---C- 4379
MUS-1.CM.01.SIVmus_01CM1085   --AC-GG-G------TC-CT------T-CG-C-CA---C-----A--T-----GGCT-C----T-A----AGAAT-CGAA--A--A--C---------CA-C--AG-A--TACCACA--GG-A--A--T--C--T--G-----------G------AAC-AA--C---C- 4389
MUS-2.CM.01.CM1246            --G-C---C--G----CC-C-------A-T-GTCAG-----CC-A--------T-C---C--T--TA-CAAGGAAT-CCAA--T-TTGCA--------TAAC--AC----CACCACA--TG-C-----------A--G------------------AA-GCA--C---C- 4487
MUS-2.CM.01.CM2500            --G-----C-------C--C------CA---C-CAG--C--C--A--------T-C---C------A--AAGGAAT-CCAA--A-TTGCA--------A--C--AG-A--CAGTACA--TG----T--T--C--A--G--C-----G---------AA-GCT-----GC- 4446
MUS-3.GA.09.09GabOI81         T-G---C--------A-TCA-----------GTCAT--C-----A--------G-CA--------TA---AAGAAT--GCT--A--AGCA-----------C--AG-A--TAC-ACA--G--A--A-----C--T-----------G---------AA-GCA--C--GC- 3955
MUS-3.GA.11.11GabPts02        --G---C-G--G----CT-TG----------C-CAT--------A--------GCCA--C-----TA----ACAAT--GCT--T---ACA-------GAAAC--TG-A--CAC-ACA--GG-A--A--T-----A-----------G---------AAC--C------C- 3976
OLC.CI.97.97CI12              G-GC-T-G-C----TA-TCA------T------T---T--C------------G-CA----TGTGAA-CAG-A-A-GCA---A-TA--C-------TTA--G--TG----T---ACG--C-----G-GG-----T-----------A-CA-----G--A-AA----G--T 4570
RCM.CM.00.SIVagi_00CM312      --G-------------C------------G---CACT-------A--T-----AGCT--------AA----AG-A--ACA--A-T-------------TCAG--AG----CACC------G-A-----------T--------G-----G--T---AG-------C-TC- 3970
RCM.CM.02.SIVrcm_02CM8081     --GC----GC------CC---------------CATT-G--C--A--------AGCA-----T--GA-CT-AG-A--ACAA--A-TT-----------TCAG--AG----CACC--C--GG-A--------C--A--------------------G--C------C-CC- 3988
RCM.GA.x.SIVRCMGAB1           --G--G---C------C--------------GG-AGC----C--A--T-----AGCA--------AA----AG-A--ACA---G-T---C--------TCAG--AG----C-CC------G-A--T--------T-----------------G---AGC---------C- 4178
RCM.NG.x.NG411                --G-------------CC-----------G---CATT----C--A--------A-CA--C--T--AA----AG-A--GCAA--A-A-------------TCA--AG----CACC------G-A-----------T-----C--------G--------A------C-TC- 4193
SAB.SN.x.SAB1                 --GC-G--G-----CA-C--------CA-C-C-CAGC-G-----A--T-----T-C---C--T--TA-CCAGCAA--AGCA--TATT--C-------GAAA---AG----CAC------GG-A--------C-----------G-----------------------GC- 5097
SMM.CI.79.SIVsmCI2            T-G--------G--TA-C--------TA-C-CCCACT----C--A--------AGC---C-----TTCCCAAGAA--G-A-ATG-TTGCA-------TA--C--AG-A--AACC-----GG-A-----T-----A-----------G--G-----G--C------C-TC- 4118
SMM.LR.89.SIVsmLIB1           --GC---------T-A-C--------TA-T-CCCATT-G--C--A--T-----TGC---C------TCACAAGAA--G-A-ATG-TAGCA--------AAAC--AG-A--ACT---A--G-A--A--------A--G-----------------GG-A-----CC-TC- 4076
SMM.SL.92.SIVsmSL92A          T-G-----------T-C---------T-----GCACC-G--------T-----AGC---------TTC-CAGGA---G-A-ATG-TTGCA--------A-----AG-A--ATCC-----GG-A--A--T-----A--------G--T--T------G-A-----CC-CC- 4078
SMM.SL.92.SIVsmSL92B          --C------A-----TC---------TA----GCATC-G-----A--T--C--TGC----------TCCCAGGA---G-A-ATG-TAGCA------TTA---G-GG-A---TCC------G-A--G-----C--A--------G------------G-C-----CTT-C- 4058
SMM.SL.92.SL92B               --C------A-----TC---------TA----GCATC-G-----A-----C--TGC----------TCCCAGGA---G-A-ATG-TAGCA------TTA---G-GG-A---TCC------G-A--G-----C--A--------G------------G-C---G-CTT-C- 4429
SMM.US.04.G078                --G--G--------CA-C--G-----CA-C-C-CACC-G--C--A--T-----TGCT--C--TG-TTCGCAAGAA--A-A-ATG-TAGCC--------A--T--AG-A--CACC-----GG-G--A--T-----A---------------------G-A------C-TC- 4258
SMM.US.05.D215                --G--G-----G---A-T--------CA-C-C-CACC-G-----A--------TGC---C--TG--TCGCAAGA---A-AAATG-TAGCC--------A--T--AG-A---AC------GG-A--T--T--C--A-----C-----------G---G--------C-CC- 4252
SMM.US.06.FTq                 --G--G-----G---A-C--------CA-C-C-CATT-G-----A--------TGC----------TCGCAAGA---G-A-ATG-TAGCC--------A--C--AG-A---ACC-----GG-G--T--T-----A--G--C-----G---------G-A------C-TC- 4280
SMM.US.86.CFU212              --G--------G---A-T--------TA-C-C-CATT-G-----A--------TGC---C--T--TTCACAAGAA--A-AAATG-TAGC-------A-AAAT--AG----CACC------G-A-----T-----A--G--C-----------G---G--------C-CC- 4244
SMM.US.x.H9                   --GC----------CA-C--------CA-C-C-CATC-G--------T-----TGC---------ATCACAAGAA--G-AAATG-TTGC---------A--R--TG-A---ACC---R-GG-A--T--T-----A--G--C---R-G-----G---G-G-----CC-TC- 4471
SMM.US.x.PGM53                --GC----------TA-C--------CA-C-C-CATC-G--------T-----TGC---------ATCACAAGAA--G-AAATG-TTGCC--------A-----TG-----ACC--C--GG-A--T--T-----A--G--C-----------G---G-A-----CC-TC- 4915
SMM.US.x.pE660.CG7G           --G-----G-----T--C--------TA-C-C-CATC-G--------T-----TGC---------ATCACAAGAA--G-AAATG-TTGCC--------A--G--TG-A---ACC-----GG-G--T--T-----A--G--C-----------G---G-A-----CC-TC- 4987
STM.US.89.STM_37_16           --G--G--G------A-C--------T--C-C-CACC-G--C--G--T------GCT--C------TCGCAGGA---A-A-ATG-TAGCC--------------AG-A--AACC------G-A--T--------A--G--C-----------G---G-C------C-TC- 4645
SUN.GA.98.L14                 --G-----GC-CTGT-CC-TG------------CAG--------A--T-----GCCA------T-A--AAAGAT--GGA---A-TA----------TTA--C--AC----TAC-ACA--G--A--T-----C--T--G--------------G---G-A-A---C--G-T 5059
SYK.KE.x.KE51                 -----GC--A------C----------A-T---CAT--------A--------TCCA--C----T-A---A-G-CT--GC---A--------------A---G--G-A--CACCACA-----A--A--------A-----C--G--A---------AAC-AA----G-C- 4554
SYK.KE.x.SYK173_COMGNM        --CA-GC--A----CA-T---------A-----CAG-----C-----T-----ACCA--C--TGT-A---AC-A-T-C-AA--A--A---------TGT--C--AG-A--CACTACA--T-----A-----C--A-----C--G--A--------G--C-AA----GGT- 4899
TAL.CM.00.266                 --C---C--G-G--CA-C--------TA---GCCACC-G-----A--------GCCA--C--TGT-TCA--AGAAA-GCA----ATGGTA------CTCAA---AG----CTCCGCA---G----A--T-----T---TCC--G----C-------AAC-AA------C- 4665
TAL.CM.01.8023                --C--GC-GG-G--CA-C---------A--G--CATC----C--A--------GCCG--C--TGT-TCG--AGA-A-GCA----A-AGCG------CT-AA---AG----CACCACA--GG-A--A--T-----A---TCC-----GTC------GAA--A---C---C- 4163
TAN.UG.x.SIVagmTAN1           --G------A-CATTA-T--G-----TA-T-C--A-T----------------GCCA----TGTTA--CAGGA---ACAAA-TATT-----------GACA-G-AG----CAC-ACA--G--C-----T--C--A---TCA------AGT------AG--------G-C- 4898
VER.DE.x.AGM3                 T-A-----GA--CT-A-T--------CA---CCCA-C-G-----A--------ACC------T--GTC-CAAGAA--AGCA--AATG----------GAAA---AG-A--CACCACA--TG-A-----T--C--T---TCA-----CTCTA----GAG----------C- 4421
VER.KE.x.9063                 ---C-----A--TTGA-T---------A--GTCCAGT----C--A--T-----ACCA-----T--TTCCCAAGA---AGC---AAT-----------GAAAG--AG----CACCACA---G-G--A--------A--GTCA--G---TC-------AGC-----C---C- 4926
VER.KE.x.AGM155               T-G------A--AT------------CA-C-CTCACC-C-----A--T-----ACCA--------TTC-CAGGAA--AGCT--TATG--C-------GAAAGG-AG-A--CAC-ACG--GG-A--A--T-----A--GTCC--G---TCTA-----AG------C---C- 4918
VER.KE.x.TYO1_patent          --A------A--CTGA-T--------TA---C-CAGT----C--A--------GCCT--C--T---TCCCAAGAA--GGCA--AATA----------GAAA---TG-A--TAC-ACA--T--A--A--T--C------TC-------TC-A-----AGC-----C---C- 4923
WRC.CI.97.97CI14              ACC--GC-GC-G---A-T--------CA-C--G-A------------------G-CA-------AGTC-CAG-AAA-GCA---TAT---C------TTA-----AG-A--A--------G--C-----T-----T---------------A-----AG--A---C-GGAG 4999
WRC.CI.98.98CI04              --CC--C--A-C-----T--------TA-G----AG--T-----G--T-----A-CA-----T-AGTCACAG-AAA-GCA-A-TAT---C------TTA-----AG-A--A--------G--C--T--T-----T---------------A-----AGC-AA----G-AG 5014
WRC.GM.05.Pbt_05GM_X02        ACCA--C--------------------------AG--T---------T-----A-CA-------AATCACAG-AAA-GGA----ATA--C------TTA--C--AG-A-----------G--------------T--G------------A-T--GAG--AA----G--T 3928
```

```
H1B.FR.83.HXB2             AAAAACAAATT...ACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAAATCCACTTTGGAAAGGACCAGCAAAGCTCCTCTGGAAAGGTGAAGGGGCAGTAGTAATACAAGATAATAGT...GACATAAAAGTAGTGCCAAGAAGAAAAGCAAAGATCATT 5033
Pol                        Q  K  Q  I  .  T  K  I  Q  N  F  R  V  Y  Y  R  D  S  R  N  P  L  W  K  G  P  A  K  L  L  W  K  G  E  G  A  V  V  I  Q  D  N  S  .  D  I  K  V  V  P  R  R  K  A  K  I  I
H1O.CM.91.MVP5180          -----------...TT-----CA-C----------C-------A--T-----G-C--TA-------------G---C-A----G----------G----------C-----------AG-A..-----T-----G--A-----------G-----A--A--C 5062
H1N.CM.95.YBF30            ---C--------...TT----G-----------------------A---------G---CA---------------C--A--T--G-------A----------G----T--------CG-G..--T----------C---C-T--G--------A--A--- 4627
H1P.CM.06.U14788           -----------...TT----------------A---C-------A---------G-C--TA---------------G-CA----G------------------------C--------C--AG-A..--T--T--G-----C--T--G-----------A--A--- 4504
CPZ.CD.06.BF1167           -----------...TTT---C---------------C-----T-----TCA---G-C---G-G--------------CCGT----G------------------------C---A-G--GGGAGAA..--A--C--------A--------G--G--T-----T--A 5141
CPZ.TZ.00.TAN1             -----------...TT----G----------------------T-----GC---G----A------G-------GCGA--A--G----------------------------A---AGGAGAA..-----T--------A--C--G-----------A-----A 4677
GOR.CM.07.SIVgorCP2135con  -----------...TT--------------------C-------A---------G-C--TA---------------G-C------G------------------T----C--T--G--C--AG-G..-----T--------T--T--G-----G-----A--A--- 4569

MAC.US.x.239               --TC-A---AC...T-----T--A-------------C-------A--AG-----G---A---G-----G-----C-GTG----AT-G--------G-----A-----CA-CT--A-G-TAGGG-CA..-----T--G-----A--C--------G--T--A--T--C 5332
Pol                        Q  Q  S  K  .  N  S  K  F  K  N  F  R  V  Y  Y  R  E  G  R  D  Q  L  W  K  G  P  G  E  L  L  W  K  G  E  G  A  V  I  L  K  V  G  .  T  D  I  K  V  V  P  R  R  K  A  K  I
H2A.SN.86.ST_JSP4_27       --GC-A---A-...T-----T-A-------------C---T---A--AG-----G---AG--G-----------T-GGG-A--A--G----G--G--C--A-----CA--G-CA-G-TAGGGGC-..----------A--A--A--------G-----T--------C 4860
H2B.CI.x.EHO               ---C-A---A-...TT----T-C--------C-----C-------A--AG-----G---A---C-----G-----T-GTG-T--AT-G--------G-----A-----CA-C---A-G-TAGGG-CA..--A--C------A-A--C--------------------A 5386
H2G.CI.92.Abt96            ---C-A---A-...T-----T--A-------------C-------A--AG-----G-C-AG--C-----------T-GTG----TT-G--------G-----A--T--CA-C---A-G-TAGGG-CA..--A--C-----T--A--C--------------A-----A 4742
H2U.CI.07.07IC_TNP3        ---C-A---A-...T-----T-CA-------------C-------A--AG-----G-C-A---C-----G--G--C-GTG----AT-G--------G-----A-----CA-C---A-G-TAGGG-CA..--------------A-A--C------------T--A-----C 4847
H2U.FR.96.12034            --TC-A---A-...T-----T--A-------------C-------A--AG-----G-C-A-T-G--------C------G-CT-A--G-----G--A-----A--G--CC-T--TA-G-TAGGG-CA..-----T--G---A-A--C--G------------------C 4894
ASC.UG.10.RT03             ..-CCA--T--...C-----T--TCG--------------------ACAG-G-AG-GC--TGAG---CT---G------C-A-----G--------A-----T-----G-----CA-GACA-C--CAGGT--GC-GTTGAC---C----A---G--------------- 4436
ASC.UG.10.RT08             ..-CCA--T--...C-----T--TCG--------------------ACAG-G-AG-GC---GAG---CT---G------C-A-----G--------A-----T-----G-----CA-GACA-C--CAGGT--GC-GTTGAC---C----A---G--------------- 4460
ASC.UG.10.RT11             ..-CCA-----...C-----T--TCG--------------------ACAG-G-AG-GC--TGAG---CT---G------C-A-----G--------A-----T-----G-----CA-GACA-C--CAGGT--GC-GTTGAC---C--------G--------------- 4445
COL.CM.x.CGU1              C----TT-CAAAATCA----T--A--------AA--CC--C-GG-AA--GCA-...-CAGG-GAG---C---------G-G-A--GG-G---------------A--G-----C--CAG-A--TC-CAAGGTAC-T-GTTT--GAAA--T----------T--------CC 4799
COL.UG.10.BWC01            C-C--AT-CAAAAT-A----T-----------AA---G----GG-AA--GCAT...-CAGG-GAG---C---------G-C--T-GG-G--------A-----A--------C--CA-GA--C-AGAAGGCTC-T-GTTT--GAAA--TC----G--G-T---A----CA 4598
COL.UG.10.BWC07            C-C--AT-CAAAAT-A----T--A-------CAA---C----GG-A---GCA-...-CAGG-GAG---C----G----GCC--T-GG-G--------A-----A--------C--CAG-AGC-CAGAAGGTAC-T-GTTT--GAAA-----G-----G-T---A----G- 4826
DEB.CM.04.SIVdeb04CMPF3061 --C--A------...T-----T--TC-GGA--CAAA--C--C-----A--AG-AGCTG----CAG----------C--------CA-----G--------A-----A-----G-----CA--TTGG-C-CAGGG--TT----------T--C---------------AG-A--A 5023
DEB.CM.99.CM40             --C--A------...T-----C--TT-GGC--CAA---C--------A--AG-TGCTG-----AC----------G-------CC--G--G--------A-----T-----G--GTGCA-GACAG-GGTAGGA--T--C-----G--C--T-----G---------AG-A--C 4842
DEB.CM.99.CM5              --C--A------...T-----C--TT-GG--ACAA---A--------A--AG-AGCTG-C---AC-------G-----C--T-CCT-A--G--------A-----A--G--G--GTGTA--ACAG-GGTAGGA--------G-----A----------------C---G-G--A 4836
DEN.CD.x.CD1_CM0580407     ..-CCA--CAA...C-----T--TCG------AA---C--------A--GG-AGCT-----G-AA---C-----------TG----G--T---------------A-----T--C--CA-GACAC-GGCAGGA-----CTTCCCG--A--T--------G-----A-----C 4920
DRL.DE.11.D3               --C--A-T----...CTC----------------------------A--GG-A---G---AGT-G--------------G-A--TG-------G--G-----A-----G--C--CA-G--AGG--C-...--T-G------A-A--T------------------------ 4507
DRL.x.x.FAO                --C--A-T----...CTC----------------------------A--GG-A---G---AGT-G--------------G-A--TG-----------G-----A--G--G--C--CA-G--AGG--C-...--T-G------A-A--T------------------------ 4600
GRV.ET.x.GRI_677_gri_1     ----C-----C...CA-------TTG------AA---C--C-----A--AG-T---G------G-G--------------GCGA---A-----------A-----C--G--G-----TA----GGGGGAA..------C--G-----C--C--G-----G--T-----T--C 5206
GSN.CM.99.CN166            -C-C-----A-...T-C---T-------A------------------C-AG-AGCT--C--T-----C----------GTA-----------------------A--------G-C---AC-C-AGCAGGT--G--T-TTAC---C--T-----G--------------C 4764
GSN.CM.99.CN71             -C-C-----G-...T-C---T-------A------------------C-AG-AGCT-----T-A---C----------GTA--T-------G---------A---------GG-C---ACCC-AGCAGGT--G--C-TCAC---C--T-----G--------------C 4746
LST.CD.88.SIVlhoest447     ------T--A-...T--------TT---A---AA------C-----AC-AG-AC--G----C-AG-----G-----T---C----T--G--------A-----T--T-----GG--A----AGGAGAC...AC---CTTCTC---C--C-----------T---C-AG-A 4310
LST.CD.88.SIVlhoest485     ------T--A-...T--------TT---A---AA------C-----AC-AG-AC--G----T-AG-----G-----T---C-AA-T--G--------A-----T--T-----GG--A----AGGAGA-...AC---CTTCTC---C--C-----------T---C-AG-A 4310
LST.CD.88.SIVlhoest524     -----TT--A-...T--------TT---A---AA------C-----AC-AG-A---G------AG---G------T--TC-A--T--G--------A-----T--T-----GG--A-G--GGGAGAG...A----TTTCTCTA-C--C--G--G--------AT-AG-G 4307
LST.KE.x.lho7              -----TT--A-...T--------TT---A---AA-------------AC-AG-A---G----T-AG----------------GC-A--T--G--------A-----T--T-----GG--A----AGGGGAG...A----CTTCTC---C--C---------------AC-AG-A 5394
MAC.US.x.251_1A11          --TC-A---AC...T-----T--A-------------C-------A--AG-----G---A---G-----G-----C-GTG----AT-G--------G-----A-----CA-CT--A-G-TAGGG-CA..-----T--G-----A--C--------G--T--A--T--C 5331
MND-1.GA.x.MNDGB1          ----T-------...T------A-TT-------AA------C-T---A--AG-A---G--AG--G-----------T-GT-TC--TT-G--------A---------------TT--A--T--C-AGAA..--G-----GA---A--T--------GTGT--A--A--A 4720
MND-2.CM.98.CM16           --C-CA-T---...CA------------------------------A--AG-A--GG-C-A--AG---------------G-A---A-T--------A-----A--G--G-----CA----GGGG-CA..---T-G--G-----A-----G-----------A-----C 5034
MND-2.GA.x.M14             --C--A-T----...T-------------------------------A--AG-A--GG-G-AG--C---------------G-A---A-T--------A-----A-----G-----CA----AGGA-CA..---T-------T-----C-----------T--A--A--C 4961
MND-2.x.x.5440             G-C--A-T----...T----------------A-------------A--AG-A--GG-C-A--AG---------------G-----A-T--------A-----C--T--G-----CA----AGGAGT-...--TT----G--G--T--T--------------A--A--- 4594
MNE.US.x.MNE027            --TC-A---AC...T-----T--A-------------C-------A--AG-----G---A---G-----G-----C-GTG----AT-G--------G-----A-----CA-CT--A-G-TAGGG-CA..------T--G-----A--C--------G--T-----T--C 4800
MON.CM.99.L1_99CML1        -C--C-------...------T--TC---A-------------T----ACAG-TGCCG-C--T-AG--------------GCGC--A-----------------T--C-----C---A--AC-G-GGAGGGC--T--CCTGAC---C--T--G-----G---------A--C 4739
MON.NG.x.NG1               -C--C-------...------T--TCG--A-------------T---AACAG--GCTG----TTCG-----G---------C-T--T-------G--G-----A--TT-G--G--TA-GACAG--CAAGGT--GG---T-AC---A---C-C--G--G------------ 3332
MUS-1.CM.01.CM1239         --TC-----AA...T-C---T---------------------T---ACAG--TCTG-C---TC------G-----T--TGCCT-GT-G-----G---------A---------CG-CA--ACAG-ACAAGGAC-GG---TCACT--A--T--------G-----A--A--C 4716
MUS-1.CM.01.SIVmus_01CM1085 --C------AA...TAC---T---------------------T---ACAG-TTCTG----GTC------G-----C--CGTCT-A--G----G---------A--------C--CA--ACAG-CTCAGGA--GG---TCAC---T--T--------G-----A--A--C 4726
MUS-2.CM.01.CM1246         ----T--C-A-...------T--TC---A-----------------AAGAG-AGC---C--TTA----CTGA-------TCGA--T--T--------A-----A--TC----G--CA--ACA--GGAAGGT--G--CGT-AC---C-----------------A--A--C 4824
MUS-2.CM.01.CM2500         ----CAC--AA........T--TC---A-----------------AAGAG-AGCG------AG----C--------T--TC-A-----------G---------A---A-------CA--ACAG-CCAAGGT--G--T-T-AC---A-----G--G--G--------T--C 4777
MUS-3.GA.09.09GabOI81      -CTC-----A-...------T----------------C--------AC-AG--TCCG----TT-G-----G---------G-CT--G--G-----G--A-----A--C-----C--CA--AC-G-CTTAGGT--GG-----AC---T--T--G--------C--A--Y--C 4292
MUS-3.GA.11.11GabPts02     ----C-------...T-C---T-----------------------AC-AG-AG-CG-C--G--G--------T--T--GGC------T--------A-----T-----G--C--CA--ACAG-CTTAGGA--A---CTCAC---C--T--G--G-----T---G-A--C 4313
OLC.CI.97.97CI12           C-C--A-T-AA...TTT----AATTTTCAGGATATCGAGTG--TTATA-A-A--A-G-AGGC-AG---TGT--C-------G-A-----G--------A-----A-----CC-GC-TA----AGGAGA-...A-GTATTTCTCT-----TCTG--G---------A--AG-G 4901
RCM.CM.00.SIVagi_00CM312   ----T-------...T---------------------------TT--A--AG---GG---AG-AG-----G-----G--C-CT---A----------A-----A-----G--G--C--G---GGACAG...--TC-G--G-----C--T--------GTGT--A--TG-A 4304
RCM.CM.02.SIVrcm_02CM8081  ----T-------...T---------------------------T---A--AG-A--GG---A--AG-----------T--C-CT--AA----------A-----A--G--------C--G--CGGACAA...--TT----G-----C-----------GTGT--A-----A 4322
RCM.GA.x.SIVRCMGAB1        ----T-------...T------------------------------A--AG-A--GG---A---G----G--T-----T----GA-----------A-----A--------C--C--G--GG-G-C-GGA---T-G--G-----C--C--G-----------A-----A 4515
RCM.NG.x.NG411             ----C-------...T---------------------------T---A--AG-A--GG---AG-AG-----G--------T--A--A-----G---A---------T--C--G--C--G---C-AGAG...--TC----G-----C--T--------GTGT--------A 4527
SAB.SN.x.SAB1              --C--A------...T--------------------C----------AG-----G-C--TG-G--------T--G--T------A----------A-----A--------T--------AC-GG--...--AT-G---AC-A-C--------------C--A--T--A 5431
SMM.CI.79.SIVsmCI2         ---CCA---AC...T-----T-CA-------------C-------A--AG-----G---A-T-G-----G--C--G--T--C--T--T--------A-----A-----CA-T--TA-G-TAGGG-C-...-----T--G---A-A--T--------G-----------C 4452
SMM.LR.89.SIVsmLIB1        --TC-A---A-...T-----T--A-------------C-------A--AG-----G---A---C------------C-GTG----GT-G--------G-----A-----CA-CC--A-G-TAGGG-CA..--T--T--------A--T--G-----G-----A-----C 4410
SMM.SL.92.SIVsmSL92A       ---C-A---AC...T-----T-CA----A---------C-------A--AG-----G-C-A---C-----G--T--T-GTG-----T-G--------G-----A-----CA-----A-G-TAGGATCA..--T----------A-A--T-----G--------------A 4412
SMM.SL.92.SIVsmSL92B       CTTC-----A-...T-----T------------------------A--AG-A---G-C-A---G-----G-----R---G-----------------A---------T--T-----TA-G-TAGGG-CA..--A--------G--A-----------------A-----C 4392
SMM.SL.92.SL92B            -TTC-----A-...T-----T------------------------A--AG-A---G-C-A---G-----G-----------G-----------------A---------T--T-----TA-G-TAGGG-CA..--A--------G--A-----------------A-----C 4763
SMM.US.04.G078             --TC-A---A-...T-----T--A-------------C----T---A--AG-----G-C-AG--G-----------C-GTG----AT-G--------G-----A-----CA-CT--A-G-TAGGG-CA..--G--C--G-----A--C-----G-----T--A--T--C 4592
SMM.US.05.D215             --TC-A---A-...T-----T--A-------------C-------A--AG-A---G---AG---------------C-GTG----AT-G--------G-----A-----CA-CT--A-G-TAGGG-CA..---------G-----A--C-----------T------TG-- 4586
SMM.US.06.FTq              --TC-A---A-...T-----T--A-------------C-------A--GG-----G-C-A---G-----------C-GTG----AT-G--------G-----A-----CA-C---A-G-TAGGG-CA..-----T--------A-----G-------------A--T--C 4614
SMM.US.86.CFU212           --TCCA---A-...T-----T--A-------------C-------A--AG-----G---AG--G-----------C-GTG-A--AT-G--------G-----A-----CA-CT--A-G-TAGGG-CA..--T-----G---------C------------T--------A 4578
SMM.US.x.H9                --TC-A---A-...T-----T--A-------------C-------A--AG-----G-C-AG--G-----------C-GTG----AT-G--------G-----A-----CA-CC--A-G-TAGGG-CA..--G--C--G-----A-----G--G-----T--A--T--C 4805
SMM.US.x.PGM53             --TC-A---A-...T-----T--A-------------C-------A--AG-----G-C-AG--G-----------C-GTG----AT-G--------G-----A-----CA-CT--A-G-TAGGG-CA..--G--C--G-----A--------G-----T--A--T--C 5249
SMM.US.x.pE660.CG7G        --TC-A---A-...T-----T--A-------------C-------A--AG-----G-C-AG--G-----------C-GTG----AT-G--------G-----A-----CA-CC--A-G-TAGGG-CA..--G--C--G-----A-----G--G-----T--A--T--C 5321
STM.US.89.STM_37_16        --TC-A---A-...T-----T-CA-------------C----------AG-----G---A---A-----G-----C-GTG----AT-G--------G-----A-----CA-CG--A-G-TAGGG-CA..-----T--------A-----G-----G--T--A-----C 4979
SUN.GA.98.L14              -G---AT--A-...T--------TT---A---AA---C--C-----A--AG-A---G----T-AG------------T--T--A--T--G--------G-----A---------C--CA--C-GGG-GAG...A----CCTG-----A--T--G------------AC-AG-A 5393
SYK.KE.x.KE51              -C-C-.........CA----T--TCG------AAA--A-----T--AC-TG-T-CT...T-TGA----C-G---------TTCT--A--G-----G--A-----A---------G--C--GACACCAGA-CAAC-GG---T--C---A--T-----G---------A-----C 4882
SYK.KE.x.SYK173_COMGNM     C--C-TC-......C-----T--TCG------A-------C---C-CC-GG-G-AG...AACGAG-----G---------TCGT--A--T----------------T---------GG-G---ACAG-GGAGGGA-----TTTT-C---A-----------G-----A--A--- 5230
TAL.CM.00.266              ----C-------...CA---TT--TCGG-----AA---C--------AA-GG-TGCT-----T------C-G---------TC-C---A-T--------A-----A--G--T--CT-G-G-ACAG-CGAAGGA--GG---TCAC---C--T-----G--G-----A--A--- 5002
TAL.CM.01.8023             ----C-------...CAC--TT--TC-G-----AA---G--C--T--AA-GG-TGCT-----T------C-G---------TC-C---G-G--------A-----A-----C--CC--AGGACAG-CGAAGGA--GG---TTAC---T--C-----G--G--T--A--A--C 4500
TAN.UG.x.SIVagmTAN1        ---C----CAA...CA-------TC-------AA---C--C-----A--AG-A---G-G---G-G---------------C-CT-----------G--A-----A--G-----C---A----AGG-GAG...--GT-G---------A-----G-----G-----A--T--C 5232
VER.DE.x.AGM3              ---CCA------...CA-------TT-------A-A--C--C-----A--AG-G---G----TG-C-------------G-C--T-AA-T--------G-----T-----G--C--CA---GAGG-GTG...--AT------A-TAC---------G---------A--T--A 4755
VER.KE.x.9063              ---CCA------...CA-------TT-------A-A--C--C-----A--AG-G---G-----G-G-----G----------C-AT-AA----------G-----T-----G--C--CA---AGGAGAA...--GC-G---A-T--C-----------------A-----C 5260
VER.KE.x.AGM155            ---CCA----C...CA-------TTG------A-A--C--C-----A--AG-----G-----G-G-----G-----T--TCGC--GA----------A-----C--G-----TC-CA-G--AGG-GAA...--AC-G--G-----T--G-----G---------A-----A 5252
VER.KE.x.TYO1_patent       ---CCA------...CA-------TT-------A-A--C--C-----A--AG-G---G-C--TG-G-----------------C-AT-AA----------G-----A-----G--CC-CA-G...G-CG-AAGT---C----G--T--A--------G-----T--A--T--- 5257
WRC.CI.97.97CI14           --C------AA...TT----T--A---------A---G--------AACAG-AGCC----AG-AG---C----------GT--------CT--------A-----A---T--------G--ACACCGGAA...-GT----T-AC---T--G----------------C-T--A 5330
WRC.CI.98.98CI04           --C------AA...TT----T--A---------A-------------AACAG-AGCC----AG-AG---C----T----GT--------T---------A-----A--CT-----G-GG--ACCCCAGAA...-GT----T-AC---T--T-A---G---------A--A--A 5345
WRC.GM.05.Pbt_05GM_X02     --C------AA...TTC---T--A---------A------C------ACAG-AGCCTCG-AG-AG---C---------GT--AT-GG------------G--A--TT----TG-G--GAC-CC-GAAGGA--TT--CTTAC---T--C--------------A-----A 4262
```

```
                              Vif start                                                                     Pol, Gag-Pol, and p31 integrase end
H1B.FR.83.HXB2              AGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGAT.........................................TAG.............................................AACATGG....... 5103
Vif                                M  E  N  R  W  Q  V  M  I  V  W  Q  V  D  R  M  R  I_ .__.__.__.__.__.__.__.__.__.__.__.__.__.R__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__T__W__.__.__
Pol                         _R__D__Y__G__K__Q__M__A__G__D__D__C__V__A__S__R__Q__D__E__D__.__.__.__.__.__.__.__.__.__.__.__.__.*
H1O.CM.91.MVP5180           --A-----------------------AC---A--A------A-------ACA--AAG-GAAAGCATGGAACAGCCTGGTGAAATACCA.............--A............................................................. 5155
H1N.CM.95.YBF30             -------------------------------G-------------G----------AA--CAGGAAATGGAA.............................---............................................................. 4702
H1P.CM.06.U14788            --AA----------------------------------------GA--CC---AGA--AAG-GAAAGCCTGGAACAGTCTGGT..................--A............................................................. 4588
CPZ.CD.06.BF1167            -A---------------G---------CG-G-G--------GA-------A--------...........................................---............................................................. 5204
CPZ.TZ.00.TAN1              -AA--G---------------------CA-G-G--A---ATGA-------A-----AC-...........................................---............................................................. 4740
GOR.CM.07.SIVgorCP2135con   ------------------------------------------GA--CC----GA---AG-GAAAGCCTGGAACAGTCTAGT..................--A............................................................. 4653
                              Vif start                                                                     Pol, Gag-Pol, and p31 Integrase end
MAC.US.x.239                -AA---------GG-GGA-AA-AG-TG---AGCA--TCCCAC-TGGAGG-TACC-GA--GGCTAGAGAGGTGGCA...........................---............................................................. 5410
Pol                         _I__K__D__Y__G__G__G__K__E__V__D__S__S__S__H__M__E__D__T__G__E__A__R__E__V__.__.__.__.__.__.__.__.__.__.__.A_*_
Vif                                M__E__E__E__K__R__W__I__A__V__P__T__W__R__I__P__E__R__L__E__R__W__H__.__.__.__.__.__.__.__.__.__.__.S__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__
H2A.SN.86.ST_JSP4_27        -AA--C------GG-AG-CAA-AGATG---AGCG--TCCAACTTGGAGGGT-CCAG---GGATGGAGAGGTGGCA...........................---............................................................. 4938
H2B.CI.x.EHO                ---A-C------GG-GGA-AA-A-TTG---TGCA---CC-ACGTGGAGG-TACCAT-C-GGCTAGAGAGGTGGCACAGTCTAAT..................--A............................................................. 5473
H2G.CI.92.Abt96             -AA---------GG-GGA-AA-AG-TG---AGCA--ACCAAC-TGGAGG-TACCAG-C-GACTGGAGAAGTGGCA...........................---............................................................. 4820
H2U.CI.07.07IC_TNP3         -AA---------GG-GGA-AA-AG-TG---AG-G--TCCAACCTGGAGGGT-CCAG-C-GGATGGAGAGATGGCACAGCTTGGTGAAATACCT.............--A............................................................. 4943
H2U.FR.96.12034             --A---------GG-GGA-AA-A-TTG---AG-A--TCCAACTTGGAGGGT-CC---A-GGTTAGAGAGATGGCACTCCCTGATCAAACACCT..............---............................................................. 4987
ASC.UG.10.RT03              -A-CCC-----CGCCAGCCAAAAT-TG-G-A--AAG-A--GT-TGGATAGCCC-AGAA-GGAGGATACAAGCCTGGAT........................---............................................................. 4517
ASC.UG.10.RT08              -A-CC------CGCCA-CCAAAAT-TG-G----ACG-A-AAT-TGGATAGCCCCAGAA-GGAGGATACAAGCCTGGAT........................---............................................................. 4541
ASC.UG.10.RT11              -A-CC------CGCCAGCCAAAGT-TG-G----AAG-A-AAT-TGGATAGCCC-AGAA-GGAGGATACAAGCCTGGAT........................---............................................................. 4526
COL.CM.x.CGU1               --ACTCC-ATATGG-G-AGAT-TG--CAG---GAA-T--TT-TCC-ATGG-C-AA----AGCCGAAACCGTTAAGGGAATGGAT..................--A............................................................. 4886
COL.UG.10.BWC01             -AA-TGA--TATGG-G-A-GAATG-TCAG-A-GGC-T--TT-CCC-ATG--C-A--A--AGCAGGAGAAGTTAAGGAAGTGGAT..................--A............................................................. 4685
COL.UG.10.BWC07             --AAT-AC-CATGG-ACAGACCA---A-TG-TCA---A--ATCTGCT--CCA---G-A-GGAAAAACCAGGAGAAACTGAGACTTTGGAT............--A............................................................. 4919
DEB.CM.04.SIVdeb04CMPF3061  -A---C---------G-TG---GTA--A-GAT-GACCCACA-GA--CC--T--A................................................--A.......................................................GTTAG 5085
DEB.CM.99.CM40              -AA--A---------G-C----ATA--A-G-T-AACACACA--A-TCCA-T--A................................................--A.......................................................ATTAG 4904
DEB.CM.99.CM5               -AA------------GGAG---ATA--A-GAC-GACACACA--A-CCCAGT--A................................................--A.......................................................ATTAG 4898
DEN.CD.x.CD1_CM0580407      -A-CCA---AATGC-G-A...T-TA-GA-G-GAGAACA-AAC-AAG--AT----AGT--GACTGACATA.................................--A.......................................GATTGGGGA-ACCAAATTAG 5009
DRL.DE.11.D3                -AA------------A-TG---ATA--C--ACCAACA---AG---.........................................................-GA...........................................GAGGTACAT-GTAGAAA 4572
DRL.x.x.FAO                 -AA------------A-TG---ATA--C--ACCAACA---AG---.........................................................-GA...........................................GAGGTACAT-GTAGAAA 4665
GRV.ET.x.GRI_677_gri_1      -AA---------G-GAGA-AAA--ATG---AG-GAG-GTAGT-TGGAGGGT-TCAGA--GGCAAATAAGCAGATGGAGGGGGATAGTGACTTACAAGATCAGGAA--A............................................................. 5314
GSN.CM.99.CN166             -A-CCA------G-GGCA-A-A---AA----TGG---GT----CCCATACAA--A-T--C..........................................-GA.......................................................GAGGA 4832
GSN.CM.99.CN71              -A-CCA------G-GGCA-A-----AA----TGG--AGT---GCCCATACAAG-A-T--CAGAGAGGAAGGTAGAATGGCTGAT..................---............................................................. 4833
LST.CD.88.SIVlhoest447      -AA---------G-GGGACCAAAG-A-AG---GG---GCCT-CA--ACA-C...................................................---..........................................................GA 4366
LST.CD.88.SIVlhoest485      -AA---------G-GGGACCAAAG-A-AG---GG--AGCCTGGA--ACA-C...................................................---..........................................................GA 4366
LST.CD.88.SIVlhoest524      -AA---------G-GGGACCAAAG-A-AG-C-GA--AGCATGGA--AC--A...................................................---..........................................................GA 4363
LST.KE.x.lho7               -AA---------G-GGGACCAAAG-A-AG---GA--AGCCTGGA--ACA-C...................................................---..........................................................AT 5450
MAC.US.x.251_1A11           -AA---------GG-GGA-AA-AGATG---AGCA--TCCCAC-TGGAGG-TACC-GA--GGCTAGAGAGGTGGCA...........................---............................................................. 5409
MND-1.GA.x.MNDGB1           -AA---------G-G...-GT-G-AAGA--AG-CAG--TAACTTGGA-AGT-TC................................................---.......................................TCAGAG--T-GAAAAGTGGC 4794
MND-2.CM.98.CM16            --A------------ACAG---ATA----CCCAAC----A-GC-..........................................................-GA...........................................TAG-CATGCAGTAGAAA 5099
MND-2.GA.x.M14              ---------------ACAG---ATA--A--CCCCACA---AGGA---------GACA-C-..........................................-GA..........................................................AA 5026
MND-2.x.x.5440              -A---C---------ACAG---ATA-CA--CCCCACA---G---A.........................................................---.......................................ACAGGACGTCAA-TTGAGAG 4662
MNE.US.x.MNE027             -AA---------GG-GGA-AA-AG-TG---AGCA--TCCCAC-TGGAGG-TACC-GA--GGCTAGAGAGGTGGCA...........................---............................................................. 4878
MON.CM.99.L1_99CML1         -AACCC-----C-CCAC-GAAAAT-TG-G--G-GACACCAACCAATATA-TCTCAGAA-GCAAGATGGATTGGCTAAT........................---............................................................. 4820
MON.NG.x.NG1                C-CCC------CTCCA--GAA-AT-TG-G-AG-AAGCCCAGTGCACAT---ATCAGA--GGTGGATGGAATGGCTGAT........................---............................................................. 3413
MUS-1.CM.01.CM1239          -A-CCA-----CGCCA--GAAAAT-TG-G-AG-AAGTCCAAT-CAG--G-CC-CAGAA-GGAAGATGGATTGGATAAT........................---............................................................. 4797
MUS-1.CM.01.SIVmus_01CM1085 -AACCC-----CGCCA--GAA-AT-TG-G-AG-AAGTCC-AT-CAG--G-CCTCAGAA-GGAAGATAGATTGGATAAT........................---............................................................. 4807
MUS-2.CM.01.CM1246          -A---------C-CCAG-CAAAAT-TG---AG-GAGCCC-ATTCAGT---T-TCAGAA-GGAAGATGGATTGGCTGAT........................---............................................................. 4905
MUS-2.CM.01.CM2500          -AA--C-----C-C-A--GAAAAT-TG---AG-GACCCCAATCCAGTTT-T--CAGAA-GGAGGATGGACTGGCTGAT........................---............................................................. 4858
MUS-3.GA.09.09GabOI81       -AACCA--CAATGCC----AAAAT-TG-G-AG-AAGTCCAAT-CAG--G-CATCAGA--GGAAGATGGACTGGTTGATCAGGGCAAC...............--A............................................................. 4382
MUS-3.GA.11.11GabPts02      -A-CCG-----CCCCA---AA-AT-TG-G--G-ATCCCC-ATTCAG-TG-C--C-GAA-GACACCTGGATTGGTTGAT........................---............................................................. 4394
OLC.CI.97.97CI12            -A-CC-------GTGG-T-GCAGT-CCA---TA.....................................................................---.........................................GCT-AATAATGGGC 4951
RCM.CM.00.SIVagi_00CM312    -AA--------C-G-A-A-AT-TG-A-AG---GACCAATATGGAGG-TGGAC-G-----AAAT.......................................-GA..........................................................AC 4372
RCM.CM.02.SIVrcm_02CM8081   -AA--------C-G-A-AGAT-TG-A-AG---GACCAGTATGGAGG-TAGAC-G---A-AGAT.......................................-GA..........................................................AC 4390
RCM.GA.x.SIVRCMGAB1         -AA--A-----C-G-A-AGAT-TG-A-AG---GGCCAATTTGGCGG-TAGAC-G-----AAAT.......................................-GA..........................................................AC 4583
RCM.NG.x.NG411              -AA--------C-G-A-AGA--TG-A-AGA--GACCAATATGGAAG-TAGAC-G-----AAGT.......................................-GA..........................................................GC 4595
SAB.SN.x.SAB1               -AA------------.......---TTG---AG-CAG-CCC-TCTGGA-GGTA-C-G-CGGACAGCAGGAGAGGTGGAC.......................---............................................................. 5506
SMM.CI.79.SIVsmCI2          -A----------GG-GGA-AA-AGCTG-G-A-CA--CCCTACTTGGAGA-TCCC-----GGATGGAAAGATGGCACAGCCTGAT..................--A............................................................. 4539
SMM.LR.89.SIVsmLIB1         -A----------GG-GGA-AA-AG-TG---AG-G--TCCCACTTGGA-G-TACCAG---GATTAGAGAGATGGCA...........................---............................................................. 4488
SMM.SL.92.SIVsmSL92A        -AA---------GG-GGA-AA-AGATG---AGCA---CC-AC-TGGA-A-TACCACAC-GGTTAGAGAGACGGCACAGCTTAGTGAAATATCTAAAGTA......--A............................................................. 4514
SMM.SL.92.SIVsmSL92B        -AA---------GG-AGACAA-AGATG-G-AGCA---CCAGT-TGGAGG-TC-G................................................---......................................--G-ATAGTGGAGC 4463
SMM.SL.92.SL92B             -AA---------GG-AG-CAA-AGATG-G-AGCA---CCAGT-TGGAGG-TC-G................................................---......................................--G-ATAGTGGAGC 4834
SMM.US.04.G078              -AA-------C-GG-GGA-AA-AG-TG-G-AG-G--TCCCACCTGGAGG-TCCC-G---GGCTAGAGAGATGGCA...........................---............................................................. 4670
SMM.US.05.D215              -AA--C----C-GG-GGA-AA-AG-TG---AG-A--TCCCACCTGGAGG-TACCAG---GACTGGAGAAGTGGCA...........................---............................................................. 4664
SMM.US.06.FTq               -AA--C----C-GG-GGA-AA-A-TTG---AG-G--ACCCACGTGGAGG-TACCAG---GGCTGGAGAGATGGCA...........................---............................................................. 4692
SMM.US.86.CFU212            -AA---------GG-GGA-AA-AG-TG---AG-G--TCCCACTTGGAGG-TACCAG---GGCTGGAGAAATGGCA...........................---............................................................. 4656
SMM.US.x.H9                 -AA--C------GG-GGA-AA-A-TTG---AG-G--TCCCACTTGGAGG-TACC-G---GGCTAGAGAAGTGGCA...........................---............................................................. 4883
SMM.US.x.PGM53              -AA---------GG-GGA-AA-A-TTG---AG-G--CCCCACTTGGA-G-TACC-G---GGCTAGAGAAGTGGCA...........................---............................................................. 5327
SMM.US.x.pE660.CG7G         -AA--C------GG-GGA-AA-A-TTG---AG-G--TCCCACTTGGAGG-TACC-GA--GGCTAGAGAGGTGGCA...........................---............................................................. 5399
STM.US.89.STM_37_16         -AA---------GG-GGA-AA-AG-TG---AG-G--TCCCACCTGGAGG-TACCAG---GGCTGGAGAGGTGGCA...........................---............................................................. 5057
SUN.GA.98.L14               -A---C------GG-G-ATCCT-CTCA-TG--GATG----G-............................................................-GAAAGAGAATGGGAAGATAGAAGACAGTATAAGATAGTTAGAATAGTATGGCTTATAGATAGAA 5500
SYK.KE.x.KE51               --CAG-G-----G--AGAG---ATA---G-AC-CACC---A--T--C-TCCA-ATCAA--..........................................-GA............................................................. 4945
SYK.KE.x.SYK173_COMGNM      -CA--CC-----G--AGA----ATA---G-TCCCAC----AG-A-GAC-CCA-GACA---..........................................-GA............................................................. 5293
TAL.CM.00.266               -A-CCC------C--GCTT---GTAACA-GAT-GACC---A----TCCA----AC-----GCAGAGATGGGTAGGGACAAT.....................---............................................................. 5086
TAL.CM.01.8023              -AACCC------C--GCT----GTAACA-GAC-GACT---A-G--TCCA----AC-----GCAGAGATGGGTAGGGACAAT.....................---............................................................. 4584
TAN.UG.x.SIVagmTAN1         -AA--A------G-GAGA-AAA-T-TG-G---CAAGACTAAC-TGGAGGGT...................................................---.................................................TGGAGAGC 5294
VER.DE.x.AGM3               -A--------A-CC-AGA-AAAG-ATG-G----GAGAGTAACTTGGA-GGT-CC-GA-GAGCTGATAAC.................................--A............................................................. 4827
VER.KE.x.9063               -AA-------A-CC-AGA-AAAG-ATG-G----GAG-GTAAC-TGGA-GGTCTCAG--GACCTGATAACCAAATGGCAAGGAATAGTCAAATATTGGATGAC....--A............................................................. 5365
VER.KE.x.AGM155             -AA--C----AGCC-AGA-AAA--TTG-G----GAGACTCACCTGGA-GGT-CA-GA-GAAGTGATCACCAAATGGCAGGG...GATAGT.............---............................................................. 5342
VER.KE.x.TYO1_patent        -A--------A-CCCA-ACAAAG--TG-G-A--GAG-GT-ACGTGGA-GGTACCAG--GATCTGATAAC.................................--A............................................................. 5329
WRC.CI.97.97CI14            -AA-T--GGAATGG-G-AGGAATG-A-AGGAG--CCTCT-AT-AG-AT......................................................---.......................................................TAAAG 5386
WRC.CI.98.98CI04            -A--TC-GGAATGG-G-AGGAATG-A-AGGAG--CCTCTAAT-AAGAT......................................................---.......................................................TAAAG 5401
WRC.GM.05.Pbt_05GM_X02      -A-AC--GG-ATGG-A-AGGAATG-A............................................................................---..............................CAGTACCACTAATT-A-ATCTCAAAAG 4318
```

```
H1B.FR.83.HXB2              ........AAAAGTTTAGTAAAACACCATATGTATGTTTCAGGGAAAGCTAGGGGATGGTTTTATAGACATCACTATGAAAGCCCTCATCCA...AGAATAAGTTCAGAAGTACACATCCCACTAGGG..................GATGCTAGATTGGTAATAACAACA 5244
Vif                         .__.__.__K__S__L__V__K__H__H__M__Y__V__S__G__K__A__R__G__W__F__Y__R__H__H__Y__E__S__P__H__P__.__R__I__S__S__E__V__H__I__P__L__G__.__.__.__.__.__.__D__A__R__L__V__I__T__T_
H1O.CM.91.MVP5180           ............................A--CA-G--TAA---G--CGC-AAC---CG------G-----T------TC-AGGA-----...-A-G-C-----G-CG--GT-T--T---G---CA....................--A---GATA-A--GG-C--C--- 5273
H1N.CM.95.YBF30             .............C----------T----------G---AAA--G--A-AA-------A-----------T-------CA-A---C---...-A-----------------T------G----T..................C-G--A-----A--G-C-GTC--T 4838
H1P.CM.06.U14788            ......................-T-T--C-AA---AGG--TAAA--GA-A--AAAG----------G-----T-----G-CAA-C-----T...--G--T-------C---TT-T--T---G----A..................ACA--A-CCA-TA-TG-G--T--Y 4715
CPZ.CD.06.BF1167            .AGCTTGG--C--CC---------------T---AAGAGTAAAC-TTG--AA-A----AAA----------T-----T-TAGAC--C--C...----AGG-AGG------G--T--A--CA--ACT..............GAGA-A--A-A-------------T- 5355
CPZ.TZ.00.TAN1              .AACATGG-C---CC----T-----T-----C-T-ACAA-CAAATGCTG--AA-AT---AAG----------T-------CTGA-ACA---...-A--G-GCAGGG--A-------A--T---ACA..............GAAAGAT-A-A---A--GG-TTT-CAT 4891
GOR.CM.07.SIVgorCP2135con   ......................-T------AA---AGAA-TAAAC----A-AACAG-----C-----G-----------GCAGTCA----T...--GG--------GG--GT----A---G----C..................CCC---GA-A-TA--G-G--C--- 4780


MAC.US.x.239                ..........CCTCA--AA-T-T-TGA-ATATA-AAC-AA--ATCT-CAA-A--TT--C-A-GTGCCC---TTTA-G-TCG-ATGGGCATGG...T-G-CCT-CAGCAG----AT-T--------CA-...............GAA-GAAGCCAT--A-A-G--CA-GGG 5552
Vif                         .__.__.__.__L__I__K__Y__L__K__Y__K__T__K__D__L__Q__K__V__C__Y__V__P__H__F__K__V__G__W__A__W__.__W__T__C__S__R__V__I__F__P__L__Q__.__.__.__.__.__E__G__S__H__L__E__V__Q__G_

H2A.SN.86.ST_JSP4_27        ..........CCT-A-CAAGT-T-TAA-ATACAGAACAGG--ATCT--AG-A--TG--C-A-GT-CCC--C--TA-G-TGG-ATGGGCGTGG...T-G-CTT-CAGCAGG---ATAT-----T--AAA...............GGA--AAG-CATC---AG---CAGG-- 5080
H2B.CI.x.EHO                ...................-T-C-TGA-ATATAGGACAAA--ACTTGCAACA--TC-CT-A-GT-CCT--C--TA-G-T-G-ATGGGC-TGG...T-G-CTT-CAGTAG----ATAT-T--C--GAAA...............GAA-GA--ACATC-A-A-G-CCA-GG- 5606
H2G.CI.92.Abt96             ..........TCTGA--AA-TTCTTAA-GT-TA-AACAAA--ACCT-CAG-A--CGGTT-A-GT-CCT------A-A-TTG-ATGGGC-TGG...T-G-CTT-CAGCAGG---ATAT-----T--ACA...............AAA--A--ACAT----A----CA-GGG 4962
H2U.CI.07.07IC_TNP3         ............................GCATAGAACAAAG-AATTGAG--A--T---T-A-GTACCC------A-A-T-G-ATGGGCCTGG...T-G-CTT-CAGCAGG--ATAT-T--CT-G-AA...............GGA--AT-ACAT--A-A----CAGG-C 5067
H2U.FR.96.12034             ..........................G-GTATA-CACAAA--A-CT-CAG-AA-TG--T-A-GTGCCC------A-G-T-G-ATGGGC-TGG...T-G-CTT-CAGCAGG--ATAT-T--T---AAA...............AAA-GAT-ACAT----AGG-GCA-G-C 5113
ASC.UG.10.RT03              ..........TCTAACTAAGT-T-T-A-GTATA-AACAAAG-AATT--AGGAA-T-CAA-GGATA...-----T--CC--TTATA-GCA-ATAGGTAT-C-CAAAACA---C-AT---------TACCTTG.............AAA-TAT-ACAGGATTATG-GGA-G-- 4662
ASC.UG.10.RT08              ..........TCTAAC-AAGC-T-T-A-ATATA-AACAAAG-AATT--AGGAA-T-CAA-GGATA...-----T--CC--TTATA-GCA-ATAGGTAT-C-CAAAACA---C-AT---------TACCTTG.............AAA-TAT-ACAGGATTATG-GGA-G-- 4686
ASC.UG.10.RT11              ..........TCTAACTAAGT-T-T-A-GTATA-AACAAAG-AATT--AGGAA-T-CAA-GGATA...-----T--CC--TTATA-GCA-ATAGGTAT-C-CAAAACA-----AT---------TACCTTG.............AAA-TAT-ACAGGATTATG-GGA-G-- 4671
COL.CM.x.CGU1               ..................................GAAGCAGTTTCC--AAT-CCA--AT-A----TT-GGAA-AACA--GTCA--CG-ATGG...T-G-C-GAAGGTA--A--G-GT-TTACT-----...............TCA-C-AT--AG--AAGC----T-TAC 5004
COL.UG.10.BWC01             ..................................GCAGCAGTTTCC-AA-T-CA---ATC-C--C-A-GGGA-AACA--GGCA---G-ATGG...T-G-C-GAAGG-AG----GTGG-----T-----...............GAA-CAATA-AGC--CAT--C-T--TG 4803
COL.UG.10.BWC07             ..................................AAAGAATTTTCC--AAT-TACC-ACCA-----A-GGAA-AACA--GTCA---G-ATGG...T-G-C-GAAGG-AG-A--GTT--TAGGT-----...............GATAGCAT--AG--AAC-G---T--TG 5037
DEB.CM.04.SIVdeb04CMPF3061  ATAGATGGTGT---GCTA-----A-GGGA-AA---AAAA-CAAAG--TTGGCTCCTATAACA-GGCA--------T----CT-AGAA--AGTTATTTT-CTCAGA--AG-A--ATAT-T--TA-CACA..............AAA--A-GA-A-G-A---G--GA-CTC 5240
DEB.CM.99.CM40              ATAGGTGGTGC---GCTA-----AGAG-A-AA--CAAGA-CAAAG--TTAGATACTGTACAA-GG-T---C--T-T----TTAAGGA-CT--TATTAT-C-CAAA--A--A--GTAT----CA-CACA..............AAA--A-GGG--G-A---G--GA-TT- 5059
DEB.CM.99.CM5               ATAAGTGGTGT---GCCA-----AGAGGA-A----AAGA-CAAAG--TTACAA-C-ATAAAG-GG-T---C--T------TTAAGAAGGG--TTCTAC-CTCAGA--A--A--GT-T----CG-CACA..............GAA--A-GA-A-G-A---G-GGA--T- 5053
DEN.CD.x.CD1_CMO580407      AACATTGG-TG--CACTT-G---TCTATA--A--CAAAA--AATG-GTTA-AA--GACAG-GAT--T--CA--T-T-AG--T-A-ATGGGA-TGGTAT-CC-TGA--A--A------------GA-A..................-CG---TACA---A-G---AGCTC 5161
DRL.DE.11.D3                GGTGGAACTCCCTG-GCAA-T-T--TA-GCATA-G-GAGA-AA-T-TTTAGAAA------A----GC--C-----TCC--T-TT-A...GGTTGGT-G-CTCA-AGTC-GTGGAC------CT-CAAA..............GACA-GAG--A-A-A---G-----GTC 4724
DRL.x.x.FAO                 GGTGGAACTCCCTG-GCAA-T-T--TA-GCATA-G-GAGA-AA-T-TTTAGAAA-----CA----GC--C-----TCC--T--T-A...GGTTGGT-G-CTCA-AGTC-GTGGAC------CT-CAAA..............GACA-AAG--A-A-A---G-----GTC 4817
GRV.ET.x.GRI_677_gri_1      ..............................................-CAATT-CCT---GAA--C--------T-GGC-GGTG-AATGG-AGTTTT-G-CCTACAGCC-GT-CATT-----CT--TCA..............AAA----A-TACA-A-A-G-G-AT-TT 5423
GSN.CM.99.CN166             AGGTAGACTGGTTGC-TAG-GCTACTA-ACACC-CA-A-GGTCAGGCAAA-CACC--TTG-A---GTC---------CC--TTG-AA--C-A-AGATTT-C-CAAAACA-GA-T-GAT-AG--A-G-AT..............ACTAGGAAG...G-A-GGGA-GA-GT- 4984
GSN.CM.99.CN71              ...................-GC-ACTA-ACATC--A-A-GGTCAGG-AAA-CTCC--TTG-G----TC--C------C--CTA-AA---A-AGGTTT-C-CAAAACA--A-T-GAT-AG--T---AT..............ACTAGGCAG...G-A-G-GAGGA-GT- 4966
LST.CD.88.SIVlhoest447      GCCTAGTAG-G-AG--CC--G-TAT-AGG-GAGCACA-AA--A--CTCAA-AT-AT---G-AAGCTT----GGAACA-GGGTAGGCTGGGA-TTTTAT-CTTA-AATA-TA--GTG--A--TT---A-..................ACA-GA-CC--A---G-G-G-TGC 4518
LST.CD.88.SIVlhoest485      ACCTGATAG-G-AA--CC--G-TATTAGG-GAGCACA-AA--AA-CTCAGGTT-AT---G-AAGCTT----GGAACA-GGGTAGGCTGGGA-TTTTAT-CTTA-AATA--A--GTA--A--TT---A-..................ATA-GA-CC--A---G---G-TGC 4518
LST.CD.88.SIVlhoest524      GCTTAGTAG-G-AA--TC--G-TAT-A-A-G-A--CA-----AA-GTAAA-AT-AT---G--AG-TT----GGTACA-G-GTAGGATGGGAGTTTTAT-CTTACAATA--T---TA--T---T---AT..................ACA-GA-CT--A---G---G--TT 4515
LST.KE.x.lho7               CCTTAATAG---AG--TT--G-CATAGTC-G--TAAACA---A-GC-AAAGAAAAT---G-AGGCTT----GGAACA-GTGTAGG-TGGGA-TTCTAT-CCTA-CATA--T--GTG--A--T----A-..................TCA-GA-CT--A---G-C-G--T- 5602
MAC.US.x.251_1A11           ..........CCTCA--AA-T-T-TGA-ATATA-AAC-AA--ATCT-CAA-A--TT--C-A-GTGCCC---TTTA-G-TCG-ATGGGCATGG...T-G-CCT-CAGCAG----AT-T--------CA-...............GAA-GAAGCCAT--A-A-G--CA-GGG 5551
MND-1.GA.x.MNDGB1           ACTGGTTAGT---AAG-CAG-TGGCATGGGCCAC--CAAATAATG-G-AAG-AT-T----GGCTGTAT-C---T-T-ATGGCTTA-A--GA-TGGTAT-CTT-CAGTA-----GTG--TAT-A--AATAGG..............--CATA-----AA--G-T-G--GC 4949
MND-2.CM.98.CM16            GATGGAATTC-CTGG--AA-T-C---A-ATATA-A-GAGA-AAAC-TTTGGACAA----GA--TA-AGGCA--T-T-C--T--T-AGGATGG...T-G-C-CAC---C--AA-T-T--T--C----AA..............GAA----AACAGA-CA----C--C-TC 5251
MND-2.GA.x.M14              AATGGAATTC-TTAG-CAA-T-C---A-ATATA-G-AAGA-AAAC-CTTAGATAAG---GAACTATTC--C----T-C--T--T-AGGATGG...T-G-CGCAC--TC--AAGATT-----CT-CAAA..............GAT-GCT-A-AGA-TA-T--T--TG-C 5178
MND-2.x.x.5440              ATTGGAATTC--TAG--AA-T-T--TA-GTATA-A-GAGAGAAAC-TTTAGAAAAT---CAA-TATAC-CC--T-T-C--T--T-AGGATGG...T-G-C-CAC--TC--AAGA-G-----CT-TAA-..............GAT-G-T-A-AGA-TA-C--C--CGTC 4814
MNE.US.x.MNE027             .......CCTC-TAAA-TATCTGA-AT---AAAC-AAAGATCTAC--AAGGTTT-C-AT...GTGCCC-----TA-G-TCG-ATGGGCATGG...T-G-CCT-CAGCAG----AT-T--------CAA...............GAA--AAGCCAG--A-A-G--CA-GGG 5020
MON.CM.99.L1_99CML1         ...................-TGCACTA-ATG-C-CA-ACTCACAGG-AG-GCACCC-TTG-C---GT---------CC-GCTG-ACA---AGAGATTCTC-CAGAACA--A-TA-AC----CT---ACCTAGGAATCACTCAAGAA-GA-AGTC--G--C--C-TATTTG 4971
MON.NG.x.NG1                ...................-GC-AC-A-ATATA-CA-A-ATACAGGGAAAGT-CC--TTG-G---GT----------C-GCTTA----C-AGCGCTTTTCGCAA-GCA-GA-TA-GC-T--TT---A-ATGGGTATCACACAARAG-RG-AACA--G--C--C-TAC-TR 3564
MUS-1.CM.01.CM1239          ...................GGCTACAA-ACATC--A-C-TTTCAGG-AAA-CACC--TTG-G---GT---C-----CC-GGTA-AA--C-A-AGATTTTCCCAAAACA-GA-CA-GC-AG--T-GACACAAAACACCTTAGAATCA-GA-AG-C-GAA--C-CCTTC-T- 4948
MUS-1.CM.01.SIVmus_01CM1085 ...................GGC-ACAA-ATAC--CA-CCTTTCAGGGAGAGCA-CC-TTACA---GTG---------C--TTGTAC----AGAGGTTCTCCCAGAGCA--A-CAGA--A---T-GACACAAAATACCCGGGAATCA-G---AGT-GAAAGT-CTTAC-T- 4958
MUS-2.CM.01.CM1246          ...................-TGCACTA-ACAT---A-C-TCTCAGGGAAAGCTCC--TTACA---GTC--C-------C--TTG-AG----AGAGATTC-CCCAGAACA--A--A-A--A--TT-GA-CATAAACAGAATAGAAGGA-GCC-A-C-GA-ACT-C-TAC-T- 5056
MUS-2.CM.01.CM2500          ...................-TGTAC-A-ACACC-CA-C-T-TCAGGGAAAGCACC--TC-CA--C-TC-----T--CC--GTG-AG----AGAGATTC-CTCAAAATC-GA-TA-G--A--CT-GACAATAAAGAAGAGAGAGGAA--CAAAGT-GAATCT-C-TAC-T- 5009
MUS-3.GA.09.09GabOI81       ...................-T-C.........CT-A-C-TCTCAGGGAMAGCCCC----ACA---GT----GCA--CC--CT--AATCA-A-AGATTCTCTCAGAATA-GA-TA-A--A--CT-G-CAATA......CAACTCAGA--CAT-GA-GGACAGCAGTGG-T- 4518
MUS-3.GA.11.11GabPts02      ...................GTGCAC-A-GTATC-CA-A-T-TCAGGGCAA-CACC----AG----GT----GCA--CC--TTA-AA----AGAGGTTCTCCCAGAACA--A-TAGA--A------ACACAG...GAAACAACAGAAAGG-AA-A-G-TTCT-C-CAC-T- 4542
OLC.CI.97.97CI12            AGGTAGAAC-TGTGAA-T-C-TGGCTA-GT-AGCCTACAA-TTTTGG-GATACCAT---A-A-GGCA-A--TGT--CA-GTA---AAGGGA-TGGT-G-C-CA-GG-A--A-TT-T--T---G--CCA...................-GG-G-GAGA-A-A---T---CTC 5103
RCM.CM.00.SIVagi_00CM312    AATGGCACTCTCTAG--AA-T-C--T--GTACA-A-GAAA-AA-G----A-A--A----GAA---GTG-CC--T-T-A--GTA---TGGGG-TGGT-GTCTCA-------A-C---G-T--CT-G--A...............AAAA-CA---AGG-TAA---T--C--T 4527
RCM.CM.02.SIVrcm_02CM8081   AGTGGCACTCCCTAG--AA-T-T--TATGTATA-A-GAAAGAAAG----A--A-A----GAA---GT--CC--T-TCA--GTA---TGGGGGTGGT-GTCTCA------G--C-----T--CT-G--A...............AATA-CA---AGA-AAAGG-G--C--C 4545
RCM.GA.x.SIVRCMGAB1         AATGGCACTCTTTGG--AA-T-T--T--GTACA-A-GAAAGAAAGC---A-AA-AG---GAG---GTG-C-----T-A--GTA--ATGGGG-TGGT-GTCGCA---------T-----A--TT---A-...............GAA-GAT-A-AG--AAA------C--C 4738
RCM.NG.x.NG411              AGTGGCATTCCTTAG--AAGT-T--T--GTATA-G-GAAAGAA-G------AA-AT---GAG---GT--CC--T-T-A--GTA---TGGGGGTGGT-GTCTC----------C--------CT-G--A...............AATA-CA---AGA-AAAGG-T------ 4750
SAB.SN.x.SAB1               ..............CC----G---T----C---C----G--CAAAC-GTG-GTTCAT---CGG----CC-C---TAC-A---TAAGATGGAATTGGTACTCCTA-CA----TGGGT---A--C--GAA-...............GAT-GA---CT-A-AAA-G-T--T-AT 5648
SMM.CI.79.SIVsmCI2          ...................-C-T-TTA-GT-CA-CACAAA--AATT-CAA-A--TG--T-A-GTACCC---A--A-G-T-G-ATGGGCATGG...T-G-CTT-CAGCAG----ATAT-T--T--G-AA...............AAA--AT-ACAC--A-A-G--CA-GGG 4672
SMM.LR.89.SIVsmLIB1         ..........CCTAA--AA-C-T-TAA-GTATA-CAC-AA--A-CTGAAC-AA-CC--C-A-GTACCC------A-G-T-G-ATGGGCGTGG...T-G-CTT-CAGTAG---GAT-T-T--C--GCAA...............GGA--A--ATATC-A-A-G--CA-GGG 4630
SMM.SL.92.SIVsmSL92A        ..................................AACAGA--ATCTGAAG-AA-T-GTC-ACGT-CCC------A-A-T-G-ATGGGCATGG...T-G-CTT-CAGCAG----ATAT----T--G-AA...............AAA--A--CCAT--A-AGG--CA-GG- 4632
SMM.SL.92.SIVsmSL92B        GCTGGCAT-GCT-CA-TAA-TTC--TA-GTATA-AACAAG--A-CT--AG-AA-CC--C-A-GTGCC-------A-A-T-G-ATGGGCATGG...TAC-C-GCC--TAGG--GATTT----CT-G-A-...............GAG-GAAG-CAC----AGG--CA-GTG 4615
SMM.SL.92.SL92B             GCTGGCAT-GCT-CA-TAA-TTC--TA-GTATA-AACAAG--A-CT--AG-AA-CC--C-A-GTGCC-------A-A-T-G-ATGGGCATGG...TAC-C-GCC--TAGG--GATTT----CT-G-A-...............GAG-GAAG-CAC----AGG--CA-GTG 4986
SMM.US.04.G078              ..........CCTCA--AA-C-TTTGA-GTATA-CAC-AA--A-CT-CAG-AA-T---C-A-GTGCCC-----TA-G-TTG-ATGGGCGTGG...T-G-CTT-CAGCAG---AT-T-----T--CAA...............GGA--AA-CCAT--A-A----CA-GG- 4812
SMM.US.05.D215              ..........TCTCA--AA-C-TTTAA-GT-CA--AC-AAG-ACCT-CAA-A--CT--C-A-GTACCC-----TA-G-T-G-ATGGGCATGG...T-G-CTT--AGCAG----AT-T-T--C----AA...............AAA--AT--CAC--A-A-G--CA-GGG 4806
SMM.US.06.FTq               ..........CCTCA--AAGC-TTTGA-ATATA-CAC-AA--A-CT-CAG-AA-CC--C-A-GTGCCC-----TA-G-TTG-ATGGGC-TGG...T-G-CTT-CAGCAG---GAT-T-----T---AA...............GGG--AT--CAC--A-AGG--CA-GG- 4834
SMM.US.86.CFU212            ..........CCTCA--AAGC-TTTAA-ATATA-CAC-AA--ACCT-CAGCA--C---T-A-GTCCCC-----TA-A-TTG-ATGGGCATGG...T-G-CTT-CAGCAGG---AT-T-T---T--CAA...............GGA--AT--CAT--A-A-G--CA-GGG 4798
SMM.US.x.H9                 ..........CCTCA-CAA-C-C-TGA-ATATA-CAC-AA--ACCT-CAG-A--CT--T-A-GTACCC-----TA-A-TTG-ATGGGCATGG...T-G-CTT-CAGCAG----AT-T----GT--A--...............GAT--G---CAT----A-G--CA-GG- 5025
SMM.US.x.PGM53              ..........CCTCA-CAA-T-CTTGA-AT-TAGAAC-AAG-ACCT-CAG-A--CT--T-A-GTACCC-----TA-A-TTG-ATGGGCATGG...T-G-CTT-CAGCAG----AT-T--------A-A...............GAT--AT--CAT----A-G--CA-GG- 5469
SMM.US.x.pE660.CG7G         ..........CCTCA-CAA-C-C-TGA-ATATA-CAC-AA--ACCT-CAG-T--CT--T-A-GTGCCC-----TA-A-TTG-ATGGGCATGG...T-G-CTT-CAGCAG----ATTT-----T--A-A...............GAT--GA--CAT----A-G--CA-GG- 5541
STM.US.89.STM_37_16         ..........CCTCA--AA-C-TTTAA-GTATA-CAC-AA--A-CT-AGC-A--C---C-A-GTGCCC--C--TA-G-TTG-GTGGGCATGG...T-G-CTT-CAGTAG---GATTT-T--CT-GCAA...............GGA--A--ACACC-A-A-G-TCA-GGG 5199
SUN.GA.98.L14               TTGCTGTAG-G-AA--TT--G-TTTAAGA-GAATGCA-AGG-AA-C-AAAGAT-AT---G-A-C--TGT--GGAACA-GT-CTGGATGGGAGTGGTAC-C-TACAATA--A-TATT--T---G--ACT..................T---GG-C-G-A---G---GG-TC 5652
SYK.KE.x.KE51               .TAAAATGC-GCA-A--T----GGTA--C-A----AAAAGCAAAC--TTAGA-AA-GCCACA----A------T---C-G-TTA-ATGGGAGTGGT-G-CTC-AG--C--TGGG-A--A---G----A..................C-A-G-GT---AA-T----A-TTC 5096
SYK.KE.x.SYK173_COMGNM      .TAGATTGC-GCA-A--A-----ACA----A----AAAAGCAAAG--TTGGA-AA-GCCACA----A---------CC---TAGAATGG-A-TGGTAC-CTTA--GCC--TGGACA--A---G----A..................----GA-CTA-ATGG--C--GTTC 5444
TAL.CM.00.266               ......................GTTT--C--AC--AAAA-CAA-T-CTTACAAAAG-T-A---GGGTT-------T-C--TTGAGA--CG-CTCCTAT-C-CAAG--A-GTGGATAC----GT---AA...GAAAAGGACACCAGA-TA-AATAT--AACC------CAG 5231
TAL.CM.01.8023              ......................GT-T--C-----CAAAA-CAA-T-CTTACAAAAG-TAA-A-GGGTC-----T-TCC--TTAAGAATGG-CTCTTTC-CGCAGG-CA--TGGAT-C----C----AA...GAAAAAGAGAATAGA-TA-AGTACA-TACC--C--TCAG 4729
TAN.UG.x.SIVagmTAN1         ATATAGAC---T-GAAG-GC-T-GTGA-ATACC-CA-GAGGAACCGGTTACA--AT---ACA---CTTATG-----CC--T-TGGATGGG-TTGGTAC-C-T--AGTAG-T-C-T------CT----A...............GGA--A-GA-A-A-A---G--GATTGT 5449
VER.DE.x.AGM3               ..........-T-GCA--GG-T-GTGAGGTAC-GGA-GAGGACT-G-AAATTA-AC---AAA---C--ATG-----CC---TTA-ATGGG--TGGTAC-C--TGAGTAG-TATG-G--A--C------...............CAAC---GA--TA-ACATG--GATCT- 4972
VER.KE.x.9063               ..........................................A-GGAACCTTAA----CAA--C-TGATG-----CC---TTAA-TGGG-TTGGTAT-CT-TG-GCAG-TACAT-------A----A...............TTA--A-GAGA-A-ACAG---GACTGC 5477
VER.KE.x.AGM155             ............................ATAT-GGA-GAATAAA-GGAA-CT-AA----GAA--C-A-ATG--T---C---T-A--TGGG--TGGTAC-CT-TGAGCAG-TATGTA--A--C--CCCA...............GGAAG--GAGA-A-CCATG-GGAT-TC 5469
VER.KE.x.TYO1_patent        ..........-T-GCAG-G--T-GT-AGATAT-GGA-GAG-CAA-G-AA-TT-AA----AAC---TATATG--T--CC---TTA-ATGGG-TTGGTAC-CC-TGAGTAG-TATGTA--A---A----A...............AAAC---GGGA-A-ATGTG--GACCT- 5474
WRC.CI.97.97CI14            GTAAATTC-GG--A-GGACCT-TT-TATAGCAA--C-AATTAAAGC--GGCA--C-AAT---CTAT-GGGATCT-GGC-TGCAG-AGCAGA-TGCTTT-CT--GCATAT----A-A--A---A--CA-...............GGAC-CTGGGAGA--CA---T-G-TGC 5541
WRC.CI.98.98CI04            GAAGGTTT-----A-GG-C-T-TTGGATAGCTC--C-AATGCAA--TAAA-A--C-A-T----TAT-GGGAACT-GGC-T-CT-AGGCAGA-TGCTTT-CT--A-ACCT-T-GAGA--A--CT--CAA...............GGGC--TGGGA---ACA---C-GGTGC 5556
WRC.GM.05.Pbt_05GM_X02      GAAGATTT-G---A-GGAGCT-TTGGATAGCAG-GC-AAT-CAAG--AAACAA-C-A-A--CCTGTTTGGCTCA-GGC-TGTA-A-GCAGA-TGTTTC-CT--ACATAT----AGA--T--T--TCAA...............GGGC--TGGGA-A--CAG----GGTGC 4473
```

```
H1B.FR.83.HXB2              TATTGGGGTCTGCATACA............GGAGAAAGAGACTGGCAT..........................................................................................................................TT 5282
Vif                          Y  W  G  L  H  T  .  .  .  .  G  E  R  D  W  H  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  L
H1O.CM.91.MVP5180           --------AT-AATGC--............--G--------GGAA--C..........................................................................................................................-- 5311
H1N.CM.95.YBF30             --------G--AACA---............------CAGTCT------..........................................................................................................................C- 4876
H1P.CM.06.U14788            --------G--CATGC-T............--G--------AGAA--A..........................................................................................................................-- 4753
CPZ.CD.06.BF1167            --------C--A--CTGT............--G----A--CT------..........................................................................................................................C- 5393
CPZ.TZ.00.TAN1              -----------AGCCTGT............---------CCA------CTAGGTCATGGC................................................................................................................ 4939
GOR.CM.07.SIVgorCP2135con   --------A--CATGC--............-------AG--AGAA--A..........................................................................................................................-- 4818

MAC.US.x.239                ------CA-T--ACAC--...............----A--GG----TCAGTACTTATGCA................................................................................................................ 5597
Vif                          Y  W  H  L  T  P  .  .  .  .  .  E  K  G  W  L  S  T  Y  A  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
H2A.SN.86.ST_JSP4_27        --C---AAC--AACAC--...............----A--GA---TCTCCTCCTATTCA................................................................................................................. 5125
H2B.CI.x.EHO                --C---AAC---ACCC--...............-----G-GA-TCTTGAGTTCCTATGCT................................................................................................................ 5651
H2G.CI.92.Abt96             ------AACT--ACCC-T...............----AG-GA---TTAAGCCAGTATGCA................................................................................................................ 5007
H2U.CI.07.07IC_TNP3         ------AA-T--ACAC--...............--G-AG-GA----T-TCAACATATGCA................................................................................................................ 5112
H2U.FR.96.12034             ------AA-T--ACAC--...............--G-AG-GA----TATCCTCTTATGCA................................................................................................................ 5158
ASC.UG.10.RT03              ACAGCCATAGAAGTACA-ATA.....TTGT-G--TCTGTCAA-TTG-AGAAGTAGGGATCCACCTGGAACAAGAGACACACAAGCCCTTACTATTAGT...TGGACTTACACC.........................................................AG 4769
ASC.UG.10.RT08              ACAGCCATAGAAGTACA-ATA.....TTGT-G---CTGTCAA-TTG-AGAACTAGGAATCCACCTGGAGTAAGAGACACACAAGCTCTTACTATTAGT...TGGACTTACACC.........................................................AG 4793
ASC.UG.10.RT11              ACAGCCATAGAAGTACA-ATA.....TTGT-G---CT-TCAA-TTG-AGAAGTAGGGATCCACCTGGAGTAAGAGACACACAAGGCCTCACTATTAGT...TGGACTTACACC.........................................................AG 4778
COL.CM.x.CGU1               -TGG-CAC-T-AACCC--...............--T---T-TCATAGAGGACCACCA................................................................................................................... 5046
COL.UG.10.BWC01             CTG---CA-T-AGGA---...............--T---T-TCATAGGGGAACAGCA................................................................................................................... 4845
COL.UG.10.BWC07             CTA---CA-T-AAGC---...............--T--GTTTCATAGAGGGAGTGCA................................................................................................................... 5079
DEB.CM.04.SIVdeb04CMPF3061  -TC---T-C--AGCGC--...............----A--GA---TTGCCAGGAATGGCA................................................................................................................ 5285
DEB.CM.99.CM40              CTA---T-CT--ACCC--...............--G-A--GA----TACCCACTATGGCA................................................................................................................ 5104
DEB.CM.99.CM5               -TC---A-C---GC-C--...............----A--GG---ATGCCCTCTATGGCG................................................................................................................ 5098
DEN.CD.x.CD1_CMO580407      ------CA-T--ACCC--...............----A--GA-AT-TGTCACAAGTAGCA..............................................................................................................A- 5208
DRL.DE.11.D3                CTG---AAC--CACCC--...............--C-A--GG---TTAAGTAAATATGCCATA............................................................................................................. 4772
DRL.x.x.FAO                 CTG---AAC--CACCC--...............--C-A--GG---TTAAGTAAATATGCCATA............................................................................................................. 4865
GRV.ET.x.GRI_677_gri_1      ---CACAAC--CACCC--...............-------GA----TCTCAAGTCATGGA................................................................................................................ 5468
GSN.CM.99.CN166             G-AGCAACATATATAGAGATAACTATCCTAT-G--C-C-AC-A-T---GGACCAGCTAGCTTATCCCGGTCCACCTATTACCAACAAGCAGTCATAATAGAATGGGTATACAAC........................................................AG 5100
GSN.CM.99.CN71              G-AGCCACATATAT-GAGATAACCATCCTAT-G--C-C-ACTA-T---GGACCAGCCAGTCTATCCCGATCCACCTACTACCAACAGGCAGTCACAATAGAGTGGGTTTACAAC........................................................AG 5082
LST.CD.88.SIVlhoest447      ---GCACACT--GCAG--...............-G-----GA---ATCTCTCAGTGGGCA................................................................................................................ 4563
LST.CD.88.SIVlhoest485      ---GCACACT--GCAG--...............-G-----GA---ATCTCTCAGTGGGCA................................................................................................................ 4563
LST.CD.88.SIVlhoest524      ---GCTCA----GCAG--...............-GG----GA---AT-TCACAGTGGGCA................................................................................................................ 4560
LST.KE.x.lho7               ---GCACA-T--GCAG--...............-G---G-GA---ATCTCACAATGGGCA................................................................................................................ 5647
MAC.US.x.251_1A11           ------CA-T--ACAC--...............----A--GG----TCAGTACTTATGCA................................................................................................................ 5596
MND-1.GA.x.MNDGB1           ------CA-T----A-T-...............--GGT--GA--CTTAAGTACTTATGCAGTAAGCATA....................................................................................................... 5003
MND-2.CM.98.CM16            CTG---AACT-AACCC--...............----A--GA---TTATCAACTTATGCC................................................................................................................ 5296
MND-2.GA.x.M14              CTA---AACT--ACCC--...............----A--GA---TTATCTCAATATGCA................................................................................................................ 5223
MND-2.x.x.5440              -TG---AACT--ACCC--...............----A--GG----TGTCTCAATACGCC................................................................................................................ 4859
MNE.US.x.MNE027             ------AA-T-AACAC--...............-------GG----TCAGTACTTATGCA................................................................................................................ 5065
MON.CM.99.L1_99CML1         G-GATAACAA-TT--TGG...............--CTT-ACAAACGTGGGACCTAGTGCACTCTCACCAACAGTGTACAACAAACAGGCCTACACAATAGAATGGGTTTATTGGATGCAAAAGGTGGAACGGGTCCCTTTAGATCCCGTG..................-G 5108
MON.NG.x.NG1                G-GATCACGA-C-T-TGG...............--C-YTAC-AATGTAGGGCCCGCGTCACTATCTCCATCAGTCAAATGGAAGCAAGCATACATCATAGAATGGGTATATTGGAGGCAGCCCAGGGTGCCACAGCAAGAGGGAGTAGTAGATCGAGTGTGGTTTCCTAG 3719
MUS-1.CM.01.CM1239          G-GATAACAG-C-TATGG...............--TGTGACTAATGTAGGACCAGCAAGCCTATCCAAAAGTACCTATTGGAAGCAGTCATACATCCTGGAATGGGTATACATG........................................................AG 5049
MUS-1.CM.01.SIVmus_01CM1085 G-ACT-ACAG-C-TGTGG...............--TGTGAC-AATGTGGGACCAGCAAGTTGCTCTAAAAGCACATGGTGGAAGCAATCCTTTATCCTAGAGTGGGTATATCTA........................................................AG 5059
MUS-2.CM.01.CM1246          G-GATAACCA-ATTGTT-...............--TGTCACAAACGTAGGGCCAGCCAGTTTATCAAGGTCCACCTATTGGCAGCAATCCTATATCCTGAAGTGGAGATACATC........................................................AG 5157
MUS-2.CM.01.CM2500          C-AATAACAA-T-TATGG...............--TGTTAC-AATGTAGGGCCTGCCAGCCTCTCAAAGTCTACTTACTGGCAGCAGTCATATATATTGGAGTGGGTTTATGTT........................................................AG 5110
MUS-3.GA.09.09GabOI81       G-GGCCACC-ATATAGA-ATAACAGTACTTT-G--TGTGAC-AATGTAGGATATGCCAGCCTGTCTAGAAGTACACAATGGAGGCAGACATATATATTGGAGTTTGTCTATCAG........................................................AG 4634
MUS-3.GA.11.11GabPts02      G-GATAACCA-A-TGTGG...............--TGT-ACTAATGTAGGCCCAGCCAGCCTCTCCAAAAGTACATATTATGAACAATCCTATATCCTAGAATGGGTATACAAG........................................................AG 4643
OLC.CI.97.97CI12            --C---CA-A-TAGACAT...............-GG-A--GA-TA--AAATTCAGGTGGGGTG............................................................................................................. 5151
RCM.CM.00.SIVagi_00CM312    ------AA---CACAG--...............----AG-GA---TTAAGTACTTATGGA..............................................................................................................G- 4574
RCM.CM.02.SIVrcm_02CM8081   ------AA---AACA---...............----AG-GG---TTGGGAACTTATGGG..............................................................................................................GC 4592
RCM.GA.x.SIVRCMGAB1         ------AA-T--ACAGT-...............----AG-GA----TAGGGACATATGGG................................................................................................................ 4783
RCM.NG.x.NG411              ------A----CACA-T-...............----AG-GA---TTAGGAACCTATGGA..............................................................................................................G- 4797
SAB.SN.x.SAB1               ------CA-T-AACCC--...............--G-A--GA---TTAGAAACCTATGCA................................................................................................................ 5693
SMM.CI.79.SIVsmCI2          --C---AA---AACAC--...............--G-A--GA----TCTCTTCTTATGCA................................................................................................................ 4717
SMM.LR.89.SIVsmLIB1         --C---AA-T-AACAC--...............--G-A--GA----T-AGCACCTATGCT................................................................................................................ 4675
SMM.SL.92.SIVsmSL92A        ------AA----ACAC--...............----A--GA----TGTCCTGTTATGCA................................................................................................................ 4677
SMM.SL.92.SIVsmSL92B        --C---AA----ACCC--...............----A--GA---TTATCAAGCTATGCA................................................................................................................ 4660
SMM.SL.92.SL92B             --C---AA----ACCC--...............----A--GA---TTATCAAGCTATGCA................................................................................................................ 5031
SMM.US.04.G078              ------AAC---ACAC--...............----A--GA----TCAGTACTTATGCA................................................................................................................ 4857
SMM.US.05.D215              ------AACT--ACAC--...............----A--GA----TCAGTACCTATGCT................................................................................................................ 4851
SMM.US.06.FTq               ------AACT--ACAC--...............--G-AG-GA----TCAGTACTTATGCA................................................................................................................ 4879
SMM.US.86.CFU212            --C---AACT-AACAC--...............--G-A--GA----TCAGTACTTATGCT................................................................................................................ 4843
SMM.US.x.H9                 ------AA-T--ACAC--...............--GRA--GG----TCAGTACTTATGCA................................................................................................................ 5070
SMM.US.x.PGM53              ------AA-T--GCAC--...............----A--GA----TCAGTACTTATGCA................................................................................................................ 5514
SMM.US.x.pE660.CG7G         ------AA-T--GCAC--...............----A--GA----TCAGTACTCATGCA................................................................................................................ 5586
STM.US.89.STM_37_16         ------AA-T-AACAC--...............----A--GA----TCAGTGAATATGCT................................................................................................................ 5244
SUN.GA.98.L14               ---G-ACAC--CACAC--...............-C--A--GG-T-AT-AACCAGTGGGGA................................................................................................................ 5697
SYK.KE.x.KE51               ---CATAA-T--ACAC--...............--G-A--G-CT-TTAAGAACAGAGGGA................................................................................................................ 5141
SYK.KE.x.SYK173_COMGNM      ---CACAA-T--GCCC--...............-------GA---TTACATATGCAGGGA................................................................................................................ 5489
TAL.CM.00.266               ------T--T-AGCAC--...............--CCAT-GA-ATACGTCCGATTGGGACACAATATCCATT..................................................................................................GA 5290
TAL.CM.01.8023              ------T-CT-AGCAC--...............--C-AG-GA-ACACATCAGACTGGGATACAATATCCATT..................................................................................................GA 4788
TAN.UG.x.SIVagmTAN1         --C---CAC--AACGC--...............---CAG-GA----TCTCTACCTATGCA................................................................................................................ 5494
VER.DE.x.AGM3               ------CA----ACAC--...............----AG-GA----TATCAACATATGCT................................................................................................................ 5017
VER.KE.x.9063               ------CA-T-AACAC--...............-------G-----TCTCTACTTACGCC................................................................................................................ 5522
VER.KE.x.AGM155             ------CA-T-AGC-C--...............A--CA--GA----TCTCAACTTATGCA................................................................................................................ 5514
VER.KE.x.TYO1_patent        ------CA-T-AACAC--...............--GCA--GA----TATCCACATATGCA................................................................................................................ 5519
WRC.CI.97.97CI14            ------AA-T--TCAC--...............----AG-GA---ATGGACACCATGGGATTT............................................................................................................. 5589
WRC.CI.98.98CI04            ------AA---AGCAC-T...............----AG-GA---TTAGACACCATGGGGTTT............................................................................................................. 5604
WRC.GM.05.Pbt_05GM_X02      ------AA-T-AGC-C--...............----T--GA---ATAGATACTTTGGGATTC............................................................................................................. 4521
```

```
H1B.FR.83.HXB2              AAGCAGGACATAACAAG.................GTAGGATCTCTACAATACTTGGCACTAGCAGCATTAATA..............ACACCAAAAAAGATA...AAGCCACCTTTGCCTAGTGTTACGAAACTGACAGAGGATAG..................... 5558
Vif                         Q  A  G  H  N  K  .  .  .  .  .  .  V  G  S  L  Q  Y  L  A  L  A  A  L  I  .  .  .  .  .  T  P  K  K  I  .  K  P  P  L  P  S  V  T  K  L  T  E  D  R# .  .  .  .  .  .  .
H1O.CM.91.MVP5180           TG---------GTC--.....................-----GA-A-------T---A--CT-GAA----G--G-G...............-A-GT-----GA-AT..------T--CC-A--C-----CCA--G-T-A-----A-----..................... 5587
H1N.CM.95.YBF30             C------G--C--AC--.....................-----CA-C--------TC-A------A----C-GGG-G..............GG-G----G----G-..---------C--A--------G--T--G--A-----A-----..................... 5152
H1P.CM.06.U14788            C-----------GTC--.....................-----C--CT-G--G--T--A--CT--AA--TGG--G--...............GGGAAGGT---R-GG..---------CC-C--------CCA--T-T-----C-A--C-T..................... 5029
CPZ.CD.06.BF1167            CTA--------TCT---.....................-----GAGC-----G-TTC-A--CT-TAA-AA-G-GC--..............GA--AG--C---CAG..-----T-A-AGA---CCAT-AC-CTC-G-A--T-TCCTGGCAGAGGA.................. 5675
CPZ.TZ.00.TAN1              --T-------C-GAC--.....................-----GA----G----T---A--TT-CAG-AAGG--G-TGAGAGCCAA...GAT-A--AGCC----GG-CCA-G-AGG--C-----ATC------A-----A--------C--..................... 5220
GOR.CM.07.SIVgorCP2135con   C-----------GT.................CAG-----GA-AT-G--------A------AG---T---G--...............GGGAA-GTC-G-G--...-GA--------T--A------CA----T-AG---CA----C..................... 5094
                                                                                                                                                                  Vpx start
MAC.US.x.239                CGAG--CT-----GT-CCAG.................---CC-AGC-----G-----A-----GAA--T-G---GCG..............................................................AT-TCAGAT......CCCAGGGAGAGAATC 5835
Vpx                                                                                                                                                                 M  S  D  .  .  P  R  E  R  I
Vif                         P  R  A  H  K  Y  Q  .  .  .  .  .  V  P  S  L  Q  Y  L  A  L  K  V  V  S  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  ##D  V  R  .  .  S  Q  G  E  N
H2A.SN.86.ST_JSP4_27        CCCA--CC-----GT-CCAG.................---CCG--A--C--G-TTC----CT---TG-T-G-GCA-CAAA.........................................................AT--CAG-C......CCCAGAGAGACAATA 5366
H2B.CI.x.EHO                C-T---CT----CAGGCCAG.................---CC-AG-T----G-TTC-A--C---AG--TTG--CA-GAAGGAAAAG....................................................AT-...GAT......CCCAGGGAGAGAGTA 5895
H2G.CI.92.Abt96             C-A---CT-----A-G-CAG.................--GCCTAGCT-G----TTC-A--CT--CA--T-G-GCA-AAAGGAC.......................................................AT-...GAT......CCAAGGGAGAGAGTM 5248
H2U.CI.07.07IC_TNP3         C-T-G-CT-----AGG-CAA.................---CC-AGCT------T-C-----T--AAGAT-G--CAGG..............................................................AT-TCAGAC......CCCAGAGAAAGAATA 5350
H2U.FR.96.12034             C-T-G-CT----GAC--CAG.................---CC----T------TTC-A-----CAG------GCA-G.............................................................AT--GAGAT......CCAAGAGAGAGAATC 5396
ASC.UG.10.RT03              C-CATTAT...TTAGGCAAA............GAGA--CCCA-AT-G------C-A------AGG---A--CA-G................................................................................................. 4993
ASC.UG.10.RT08              C-CATTAT...TTAGGCAAA............GAGA--CCCA-AT-G------C-A------AGG---A--CA-G................................................................................................. 5017
ASC.UG.10.RT11              C-CATTAT...TTAGGTAAA............GAG---CCCA-AT----------A------AGG---A--CA-G................................................................................................. 5002
COL.CM.x.CGU1               -G-T------CCTAGGA....................--GTTGAAGT-G---G-AC-ATGT--TCTG--CC---G-G................................................................................................ 5257
COL.UG.10.BWC01             -G-TG--G---CTAGGA....................---A-----------G-GC-ATGT--GAGCAAGC-CCA-A............................................................................................... 5056
COL.UG.10.BWC07             -G-T------CCTTGG-....................---CT-A-CT-G---G-GC-ATGT--CAAC--CA--GACG............................................................................................... 5290
DEB.CM.04.SIVdeb04CMPF3061  -TCATA-GGGAGCAGGAGGT.........AACGGACC-CCT--AT----G---A-TTGCT--AGG-T-GCC---AAAT.............................................................................................. 5508
DEB.CM.99.CM40              -CCATA-GGGACCAG-AGGT.........AAAGGACC-CCT--A-----G---A-A-----TAG--T-GC-G--AAAC.............................................................................................. 5327
DEB.CM.99.CM5               -CCATA-GGGACCAG-AGGT.........AAAGGACC-CCC--AT----G-----A-----AGGA--GC-G--AAAT.............................................................................................. 5321
DEN.CD.x.CD1_CMO580407      -GAAG-C----T--C-TCCA......AGGACAGGTCC---TAGGGA--CGCCA-CATT--AGTATAT-G--T--......CAAGCAGCC-TTAG-T-TGGA-AGGTT-CCA-CATCACA--C-CCAGC-TT-GTT-C-AGA.............................. 5495
DRL.DE.11.D3                CTAG---G--C--G..............GGCCAG-----GAG---T--G--TC-T--CT-GCA----CAC---AAGTACCTTAGGC.......................................................AT--CAGAAGGACAGCCAGCGGAGAGAGCT 5022
DRL.x.x.FAO                 CTAG---G--C--G..............GGTCAG-----GAG---T--G--TC-T--CT-GCA---CCAC---AAGTACCTTAAGC.......................................................AT--CAGAAAGACAGTCAGTGGAGAGAGCT 5115
GRV.ET.x.GRI_677_gri_1      -G-A---------A..............GGTCAG---CA---A--G---CTT----------TT----AT-C-A................................................................................................ 5682
GSN.CM.99.CN166             CG-A--CT--CCTG.................CAG------A-A-CT--GACGC-T-AG-AGTT---C--CT-TGCATATGTCAAACATT................................................................................ 5336
GSN.CM.99.CN71              CG-A--CT--CCCA.................CAG------A-A-CT--CACGC-T-A-AGGCT---C--CT-TGCGTACGTTAGATACT................................................................................ 5318
LST.CD.88.SIVlhoest447      CCA-T--------A..............GGTTGT--TCTC--AT-G--G-T-C-CTGCT-GAG-CA-C--CA-C................................................................................................ 4777
LST.CD.88.SIVlhoest485      CCA-T--------A..............GGTTGT--TCTC--AT-G--G-TTC-CTGCT-GAG-CA-C--CA-C................................................................................................ 4777
LST.CD.88.SIVlhoest524      CTA-T--------A..............GGTTGT--TCTC--AT-G--G---C-CTGCT-GAG-CA-C--CA-C................................................................................................ 4774
LST.KE.x.lho7               CTTT---------A..............GGTTGT--TCTC---T-G--G-TTA-CTGCT-GAG-CA-C--CA-C................................................................................................ 5858
MAC.US.x.251_1A11           CGAG--CT-----GT-CCAG.................---CC-AGC-----G-----A-----GAA--T-G---GCG..............................................................AT-TCAGAT......CCCAGGGAGAGAATC 5834
MND-1.GA.x.MNDGB1           C---T--C-----AGCACAA............GGT--TCTC---T-G--G-TTC-CTGCT-GAG--TCA-CT................................................................................................... 5226
MND-2.CM.98.CM16            -CTGG--------A..............GGGAAG-----GAGC--G--G---------TT-GCT-AGT-AT-C-GCCTACTGTAACA...................................................AT--CAGAA......GGGGCCCCAGAGATC 5543
MND-2.GA.x.M14              CCTGG--------A..............GGTCAG-----GA--------CT------TT--AG---C-AC--CAAGTTTTGCAGAA....................................................AT--CAGAA...............AGAGCA 5461
MND-2.x.x.5440              CCTGG-------GA..............GGTCAG-----GA--------G-TT-----CT--CA---C-ATC-CAGAGGTAGAAAGC...................................................AT--CAGAG...............AGGGCA 5097
MNE.US.x.MNE027             CGAG--CT-----G--CCAG.................---CC-AG------G-----A-----GAG--T-G---GTT..............................................................AT-TCAGAT......CCCAGAGAGAGAATC 5303
MON.CM.99.L1_99CML1         T----AC---CT--CC-AAG.................-----GA-C-CGTTGAC---A-AGAA-TT---C--CC-TGCCCTAGTGCGATATGG-G---CCCCC-C-AAGTCAT-CG--CCATTGGACC-CC--G.................................... 5389
MON.NG.x.NG1                T-CAGAC---CT----TAAG.................------A---CT-C-AC-C-A-AGAA-TT---C--CC-TGCTTGCTTCAAGTATGG-AA-TGC-CC-CCGAGCCAT-CT-CACATTGGACC-CC--C.................................... 4000
MUS-1.CM.01.CM1239          C--G--C-----CA...AAG.................------C-C-CT-CCACTC-ACAGG--TT----C--T--ACATTAAAGAAACATGG-AAG-GT-CC-CCCAGCCA-AGA-C-CC............................................... 5314
MUS-1.CM.01.SIVmus_01CM1085 C--AC-C---C-CA---.....................------C---CT-CT-CT--ACA-G-CCT---T---T--GCATTAAAGAAGC............................................................................... 5292
MUS-2.CM.01.CM1246          TCAAGAC---CT--CC-AAG.................------A---CT-T--CTC-A-A-AAGCT------CT-TGCATACATCAAACAT---AATGG-GCA-GCGCCTCCAAGT-ACCCT--CTC......................................... 5431
MUS-2.CM.01.CM2500          -GA---C---CCT-CCCAAG.................------A-C-CT-CT-C-C-A-AGAAGTTG--C--CT-CGCATATATCAAACAGGG-GATGG-GCA-C-CCCTCCAAGT-ACCCT--CCC......................................... 5384
MUS-3.GA.09.09GabOI81       C-AA--CG--CTTTCCAAAG.................------A-A-C--TC-CA--ACA--A-CT---C--CT-CCAGT.............................................................................................. 4861
MUS-3.GA.11.11GabPts02      -GCA--C---CTT--GTAAG.................-----CA---CC-TCACAC-A-AGAAGTT---T--CC-TGCATACATAAAGCACT................................................................................ 4882
OLC.CI.97.97CI12            -GA----T---CTTGG-....................--CCTG--GT----G-TGC----TT-GAG----G-C-G-A................................................................................................ 5359
RCM.CM.00.SIVagi_00CM312    -GCG------C-GAG-TCAG.................-----GA----T--G---C-T--TT--CA---T-GGG-TAAATCCCAGTTA...CAGAA--.............................................AT--CAGAA...............AGAGCT 4828
RCM.CM.02.SIVrcm_02CM8081   -GAGG-------GAG-TCAG.................-----GA-A--C------C-A---A--CA---T-GGGCTAGAGAGCAAGTA...-AGAA-C.............................................AT--CAGAA...............AGAGCG 4846
RCM.GA.x.SIVRCMGAB1         ---GG-------GAG-TCAG.................-----GA----G----TT------A--CA---G-GGGCTAAAAGTCAGTTA...GATAGGT.............................................AT--CAGAA...............AGAGCC 5027
RCM.NG.x.NG411              -GAA--------GAG-TCAG.................-----GAGC--------TC-----A--AA---T-GGGCTAGGCAGCAGCTA...-AGG................................................AT--CAGAG......GGCAGAGAAAGAGTG 5057
SAB.SN.x.SAB1               -G-TG------C-GCCAACA.........GGGCAA---CC---AT----G--T--A-----TAGG-T--AT-C-A................................................................................................ 5910
SMM.CI.79.SIVsmCI2          C-T-G-CC-----G-G-CAG.................---CC-AGCT-G----TTC----T--CAGGATTC-CCACAAACAGA.......................................................AT--GA--T......GCAAGAGAGAGAATC 4961
SMM.LR.89.SIVsmLIB1         CGAA--CT-----AT-TCAG.................---CC-AG-T-G--G-T---A---T--AA--T-G--GAGC.............................................................AT-TCAGAT......CCCAGAGAGAGAATC 4913
SMM.SL.92.SIVsmSL92A        C-T-T-CT---C-AT-TAAG.................---TC-AG-T------T------TT--AG---C--GCA-G.............................................................AT-...GAT......CCAAGAGAACAGATC 4912
SMM.SL.92.SIVsmSL92B        CG-T--CC---C--TCTAGT.................--CCA-AGC-----GCTG-----C---AAG-T-G--T--CAAA..........................................................AT-ACAGAC......CCAAGGGAAAGAATC 4901
SMM.SL.92.SL92B             CG-T--CC---C--TCTAGT.................--CCA-AGC-----GCTG-----C---AAG-T-G--T--CAAA..........................................................AT-ACAGAC......CCAAGGGAAAGAATC 5272
SMM.US.04.G078              CGAA--CT--C--A--CCAG.................---CC-AGCT-G--G--TC-A---T--AA--T-G-G-GTT.............................................................AT-TCAGAT......CCCAGAGAGAGAATC 5095
SMM.US.05.D215              CGAA--CT----CG--CCAG.................---CC-AG-T-G--G--T--A-----AA--T-G-GGA-C.............................................................AT-TCA-AT......CCCAGAGAGAGAATC 5089
SMM.US.06.FTq               CGAA--CT-----G--CCAG.................---CC-AGCT----G--TC-A---T--AG-AT-G---GCC.............................................................AT-TCAGAT......CCCAGAGAGAGAATC 5117
SMM.US.86.CFU212            C-AG--CT----GA--TCAG.................---CC-AG---G-----TC-A-----GAGT-T-G---GTC.............................................................AT--CAGAT......CCCAGGGAGAGGATC 5081
SMM.US.x.H9                 CGAA--CT-----G--TCAG.................---CC-AGC-----G--TC-A------A---T-G---GTC.............................................................AT-TCRSRT......CCCAGGGAGAGRATC 5308
SMM.US.x.PGM53              CGAA--CT-----G--CCAG.................---CC-AGC-----G--TC-A------A---T-G-G-GTC.............................................................AT-TCAGAT......CCCAGAGAGAGGATC 5752
SMM.US.x.pE660.CG7G         CGAA--CT-----A--TCAG.................---CC-AGC-----G--TC-A------A---T-G---GTC.............................................................AT-TCAGAT......CCCAGGGAGAGGATC 5824
STM.US.89.STM_37_16         CGAA--CT-----G--CCAG.................---CC-AGCT----G-----A--TT--AA--T-G-GGA-C.............................................................AT-TCAGAT......CCCAGAGAGAGAATA 5482
SUN.GA.98.L14               ---T--CT--C--A..............GGCCTG--TCTC---T-G--G-TTC-CTGCT-GAG--TTC-CC................................................................................................... 5908
SYK.KE.x.KE51               TCCATAA-G-AC-G.......................---CTT--AT-G-----T-----TT--CA-AAG-ATT--AGCAAG..G...................................................................................... 5358
SYK.KE.x.SYK173_COMGNM      C-CATACTG-CC-G.......................---CCC--A--T--G--T--A---T--CA--TG-ACC--AAAG.............................................................................................. 5700
TAL.CM.00.266               G-CACAA-GG--GG..............GACACGCC-CCG--------G-T-C-A--CT-GA-----GC-G-TAGAT............................................................................................. 5517
TAL.CM.01.8023              G-CATAA-GG--GG..............GACTCGCC-CC---CT----G-T-C-A--CT--AAG-T-GC-G-TAGGC............................................................................................. 5015
TAN.UG.x.SIVagmTAN1         -G-A------CGTAGCAGAAAACCATTGGGGTCAG---C-G--C--G--G-TT--A--C--GA---T--AT-C-GACTTCCTAAGAA.................................................................................. 5738
VER.DE.x.AGM3               CC-AG-----C--G--AACA.........GGACAG---CCC---T-G--------A--T--CCTT---CACCA-A................................................................................................ 5243
VER.KE.x.9063               CC-AG--------A-CCACA.........GGAAAG---CCC--A--G--G-----A--C--TCTG--CCATCAGA............................................................................................... 5745
VER.KE.x.AGM155             C--G---------ACTAACA.........GGTCAG---CCC--CT-G-----T--A---T--CT---CCATCA-A................................................................................................ 5740
VER.KE.x.TYO1_patent        C--A---G-----A-GCACA.........GGACAG---CC----T-G--G---C-A--T--GCTC---CATCA-A................................................................................................ 5745
WRC.CI.97.97CI14            -CAA-------GCAGGA....................--GCTTA-CT-G--GAC----TGCT--CAG-TC--GCA-CATGAGCAACTCCAG-G-G--.......................................................................... 5826
WRC.CI.98.98CI04            -CAA-------GCAGGA....................--GCTTA-CT-G--GAC----TGCT--CAG-TC--GCA-CATGAGCAACTCCAGGG-A--.......................................................................... 5841
WRC.GM.05.Pbt_05GM_X02      -TCA-------CTAGGA....................--GCTTA-CT-G--GGC---ATGTT-GCAG--C--GCA-CATGAGCAACTCAGG-AGAG-.......................................................................... 4758
```

Draft of 30 September 2015, not for public release
plvdna_main.tex r8876–10546

Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

```
                                                                                                                                    frameshift insert in HXB2
                                                                                                                                          Vpr premature end (HXB2 only)
H1B.FR.83.HXB2              ATTTGGCTCCATGGCTTAGGGCAACATATCTATGAAACTTATGGGGAT........................ACTTGGGCAGGAGTGGAAGCCATAATAAGAATTCTGCAACAACTGCTGTTTATCCATTTTCAGAATTGGGTGTCGACATAGCAGAATAGGCGTTACTC 5812
Vpr (frameshifted)                                                                                                                                           F  R  I  G  C  R  H  S  R  I  G  V  T
Vpr                         _I__W__L__H__G__L__G__Q__H__I__Y__E__T__Y__G__D__.__.__.__.__.__.__.__.__.__T__W__A__G__V__E__A__I__I__R__I__L__Q__Q__L__L__F__I__H__F__Q__N__W__V__S__T__*
H1O.CM.91.MVP5180           CC----T-A--A-C--GT-----GT-C--T-----G--------A--C.........................-------A------TATG--A--T--------CT-A--------A-------C----.-AT--------A--C-A-----T--------AA--CTC- 5840
H1N.CM.95.YBF30             --------A-----G-----A-----C------A-C--A-----A--C.........................--C----AG--G--A--G--A--T--C--G--A--A------T-A------------.-AT--G-----C--C-AG--C-----------GA-C---- 5405
H1P.CM.06.U14788            CC----T-A--CTCA-----A--GT--------A-T--C--------C.........................--C----AG-----AACT--A--T--T--G--C--A------T-AA-Y----------.-AT--------A--C-A------T--------TA-CTTGA 5282
CPZ.CD.06.BF1167            -A-ACA--A--AAA-A----CA-CTGGG-G-T-C-G-AC-----A--.......................T-A----A------AC---TGC-T----CTC-CT-A---A--GCC-------CT---.-AT---TAC------GCC-------------G....... 5924
CPZ.TZ.00.TAN1              CCAATAT-A--G--GG----AA-TTGGG---TCACC-T------A--C.......................T-C----AG-----AC-G-AAT----C-AG--CT----GAG-GCTT------C---C.-ATC-CCA---T---ATC--C-----------ATCA.... 5469
GOR.CM.07.SIVgorCP2135con   CCC---T-G--CCAA--G--A--G---G-T-----T--------A--C.........................-------A-------AC---A-----C-----CT-A--G---G-AA-C---G-T---.-AT-----------C-AT-------------GA--TTG- 5347
                                                                   Tat exon 1 start

MAC.US.x.239                CGC-T---AAC--CAC-T--TA-T---------A-T-GAC----A--C.........................--CCTT-AG----CA-G--AAC-C--T-----C--C----G-GC---C--C--G---.------GGC--A--CATC--CTC------C---CAAC--G 6406
Vpr                         _R__L__L__T__A__L__G__N__H__I__Y__N__R__H__G__D__.__.__.__.__.__.__.__.__.__T__L__E__G__A__G__E__L__I__R__I__L__Q__R__A__L__F__M__H_#F__R__G__G__C__I__H__S__R__I__G__Q__P
Tat exon 1                                                           M__E__T__.__.__.__.__.__.__.__.__.__P__L__R__E__Q__E__N__S__L__E__S__S__N__E__R__S__S__C__I#__S__E__A__D__A__S__T__P__E__S__A__N__L
H2A.SN.86.ST_JSP4_27        CGC-T---AAT-ACTC-T--CA-CT---------CT-GAC----A--C.........................--CCTT-A---C-CCAG--GGC-C--T--G--C--A---G-GCC--CC-CT-G--C.------GCA--A--CG-C-GCTCA--G--T--TCAGC-CA 5948
H2B.CI.x.EHO                CGC-TAT-AAC--CAC-T--CA-TTT------CA-T-GGC----AA--.........................--CCTT-A-----CA-G--AGC-C--T-A-C-C--C--G-GGGCA--C--CC----C.------GGG--T--C-A---CTCA--G--T--ACAAC-GG 6477
H2G.CI.92.Abt96             CGC-T---AAC--CGC-T--TA-CT----------T-GGC----A--C.........................--CCTC-AG----CA-G--AAC-C--T-A---C-----GAG-GCT--C--CC-G--C.--TC--GGA-AC---A--TGCTC---G--T--A-GG-ACG 5822
H2U.CI.07.07IC_TNP3         -GC-T---AAC--CAC-T--TA-CT----------T-GG-----A--C.........................--CCTT-A-----CA-G--AGC-T--C-A---C------AG-GCC--C---G-G--C.---C--GGA--A---GCG---TC------C--G-GACAAA 5921
H2U.FR.96.12034             CGC-T---AAC--CGC-T--TA-CT--------A-T-GGC----A--.........................--CCTT-A-----CA-G--AGC-C--T-A---C------AG-GCAA-C--CC----C.------GGG--A-----G--CTCA--G--C---AGC-GCA 5966
ASC.UG.10.RT03              GAAGAA----G---AG-TT----G--CTG-CTCTCCCTCCC--CA--CAGTAATTGGACAGAACAGCAGGCA-TGGCA--TACT-CAAT--ATTAC-----G--CG-T---A-GT-AT---GGGAT---.-AT-AG-AA--A--CTAT--C--ATACCAT-CTAGAGTAA 5285
ASC.UG.10.RT08              GAAGAA----G---AG-TT----G--CTG-CTCTCCCTCCC--C---CATTAATTGGACAGAACAGCAGGCATGGGCA--TACT-CAAT--ATTAT----AG--CA-T---A-GT-AT---GGGAT---.-AT-AG-AA--A--CTTT--C--ATACCAT-TTAGAGTAA 5306
ASC.UG.10.RT11              GAAGAA--------AG-TT----G--CTG-CTCTCCCTCCC--CA--CAGTAATTGGACAGAACAGCAGGCGTGGGCA--TACT-CAAT--ATTAC----AG--CA-T---A-GT-AT---GGGA---C.-AT-AG-AA--A--CTAT--C--ATACCAT-CTAGAGTAA 5294
COL.CM.x.CGU1               -C-CACA--TCA-TGA--AT-TGGGG-TGTAC-AG--GGC-CT---GGAGA.....................C-G---A-G-CT-GTCTG--TTATG-G-----CA--G-GTGTGCC--C-ACT-G--C.G-TCATGCG---CAAT-T-GCTTTCCCTCTACGCAA-AC- 5511
COL.UG.10.BWC01             -C-CATA--TCA-TGA--AT-ATGGGATGTGTAAG--GGC--T--CCAGAC.....................CAA-----T-CT-CTAT---TTATG-G--G--CT--G--AGTGCCT---ACT-A---.A-TCATGCA---CAAT-TA-ATTTCCTTCCA-AACG-ACG 5313
COL.UG.10.BWC07             -CCCATA-AGCG-T-A-TT--TGGGG-TGTAC-AGG-GAC--T--CCAGAC.....................CAA------ATT-G-ATG--TTATG-GC-G---A--G--AT-GCC--C-A-T-A---.C-GCATGCG--TCAAT-CA-ATTTCCTTCCA-AACC-ACA 5538
DEB.CM.04.SIVdeb04CMPF3061  GACGAAT-AAT----G-CT--A-TT--TGTGTA--GTTACCA-CA--GCCAGGGTGGACAGGAGAACAGGCTTGGGCA--CTC---AAT--A-TAC----AG-AGG-C--GA--A--A-T-GGC-A---.C-GGC-GAA------TAT-----A--GGC--CA-C-T-AA 5800
DEB.CM.99.CM40              GAAGAA--ATT---AG-TT--A-T---TGTGTA--CTTACCC-CCC-ACCAGATTGGACGGGAGAACAAGCATGGGCT---TCG--CATT-A-TAC----A---AG-T---A--A--T-A-GGT-G--C.C-G---GAA-CT--CTTT--C--GGAG-G--AG-C---CA 5619
DEB.CM.99.CM5               GAAGA-T-ATGG--AG-CT--A-T--CTGTGTA--CCT-CC--CCC-ACCAGATTGGACAGGAGAACAAGCATGGGC---GTC---AATT-A-TAT----A---AG-A---AG-A----T-GGC-T---.C-G---GAA-CA--CTTC-----AGAG-GG-AA-C----A 5613
DEN.CD.x.CD1_CMO580407      GAAGA-A--AC---AA--T--G-T--CTG--GGTC-TTACCA-C-CTCCCACATTGGACAGAAGGACAAGTC-TGGCA--C-CC--CAT--ATT-C---------CA----GA--GAAA-A-GG-AG---.-AT---G-A--A---TTC-----GGA-GC--AAAGAGT-A 5742
DRL.DE.11.D3                CGC-TAT-G---AATG----ATC-TGGG---------GA-----A---........................T-A-TA-A-------C-GCAGC-T--------CT-----AG-GCA--T--CT-G---.--T---TA---A--C-AGG-AG-T---T-T--A...C-AG 5606
DRL.x.x.FAO                 CGC-TAT-G---AATG-----TC-TGGG---------GA-----A---........................T-A-TA-A-------C-CCAGC-T----T-C-CT-----AG-GCA--T--CT-G---.--T---TA---A--C-AGG-A--T---C-T--A...C-AG 5702
GRV.ET.x.GRI_677_gri_1      -ACAT---AATCC-AG--T--A-TT-CTGTGTA--GGAGGGAA--AGA............CATAATACCCCATGGAAT-AGAT--GCT-CAAGTACTAT------G-T---A-GTCTA-----G-A---.------TG---T---A--AGG--AG--CCTTTTTCCC--T 5977
GSN.CM.99.CN166             GACGA-----G---AA-TT----TG--G-GACG--GCTCCC--CC---CCAGATTGGACAGTGGATCAGGCAG-AATA---T-T-CAAT--ATTAC--C...GGG----AC---T--C----GA--C.--T-A-GA---A--CTTT--C--GGTC-ATTTATCA...G 5622
GSN.CM.99.CN71              GAAGA-----G---AA-TT----TG--G-AACG--GCTCCC--CC---CCGGATTGGACAGTGGATCAGGCAG-AATA---T-T-CAAT--ATTATG-C--G-GAG---GAC-T-A--C----GA--C.--T---GA-------TTC--C--ATAC-ATA-AA--GTCA 5610
LST.CD.88.SIVlhoest447      -CAATTT-GGCCCAAG-CTCAG--T-CTGTGTCAGC-G-ACC---TCAGAG.........................GA-AA--CCTGCATGAAAT-C--T-CC--AG-AA-TAG-GCA--C-----T---....TTGCC-ACC----CCAGCTATGC-GC--CAAGA.... 5050
LST.CD.88.SIVlhoest485      GCCATTT-GGCCCAAG-CTC-G--T-CTGTGTCAGC-G-ACC---TCAGAG.........................GA--A--CCTGCATGAGAT-C--T-CC--AG-AA-TAG-GCA--C-----T---....TTGCC-ACC----CCAGCTATGC-GC-ACAAGA.... 5050
LST.CD.88.SIVlhoest524      GCAATTT-GGCCCAAG-CTC-G--T-CTGTGTCAGC-G-ACC---TCAGAA.........................GA--AG-CTTGCATGAGAT-C--C-CTT-AG-TA-TAG-GCT--C-A-T-A---....CTGCC-CAA----CCAG-TATGC-GC--C-AGA.... 5047
LST.KE.x.lho7               G-AAT-T-GGCCCAAG-CTC----T--TGTGTA--C-G-ACA---TCAGAA.........................GA--AG-CCTGCATGAAAT-C--C-CCC-AG-CA-TAG-GCTT-----C-T--C.C--CC-...ACT----CCAGCTATGC-...-CA-CC-GA. 6131
MAC.US.x.251_1A11           CGC-T---AAC--CAC-T--TA-T---------A-T-GAC----A--CACC.......................CTT-AG----CA-G--AAC-C--T-----C--C--G-GGGCA--C--C--G--C.------GGC--C--CATC--CTC------T---CAAC--G 6405
MND-1.GA.x.MNDGB1           ---ATC--AA---CAG--TCAG--T--TGTGTGC---ACACA---AG-GAG.....................GAAGA---CT-T-A-A--TTT--T-CC...T-AA--A-TAG-GCCA-T-GGG----C.C-AGCTCAA------GATGGA-C-TTC...........A 5482
MND-2.CM.98.CM16            CGC-TAT-G----CAG----AAGCTGG-----CA----C---------.........................--A-TA-A--G---CG-CAGC-T--CTC------T---A--GCT----C--G--C.-AT---TA---A--CA-GG-A---C-T--C--ACAAGAAG 6086
MND-2.GA.x.M14              CGC-TAT-G----CAG----AAGCTGG--T--------A--------C.........................--C-T--AG--G---C-GAAGC-T--C-CCC-CT-A--GAG-GCT--A---C-G----.--T---CA---A--CA-GG-A---C----T--ACAAGAAG 6013
MND-2.x.x.5440              CGC-T---A----CAG----CTC-TGGG-G-----GCAAC-A--A--C.........................----TA-A---T--CC--CAGC-------C---T-----AG-GCA-----CT-A--C.--T---TA---A---A-GGGA--TC-C--C--ACAAGAAG 5651
MNE.US.x.MNE027             CGC-T---AAC--CGC-T--TA-TT----------T-GAC----A--C.........................--CCTT-AG----CA-G--AAC-C--T-A---C--C----GGGC---C--C--G--C.------GGC--C--CAAC--CTCT-----C---CAAT--G 5874
MON.CM.99.L1_99CML1         GAAGA-----G----A-TT--G----GG-AACG--GCTGCC--CT--CCCGGCCTGGAATGCAGACCAGGCATGGGCA---T-T-CCATT-ATTAT-CT---TGGG-C---ACCA-C--T-A--GG--C.-A---GGAA--C---TAC-----ATATGC--AGCAA-TAA 5636
MON.NG.x.NG1                GAAGA-----G----A-CT--G-T--GG--AC----TTGCC--CC--CCCAGGGTGGAACGCAGACCAGGCATGGGCA---T-Y-CTATT-ATTAC-CC--GCGGG-----AC-A-RA-A----GG--C.A-G--GGA-------TAT--C--GCATGCCC-A--GRT-- 4247
MUS-1.CM.01.CM1239          GAAGA-----G---AA-CT--A-TG--G-AACG--GCTACCA-CC---CCTAATTGGACAGTGGATCAGGCTG-AATA---T-T-CCAT--ATTAC--T--G-GAACA---AC---C--T--C-GG---.-ATC--GAG--A---TAT-----GTACTCTAA-ACA-TCA 5574
MUS-1.CM.01.SIVmus_01CM1085 GAAGA-----G---AA-TT--A-TG-GG-AACG--GCTGCCA-CC---CCTGATTGGACAGTGGATCAGGCGG-AATA---T-T-CCAT--A-TATG-T--G-GAG-----AC---TG-C----GG---.------GAC--A--CTAT-----GTAT-ACA-A--A-TGA 5584
MUS-2.CM.01.CM1246          GACAA-----G---AA-TT--G-TAGAG--ACA--GCTCCCA-CA---CCACAATGGTCTGTAGATCAGGCAG-CATAA--T-T-CCATT-ATTAC--T-A-TCAG-C---ACC--CT----C-GA--C.-AT--GGAC--A---TAC------TATGCTCAAACA-TCA 5691
MUS-2.CM.01.CM2500          GACGA-----G---GA-TT--G-C-GAG--ACG--GCTCCCA-CA--CCCTAATTGGACTGTTGATCAGGCAG--ATA---T-T-CCAT--ATTAC--T---TGGG-A--GGGT--C--C--C-GG---.---C--GAG--A--CTAC-----GTAT-GTCAG----T-A 5644
MUS-3.GA.09.09GabOI81       GAAGA-----G---AA-TT--G-TAGAG-AACG---CTCCC--CT---CCACAATGGACAGTGGATCAAGCTG-AATC---T-T-CAAT--ATTAC--T---TGGG-C---ACCA-TA-C----GA---.-A----GAG----CTAC-----ATATGC-AA-ACC-TA- 5165
MUS-3.GA.11.11GabPts02      GAAGA-----G---AA-TT--A-CG-CG-AACG--GCTACCC-CT--CCCAGATTGGTCAGTGGACCAAGCAG-AATA---T-T-CTAT--ATTATG------GAG-AA-CACTT--A-A--C-GA---.--T--GGAA--A---TAT--C--ATAT-ACCAGA-C-TAA 5174
OLC.CI.97.97CI12            GAG-TAT-GGTCC-AG-TTTT-T-TTCTGTGC-C---GAC--ATCA--.....................GATGAGAT-AGTACCTCTTTGCTGT-C--T--GC-A--A---AG-GCCT-----CAA---.-AT----A---A--C..................CCAGGAG 5600
RCM.CM.00.SIVagi_00CM312    GAGCTTT-G--G-CAG----AA-TT-C-GT-----GCAGC----A--A.......................-G-ATA-A---G--TA-GAAA--G----CTT-GT--A-TAG-GCTT-----T-G--C.-AT---CA---A--GTC-GGA-----GG-T--AACGG-AA 5405
RCM.CM.02.SIVrcm_02CM8081   GAGCTTT-A----CAG-----A-TT-C--A-----GCAA---------........................-GC-TA-AG-----TAG--TAC----T-CTT-G--T---AGGGCT-----CT-G--C.-AT---CA---A---ACGGGC--T--G--C--AACG-A-A 5423
RCM.GA.x.SIVRCMGAB1         GC--T---AACGCAAA----AG-TT--G-A-----GCAAC--------........................-GCATA-AG------C-G--A--G--TC-GC-G--T---AG-GCCT-----C-T---.--T----A---A---GC-GGG--T--G--T--AACAT-CA 5602
RCM.NG.x.NG411              GAA-T---TACACAAGA---AA-TT----A-----GCAGC----A---........................-G--TA-AG-----AA----T--G----CCT-G--TA-TA--GCA--T--CT-G--C.-----GCA---A---GA-GGG-----G--T--A-GAG-GA 5637
SAB.SN.x.SAB1               -A-CTC--TTTCC-TC-TT--TGGA----TGTG---GAACCA-CCAT-.........GATCATGGGCAAACC-GGCT--A---CTG-T-TAAATATTGT-----------A-GGCTT-A---G-G---.A-G-A-GG-A-A---T-TA-GCCA-AG-CTCA-CC-G-GT 6253
SMM.CI.79.SIVsmCI2          CG--T---AACC-CGC-T--TA-TT----------T-GGC----A--C.........................--CCTT-A-----CA-GG-AAC-T--C-AG--C-------G-GCA--C--------C.------GCA--A----TT--CTC------T--GCAC-G-G 5532
SMM.LR.89.SIVsmLIB1         CGC-T-T-AA-C-CGC-T--CA-TT----------T-GG-----A--C.........................--CCTC-A-----CA-GG-AGC-C--T-AG--C-------G-GCA--C--C--G--C.------GGA--C---A-G--CTC------T---CAGT-AG 5480
SMM.SL.92.SIVsmSL92A        CGC-T---AAC--CTC-T--TA-CT----------T-GG-----A--C.........................--CCTT-A-----CA-G--AAC-T--T-A---CT---GAG-GCC-----CC----C.------GGA------AC-A-CTCA-----T--ATGC-GAA 5482
SMM.SL.92.SIVsmSL92B        CGC-T---TAC--CT-----AA-TT-C--A-----T-----C------.........................--AATA-A---G-CA-GG-AA-----T-AG--AT-A---A-GGCAT----CT-A--C.--T--GCA---A--CACC---TCA--G--T--ACAGC-AA 5468
SMM.SL.92.SL92B             CGC-T---TAC--CT-----AA-TT-CG-A-----T-----C------.........................--AATA-A---G-CA-GG-AA-----T-AG--AT-A---A-GGCAT----CT-A--C.--T--GCA---A--CACC---TCA--G--T--ACAGC-AA 5839
SMM.US.04.G078              CGC-T---AAC--CGC-T--TA-TT----T-----T-GAC----A--C.........................--CCTT-AG----CA-G--AGC-C--T-A---C-------G-GCC--C--C--G--C.-------A---A---GTC--CTC-C----T---CAA-GCG 5666
SMM.US.05.D215              CGC-TA--AA---CGC-T--TA-TT----------T-GAC----A--C.........................--CCTC-A-----CA-G--AAC-C--T-AG--C--T----G-GCT--C--C--G--C.-----GGGC--A--C-A---CTC------T---CAAC--G 5660
SMM.US.06.FTq               CGC-T---AAC--CGC-T--TA-TT----------T-GGC----A--C.........................--CCTT-AG----CA-G--AGC-C--T-A---C--A---AG-GCA--C--CC----C.------GGA--C--CACC---TCA-----C---ACA-A-A 5688
SMM.US.86.CFU212            CGC-T---AAC--CGC-T--TA-CT-C--T-----T-GGC----A--C.........................--GCTC-AG----CC-G--AGC-C--T-AG--A--C--G-G-GC---C--C-----C.------GG---A---GCC--CTCA-----C--TCACC-AA 5655
SMM.US.x.H9                 CGC-T---AATY-CGC-T--TA-CT----------T-GAC----A--C.........................--CCTT-A-----CA-G--AGC-C--T-----C--C--GGG-GCA--C--C------.------GG---A--CA----CTC------C---CAGT--G 5880
SMM.US.x.PGM53              CGC-T---AAC--CGC-T--TA-CT----------T-GAC----A--C.........................--CCTT-A-----CA-G--AAC-C--T--G--C--C----G-GC---C--C--A--C.------GGA--A---GCT--CTC------C---CAA-G-A 6325
SMM.US.x.pE660.CG7G         CGC-T---AAC--CGC-T--TA-CT----T-----T-GAC----A--C.........................--CCTT-A-----CA-G--AAC-C--T-----C----GAG-GCA--C--C------.-------G---A---GCC--CTC------C---CAAT--A 6396
STM.US.89.STM_37_16         CGC-T---AAGC-CGC-T--TA-CT----------T-GAC----A--C.........................--CCTT-A-----CA-G--AGC-C--T-AG--C--C--G-G-GCC--C--C--G--C.------GGC--T--C--C--CTC---G--T---CAAC-AG 6054
SUN.GA.98.L14               GCAATAG-TGCCCAAG-CTT----T-CTGTGCG--CT--ACA---TCA.......................GAAAA--A---GTGTATGAGAGC---C-CTT-G--CA-TAG-GCCT----C--T---.C-GCC-...AT-----CT....................... 6157
SYK.KE.x.KE51               GAAGAAA-ATT---AG-TT--A-T--ATG--TA--GGTGG-A-CA-G-CCAGAAGAAACACCCACCATGGCTTGGATTAG-T-CA--CT--AT--GCAT---GCA--CA-CTTTA----T---GAA---.--TGCTGCA--A----C---G--G-C--AGTA--CG-GA- 5680
SYK.KE.x.SYK173_COMGNM      CAAGAAA-AT----AG-TT--A-T--ATC--TAA-CGAAG-G-CA-GGACTGAT...TCCCCGACAATGGCTTGGGAAAG-ACCA--TTG-A---GG-----GCA--CA-T-TTA----C---GAA---.--TGCTGCA--A----CT--AC-G-CC-G-TAT-CC-GA- 6016
TAL.CM.00.266               GACGA-G--AT---AG-CT--A-T---TGTGTA---TTGCCA-CA--CCCTGATTGGACAGGAGAACAGGCATGGGCT--CTC---CAT--ATT-TGCT--G-AGGGTA-T---A--T-AC-CC-A---.---CAGCAA-----CTAT---GAGCAGCA-A-AAGA---- 5809
TAL.CM.01.8023              GAAGA-G--AT---AG-TT--A-T---CTGTGTA--GTTACCA-CA--GCCTGATTGGACAGGCGAACAAGCATGGGCT--TTC---CAT--ATT-TGCTC-G-AGGGTA-CG-GA--T-A--CC-A---.---CAGCAA-----CTTT---GAGCAGCA-A-AAGA---G 5307
TAN.UG.x.SIVagmTAN1         GAACTC---TTCCAAG--T--A-CTTCTGTC-G---GAAGGG-AAAGG............AATGGGGCACCA-TGATA-A-A-G-CTT-TAGATATTAT--GT-GG-A--GA-GGCTT-A---G-G---.--TC--TG---C--CA-GAGA--A-CTCCTTTT-AGC-CT 6033
VER.DE.x.AGM3               GAGCTC--GTTCCAAG--T--A-CT--TGTC-G--GGAGGGA-AAAGG............AATCGCACTCCC-TGCTA-A-A-G-CTT-TAAATATTAT-A-T-GG-----A--GCT--C---G-G---.---C-GTG---A--C--CAGA--ACA-CCCTTT-AAC-AT 5538
VER.KE.x.9063               GAGCTC---TTCCAAG-TT--A-TT--TGTC-G--GGAAGGA-A-AGA............CGGGGAGCCCCA-TGAC--AGA-G-CCT-TAGATACTAT--GT-GG-T--GA--GCA--C---G-T---.---C-TTG---T-----TAGA--ACAGCCCTT--AAC--T 6040
VER.KE.x.AGM155             GAACTC--AT-CCAAG--T--A-TT--TGTC-G--GGAAGGG-A-AGA............CAGGGAAGACCC-TAGC--A-A-G-CAT-TAAGTATTATC-CT-AG-T--GA--GCT--C---G-G---.---C-GTG---A-----CAGG--ACA-CCCTTT-AGC-AT 6035
VER.KE.x.TYO1_patent        GAACTC--GTTCCAAG--T--A-CT-CTGTC-G--GGAAGGA-AAAGA............CATGGTACTCCC-TGAT--A-A-G-CCT-CAAATATTAT--GC-AG-A---A-GGCT--C---G-G---.--TC--TG-----CA-GAGA--GCAGCCCTTT-AAC-AT 6040
WRC.CI.97.97CI14            GAAGTC--TGTAA-AG--TTT-TTT--TG--C-C--.....................CAGAGAGAGTATGAGGAAGTCC-TACTA--C-TTATT-CCAC--GG-A------AG-GG---C--C--G---.C---CCCAA-------CTA-ACAACA-GAGAAAAGGGAA- 6062
WRC.CI.98.98CI04            GAAGTC--TGTAA-AG--TTT-TCT-CTG--CGC--.....................CAGAGAAATTATGAAGAAG-TC-CACTA--C-TTATT-TCAC--GG-A--A---AG-GG---C--C--G--C.C-T-CCCAA-------C-A-ACAACA-GA-AAGAGAGAA- 6077
WRC.GM.05.Pbt_05GM_X02      GAAGTC--TGTAA-AG--TTT-TCT--TG--CGC--.....................CAGAGAAATTATGAAGAACA---CACTA--C-CTATT-TCAC--GG-C--T---AG-GGAT-A-A-C-A--C.C---T-CAA--A----CCA-AGAA-A-GAGAA-AGAGAAA 4997
```

Draft of 30 September 2015, not for public release

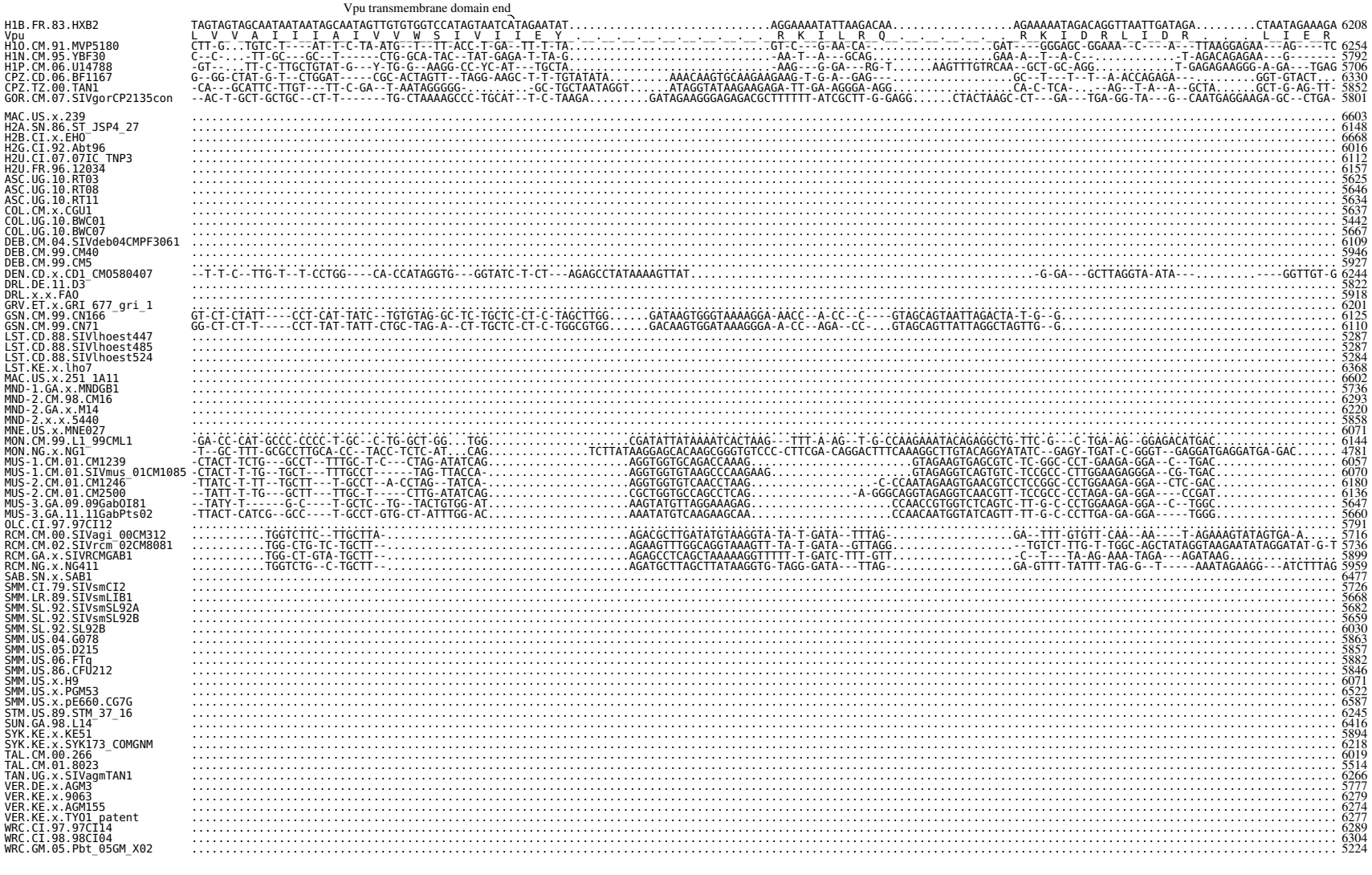
plvdna_main.tex r8876–10546

```
                                                                      Tat exon 1 start  Vpr end
H1B.FR.83.HXB2              GA..............CAGAGGAGAGCAAGAAATGGAGCCAGTAGATCCTAG....................................................................................ACTAGAGCCCTGGAAGCATCCAGGAAGTCAG 5881
Vpr (frameshifted)          R__.__.__.__.__.__Q__R__R__A__R__N__G__A__S__R__S__*
Tat exon 1                                                   M__E__P__V__D__P__R##__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__._L__E__P__W__K__H__P__G__S__Q_
H1O.CM.91.MVP5180           C-TCT...AACACAAGAGGA--AG--AG----------T-----------GA....................................................................................GA-GCCC--T---C-T--C--T--G--CA-- 5921
H1N.CM.95.YBF30             CT..............--A...--GAG---G------A--------------....................................................................................-T-------------T-----------C--A 5471
H1P.CM.06.U14788            C-CCC.........TCT-GA--AG--AGG---C-----C-----------GA....................................................................................T--GCCT--A---C-A--G-----G------ 5357
CPZ.CD.06.BF1167            ................ATAG-ACC-AGG--G------------------AGA....................................................................................T--T-----T---TTA--C---A-CTC-ACA 5991
CPZ.TZ.00.TAN1              .....................................................................................................................TGAATCCCATAGATCCTCAGG---CA--A---G-A---------GC-GCA 5519
GOR.CM.07.SIVgorCP2135con   C-.........CCCTCTAGA--AG--AG----------C---AC------GA....................................................................................G--GCCA--A---C----G-----G--C--- 5422
                                                                                 Vpr end
MAC.US.x.239                -GGGAGGA.........A-TCCTCTCT--GCT-TACCGC--TC----AG-AT-CTATAA............................................................................................................ 6456
Vpr                         G__G__G__.__.__.__N__P__L__S__A__I__P__P__S__R__S__M__L__*
Tat exon 1                  _G__E__E__.__.__.__I__L__S__Q__L__Y__R__P__L__E__A__C__Y__N##__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__._
H2A.SN.86.ST_JSP4_27        -GGGACGA.........A-TCCTTT-T--GCT-TACC-A--CC----GG-AT-CGATAA............................................................................................................ 5998
H2B.CI.x.EHO                -GGGAGGA.........A-TCCTCTCT--GCT-TACCGC--TC-TA-GG-AT-CAGTAA............................................................................................................ 6527
H2G.CI.92.Abt96             C-GGCGGA.........A-TCCTCT-T---CT-TACCGC--TC----GG-GT-TTCTAA............................................................................................................ 5872
H2U.CI.07.07IC_TNP3         --GGAGGA.........A-TCCTCTC---GCTGTACCGC--TC----AA-GT-TTATAA............................................................................................................ 5971
H2U.FR.96.12034             -GGGAGGA.........A-TCCTCTCA---CTGTACCGC--TC-T--GG-GT-CTCTAA............................................................................................................ 6016
ASC.UG.10.RT03              -GAGGTTTCCACAACTGAGACCACTGAG-G-T-CA-C--GACAGGACA-GCCC......CATGACAGTGATCCGCCACAGGTGCTAAGACCCTCTCGTTATCATCCAGATAACTAATGGATCCATCTCAAGAGGGGTTAGAA-A-T...G-C-TAGGG-GCCTGCACA-C 5446
ASC.UG.10.RT08              --AGATTTCCACAGCTTAGACCACTGAG-G-T-CA-C--GACAGGA-A-GCCC......CACGACAGTGATCCGCCACAGGTGCTAAGACCCTCTCGTTATCATCCAGATAATTGATGGATCCATCCCAAGAGGGGTTAGAA-A-T...G-C-TAGGG-GCCTGCACA-C 5467
ASC.UG.10.RT11              AGAGGTATCCACACCTAAGACCACTGAG-G-T-CA-C--GACAGGATA-GCCC......CATGATAGTGATCCACCACAGGTGCTAAGACCCTCTCGTTATCATCCAGACAATTGATGGATCCCTCCCAAGAGGAGGTAGAA-A-T...G-C-TAGGG-GCCTGCACA-C 5455
COL.CM.x.CGU1               CT.................................................TAG................................................................................................................. 5516
COL.UG.10.BWC01             CCCCC..............................................TGA................................................................................................................. 5321
COL.UG.10.BWC07             ACCCT..............................................TGA................................................................................................................. 5546
DEB.CM.04.SIVdeb04CMPF3061  --CAT...TACCCTAACATC--ACC-CTGTC-GGAA--AATCAGGAGGTTAG-GATCAGGAATAA.................................................TCATTTTCCTCTTACAGAAG.ATGGAG---GAAAT-G-C-TATTCCA-G-GAGA 5915
DEB.CM.99.CM40              --GGG...TATCCTAACATT---CC-TTG-C-GGAA--AAT--AGA-GTTAG-GACGGAGAATAA.................................................CCTTTCACCTTTGCTTCAG..ATGGAGAGCATAGATCCATT-GA--A---AGCT 5733
DEB.CM.99.CM5               -GAGA...TACCCCAATATT--ACCCCT--C-GGAA--AA---GGAGGT-AG-GATGGAGAATAA.................................................CCTCTCTTTTACATCATAG..ATGGA---AATAGATCCATTCAAG-A--CAGCT 5727
DEN.CD.x.CD1_CMO580407      --CAC........TAT-CA-AC-T-AG-CCCTTAC-CC-TC-C---GAAG--CCATGA................................................................................AC-C-GA-A-C-TAG-C--CTTCGC-GGC 5821
DRL.DE.11.D3                C-GGACGCTGG......A-TCCATTTCG--CCTT-CC-AGGGCG-ACCAAC-ATTGTAA............................................................................................................ 5659
DRL.x.x.FAO                 C-GGACGCTGG......A-CCCATTTCG--CCTT-CC-AGGGCG-ACCAAC-ATTGTAA............................................................................................................ 5755
GRV.ET.x.GRI_677_gri_1      ACGAAGAGAGGAGAAATGGACAAG---G-G--....-CCC-ACCCCCT...............CCTCCAGGACTTGCATAG...................................................................................... 6039
GSN.CM.99.CN166             -GAGATACCCAGCAATTAG-CCCTC-AGGG-C-CC-CCC-ACCAGACAGTA-CTCTGTTTCTCATGCCGATCCTGAGCAGCCTAGGAGACCCTCAAGATATCGCATGGATGAATAG.....AGACCCGGAAGAGTGTCTCT-GCCCAA----C----G------GCGGC- 5787
GSN.CM.99.CN71              --AGGTACCCAGTAATCAG-CCCTT-AGGG-T-CC-CCC-ACCAGATAGTA-CTCTGTTCCTCATGCTGATCCTGAGCAGCCCAGAAGACCTTCAAGATACCGCATGGATGAATAG.....AGACCCGGCAGAGTGTCTTT-GCCTAAG---C-A--G------GCCGC- 5775
LST.CD.88.SIVlhoest447      .....GCAGGAACAGCT--CCC--C--A-GC-GCACC-AGACCC-CTAAAAGAGACATACAAAGAGGCCGTTACTAA.......................................................................................... 5122
LST.CD.88.SIVlhoest485      .....GCAGGAACAGCT--CCC--C--A-GC-GCACC-AAACCC-CTGAAAGAGATATACAAAGAGGCCGTTACTAA.......................................................................................... 5122
LST.CD.88.SIVlhoest524      .....GCAGGAACAGCA--CCCA-C--A-GC-GCACCTAGACCA-CTGGAAGAGATATACAAAGAGGCCATTACTGA.......................................................................................... 5119
LST.KE.x.lho7               .....GCAGGAACAGCC--TCCA-C--A-GC-GCACC-AGACC--ATGAAGGAGACATACAACGAGGCAGTGACTAA.......................................................................................... 6203
MAC.US.x.251_1A11           -GGGAAGA.........A-TCCTCTCT--GCT-TACCGC--TC----AG-AT-CTATAA............................................................................................................ 6455
MND-1.GA.x.MNDGB1           -GGAAAGAAGACCACAA-T-CCCCCCT--G--TTCA-GC-A--AG--GATAGATTATAA......................................................................................GCAACT--TAGA--AGTA-T-T 5563
MND-2.CM.98.CM16            --GGGGA......TGTT-CCCCTT-CG-TCCTT-CCGAGATCGGACAA-CCCCTGTGA............................................................................................................ 6139
MND-2.GA.x.M14              --GGCAAA......TATA-CCCCCTTAG-TCCTT-CC-AGGCCG-ACAA-CCCTTGTAA............................................................................................................ 6066
MND-2.x.x.5440              --GGAAGA......TATA-CCCACTTCG-TCCTT-CCGAG-CCG-ACAA-CCCTTGTAA............................................................................................................ 5704
MNE.US.x.MNE027             --GGAGGA.........A-TCCTCTCT---CT-TACCGC--TC----AG-AT-CTATAA............................................................................................................ 5924
MON.CM.99.L1_99CML1         --AGATACCCTGTTCTTAG-CCA-TGAG-G-G-C--CTC--G-CCC-A--AGTTCTGTTCCACAGGCCGATCCAGACAATCCCAGAAGACCTTCCAGATATAGAATGGATGAATGA....TGGAACCTGTTGACCCAGACT--CCCAAAGAAC-A--C--CCC-GCAACC 5802
MON.NG.x.NG1                -CAGGTACCCTYGRCTG-GACCACT-AG-G-C-C--CCC-AG-ACCCA--AGTTCTGTCCCTCATGCAGACCCAGAGAGGCCTTTGCACCCGTCCAGATACCACCACGATGAATAG..TGCAASGATAGTGAACCCAGC-T--CCC---GA-C-------CTC-GCGACA 4415
MUS-1.CM.01.CM1239          -GAGATACCCCAACATAAGACCCTTGAG-G-G-CACA---TCC-CCT--TA-TTCAATGCCTAATGCTGATCCTACTCCTCCTTTGAGACCCTCTAGATATAGGATGGATGAGTAA....TGGATCCTTCAGTAGAAAAC--TCCC--GGAAC---GC--C--CGCAGC- 5740
MUS-1.CM.01.SIVmus_01CM1085 --AGATATCCAACCATAAG-CCTTTGAG-G-C-CACA-C--CC-CCT--TA-TTCAACCCCAAATGCTGACCCTGTACCTCCTTTGAGTCCTTCTAGGTACAGGATGGATGAATAA....TGGATCCATCAGTAGAGGAGT-GCCAAAGGAAC-A-GA--G--TGCCGCA 5750
MUS-2.CM.01.CM1246          -GAGATACCCACCGTTG-GCCCATTGAG-G-G-CGCAGC--CCAGAC--TA-TTCCATGCCAAATGCTGACCCCACACCATCTTTGAGACCCTCCAGATATAGAATGGATGAATAA....TGGACCCTTCAGTTGAAGGC--GCCT--GGACC---G---C---GCGGCA 5857
MUS-2.CM.01.CM2500          --AGATACCCACAGGTG-GCCCCTT-AG-G-G-CGCA-C-TCCAGAC--AA-TTCAACACCAAATGTTGACCCCTTTCCTCCACGAAGACCCTCCAGGTACCGCATGGATGAGTAA....TGGATCCAGAAGTAGAGAATT-GCCC--T---C--AGA---TCCTCAGCT 5810
MUS-3.GA.09.09GabOI81       AGAGATACCCCATACTTAGACCACTGAG-G-G-CACAGC-ACCAGA---TA-TTCTATGCCTCATGCAGATCCGACTCCCCCTTTGAGACCATCCAGGTATCGCATGGATGAATAA........CCTAGAGGAAGAAAAC--GCCT--G---C--AGAGGCACTGC-GCA 5327
MUS-3.GA.11.11GabPts02      --AAGTACCCCATAATAAGACCATT-AGGG-C-C-ATGC-TCCAGATA-TA-CTCCACCCCTCATGCAGATCCCGAAATTCCAAGAAGGCCATCAAGATACAGAATGGATGAGTGA....TGGAAGTGGTACCGCCAGATAC-CCTAA----C-A--C--CCCGGCAACA 5340
OLC.CI.97.97CI12            TGGAGCTTAGAAGA...............-ACCC-CACAGA--ATCTGAAGGAATGC.............................................................................................................. 5642
RCM.CM.00.SIVagi_00CM312    --GGAAGA.........A-CCCTCT-CA-TCC-T-CC-AT-------AA-ACCTTGTAA............................................................................................................ 5455
RCM.CM.02.SIVrcm_02CM8081   --GGAGGA.........A-CCCTCT-CA-TCC-T-CC-TTATCA---AA-ATCCTGTAA............................................................................................................ 5473
RCM.GA.x.SIVRCMGAB1         --GGAAGT.........A-TCCTCT-CG-TCC-T-CCGCAA-CG---AA-ATCATGTAA............................................................................................................ 5652
RCM.NG.x.NG411              --GGAGGA.........A-TCCTCTTAG-TCC-T-CTTCA-TCG---AA-ATCTTGTAA............................................................................................................ 5687
SAB.SN.x.SAB1               ATGGACCAGGAGCAGGAGGCCC-CCCC--G-TCTG----GAGC-TC-GGAGGAGCTGCATCGGCCGCTCCAGGCCTG...TGA................................................................................... 6333
SMM.CI.79.SIVsmCI2          -GGGAGGA.........A-TCCTCTCT---CTGTACCGC--TC-C--GGAGT-CTCTAA............................................................................................................ 5582
SMM.LR.89.SIVsmLIB1         --GGAAGA.........A-TCCTCTG----CT-TACCGC--TC----GGAGT-CTTTAA............................................................................................................ 5530
SMM.SL.92.SIVsmSL92A        --GGAGGA.........A-TCCTCTC----CT-TACCGC--TC----AA-GT-CTGTAA............................................................................................................ 5532
SMM.SL.92.SIVsmSL92B        -GGGACCA.........A-TCCCCTG-G-TCT-TACCGT--GC----AA-GT-TTGTAA............................................................................................................ 5518
SMM.SL.92.SL92B             -GGGACCA.........A-TCCCCTG-G-TCT-TACCGT--GC----GA-GT-TTATAA............................................................................................................ 5889
SMM.US.04.G078              --GGAGGA.........A-TCCTTTG----CT-TACCGC--TC----GAAGT-CTGTAA............................................................................................................ 5716
SMM.US.05.D215              --GGAGGA.........A-TCCTCTG----CT-TACCGC--TC----GG-GT-CTATAA............................................................................................................ 5710
SMM.US.06.FTq               --GGAGGA.........A-TCCTCTTT---CTGTACC-A--TC----AA-GT-CTGTAA............................................................................................................ 5738
SMM.US.86.CFU212            --GGAGGA.........A-TCCTTTG----CT-TACCGC---C----GATGT-TTGTAA............................................................................................................ 5705
SMM.US.x.H9                 --GGARGA.........A-TCCTCTCT---CT-TACCGC--K-----GG-GT-TTATAA............................................................................................................ 5930
SMM.US.x.PGM53              --GGAGGA.........A-TCCTCTGT---CT-TACCGC--TC----GG-GT-CTCTAA............................................................................................................ 6375
SMM.US.x.pE660.CG7G         --GGAGGA.........A-TCCTCT-T---CT-TACCGC--TC----GA-AT-CTATAA............................................................................................................ 6446
STM.US.89.STM_37_16         --GGAGGA.........A-TCCTTTG----CT-TACCGC---C----GGAGT-TTGTAA............................................................................................................ 6104
SUN.GA.98.L14               ..............TCTT-C-TC---ATC----G----T-AG-A-ATGT-A-CCCAAGGCCACCAACAAGACCAAGACCAGGGCAAGGGGACATTAGAAGAGGCCTACAAGACTAA................................................... 6259
SYK.KE.x.KE51               ACAGGGGATACCCACAT-CTTCTTAG.........................................................................................................................AAGGAACATTTCTCTCAA-T 5726
SYK.KE.x.SYK173_COMGNM      ACAGAGGGTACCCCCAT-CTTCTTAG.........................................................................................................................AAGGAACATTCCT-GAGA-A 6062
TAL.CM.00.266               CGCACTACCCCAATATC-GACCCCTTT--G---GA--G-AG--G-CCATGC-ATAA............................................................................................................... 5865
TAL.CM.01.8023              CCCACTACCCCAACATA-GACCCCTTT-CG-C-GGA-G...-AC-CCATGC-ATAA............................................................................................................... 5360
TAN.UG.x.SIVagmTAN1         ACGAAGAAAGGAGAAATGGAGTCG---GGG--CGG-AT-G---AGAG.............CCTCCTCCAGGACTTGCATAG...................................................................................... 6101
VER.DE.x.AGM3               ACGAAGAAAGGAGGGATGGACAAG-G-G-G--CGA-C--GGC-CGT-................CCACCAGGACTTGATTGA............................................................................-CA-CTCA-A 5613
VER.KE.x.9063               ACGAGGAAAGAAGAGATGGACAAG-G-G-G---GA-CGAA-C--GCT................CCACCAGGACTTGATTAG............................................................................-CA-TA-A-- 6115
VER.KE.x.AGM155             ACGAGGAGAGGAGAAATGGACAAG-G-G-G---GACC--GAC--GTC................CCACCAGGACTTGATTAA............................................................................-CA-TACAG- 6110
VER.KE.x.TYO1_patent        ACGAGGAGAGGAGAGATGGACAAG-G-G-G-C-GA-C-AATC--GTC................CCACCAGGACTTGAGTGA............................................................................--ACTA---- 6115
WRC.CI.97.97CI14            A-CAATTTCAAAGAGCAG-ACAATTGTAT-----CTTG-AG--A-C-G-AGGAGATCAGGGAGAAGTTGAGCAAAGA............................................................................................ 6139
WRC.CI.98.98CI04            A-CAATTTCAAAGAGCAG-ACAATTCTTT-----CTTG-AA--AGC-G-AGGAGATCAGGGAGAGGTTGAGCAAAGA............................................................................................ 6154
WRC.GM.05.Pbt_05GM_X02      --GATTTTCAAAGAGCAG-ACAATTTTAT------TTG-AA--GGCT--AGGAGATCAGGGAGGAGTTGAGCAGAGA............................................................................................ 5074
```

Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

```
                           Tat Rev exon 1 end intron start                                                                                        Vpu start (ACG in HXB2)
                                                                                                                                                  Vpu transmembrane domain start
H1B.FR.83.HXB2             TTCTCTATCAAAG...CAGTAA...............................................................GTA.........GTACATGTA.....................ACGCAACCT.........ATACCAATAGTAGCAATAGTAGCAT 6095
Tat exon 1                 __S__L__S__K__.__Q
Vpu                                                                                                                                       _T__Q__P__.__.__.__I__P__I__V__A__I__V__A__
Rev exon 1                 L__L__Y__Q__S__.__
H1O.CM.91.MVP5180          -C--G--C--G--...------...............................................................---ACGCTG.................................-T---T-AAGAGAACCTGC--G-CT--A----TT--...AGTG 6135
H1N.CM.95.YBF30            -CT-A--C--G--...------...............................................................---AAACCTGTATA-...........................-T-...-TG...TCATTGGG-TTC----CGTT-......-G-G 5685
H1P.CM.06.U14788           -C--G--------...T-----...............................................................---GAAATA..ATAAGYTGTGTTA..................-T---T---AGGGATGAGGC-GT-T--A--AT-GCT...-GGG 5587
CPZ.CD.06.BF1167           -C--G--CG-G--...------...............................................................---TTA....................................-T--TGTGGCAATTTTTACAGTGGT-GCA-TATT---GCTGGG 6199
CPZ.TZ.00.TAN1             -C--G--CG-C--...------...............................................................---CCC....................................-T-AT-AAA......ATAG--GTGGG-AGT-TGTC-ACTAATG 5724
GOR.CM.07.SIVgorCP2135con  -C--G--------...------...............................................................---GAAGCTTTTAC-GT-A.......................-T---T--AAGAGATATA---GTT---A-TAT-......-G-A 5646
                           Tat Rev exon 1 end Tat Rev intron start
MAC.US.x.239               ----GC------C...A-----...............................................................--................................................................................. 6603
Tat exon 1                 __S__A__S__N__.__K
Rev exon 1                 L__L__H__Q__T__.__
H2A.SN.86.ST_JSP4_27       ----GC----G-C...A---G-...............................................................---AG.............................................................................. 6148
H2B.CI.x.EHO               ----GC-C----C...G---G-...............................................................--................................................................................. 6668
H2G.CI.92.Abt96            ----GC------C...G-----...............................................................--................................................................................. 6016
H2U.CI.07.07IC_TNP3        -C--GC------C...A-----...............................................................--................................................................................. 6112
H2U.FR.96.12034            -G---C------C...A-----...............................................................--................................................................................. 6157
ASC.UG.10.RT03             C---G--CG-C--...------...............................................................--C................................................................................ 5625
ASC.UG.10.RT08             C---G--CG-C--...------...............................................................--C................................................................................ 5646
ASC.UG.10.RT11             --T-G--CG-C--...------...............................................................--C................................................................................ 5634
COL.CM.x.CGU1              .......CA--CA...------...............................................................--................................................................................. 5637
COL.UG.10.BWC01            .......CGCTGC...------...............................................................--................................................................................. 5442
COL.UG.10.BWC07            .......CA--GC...A-----...............................................................--................................................................................. 5667
DEB.CM.04.SIVdeb04CMPF3061 A--AAAGCA.........--G-...............................................................---CAT.TGGTTGTA-AT................................................................... 6109
DEB.CM.99.CM40             --T-G--C--G--...------...............................................................-C-AATCTGCATGTA-ATGC--AAGGGAACATT.................................................... 5946
DEB.CM.99.CM5              G-T-A--C--G--...------...............................................................---AAGCTGCATATGTATATAGTATCACA........................................................ 5927
DEN.CD.x.CD1_CMO580407     AG--G--------...------GTTTGATGGATCTCCTCTCATCACTGATTTTAAGCTTTTCAATAATCATTATTTGCTTAGTTT--TATATTGTTATTATT-CAGTGATAGTTCTTGACCCATAGG-T--TTAAGATCAAGTTAGG--------A-TACG-TTGCTTGG 6134
DRL.DE.11.D3               CA-G-CTGTGGCA...AG----...............................................................---GAAAAG............................................................................ 5822
DRL.x.x.FAO                CA-G-CTGTGGCA...AG----...............................................................--GAAAAAG............................................................................ 5918
GRV.ET.x.GRI_677_gri_1     CC--A--CGGC--...------...............................................................--................................................................................. 6201
GSN.CM.99.CN166            C---GC--AGGTC...------...............................................................---AGGC...................................-T---T---...............GC--C--TTTGGTGGTGGG 6002
GSN.CM.99.CN71             CC--A--CA--TC...------...............................................................---AGGC...................................-T-AGTG--...............GC--CTCTCTGGTGGTGGG 5990
LST.CD.88.SIVlhoest447     AGA--CTGTGG-C...GG----...............................................................-A.................................................................................. 5287
LST.CD.88.SIVlhoest485     AGA--CTGTGG-C...GG----...............................................................-G.................................................................................. 5287
LST.CD.88.SIVlhoest524     GGA--CTGTGG-C...AG----...............................................................AA.................................................................................. 5284
LST.KE.x.lho7              GGA--CTGTGG-C...AG----...............................................................-G.................................................................................. 6368
MAC.US.x.251_1A11          ----GC------C...A-----...............................................................--................................................................................. 6602
MND-1.GA.x.MNDGB1          AGAAGCTAT-CGA...AG----...............................................................--................................................................................. 5736
MND-2.CM.98.CM16           AG-G-CTGTGGGA...AG----.................................................................................................................................................... 6293
MND-2.GA.x.M14             AG-AGCTGTGGGA...AG----.................................................................................................................................................... 6220
MND-2.x.x.5440             AG-AGCT-TGGGA...AG----.................................................................................................................................................... 5858
MNE.US.x.MNE027            ----GC------C...A-----...............................................................--................................................................................. 6071
MON.CM.99.L1_99CML1        -C--GC-AG----...----G-...............................................................---.......................................-T-A-TTA-...TGGTGGTC--TGG---C-AT--CTTACAGCC 6016
MON.NG.x.NG1               -AA--CGGG-GTC...A-----...............................................................--G.......................................-T-A-TA-CTATTGGATAGC-G-C---TCT-TTTGGTGT-TTA 4635
MUS-1.CM.01.CM1239         -C--G--C---TC...A-----...............................................................--GAT.....................................-T-A-TTA-...TGGTATC--G--GC--CTTT-G--AC--GCA 5950
MUS-1.CM.01.SIVmus_01CM1085 -C--G--CA--TC...A-----..............................................................---GC.....................................-T-A-CTA-...TGGTATT--G--GCT--CATTG--AC--GCA 5960
MUS-2.CM.01.CM1246         -C--G--CG-TTC...------...............................................................---AT.....................................-T---TTAC...TGGTATC--GG-GCT-CCAT-G--AC---TA 6067
MUS-2.CM.01.CM2500         -C--A--CA-CTC...A-----...............................................................--GAT.....................................-T-A-TTGG...TGGTGGT-TG--GC--C--T-G--ACT--CA 6020
MUS-3.GA.09.09GabOI81      -C--G--C--TTC...A-----...............................................................---GT.....................................-T-A-TGTG...TGGTATC-TG--GCT-CCAT-G-CAC---TA 5537
MUS-3.GA.11.11GabPts02     ----G--CAGCTC...A-----...............................................................---GA.....................................-T-A-TGTG...TGGTATC--G--GC--C--T-G--AC--GGA 5550
OLC.CI.97.97CI12           AGA-TCTAGGG-C...AC----...............................................................--................................................................................. 5791
RCM.CM.00.SIVagi_00CM312   ACT-GCTA--CCTGACA---G-...............................................................---TGGAATGTCC........................................................................ 5616
RCM.CM.02.SIVrcm_02CM8081  ACT---TG-T-CAGAG------...............................................................---TGGACTGCCC........................................................................ 5631
RCM.GA.x.SIVRCMGAB1        -C--G--C-T--T...------...............................................................---TGGAGTGCCC........................................................................ 5810
RCM.NG.x.NG411             -GA-AC---TG--...------...............................................................--...................................ATGGA-T-TCC..................................... 5854
SAB.SN.x.SAB1              AGT-TCTCT-C-C...A-C--G...............................................................TA................................................................................. 6477
SMM.CI.79.SIVsmCI2         ----GC------C...------...............................................................--................................................................................. 5726
SMM.LR.89.SIVsmLIB1        --T-AC------C...------...............................................................--................................................................................. 5668
SMM.SL.92.SIVsmSL92A       -C--TC------C...T-----...............................................................--................................................................................. 5682
SMM.SL.92.SIVsmSL92B       -GA-AC----G-C...------...............................................................--................................................................................. 5659
SMM.SL.92.SL92B            A-T-GATA--TCAAAC------...............................................................--................................................................................. 6030
SMM.US.04.G078             -GT-GC----G-C...T-----...............................................................--................................................................................. 5863
SMM.US.05.D215             ----TC-CGC--C...A-----...............................................................--................................................................................. 5857
SMM.US.06.FTq              ----GC-C----C...A-----...............................................................--................................................................................. 5882
SMM.US.86.CFU212           --T-AC----G-C...A-----...............................................................--................................................................................. 5846
SMM.US.x.H9                A-T-TCTG--TCAGACA-----...............................................................--................................................................................. 6071
SMM.US.x.PGM53             ----GC-C----C...G-----...............................................................--................................................................................. 6522
SMM.US.x.pE660.CG7G        -C--GC------C...A-----...............................................................--................................................................................. 6587
STM.US.89.STM_37_16        AAT-TCTG--TCAAACA-----...............................................................--................................................................................. 6245
SUN.GA.98.L14              AGAGGCTGT-CGA...AG----.................................................................................................................................................... 6416
SYK.KE.x.KE51              -AGCTC-A---CT...-G----.................................................................................................................................................... 5894
SYK.KE.x.SYK173_COMGNM     -AGCGC-C-T-CA...GG----.................................................................................................................................................... 6218
TAL.CM.00.266              -C--G--------...------.................................................................................................................................................... 6019
TAL.CM.01.8023             -C--G--------...------.................................................................................................................................................... 5514
TAN.UG.x.SIVagmTAN1        -CT-G-----C--...------...............................................................--................................................................................. 6266
VER.DE.x.AGM3              -C--GA-CA---A...----G-...............................................................--................................................................................. 5777
VER.KE.x.9063              A----A-CAGG-C...------...............................................................--................................................................................. 6279
VER.KE.x.AGM155            -C--G--C-GC--...------...............................................................--................................................................................. 6274
VER.KE.x.TYO1_patent       -C--T--CAGC-C...------.................................................................................................................................................... 6277
WRC.CI.97.97CI14           ....GCTG-TTCA...AG----...............................................................--................................................................................. 6289
WRC.CI.98.98CI04           ....GCTA---CA...AG----...............................................................-A.................................................................................. 6304
WRC.GM.05.Pbt_05GM_X02     ....GCT----CA...AG----...............................................................-C.................................................................................. 5224
```



Draft of 30 September 2015, not for public release
plvdna_main.tex r8876–10546

```
                                                                         Env start                                                                                      Vpu end
H1B.FR.83.HXB2                GCAGAAGAC...AGT...GGCA........................................ATGAGAGTGAAG.........GAGAAATATCAGCACTTGTGGAGATGGGGGTGGAGATGGGGCACCATGCTCCTTGGG...ATGTTGATGATCTGTAGT......G 6312
Vpu                           _A__E__D__.__S__.__G__.__.__.__.__.__.__.__.__.__.__.__.__.__.N  E  S  E  .  .  .  G  E  I  S  A  L  V  E  M  G  V  E  M  G  H  H  A  P  W  .  D  V  D  D  L  *
Env                                                                                          M  R  V  K  .  .  .  E  K  Y  Q  H  L  W  R  W  G  W  R  W  G  T  M  L  L  G  .  M  L  M  I  C  S  .  .
H1O.CM.91.MVP5180             AGG--T---...---...-A-T........................................----A---A-T-AAG...AAGA-C--C...AG-A-G-CA..........A-CTT-...TA-......A-AGCCAT-GCTT--C-A--CCCA---TTGAGT...T 6349
H1N.CM.95.YBF30               -A------T...---...----........................................----A-----T-...GGGATGC---GTGG-TG-ATGGG-AT--AGA-T--T---TT-CTCTT-TATC-T--AG-AA-C...T--A-C-A-G-AAT-G-G....... 5901
H1P.CM.06.U14788              -GGT-T--A...---.....-.........................................----AG.....AAG...AAGA----C......AG---A......G--AAC-T-GA-ATCT--TTGGCTT-GA-CAT-TACT-TAATGCT--TA-CTA-ACTAGT. 5803
CPZ.CD.06.BF1167              ......--T...---...--A-TAGAAGAAG...............................----A-A--GCC.............................................TTATTA--TGGATGGATCCTG-CAA-T--TT-----T--AAGCTA- 6407
CPZ.TZ.00.TAN1                AGCAT---...---...--AGTAGAAGAAG................................----AGAATTTA........................................ATTGGA--AACTT-GATCCTC--AA-T-CA---CTAG-GATT...- 5932
GOR.CM.07.SIVgorCP2135con     CTTATG--A...--G...--T-........................................-----CT--.................................................................--T...T--A-TT-A-CT--C-T-GAG.... 5850
                                                                                         Env start

MAC.US.x.239                  ..............................................................---G--TGTCTT..........................G-GAATCA-CT-................--TA-C-CCATCT--C-TT-A-GTGTCTA-GGG...A 6661
Env                                                                                          M  G  C  L  .  .  .  .  .  .  .  .  .  .  G  N  Q  L  .  .  .  .  .  L  I  A  I  L  L  L  S  V  Y  G  .
H2A.SN.86.ST_JSP4_27          ..............................................................---...TGTGGT.........................--GAATCAACTA............T-TG-TGCCA-CTTGC-AGCT-GTGCT--CTTA......A 6203
H2B.CI.x.EHO                  ..............................................................---GC-CATGTT..........................-ATAATTACCTA................--TG--ACACTCC--C-T--A-GTATCTA-GGG...T 6726
H2G.CI.92.Abt96               ..............................................................---GC-TATCTT..........................G-CAATCA-CT-................--TA-C-CAACCT-RC-CT-A-GTAC-TA-GGG...T 6074
H2U.CI.07.07IC_TNP3           ..............................................................---GC-TACCTT..........................G-CAATCAACT-................--TA-C-CTCTCT-A-GCT-A-ATGCAT-CGTG...A 6170
H2U.FR.96.12034               ..............................................................---GC-TGTCTT..........................G-CAATCA-CT-................--TA-C-CTCTAT---GTT-A-ATGCAT-CGTG...A 6215
ASC.UG.10.RT03                ..............................................................----C-AA-T-TAAGCTTATA-TATT-GG-ATAGCAA-TATA-T-G-TTTAATA-C-ATA-ATTTGCAA-AGGGG-ATGAA-ATGATGGA............... 5715
ASC.UG.10.RT08                ..............................................................----C-AAAC-TAAGCTTATA-TATT-GG-ATAGCAA-TATA-T-G-TTTAATA-T-GTA-ATTTGCAA-AAGGG-AAGAA-ATAATGGA............... 5736
ASC.UG.10.RT11                ..............................................................----C-AA-T-TACGCTTATA-CATT-GG-ATAGCAA-TATA-T-G-TTTAGTA-T-ATA-ATTTGCAA-AAGGG-AAGAA-ATAATGGA............... 5724
COL.CM.x.CGU1                 ..............................................................---CTTAGAT-T...................................................---T--ATAAACT--C-AG-AT-AG---TAGTG...T 5683
COL.UG.10.BWC01               ..............................................................---TTT-AT--A..........................-A--ACATTA-A............GG---AA-GATAGGG--AA-ATGC--AGAA-TAGTG...- 5503
COL.UG.10.BWC07               ..............................................................----A-C-T-GT.................................................G-AA-CTTAGTT--C--AGGA--AA---TAGTT...A 5713
DEB.CM.04.SIVdeb04CMPF3061    ..............................................................---CTT-CTTTC.................................................A-AT-A-CTATT--AG-A--T----TAG-ACAA...A 6155
DEB.CM.99.CM40                ..............................................................-----GC-T-TA..........................CAGCTTTT-CTA.........CTT--TT-AGGGTTAATAG--A-AGGT-CACAGG-GAAA...C 6010
DEB.CM.99.CM5                 ..............................................................---CTTAGCTTA...........................................TTA--AA-AA-A-CAAGTC--ACCC--G-AG-GGCAGCA...C 5979
DEN.CD.x.CD1_CMO580407        AGTAT---T...---...--AT........................................-----TCAAGTC.................AG-AGGACCCTT............-AGGTCCT-CT-C-TA----GT--......ACA--TGGAATA-TATGT...A 6327
DRL.DE.11.D3                  ..............................................................----TTAGAC-T..........................-T-.................GTT--AGG-T-AC-CATTT-AGGAT-TT-AG-A-TAAGT...A 5877
DRL.x.x.FAO                   ..............................................................----TTAGAC-T..........................-T-.................-TT--AGG-T-AT-CATTT-AGGAT-TT-AG-A-TAAGT...A 5973
GRV.ET.x.GRI_677_gri_1        ..............................................................---G-GAGATT-.........................CTTATAAAAATA...........C-AA-AA-A-CAATAGG-A-A...-GTATA...GGA.... 6255
GSN.CM.99.CN166               ..............................................................-----GAA--CA.......................GT-GA--CTAT-AT--TGCTA-CAGTG-GCT-A--G-CTTTA---GCTT--C-AATCCCGGGT...T 6201
GSN.CM.99.CN71                ..............................................................----TGAG---A.......................GT-G---TTTT-AA--TGCTA-TAGTG-GC--AACGCTTATGGCT---C-A---CCGG--TCG...A 6186
LST.CD.88.SIVlhoest447        ..............................................................----ACTGCCCT..........................G--CTAATACTT...................--CTT-ATC--T--A-GAG-ACAAG-GGAG...A 5342
LST.CD.88.SIVlhoest485        ..............................................................---GACTGTCTT..........................-AGCTCCTACTT..................T-CTT-CTCT-T--AGGAGCACAAG-GGAG...- 5342
LST.CD.88.SIVlhoest524        ..............................................................---GC-TGTCCA..........................G-TAACATA-T-................--AT-GCT-CTCTGC---GGAC-AACAGCAGGA...- 5342
LST.KE.x.lho7                 .....................-GG......................................---GC-TGTCCA..........................G---TA--A.......................ATTCTGC-T---C-TC-TG-C-TAATA.... 6419
MAC.US.x.251_1A11             ..............................................................---G--TGTCTT..........................G-GAATCA-CT-................--TA-C-CCATCT--C-TT-A-GTGTCTA-GGG...A 6660
MND-1.GA.x.MNDGB1             ..............................................................----AGTGTCCA..........................G-TCTTTTTCTT..............A-A--AT-TTTAGCC.-AC-AG-A......GGT...A 5787
MND-2.CM.98.CM16              ............GTATTAACAGTAAGT...................................---CTTAG-T-C.......................................................ATAATAT-AGGA--A--TGTAG-ACTA...- 6348
MND-2.GA.x.M14                ............GT.ATAATAGTAAGT...................................---CTTAGAC-T...................................................T-AATCACAGCAA-AG-A--AATAG--ATA...- 6277
MND-2.x.x.5440                ............GT.TAAACAGTAAGT...................................---TT-AGAT-T.........................................C-TAGGTA-ATAGTCT-AGGA--A--AGTA---GTA...A 5921
MNE.US.x.MNE027               ..............................................................---G--TGTCTT..........................G-GAATCA-CT-................--TA-C-CCATCT----TC-A-GTGCCTA-GGG...A 6129
MON.CM.99.L1_99CML1           ............TCA...--GGTT...GACACAGAATCTGAATCAGAGCAAC..........-----GAAACTC.............ATGGCT-T-TC-ACCCT---TTT-AT-ATG--TT--G-GAA----............................ 6235
MON.NG.x.NG1                  ............CA-...AAT-GTTTTGACAATCCTCTGTTTGATG................---G-GACCCT-.............AC-AGTG--TCTRATTCAG-A-AC............................................. 4848
MUS-1.CM.01.CM1239            ............--C...--A-TCTTTGAGG...............................---CCGAG--T-..........................--C...CTAAT-G-GAT--------C---TT-GCTAACCC--C-TT---G-A-----ATA...A 6143
MUS-1.CM.01.SIVmus_01CM1085   ............--C...----TCTTCGAAG...............................---CCGAG--T-..........................-C----CT-A-AGCG...A------C---TT-GCTAACCC--C-TT---G-A-----ATA...A 6156
MUS-2.CM.01.CM1246            ............--C...----TCTTTGAGG...............................---CCT-C--T-...................GA-A---CGA--ATTCTC-CCGC-CTTT-GC-AAC---A-CTTTG-GCC-T-GTC-AATATTGAGA...A 6275
MUS-2.CM.01.CM2500            ............--C...--T-TCTTTGAGGACGCGGAAG......................---GC-----T-......................................-CC...A--G-----G-TT-CTTAACCC--C-TT---G-------CTG...A 6219
MUS-3.GA.09.09GabOI81         ............TA-GAA-AAGTTTTTGAGGACGCCCCCCAAGAACCTGGGGATCCTCTTATTA---CTTT---CA..........................-TCCT-CCT...........................TTG-GCAGTGAAT-TCCCTTACCCAAGT 5748
MUS-3.GA.11.11GabPts02        ............TA-GAG-AGGTATTTGAGG...............................---CC-TG--CC.............A-GA-AACATT-ACCAACCTCT-G-T-ACA---C-CA-GGC-TT-ATA-CCC--CCTT-GA-C-A-GAGGCAT 5764
OLC.CI.97.97CI12              ..............................................................------AA----........................................C-AG-AT-AT-CTTGC-A-G-T-A-GTATCTTAGGAATCA 5843
RCM.CM.00.SIVagi_00CM312      ..............................................................----A-AGTGCA..................................ATA--A-TAG--A-AA-A-CT......AGTC-A--AA--GC-ATA.... 5767
RCM.CM.02.SIVrcm_02CM8081     TATA-GATAAGA-TATAGATA-GAGGACAAC...............................----AGTGCCTA...............................A-CTT-ATA-TGTTA-GTAGT--GATT......C-TT-AGCT-TAGCAAAA.... 5824
RCM.GA.x.SIVRCMGAB1           ..............................................................---GATAA---A..............................TTAGTA-T-GTATTA-TAG-AG-AA-A---.......A-A--AC-AGTACAAGGA...T 5960
RCM.NG.x.NG411                AT............................................................---GCTAA-TTA.................................-T-A-TATAGGA--TA-AA-A--A......---T-AGGTATA-T-ATA...- 6016
SAB.SN.x.SAB1                 ........................AGT...................................----AGC-TCTT..........................-C-GTACT-CTT.........TGGC-AAGTGGGT-TTGG-GC--AG--TGG-TAGTCCAG.... 6543
SMM.CI.79.SIVsmCI2            ..............................................................---GA-TGTCTT..........................G--AATCA-CT-................--TA-C-CTGTCT--C-T-GA-GTGT-T--GGG...A 5784
SMM.LR.89.SIVsmLIB1           ..............................................................---GCCTGTCTT..........................G-CAATCAACT-................--TA-C-CCATCT--C-AG-A-GTGC-TA-GGG...A 5726
SMM.SL.92.SIVsmSL92A          ..............................................................---G-GTGTCTT..........................G-GAATCAACT-................--TA-C-CCATCT--C-AT-A-GTGC-TCAGAG...A 5740
SMM.SL.92.SIVsmSL92B          ..............................................................---GCGTGTCCT..........................G--CTTCACCT-................--TA-A-ACATCT----TT-A-GTGTGTTAGGG...A 5717
SMM.SL.92.SL92B               ..............................................................---GCGTGTCCT..........................G--CTTCACCT-................--TA-A-ACATCT----TT-A-GTGTGTTAGGG...A 6088
SMM.US.04.G078                ..............................................................---G--TGTCTT..........................G-GAATCA-CT-................--TA-C-CCATCT--C-TG-A-GTGCCTA-GGG...A 5921
SMM.US.05.D215                ..............................................................---G--TGTCTT..........................G--AATCA-CT-................--TA-C-CAATAT--C-AG-A-GTGC-T--GGG...A 5915
SMM.US.06.FTq                 ..............................................................---G--TGTCTT..........................G-CAATCA-CT-................--TA-C-CCAGCT--C-AC-A-GTGC-T-CTCG...A 5940
SMM.US.86.CFU212              ..............................................................---G--TGTCTT..........................G-GAATCAACT-................--TA-C-CCATCT--C-AG-A-GTGC-TA-GGG...A 5904
SMM.US.x.H9                   ..............................................................---G--TGTCTT..........................G-GAATCA-CT-................--TA-C-C-CTCTC-C-ATYA-GTGC-TCAGGG...A 6129
SMM.US.x.PGM53                ..............................................................---G--TGTCTC..........................G-GAATCA-CT-................--TA-C-CCATCT--C-TT-A-GTGC-TA-GGG...A 6580
SMM.US.x.pE660.CG7G           ..............................................................---G--TGTCTT..........................G-GAATCA-CT-................--TA-C-C-CTCT--C-AG-A-GTGC-TTAGAG...A 6645
STM.US.89.STM_37_16           ..............................................................---GCCTGCCCT..........................G--AATCAACT-................--TA-C-CTATCT--C-AT-A-GTGC-T-CTTG...A 6303
SUN.GA.98.L14                 ..............GT..AAG.........................................---C-GTGCCCT..................................................GAAA-ATT-ATAGGT--CTCTT-GCTGTTAGGT...A 6467
SYK.KE.x.KE51                 ......................GCACA...................................----A-C-TTTT..........................-ATATCATAA-T............G-AG-GT-A--TATAG-A--A-GTT-AG-C-TAGTA...- 5960
SYK.KE.x.SYK173_COMGNM        ..................GTA.AAA.....................................---GC--CTTTT..........................---ACTTACATA...............G-TTG-CTTTTT-GC--A--A-GT-TAG-ATTT...A 6282
TAL.CM.00.266                 .........GT-TATGTTTAGCAGCA....................................----A-TACTTA..........................GC-GTTATT........................TTTTGCT-AGCTTACT-AG-A--GTTT...A 6085
TAL.CM.01.8023                .........GTT-CATGTTTAGCCACA...................................----A-TGCTTA..........................CA-GT-CTA--CTT-GCTATTTGTG-AA-ATGG-TAGCA-ATA-AGGA-AAA--CT-ACC...T 5605
TAN.UG.x.SIVagmTAN1           ..............................................................---G--CCATTA..........................--GG-AAAAG--.........GTAT-AT-AG-AATTTTGGGA--A-GCT-AATAG-ACTG...T 6330
VER.DE.x.AGM3                 ..............................................................----AGC---CA..........................TT-CT-ATAG--................A-A--ATTAATAGG-A-A...GGAGTAGTGCTT...A 5832
VER.KE.x.9063                 ..............................................................----A-CCTCTT..........................TT-G-TATAGTT............C-TA-AGC-ATCATAGGAA-A...-GT-CCTTA......T 6334
VER.KE.x.AGM155               ..............................................................----C-AA-TTC..........................TT-G-AATT-TT............--AG-AT-A--AATAGG-A-A...GGAATAG-GATA...A 6332
VER.KE.x.TYO1_patent          ........................GT....................................-----GTAT-CA..........................-T-ATAACC-TA............GGAA-AA-A-T-ATAGGAA-A...GGGATA..........- 6331
WRC.CI.97.97CI14              ..............................................................----TTT-T--A...................................................T-A-TACTAGGA--A--AGGAATAGC-ATC...A 6332
WRC.CI.98.98CI04              ..............................................................---CTTCGCTTT..................................................A--CTTATA-GTA-AG-T---ATAG-ATTA...- 6347
WRC.GM.05.Pbt_05GM_X02        ..............................................................----ATACTTTT................................................A--T--ATTGCAG-AG-A--AGG-ATAG-ATTA...- 5270
```

```
                      Env signal peptide end Env gp120 start
H1B.FR.83.HXB2        CTACAGAAAAATTG..............TGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAGGAAGCAACC.........ACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCC...ACA 6455
Env                   A  T  E  K  L  .  .  .  .  .  W  V  T  V  Y  Y  G  V  P  V  W  K  E  A  T  .  .  .  T  T  L  F  C  A  S  D  A  K  A  Y  D  T  E  V  H  N  V  W  A  T  H  A  C  V  P  .  T
H1O.CM.91.MVP5180     A--GTA--C-----..............-AT-C------T----C-----------A---G-A--G---G-A.........C-AGTA-----C-----T----------CCT-ACAAGC--T--ACAG------A-------AT----A-----C--T--T...--- 6492
H1N.CM.95.YBF30       ..T-T---C--CAT..............-----A-----G--C-----------A--A----GA------GAG.........--A-----T--C-----T-------------CC--AG--------CT--C--CA-C-----------A--A-----T--T...--T 6042
H1P.CM.06.U14788      ..GG-AC-CGT.................-AT--T--T--A-----------T--A--T----GA--T--T-AA.........-TT--CT----C------G-------ATCGCT-AC-AGC-RW---CAG------A----------C--G--------G---...--- 5941
CPZ.CD.06.BF1167      GA-TT-G--C--AT..............-ACAC------------------G--A--A-----A--G-----A.........C-A--AT-G--C-----T--CTCA--A--G-T-GCAAGC---C-ACC-------A--------------AAT-----G--T...CTT 6550
CPZ.TZ.00.TAN1        GATTTAGC-CT-AT..............-ACAC------G-T------A--------T-----A--G--CCAA.........C-A--CT-G--------C--T------G-TATTAC-AG--G---TAA---C--CA-A-----A------AA------G--T...TT- 6075
GOR.CM.07.SIVgorCP2135con .....-CT-CGCAT..............-AT-CT--T--G--------A--G--A--C---C-A--G----AT.........GTA---T----C------G---------TTATGTGAGCCAG--ACA---C---A-A-----A--G--G--A-----G--T...--C 5988
                          Env signal peptide end Env gp120 start
MAC.US.x.239          TCTATTGT-CTC-A..............-AT----------T------T-----A-CT---G-A-T--G--A.........-TTC-C--C---------.......................--CA--AATAGGG--AC-----GA---AC-CAG--CC----A...GAT 6783
Env                   I  Y  C  T  L  .  .  .  .  .  Y  V  T  V  F  Y  G  V  P  A  W  R  N  A  T  .  .  .  I  P  L  F  C  A  .  .  .  .  .  .  .  T  K  N  R  D  T  W  G  T  T  Q  C  L  P  .  D
H2A.SN.86.ST_JSP4_27  TATATTGCGTCCAA..............-AT--G--T--T-TC-----C--G--C-------GAA-T---T--.........-TTC-C--C---------.......................--TA-AAATAGAG--AC-----GA--CATACAG--CT-G--A...GAC 6325
H2B.CI.x.EHO          A--TG-GC--GAAC..............-TT-----T----TC-----TA----C-CA-----AA-T---T-A.........-TTC-C--C--------T.......................--CAGAAACAGAG--ACC----GA--TGTACAA--CC-C--A...GAT 6848
H2G.CI.92.Abt96       TCTATTGT-CCAAA..............-AT-----T--T-T------CA----C-CA----G-A-C--G-G-.........GTAC----C--C------.......................--CA-AAATAGGG-CACA----G---TACACAG--CC-T--T...GAT 6196
H2U.CI.07.07IC_TNP3   TATATTGT-CTCAA..............-AT-----T--T-TC-----CA-C----CA----G-A-C-----A.........GTAC-C--C--C------.......................--CA--AATAGAG-CACA----GG--TACACAA---C-C--A...GAT 6292
H2U.FR.96.12034       TATACTGT-CTCA-..............-AT-----T----T------CA------CA----G-A-C--G--A.........-TTC-C--C---------.......................--CA-CAATAGAG--ACA----GA--CACACAG--CC-T--T...GAT 6337
ASC.UG.10.RT03        ............................-----A-----T-TC-----------A------C-TA-T--C---.........C-TC--T-G-----CA--...................G----T-C-AGAC--ACA----T---TAC-AAT--CT-G--A...--- 5823
ASC.UG.10.RT08        ............................-----A-----T-TC-----A-----A------C-TA-T--C---.........C-TC--T-G-----CA--...................G----T-C-AGAC-CACA----T---TAC-AAT--CT-G--A...--- 5844
ASC.UG.10.RT11        ............................-----A-----T-TC-----------A------C--A-T--CT--.........C-TC--T-G-----CA--...................G----T-C-AGAC--ACA----T---CATGAA---CT-G--A...--- 5832
COL.CM.x.CGU1         TA...AGT--TAAA..............-----A-----A---C-A--A-------CA---G-A------GATGTAAATGATCAGCAGT-C--C--CTTT--TTCCT-CCC--AG......................A--CAACAGGTGTTG-GG---T----ACCTC-G 5814
COL.UG.10.BWC01       TA...-G--GTAA-..............-AT--A-----A---C-A----------CA---G----G--GG-AGCAGTAGATCAGCAGA-C--C--CTTC---TCAT-AGT--AC......................A-CCA-CAAGTGTTA-GA---C-C--GCCGC-T 5634
COL.UG.10.BWC07       T-...AGT--GAA-..............-----A-----G--CC-G--A--G--A-CA---G----T---GAAATAACAGATCAACAAT-C--C--CTTC--CTCCT-G-T--A-......................A-GCA-CAGGTCTTA-GG--CC-T--ACCTC-T 5844
DEB.CM.04.SIVdeb04CMPF3061 AG------C-GG-TAGCCAGCAACAACAG-AT--G-----A-T---------T--A--T-----AA-T--CT-A.........GTGGGAT-G-----CA--G--A---G-TC--AG...................TCA-GA-----AGTGACA-----CC----ACACT-C 6298
DEB.CM.99.CM40        AG-A-A-GC--CAA..............-ACT-G-----A--------A-----A--C---GTA--T--C-AA.........GTTGA-T-G--C---A--G-CA-CT-ATC--AG...................TCA-GA-----AGTGACA--A--CC-G--ACATG-C 6138
DEB.CM.99.CM5         -A---A-GC--CAA..............-----G-----A--------A--G--C--A---C-CA-T--C--T.........GTTGGCT----C---A--G-TA--AG-TC--A-...................TCA-GA-----TGTGACA--A--CT-G--ACATG-- 6107
DEN.CD.x.CD1_CMO580407 -A--------GCA-..............-AC--------A-TC-----CA------CA----G-A-CAGTG-A.........G-TC-G--G--C--CA--...................G-GAT--ACAGAGGA--A------AG--GCAT-A-----T--T...GA- 6449
DRL.DE.11.D3          TAGG-...---CAT..............-----A-----G-TC-----AACC--CAA-----GACC----T--.........--ACA---GA-C-----T.......................-----TAAT---TCAT-C----TA--TACAAGT--CA----A...-GT 5996
DRL.x.x.FAO           TAGG-...---CAT..............-----A--G--G-TC-----AACC--CAAA----GACC----T--.........--ACA---GA-C-----T.......................--G--TAAT---TCAT-C----TG--TACAAGT--CA----A...-GT 6092
GRV.ET.x.GRI_677_gri_1 ..-T--GT--CC--..............-AT--G-----G-T------AA-C--A--A-----AA-TT----A.........GTTCAGGC------CATG.......................--GCCCAATACC---A-G-----A--CACCAA---CA----A...GAT 6375
GSN.CM.99.CN166       TG-AGTGTG--AAT..............---ACG-----G--C-----------G--A----G---C-----T.........C-AC-AT-G-----C--C------C-AG-TAT-GC-AG--AT--ACC-GGA--CA-----ATTT-TACA--A---T-G-----...T-G 6344
GSN.CM.99.CN71        GGT-TA--GGGAAC..............---ACG-----G-TC-----------G--A----GA--T----AA.........C-AC-AT-G--C-----C--------AG-CAT-AC-AG--AT--ACC-GGA--CA-----AT-T-TACG--A--CT-G-----...T-- 6329
LST.CD.88.SIVlhoest447 AA---A-G--GCA-..............-AT--A-----G-TC--------G---AAA---G-A------GTA.........GTGC-AT--A-------T.....................T---CTAATA---G-T-A----T----ACAT-A--CA----A...GAC 5464
LST.CD.88.SIVlhoest485 -A...A-G--GCA-..............-AT--A-----A-T---------G---AA----G-A--T--TGTA.........GTGC-AT--A-------T.....................T---CTAATA---GCT-A----T----ACAT-A--CC----A...GAT 5461
LST.CD.88.SIVlhoest524 AA-AG--G---CAA..............-AT--A-----A-TC-----A-----AAAA---G-A--T--TGTA.........GTAC--T--A-C-----T.....................T---CTAATA---GCT-G----T---TACAT-T---C-C--A...GAT 5464
LST.KE.x.lho7         ..TGG-G----CAA..............-AT--A-----A-T---------G--AAAT---G-T--TAATGTTTCA.....GTGC-CT-GA-------T.....................T---CCAATAC--GCC-C----T---GACAT-T--CT-G--A...GAT 6542
MAC.US.x.251_1A11     TCTATTGT-CTC-A..............-AT----------T------T-----A-CT----G-A-T--G--G.........-TTC-C--C--C------.......................--CA--AATAGGG--AC-----GA---AC-CAG--CC----A...GAT 6782
MND-1.GA.x.MNDGB1     T--T--G----CAA..............-AT--G-------TC-----A-----A--A-----------T-AA.........--ACA-T-GA-------T.......................-----TAATTCA-G-C-C----TA--CAC-AAT--CA----T...... 5906
MND-2.CM.98.CM16      GGTT--G---CCAA..............-----G-----A--------CAC----AAA---C-TA-G---GAG.........--ACAC--G--C------.......................-----AAATG--TCCT-A----TA---AC-AGT-----T-----...-GC 6470
MND-2.GA.x.M14        G--T--G----CAA..............-----A-----A--------CAC----AAA---C-CCC-----GA.........--GCA---T---------.......................--T--CAACA-CTCCC------TA--CACAAGT--------A...-GC 6399
MND-2.x.x.5440        TAGT--G-G--CAA..............-----G-----G--C-----CAC---CAAA---C-TA-G----GG.........--ACAC--G---------.......................-----CAATA--TCAT-C----TG--CACAAG------C-----...-GT 6043
MNE.US.x.MNE027       TCTATTGC-TTCAA..............-AT----------T------T-----A-CT----G-A-T--G--A.........-TTC-C--C--C----T-.......................--CAG-AATAGGG--AC-----GA---AC-CAG--CC----A...GAT 6251
MON.CM.99.L1_99CML1   ........TCTAAC..............---AC---C-----C--C--T--T--A--A-----ACCG-----T.........C-TC-----C----------C--T---C-A--TTACGGGTC--A-----C-GGG--CAA----CTGG-GTC-T----CC-C--G...--- 6370
MON.NG.x.NG1          ........----CA..............-----------A-TC-----A--G--G--C-AT--ACCT--G--G.........C-TC-AT-G--C-----T-----CC--G-YTATGGCTCA-A-----C-GGG---AA----CTAG--TC-T----CC-C--T...T-- 4983
MUS-1.CM.01.CM1239    TATGG-GC--TCA-..............-----------G-TC------------G--A-----A--T-----T.........C-AC-A--C--C-------------------CAT-GCAA-C-GG--ACC-GGG-----G---AT---CACA--A---C-T-----...--- 6286
MUS-1.CM.01.SIVmus_01CM1085 T-TGG-G--G-CAA..............-------------T------------G-------A--T--T---.........C-AC-A--C--C---------------C--CAT-GCAA---G---ACC-GGG-----A---ATA--CACA--A--CC-C-----...--G 6299
MUS-2.CM.01.CM1246    AACAT-CC--GCAA..............--------G--G-T----A-------G--A-CA---A-T--C--A.........C-AC----C--C----------A-------TAT-GCAA-C-G----CC-GGG--C--G-T-ATA--GACA--A--CC--------...--G 6418
MUS-2.CM.01.CM2500    TATGG-G----CA-..............-----------G-T------A--C--G--C--------T-----A.........C-AC-AT----C-----C------------CAT-GCGA-C-GG---CC-GGG-----G---ATA---ACA--A--CT--------...--- 6362
MUS-3.GA.09.09GabOI81 -A-A----TC-ACAGAA...........---AC------G-TC--------C--A--T-----A--T----AG.........C-TC-A--C-------------T----TA--TAT-GCCTC---GAGACC-GGG--CA-A---ATAG-CACA--A--CT--------...--- 5894
MUS-3.GA.11.11GabPts02 -AGAGC-G---AA-GAA...........-----A-------T------C--G--A--A--------T----AA.........C-GC-A--G--C------C-------TC--CAT-GCCTCA---AG-CCTGGA--C--G---ATTT-CACG--A---T-G------...--G 5910
OLC.CI.97.97CI12      G--T--GG...CAA..............-AT--G-----G-TC-----AAC---AAA-----GA--TAGT-AT.........GAT--GT--A-G----TG.......................----CTAACA-A-G-C-G----T----ACA--T--CC-------...T-T 5962
RCM.CM.00.SIVagi_00CM312 ..GT-AGT---CAA..............-----A-----A-TC---------T--G--A-----AA-TAGT--A.........G-AC-AA-G-------TT.......................-----TAATGCA---TCA----G---T-TAAAT---A----A...GA- 5887
RCM.CM.02.SIVrcm_02CM8081 ..CAGA-GCC-CAA..............-AT--G-----A-T------A-----A--A---GTCA-CAGT--A.........GTAC--A-G-------TT.......................--T--CAACACA--CTCA----GT--TTTAAAT--CA-C--A...GA- 5944
RCM.GA.x.SIVRCMGAB1   -ACA-A--CCGCAA..............-ATA-A-------TC-----T--C--A--A----GAA-CAGC--A.........GTGC--A-G-------TG.......................--T--TAATAC-C-ATCA----GA--T-TAAAT---A----A...GAG 6082
RCM.NG.x.NG411        TAGG---G---CAA..............-AT--G-----A-T------T--C--A--A------A-CAGT--A.........GTTC--A-G--------T.......................--T--TAACGCAC-ATCA----GA--GTTAAAT---A----A...GA- 6138
SAB.SN.x.SAB1         ............................-AT--T-----G-T------TA-C--A--A-----AA-TAGTT-A.........GT-CAAGCC------AAG.......................--CCCCAATACG--CC-A------T-TAC-AAT--CA-C-----...GAT 6651
SMM.CI.79.SIVsmCI2    TCTGGTGC-CTCA-..............-AT-----T--T-TC-----CA-C--C-CA----G-A-T--G-G-.........GTAC-A--C--C-----G.......................---A--AATAGGG--ACC----GG--TACACAA--CC-C--A...GAC 5906
SMM.LR.89.SIVsmLIB1   TATACTGT-CTCA-..............-AT-----T--T-TC-----T--T----CA----G-A-C--G--A.........-T-C----C---------.......................--CA--AATAGGG-CACC----GT---ACACAG---C----A...GAC 5848
SMM.SL.92.SIVsmSL92A  TCTTCTGC-C-CA-..............-AT--A-----T-TC-----T--T----CA----G-A-C--G--A.........-TAC-C--C---------.......................--CC--AATAGGG--AC-----GA---ACACAA--CC-G--A...GAT 5862
SMM.SL.92.SIVsmSL92B  -CTGGTGTGC-CA-..............-AT--A---A---T------TA-T--C-CA----GAA-C--T--G.........-T-C-C--C--C-----G.......................--CC--AATAGAG-CACA----GG--CGT-CAG--CT-G--A...GAT 5839
SMM.SL.92.SL92B       -CTGGTGTGC-CA-..............-AT--A---A---T------TA-T----CA----G-A-C--T--G.........-T-C-C--C--C-----G.......................--CC--AATAGAG-CACA----GG--CGT-CAG--CT-G--A...GAT 6210
SMM.US.04.G078        T-TATTGC-CTCAA..............-AT--------A-TC-----T-----A-CA----G-A-T--G--A.........-TTC-C--C--C------.......................--CA--AATAGGG--AC-----GA---GT-CAG--CC-G--A...GAT 6043
SMM.US.05.D215        T-TGGTGC-CTCAA..............-AT--A-------T------T--T----CT----GAA-C--G--G.........-TAC-C--C---------.......................--TTC-AATAGAG--AC-----GA--TAC-CAG--CT-G--A...GAC 6037
SMM.US.06.FTq         TATATTGC--GAAA..............-AT--A--T--T-TC-----CA-T--C-CA----G-A-C--G--A.........GTTC-C--C--C------.......................--CA--AATAGGG--AC-----GA---GTCCAA--CT-G--A...GAT 6062
SMM.US.86.CFU212      TCTATTGTGCTCAA..............-AT--------A-TC-----T--T--A-CA----G-A-C--G--A.........-TTC-C--C--C------.......................--C---AATAGGG-CAC-----GT---AC-CAA--CT-G--A...GAT 6026
SMM.US.x.H9           T-TATTGTGTTCAA..............-AT--A-----A-TC-----T-----A-CA----G-A-T--G--A.........GTTC-C--C--C------.......................--CAGAAATAGGG--AC-----GA--CACACAA--CT-G--A...GAT 6251
SMM.US.x.PGM53        TCTATTGC-CTCAA..............-AT--------T-T------T-----A-CA------A-T--G--A.........-TTC-C--C--C------.......................--CA--AATAGGG-CAC-----GA---AC-CAA--CT-G--A...GAT 6702
SMM.US.x.pE660.CG7G   T-TATTGT-TTCAA..............-AT--A-----A-TC-----T-----A-CA------A-T--G--A.........-TTC-C--C--C-----G.......................--CA--AATAGGG-CAC-----GA---ACACAA--CT-G--A...GAT 6767
STM.US.89.STM_37_16   -CTATTGC-CTCA-..............-AT-----G--A-T------T-----A-CA----G-A-T--G--G.........-TTC-C--C--C------.......................--CA--AATAGGG--AC-----GA---AC-CAA---C----A...GAC 6425
SUN.GA.98.L14         G--T--C--T-CA-..............-AT--A-----A-TC------ACT---AA----G-ACC----GTA.........GTGC-AT--A-A-----T.....................T---CTAATA---GCT-G----TG--TAC-T-A--CT-G--A...GAT 6589
SYK.KE.x.KE51         ---GGTT-G--CAA..............-AT--------A-T------A--T-CAAT---G-A--T--TCAA.........GTAC--A-G---------.......................--CCCCCATACGGGA-G------A---A-GAA---------A...T-T 6082
SYK.KE.x.SYK173_COMGNM TGGA-A--C--CAA..............-AT--A-----A-T-------A----ACAT---G-A--T---TAT.........G-AC----C-------A-C.......................--CTCCCATA-AGGA-GA-----T--TA-GAAT-----G--G...T-- 6404
TAL.CM.00.266         TA-T--GG--CCA-..............-AT-----CA----------A-----A--------AA-T-----A.........C--C-AT-G--------C.......................--CAT--C-AGGG-CCAA---T-A---TCCCAA--C--C--A...T-- 6207
TAL.CM.01.8023        -CG-----C-GCA-..............-TCC-------G--C--C--A--G--A--A-----AA-T-----G.........C-TC-AT-G--C--C--C.......................--TAT--C-AGGG-CCAG---T-T---TCCCAA--------A...T-- 5727
TAN.UG.x.SIVagmTAN1   TATAT-GG-C-CA-..............-ATA-T-----G-T------TA-C--A--A-----AA-CAGCT-A.........GTGCAAGCT--C---ATG.......................---CCTAATACC--CC-------A---AC-AAT---A----A...GAT 6452
VER.DE.x.AGM3         A----AGGC--CAA..............-----------A-T------A-----A--A-----AA-CAGCT-A.........GTACAGGCT--C--CATG.......................---CCCACCACC-GAC-A-----A--TAC-AA--CGA----A...GAT 5954
VER.KE.x.9063         TA-GT----G-CAA..............-----------G-TC------A-------A------A-CAGCT-A.........GTACAAGCC--C---ATG.......................--TCCTACTACC-GGC-A-----T---AC-AAT--CA----A...GAT 6456
VER.KE.x.AGM155       G----A--C-GCA-..............---A-A-----G-TC-----A-----A--A-----AA-CAGCT-A.........GT-CAAGCT----CATG.......................---CCTACTAC--GGT-G-----A--TAC-AAT--CA----A...GAT 6454
VER.KE.x.TYO1_patent  TGTT-AGT--GCAA..............---A-A-------T-----AA----A--A-----AA-CAGCT-G.........GTGCAGGCT--C--CATG.......................--TCCCACCACA-GCT-A-----T--TAC-AAT--CA----A...GAT 6453
WRC.CI.97.97CI14      G-TT--CTGT-CAA..............-AT--A-----G--C-----A----ACA---G----CAGT--A.........GA-TTC---C-G------.......................-----AAATAGATCAT-A----TA-GCACACG----A-T--A...CAT 6454
WRC.CI.98.98CI04      GG-T--GG---CAA..............-ACA-A-----G--C------------CAA---G---GAGT-AA.........GATTTCT-GC-C------.......................-----AAACAGATC-C-G----T--GCAC-CA---CA----A...CAT 6469
WRC.GM.05.Pbt_05GM_X02 GATT--G---GCAA..............-ACA-A-----G-----------G--ACA----G---------AG.........GA-TT-A-G--G------.......................------AATAGATC-C------TA---ACACAT---C----T...CAT 5392
```

```
H1B.FR.83.HXB2              GACCCCAACCCACAAGAA........GTAGTATTGGTAAATGTGACAGAAAATTTTAACATGTGG......AAAAATGACATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACCCCACTCTGTGTTAGTTT 6610
Env                         _D__P__N__P__Q__E__.__.__.__V__V__L__V__N__V__T__E__N__F__N__M__W__.__.__K__N__D__M__V__E__Q__M__H__E__D__I__I__S__L__W__D__Q__S__L__K__P__C__V__K__L__T__P__L__C__V__S__L
H1O.CM.91.MVP5180           --T-----T-----T---.........T-TCC-C-A-GC----------T--C---G-T--A---......------T-------G--C--A--------A--C--C--T-----G-----A--G--TT-------T----AG---A-G--TTTCT-A-----ACAAA- 6647
H1N.CM.95.YBF30             --T-----T---------.........--GC----ACCC-----A--T-----A-----T------GAA.........---A-A----C---C--A-----A--------T------C-G-----A--G---T-------C-----T-------------T-A-----A-C-A- 6197
H1P.CM.06.U14788            -----T-GA---AT---G.........--CAGGA-A-AT-----A-----GTC------T--T---GAC.........---T-T-----GAC---A-----G--A--C--C--T--C-----------G-----T--A--T-----------G--AGTT--A-------C-A- 6096
CPZ.CD.06.BF1167            --T--TC-G---T-T---.........A--CCTA-T...--CA-T--TACTG-A-----T---GAAACT.........---T-T---------G-----A-AA-A---T-----TC-C-G-TTC-A--G--TT-G--A--T--------G--G--A--CT-T-----A-AAA- 6702
CPZ.TZ.00.TAN1              --T-----T--TT-T---.........---ACCC-A-CC------T--AT--GG-----T---GAAGAA.........---T-------GC--G-----A-A--A------T-ATCAC-T-TTC-A--G--TT-T-----T---------------A---T-T--CA-A-AGA- 6230
GOR.CM.07.SIVgorCP2135con   -----A-G----AT---G.........T-TCC-----A--R---A-----G-CC---G-T--C---AAC.........---A----------T--A-----G--T-------T--C-----------G---T-------T---------C----AGTTA-G-----C-C--- 6143

MAC.US.x.239                A-TGGTG-TTATTC---...............--GGCCC-T-----T-------GC---G-TGCC---AAT.........---ACAG-CAC-------GCAATA------G--TGGCAAC-C-TT--GACCTCAA-------T-----------T-----T-A--CA---C-A- 6935
Env                         _N__G__D__Y__S__E__.__.__.__.__V__A__L__N__V__T__E__S__F__D__A__W__N__.__.__N__T__V__T__E__Q__A__I__E__D__V__W__Q__L__F__E__T__S__I__K__P__C__V__K__L__S__P__L__C__I__T__M
H2A.SN.86.ST_JSP4_27        A-TGATG--TAT--G---...........A--GCTT------------GGCC--CG--GCA---AAT.........---ACAG-AAC------AGCAGTA--A---G-CTGG---C---TT--GAC-TCAA---A-----C--C---C----A--CT-A-----AGCAA- 6477
H2B.CI.x.EHO                A-TGATG--TATACC---...........A-CCAAT-----A-A-----GGC----G-TGCA---GAT.........---ACAG--AC---T---GCAACAA-----G-GTGG---C-C-TT--GACCTCAA----A---------T---------------G-----G-CAA- 7000
H2G.CI.92.Abt96             A-TGATG-TTACAGT---...........C--GCAT-G-----A------GC----G-TGCC---GAT.......--CACAG-AAC-------GCAGTA--A---G-TTGG-A-C-G-TT--GAC-TCTAC---A--T-----T--------A--T--G--CA--GCCA- 6348
H2U.CI.07.07IC_TNP3         A-TGATG-TTATAGT---...........T--GCTT-----A-C-----GGC----G-TGCT---GAC.........---ACAG-AAC---G---GCAATA------G--TGG-A-C-T-TT--AACCTCTA--------------C--GC-G--A--C------A-AGCAA- 6444
H2U.FR.96.12034             A-TGGTG-TTATTC---G...........T--GCAT-------C-----GGCC---G-TGCC---AAT.......--CACAG-AAC---G--AGCAGTG--A---G-CTGG---C-T-TT--GAC-TCTA----A------------G--G--A--C--T---A-AGCAA- 6489
ASC.UG.10.RT03              --T--AGC---TAT---G.........AC-CCTA--AAT-TCTCAGG----TGG--C--TG-ACACACC.........---T-T-----G--C--T---G-CA-A-----GGCAGCA--G-TTCTA--GGCTAC----------ACC-----G-----TA-G-------AAA- 5978
ASC.UG.10.RT08              --T--AGC---TAT---G.........AC-CCTA--AAT-TCTC-GG----TGG--C--TG-ACATACC.........---T-T-----G--C--T---G-CC-A-----GGCGGCA--G-TTCTA--GGCTAC----------ACT-GG--G-----TA-G-------AGA- 5999
ASC.UG.10.RT11              -----AGC---TGT---G.........AC-CCTA--AAT-TCACAGG----TGG--C--TG--CACACC.........---T-T-----G--C--C---G--C-A-----GGCAGCGC-G-TTCTA--GGCTAC----------ACC-----G-----TA-G-------AGA- 5987
COL.CM.x.CGU1               CCTGGA--G---GT---..........CA-AACA--CC------A------GC----G-TC-T-TTAAG.........---AG-T-CAGT-GGG--G-ATGGATAATA-C-CAG-C-ACC-TA--A-----G----GA-------CT--G--G--AG--TA-----C-CCAA- 5969
COL.UG.10.BWC01             CCAGGG--A---GT---G.........CA-AACA--CC----A-C-----GTC----G-TC-T-TTAAG.........---AGTT-TAGT-C-G--G--TGGAGTATA-CTCAA-CAAC--TA--A-----GT---GA--T----C-------T-TG-CTAT----CACC-A- 5789
COL.UG.10.BWC07             CCAGGAG-T---GT--T-.........CAGTC-A---AG--CA-AT------CA--C--TGCA-TTAAG.........---CCAT-CTCCATGG-AG--TGGACA-T--C-CAGTC-G-T-TA--A--G--AA---GA--T----CC--G---T-AG-TTAT---TCAT-GA- 5999
DEB.CM.04.SIVdeb04CMPF3061  -TGATA-GAGA-............CAG---CACA--CC------AT--C-GT-C---G-TGCA-TCAAGAATCCC.......--A-A---C--AT-ATGG--A-----G-CAG-G----ATA-A----CATAC-GA--C--------G-----A--TTAT-----A-AGA- 6450
DEB.CM.99.CM40              TTAGTG-GAGA-G--............--CCC-A--CC---------CC----C-----TGCT-TTGATAACCCC.......--A-AG--G---C-ATGGC-------G-CT-----G-ACA-A--GTCAT-C--A-----------GC----A---TAT---------CA- 6290
DEB.CM.99.CM5               ATGGTT-GAGA-............CAA--CCCTA--CCT------T-CC-GT-----G--TCT-TTTCAAATGTT..........-A-C----AT-ATGG--C--C--G-CT--C----ACAGA---TCTT-T--A--C--------G----------TAT-----GC-GA- 6259
DEN.CD.x.CD1_CMO580407      ------T-T-A-AGG--G..........--CTGGA---A-----CC-----TT-C---G----C-ATCACAATCC--TA-TAGACCAG-T-GT-CA-A--AT-A-AGAT......C-G-ATA-A--G-C-T-C--A--T-----C---------A--TG-A-----C--AA- 6604
DRL.DE.11.D3                TTATTAC-TTATG-G---..........AC-AA-A-T-A----A-AGAG----------CTG-CCCTATGACAG-------AG--A--A---AGCATGG-GT-CGC-GTCATC-A-GATA---GC-GTG--T------------T---A-C-AT--TTA------G-AAA- 6157
DRL.x.x.FAO                 TTATTGC-TTATG-----.........AC--A-A-TAAC--A-AGAG----------CTG-CCCTATGACAG-------GG--A--A----GCATGG-GT-CGC-GTCATCCA-GATA---GC-GTG--T--A---------T---A-C-AT--TTAT-----G-AAA- 6253
GRV.ET.x.GRI_677_gri_1      --T-ATG-TAATAC---G.........--GCCTC-A...--CA-T------GC---CG-GGCT---GAT.........--CCGC-----A---AGCAG-GAGTA-----CATCTAC-C-TT--A----CGA-G-G---T-----T--GC-CT----CA-A---A---AAA- 6527
GSN.CM.99.CN166             ------TCA--GGC---G.............--GCC-T----CA-A-----G--A--C-----C-ATAAA.........-T---------TG--G---GA--T--C--GG-ATCAC---TTA----GGC---G--A--------C-----G--T---A-G-----GG-GA- 6496
GSN.CM.99.CN71              ------TCA--GTCT--G.............--GCCAC---CA-A-----GG-A-----T--C-ATAAA.........-T-T--------TG--G--AGA--T--C--GG-ATC-C---TTA----GGC-----GA--------C---C-G--T--GA-G-----AC-AA- 6481
LST.CD.88.SIVlhoest447      TTG-AGTCTTATGC-C-G.........A-TCCTA-TCAC--CA-A-GTTT------C-CT-AAGA-ATAAAAG-T---C-A--TA--C-G--AGCATGGAGT-C---G-C-----A-GGT----ACT-T-A-G-----T---------G---AT---TA-------G--GA- 5625
LST.CD.88.SIVlhoest485      TTA-AGTCTTATGC-C-G.........A-TCCTA-TCAC--CA-A-GTTT------C-CTCAAGA-ATAAAAG-T---C-A--CA--C-G--AGCATGGAGT-C---G-C----A-GGT----ACT-T-A-G-----T--------GA---AT---TA-------G--AA- 5622
LST.CD.88.SIVlhoest524      CTG-AATCTTATGCTC-G.........A-CCCTA-TCAT---A-T--TCT---------CTCAAGAAATTAAGG-C---C-A--AA--C-G---GCATGG-GA-C---G-CG--CA-GGT----GCT-TTA-G----------AC-----A---AT--TTAT-------C-CA- 5625
LST.KE.x.lho7               TTG-AATCTTATGC---G.........---CCTA-CTAT---A-TT-------C----CA--ACCAGTAAAAG-C---C-GG-TA--C-G--AGCCTGGTCT-C---G-ATGC-A-GGTA--CAGT-TTA-G--A-------------A-C-AT---TAT-----G--AA- 6703
MAC.US.x.251_1A11           A-TGGTG-TTATTC---..............T-GGCCC-T-----T-------GC---G-TGCT---GAG.........---ACAG-CAC-------GCAATA--A---G--TGGCAAC-C-TT--GACCTCAA-------T-------------------T-A--CA---C-A- 6934
MND-1.GA.x.MNDGB1           TCATTGCCAGATT-T--T......GAG----A-A-TCCTG--A-A-AG----------CAGGACTTATAAGGG-----C-G--A--TT-T--AGCATGGC-T-C---GGGA---A-G-TA---ACC-TA--T---------------GA-T-A----TAT-------AGA- 6070
MND-2.CM.98.CM16            TTGTTGC--TATG----G.........CA-CC-A-TCCG--CA-A---TGG-----C-CAGGTCCCATGGAAG-------GG-A---ATG--AGCATGG-GA-C---CTC-TCCA-GATA--CGC-GTAT----A--T-----G---C-G------TA------C-AGA- 6631
MND-2.GA.x.M14              TTATTGC-TTATG-G--G.........CA-C-CA-CCCC---A-A--------------CAG-CCCTATACAAG-G-----AG-----AGG--AGCATGG-GA-C---TTC-TC-A-GATA---GC-GTT-----A--G-----------G--A---TA------C-AGA- 6560
MND-2.x.x.5440              CTGTTGC--TATG-----.........CAGCATA-CCCC---A-A-----G-------CAGGCCCTATAGAAG-G-----A--A--GAC---AGCTTGG-GA-CA--TTCTTCCA-GATA---GC-GTAT----A-----------GG-----A--TTAT-----G-AGA- 6204
MNE.US.x.MNE027             A-TGATG-TTATTC---..............T-GGCCC-T---A-T-------GC---G-TGCT---GAG.........---ACAG-CAC-------GCAATA------G--TGGCA-C-C-TT--GACCTCAA-------T-------------------T-A--CA---C-A- 6403
MON.CM.99.L1_99CML1         -----T-C---G---TCT...........C---ATT-G--CA-C--T--GG-G--C--TGCT-ATCAA.......--CT-------G---G-A---GTA--------G-AGTC-C-G-TCTCG--GGCTT-------T-----G---------A--CA-G---------GA- 6522
MON.NG.x.NG1                --T--TGTT----GC-CT...........C-CCCTC-C---A-A-------GC-----TGCA-ATAAC.......--CT-------G---G-----GTG--A--C----AATCAC-C-TCTCA--GGC-T-----------------G--------TA-G-----A--GA- 5135
MUS-1.CM.01.CM1239          --T---GCA--C--G--...............---CAAT----A-CT-T-GGG-A-A-TTT-ATGTCTATAAA...--T-------GACC------GT---T---G---GGTCCA---TCA-C--GGC------A--C-----C-----G-----TA-G-----C-AAA- 6441
MUS-1.CM.01.SIVmus_01CM1085 --T--TGCA--C--G--...............--GCA-T-G---A-CT-T-GGG-AAAGTTTGACATCTATAAG..--T-T--------C------AC---T---G-G-GGTC-A-C-TTA----GGC-T----A-------C------------TA-G-----A-AAA- 6454
MUS-2.CM.01.CM1246          ------GCG-----G--G...............---CAAC----CA-CT--A-TG-GCACTTT-ATATATATAAG...--CT-T--------T------AC---T---G-G-AGTCCA-C-TTAC---GGC---T-----G-----G---C----A--TA-G-----A-AAA- 6573
MUS-2.CM.01.CM2500          -----AGC---T-----G...............---CAAT----CA-C----GGG-A-A-TTTGATATCTATAAA...--CT-T-----G--T--A---AC---T---G-GCGATCCA---TCA----GGCT--T--A--C----------C----T--CA-G-----G--AA- 6517
MUS-3.GA.09.09GabOI81       -----TTC---G------...............---G--C-T---A-TT-----T-----...GATATATATAAG...--T-T--------TG-AG-AACA--T-----T-GG--CA-T-TTA-----GCATAT--------------GC--------TA-G-----A--GA- 6046
MUS-3.GA.11.11GabPts02      -----TTC------G--...............--CCCTT-----A-C-----GG-A---...-ATATTTATAAA...--CT-T-----T--TG-AG-AACA--T-----C-AGTC-A-C-TTC-G---GCTTAT-----T-----T--GC-T------A-G---------AA- 6062
OLC.CI.97.97CI12            TTTGATG-AGATG-GAT-......GAACA-CCCA-CCCTGGA--AGA---GGTA---T-T-ATGT-TCA...G-C-----AG-A---AG---AGCATGGTTT-GC--TGCTTCC--G-TA---GC-GTT--G--A--------G--G-----T--TT-T-------AAA- 6123
RCM.CM.00.SIVagi_00CM312    -GGGGACTGT---C---G..............---CCT--------CT----GC-G---G-TGCA---GAA......--CAGTT--TAT------GCCA-G--CA--G-GTGGCA-C---AT---TCC-C---G-----T-----T-G-C--T-A--CT-A---A-A-CCA- 6039
RCM.CM.02.SIVrcm_02CM8081   -GAGGA-TAT---C----...............---CCA--T------T-------A---G-TGCT---GAA.........---C-C--TAT-----AGCAA-A--TA--G-GTGG-A-C-T-AT---TCT-C------A--T-----T-GG---T----TT-A-----A-CAA- 6096
RCM.GA.x.SIVRCMGAB1         -GAGGA-TAT-T-C---G..............--T-CAA--------T-----CGA---G-TGCT---AAT.........---AG-T-ATAT-----AGCAA-A--TA--G-GTGG-A-C-T-AT---TCC-CT-----A---------C-G-----G----T-A---A---CCA- 6234
RCM.NG.x.NG411              -GAGG--TAT---C----...............---CCA---------T----G--A---G-TGCT---GAA.........---AG-T--TAT-----AGCAA-GTCTA--G-GTGGCA-C-T-AT---TCC-CTT-G--A-----------GG-----A---A-G---A-A-CAA- 6290
SAB.SN.x.SAB1               --TGAACCAGA-GG-AC-ATAGCAGAA---CCCA-ACCT---A-C---------A---G-TGCT---AAAAATCGG--CCCTC-A----GC--AGCAG-GAGTA-C---CA-CT---G-TT--AAGT-CTT-------------T--GC-CT-A---A-G--CA-A-AAA- 6821
SMM.CI.79.SIVsmCI2          A-TGGGG-TTAT--G---..............A--AGAA-G-----------GTCC---G-TGCT---GAC.......--CACAG-AAC-------GC-ATA------G-TTGG-A-----TT--AACCTCTA-C-----------T-----G--G---T-G--CA--GCAA- 6058
SMM.LR.89.SIVsmLIB1         A-TGATG-TTATTC---..............A--AGTA-C-----C-----GGC----G-TGCC---GAC.......--CACAG-TAC---G--AGCAATA------G-TTGG-ACC-C-TT--AAC-TC-A-T-----G---------GC----T---T-A---A--GCAA- 6000
SMM.SL.92.SIVsmSL92A        A-TGATG--TACTC---...............T-GGCTA-T-----C-----GTCA---G-TGCA---GAT.........---ACAG-AAC---G---GCAATA------G-TTGG-A-C-T-TT--GACCTCTA-T--A--C-----------G------T-G---A--GCAA- 6014
SMM.SL.92.SIVsmSL92B        A-TGGAG--TATTC---...............T-GGCCC-T-----T-----GGCC--CG-TGCT---GAT.........---ACAG-AACT----AGCAATA--A---G-TTGG-A-C-C-TT--AAC-TCTA-T--A--T--------G-----T---T-A---A--GC-A- 5991
SMM.SL.92.SL92B             A-TGGAG--TATTC---...............T-GGCCC-T-----T-----GGCC--CG-TGCT---GAT.........---ACAG-AACT----AGCAATA--A---G-TTGG-A-C-C-TT--AAC-TCTA-T--A--T--------G-----T---T-A---A--GC-A- 6362
SMM.US.04.G078              A-TGATG--TATTC---...............T-GGCCA-T-----C------GCA---G-TGCT---GAC.........---ACAG-CAC---G---GCAATA------G--TGG-A-C-C-TT--GACCTCTA----A--C--------G-----T--CT-A--CA--GCCA- 6195
SMM.US.05.D215              A-TGGAG-TTACTC---...............A-GCCTA-T-----T-----GGC----G-TGCT---GAT.........---ACAG--AC------AGCTATA------G-GTGG-ACC-C-TT--GTC-TC-A-T-----T-------------------T-A--CA-AGCCA- 6189
SMM.US.06.FTq               A-TGATG-TTACTC---...............T--GCTC-C-----C-----GGCC---G-TGCT---GAG.......--CACAG-CAC-------GCAGTA------G--TGG-ACC-C-TTA-GACTTCTA----A--C-----T---C----A-----A---A-AGC-A- 6214
SMM.US.86.CFU212            A-TGATG--TATTC---...............C--GCAA-C-----C-----GGC----G-TGCC---AAC.........---ACAG-CAC---G--AGCAATA------G-CTGG-A-C-C-TT--GAC-TC-A----A---------------------T-A---A-AGCAA- 6178
SMM.US.x.H9                 A-TGGTG-TTACTC---...............T-GGCAA-T-----C-----GGC----G-TGCT---GAT.........---ACAG-CAC-----AGCAATA------G--TGG-ACC-C-TT--AAC-TC-A-T-----C---------C-C---------A---A-A-CAA- 6403
SMM.US.x.PGM53              A-TGATG-TTATTC---...............T-GGCCA-------T-------GA---G-TGCC---AAC.........---ACAG-CAC-------GCAATA------G--TGG-A-C-C-TT--GACCTCTA----A--C------------------CT-A--CA---CCA- 6854
SMM.US.x.pE660.CG7G         A-TGATG-TTACTC---...............T-GGCAA-C---A-C-----GGC----G-TGCT---AAT.........---ACAG-CAC-----AGCAATA------G-GTGG-ACC-C-TT--AAC-TC-A-T-----C----------C-C---------A---A-AGCAA- 6919
STM.US.89.STM_37_16         A-TGGTG-TTACTC---...............T-GGCAA-------C-----GGC----G-TGCC---GAC.........---ACAG-CAC-----AGCAATA------G-CTGG-ACC-T-TT--AAC-TC-A-T--A--T-----------G--T-----A---A--GCAA- 6577
SUN.GA.98.L14               TTA-AA-CTTATGC---............--GCCCA-AAC-GGGT-AGA---G-------CAGA-G-AATAAGC--T---C-A--A--GC-G--AGCATGGC-A-C---G-C-TC-A-GGTA---GCC-TTA-G-----T--------GA-C-AT--CTAT-----C-AAA- 6750
SYK.KE.x.KE51               TCTGAAC-GAT-G---TC............AG-G-TAAT-T-AGTGGG---T----C-CTGC----AATTCTTCCC---G---AAG-C-----C-TTTGC----C--GTCTGA-C-T-TCCTG--GGCTGAT-G---T------------T-T--TA-G---A-C--AAC 6240
SYK.KE.x.SYK173_COMGNM      -CAGATC-AAT-G---T-............AGGG-CAAC-T-ACAGG----T-C---CCAGCA---AATAGC-GTC-CATG--AAGGC-G--A--ATTA-----C--GTCAGC-----TCTTG---GC-AAT-GA--C-----G--GC--G-A--CA-G---A----GA- 6562
TAL.CM.00.266               --T--TC-GAAT-TC--G..............---CCA--G---C-C--T--GT-C--CG-TG---ATAAA.......---T-------G--C--C--AGCA--A--C--GGAT--CC-C-TTTTGAC-GC-AC--GA------ACT--G-----A---A-A-----A--GA- 6359
TAL.CM.01.8023              -----TC-GAAT-T---G..............--GCCA--C---T----T--GT-C---G-TG---ATAAA.......--CT-------G--C--C--AGCA--T-----GGA-GCA--G-TCTTAAC-GCAACC-----T--CACC---C-----T--TA-G-----G--GA- 5879
TAN.UG.x.SIVagmTAN1         --T-ATG-TTATACT---...............---CA-T-------CT-T--G--A---G-AGCA---AAGGAC-GG---CCAT-A--G-C----GC-G-GAGTA-C---CATTTGC-C-TT--AAGC-CT--G--------C------C-G--A--TA-G--CA-A-AGA- 6610
VER.DE.x.AGM3               --T-ATG--TACAC---G..............---CCAT----CA-C--T---CCA---G-AGCA---GCTGAC-G---CCC-T-A----C---AGCAGGAAGTA-----CA-CTGC---TT--G--G-CT--G------------------T-A--TT-G--CA---AAA- 6112
VER.KE.x.9063               --T-ATG-TTATAC---G..............--GCCAT----CA-T--T-GCCC---G-AGCC---GCAGAT-GG---CCAT-A--G-C----GCAGGAAGTA----TCATTTG--G-TT--A----CAT-G--A-----C---------T-T--T--A---A-A-AAA- 6614
VER.KE.x.AGM155             --T-ATG--TATAC---...............---CCAC-G---A-A-----GCCA---G-AGCA---GCAGAC-G----CC-T-A----C---AGCAGGAAGTA-C--TCA-CTGC-G-TT--A--G-CAT-------C--------GC--T-A--T--A---A-C-AAA- 6612
VER.KE.x.TYO1_patent        --T-ATG--TATAC---...............---CCTC-----A-C--T---CCA---G-GGCA---GGTGAT-G---CCCAT-AA---C---AGCAGCCAGTA-C--CCATTTAC-C-TT--G----CTA-G--A--T-----G--G---T-A-----A--CA-C-AGA- 6611
WRC.CI.97.97CI14            TTGGAGC-TTTGG----G.........CAGTACA--CCTG-GA-A-A---------C-CA-GAAAAGTAGCAG-----C-A--A--GC-G---GCATGGACA-GAG---CTTCAA-GATT--CACCTTAT------------------A-C--A--CTAT-----A-AAA- 6615
WRC.CI.98.98CI04            ATAGAGC-TTGGG-----.........CA-TATC-TCCT-----A-A-----------CT-G-AA-GTAGCAG-C---C-G--AA-CC----AGCATGGACA-GGG-T-CATC-A-GATT---ACT-TAT----A--T-----T--GA----A---TA------G-AAA- 6630
WRC.GM.05.Pbt_05GM_X02      ATAGAAG-AAT-AC----.........CA-TATA---A-GG-A-A-A-----------CA-AAAA-GTGGAAG-------GG-A---C-G--AGCATGGACA-GAG-G-CT---A-GATA---ACTCTGT-G--A--C---------A---AT---TAT-----G-AAA- 5553
```



Draft of 30 September 2015, not for public release
plvdna_main.tex r8876–10546

Draft of 30 September 2015, not for public release
plvdna_main.tex r8876–10546

```
                                                 V1 loop end┤    ├V2 loop start
H1B.FR.83.HXB2               ..ATAATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGC...ACAAGC...ATAAGAGGTAAGGTGCAGAAAGAATATGCATTTTTTTATAAACTTGATATAATACCAATAGATAATGATACT........................................ 6788
Env                          . I  M  E  K  G  E  I  K  N  C  S  F  N  I  S  .  T  S  .  I  R  G  K  V  Q  K  E  Y  A  F  F  Y  K  L  D  I  I  P  I  D  N  D  T  .  .  .  .  .  .  .  .  .  .
H1O.CM.91.MVP5180            TA-ATGA-ACA-T-AAT-----G-G---T--TAG---T---G-A-CT...---GT-...C-C-C--AC--AAA-G--C--A--C-G--TC-A--C---GT-TCA---C-G-GTAAGG-TA--G-CTCA............................AATGCAGTAAAT 6819
H1N.CM.95.YBF30              ..GAC--A-GAT-CAA-C-A--G-----T-----A------GCA-C-...--TGAG...C---C--A---AAA-A--C---TT--CT-TC-G------GT-GAA---G--G-------CA--GCC...................................TAT 6378
H1P.CM.06.U14788             GC-A--CCA-TCC-TT---AT--TTTC-G---AG---T----CA-C-...----TA...G---A--A---ACA-----C-C-GC-A---C-C-------G-GCA--CC---C-AA-T-G--------C-A-AGT.................................... 6304
CPZ.CD.06.BF1167             ........ACA-GTCT---A--GT-TG-A----T---T---CAA-CT...---GAA...T-T----AC--AAA-A--C-GAT----T----G--C------GAG--------GAA-GACA-A--C......................GACAGTACT 6913
CPZ.TZ.00.TAN1               ..TCT--ACCT--CTAT---G-GT-T--T--T--A--T---CAG-CA...--TGAG...T-T----A---AAAA-AC--AT----T-C--G-------G-GAA----TG--AA-GAG---GG-A-C..................................... 6429
GOR.CM.07.SIVgorCP2135con    ..TAC--CCCAC-GTT---C--TT-C-----TAGC--T---G-T-C-...--GGTT...C---A--A---AAAACA--GC-GC-A--C--A--------G-AA--------AA--C----G-GACG.........................AATGAGACCAAAGAG 6342
                                                 V1 loop end┤    ├V2 loop start

MAC.US.x.239                 ..GGCT----AC---AGC-A--G-T--G---TAAA-----C--G-CA...GGGTTA...-A-----AC---AAAA-AG-GT-CA---A-AC--GG--CTCTGCA---T-GG--TGTGA-C-A.............................GGGAATAACACTGGTAAT 7209
Env                          . G  L  E  Q  E  Q  M  I  S  C  K  F  N  M  T  .  G  L  .  K  R  D  K  K  K  E  Y  N  E  T  W  Y  S  A  D  L  V  C  E  Q  .  .  .  .  .  .  .  .  .  .  .  .  G  N  N  T  G  N
H2A.SN.86.ST_JSP4_27         ..GGGT---GAG---A------GGTCG----TCAG--------G-CA...GG-TTA...GAG--G-A----AAAA-ACT-T-TA---A-ACA-GG--CTC-AAA---G--G-CTGTGA-TCA......................AATGACACCAAGAAA 6718
H2B.CI.x.EHO                 ..GGG--A-GACT--A---A--G-T-G----TCAA-----A--G-CA...GG-TTA...-A-----A-G--TCAA-AC--T-TA-G-ACACC-GG-------AA---T--G-GTGTGA-A-GGGGACA-GGAGTAAT............................... 7253
H2G.CI.92.Abt96              ..GGT--AC--TC--A-TCC--GGTTGG----AAA--T--C--G-CA...GGGCTA...-A---G-A-C--AGAA--G--T-CA---AGAC--GG---TC--AA---T-GCA-TGYGA-C-G..........................TCCAATAAGAGTGAGAAT 6622
H2U.CI.07.07IC_TNP3          ..GG-T-A--ATC--AGCCA--GGT-GG---TCAG--------G-CA...GGGCTA...-G-----A-C-A-A-A--C-GT-TAG--AGACC-GG---TCT-G---C---G-CTGTGA-C-G......................AGTAATAATACTGATAATAAG 6727
H2U.FR.96.12034              ..GG-T-A--AC---A-CCAT-GGTG-GT--TAGA--T-----G-CA...GGCTTA...-A---G-A----AAA--AG-GT-TA-A-A-ACA-GG---TCT-A---C---G-TTGTGA-C-G..........................AATACTAGTGAAAGT 6751
ASC.UG.10.RT03               ..GG--AA--TCCTCA-TCTT--CTC--T---A----T--C--G-C-...C--G-T...T-T-A--A--G-AAAGCTC-TT-T-GG--TCC--------C-T-G---T--TGG-A-CA-TC-----G-.........................ACTAATGGA 6219
ASC.UG.10.RT08               ..GG--AA--TCCTCA-TCTT--CTC--T--------T--C--G-C-...C--G-T...T-T-A--A--G-AAAGCTC-CT-T-GG--TCC----------T-G---T--TGG-A-CATA-G------...................AGCGAAACTAATGTCACAGGG 6249
ASC.UG.10.RT11               ..GG--AA--TCCTCA-TCTT--CTC--T--T-----T--C--G-C-...C--G-T...T-T-A--A--G-AAAGCTC-TT-T-GG--TCC----------T-G---T--TGG-A-CATA-----...........................GGTACTAATGGC 6207
COL.CM.x.CGU1                ..-ATTG-A-AG--TATCC-TGGT-T--T---AGAA-G---TCA-CAGCTTTCTTA...C-T-A--A-...AGAA--G-GTT-G-ATT-GGA----CAGT-GAA---T---CTGTGC-T-GA---A-A-A-GACAGTAATAGT......................... 6195
COL.UG.10.BWC01              ..CAGTG--TA----AGACTTGGT-TG-G--TAAG--T--CTC--CAGCATTCTTA...T---A--A-...AG-A-AG--TT-AG-TT-GGG----C--T-GAG---T--GATATT-GCAGC---AGG-GCAACAATACA.............................. 6021
COL.UG.10.BWC07              ..-ATTG-CCT----AT---TGGT-TC----TAGAA-G--C-GT-CTGCATTTCTG...C-T-A-AA-...AG-A-AG-GCT-G-ATTTGA--A-AGA-T-GAA--CT---C-ATTGATA-G----GA-A-GGTAGTAGT............................. 6222
DEB.CM.04.SIVdeb04CMPF3061   ..-AC-ACACC-TGTC-C-ATATTGG--T-----------G--TCA...GGTCCA...TAT----A----AAAG-AG-G-G-C----CA-G-GGCTA-C-GA----C-GGA-TGGGC-A--GGAA--GAGACAGGGGATACT........................... 6658
DEB.CM.99.CM40               ..-AC--AACC-TGAC-C-ATATTGG--T--T--A--T---G--TCA...GGTCCT...TAT----A----AAAG-A--GAGC-CA---G-A-GGCTAG-GGA------CCAGTGGGCG------A-AGA-GGTAGTGGGAAC.......................... 6528
DEB.CM.99.CM5                ..-AT-A-ACAGTGAC-C--TATTGG--T--TA-C------G-T-CA...GGGCCA...TAT--G-AC---AA-G-A-TGC--AG----G-G-GG-TAG-GGA---C---CAGTGGGAGACAG-AA-CCAAACAAATACTAAC.....................AAAAGT 6476
DEN.CD.x.CD1_CMO580407       ..-AT-A-ACT-CC...ATCT-CCTT--G--TAA---------G-C-ACAGG--AAGGCT-T--G-A----AGAGTA-TGT-TG-A--TCA--G-AGA-T-GAA--G----GGA-......................................GATAGGGAAACTAAC 6836
DRL.DE.11.D3                 ........-TA-C--A--CA-AC-G-GTA--TAAG-A----G-G-CA...GG-TTG...TGC--G-ACTGTAA-G-AG--ATCACA-A-AA----AGGT-TGA----G--G-TTGT-A-TG-GGCA--GTAACTGAATCACCAGAAAAT...AATAATTGTAATGGT... 6458
DRL.x.x.FAO                  ........-TA-TG-A-ACA-AC-G-GTA--TAAA-A----G-G-CA...GG-TTA...TGC--G-ACTGTAA-G-AG-GATCACA-A-AAC---AGGT-TGA----G--G-TTGT-A-TG-GGCA--GAAACTGAATCACCAGAAAGC...AATAACTGTACTAATAAA 6548
GRV.ET.x.GRI_677_gri_1       ..GA-----CAG-G-A-CCAGC-TCC------A----TGCA--TGCA...GG-TAT...CA-----A-GTAAAAA----TT-TAGCATGACC-GG---G-T-AG--GT--G-CTGC-ATA----AACAGGAAGTGAAAAGGGA........................... 6768
GSN.CM.99.CN166              ..-ATGGAACAGG-CA-CCCT--T-T-----T--C--T---CAG-C-...---GAA...T-T----ACC-AAA-A-AC-GATG---AGCC-C--C-GG-G-GAA--C--T--GGA-GAGACCC---G-..........................AAC 6719
GSN.CM.99.CN71               ..-ATGGAACAGG-CAGCCA---T-T-----T--C--T--CCAG-CT...---GAA...T-T--G-AC--AAA-AGAC-GATG---AGCC-C----GG---GA------T--GAGGGC-C-GG---G-..........................AAT 6710
LST.CD.88.SIVlhoest447       ..-ATGA-ACT-CT-AGATT-AT-G-G-G--TAGA-AT---G-T-CA...GG-CTA...TGT--G-ACTGTAG-ACAG--AT-A-GCAGAG----AGGT-TGA---G-C--C-TGTGATA-G----GA................................. 5908
LST.CD.88.SIVlhoest485       ..-ATGA-ACT-CT-AGATT-AT-G-G-A--TAAA-AT---G-A-CA...GG-CTA...TGT--G-ACTGTAAAG-AG--AT-A-GCAGAG----AGGT-TGA---G-C--CCTGTGATA-G----GA................................. 5905
LST.CD.88.SIVlhoest524       ..-ATGA-ACT-C--A-ATT-AT-GGG-A--T--A-AT---G-G-CA...GGGCTT...TGT----A-TGCAA-G-AG-GAT-A-GCA-AA----AGGT-TGA---G-C--C-TGT-AT--G----G-.................................. 5911
LST.KE.x.lho7                ..-ATGA-ACC-C--AGAG--ATGC-G-A---CAG-AT---G-A-CA...GGGTTA...TGC----ACTGCAG-AC-G--AT-A-ACA-AGC---AGAT-TGA----G---CGTGCTC--GAG-GAGG..............................GAG 6971
MAC.US.x.251_1A11            ..GGCT----AC---AGC-A--G-T-GG---TAAA-----C--G-CA...GGGTTA...-A-----AC---AAAA-AG-GT-CA---A-AC--GG--CTCTGCA---T-GG--TGTGA-C-A.............................GGGAATAGCACTGGTAAT 7205
MND-1.GA.x.MNDGB1            ..GA-GAAA--GT--AGAG--ATC-TGTA--TAGG-AT--C--A-CA...GG-CTA...TGC--G-A-TC-AA-G-AG--AT-GTAA--AA----AGAGGGGA----G-G-A-TGTGA-A--.................................... 6305
MND-2.CM.98.CM16             ..GA--CTACA-C-CA-C---AC-GGGT---TAAA-------CT-CA...GG-TTA...TGT--G-A-TGCAAATT-G-GAT-G-A-A-AA----AGAT-TGAA---G---CTTGT............................ACAAATGGCACAAATAAC 6878
MND-2.GA.x.M14               ..GAG-GTACC-C-CA---A-AT-GGGTA--TAAA--T--C-CA-CA...GG-TTA...TGT--G-A-TGTAAAATAG-GATTA-A-AGAGC---AGAT-TGAG---G-G-CCTGT-CTA-A.........................AAGGAAAATGACACTGAA 6804
MND-2.x.x.5440               ..GAG-CAACA-C-CAGC-A-AT-G-GTA--TAAA-----C-CG-CA...GG-TTA...TGT--G-ACTGCAAATT-G--AT-G-G-A-AAC---AGAT-TGAA---G---CTTGC-C-A-GCT-A---AAACTGGTAGTGCT.........ACTAATAGTACTGAACCT 6505
MNE.US.x.MNE027              ..GGCT----AC---AGCCA--G-T--G---TAAA-----C--G-CA...GGGTTA...-A-----AC---AAAAGAG-GT-CA---A-AC--GG--CTCTGCA---T-GG-TTGTGA-C-A.............................GGTAATAGCACTGAAGAT 6680
MON.CM.99.L1_99CML1          ..-ATGGAACAGG-CAGCCAG--T-T--T-----A--T--CCAG-CT...---GAA...T-C--G-A---AAA-A--C--ATG---AGC--A----GG--GGAA--C-----GAAGGAG--AGGAAGC-A-GGC.............................. 6733
MON.NG.x.NG1                 ..-ATGG-AC-GG-CAGCCTC-TT-C--T--T--A--T--CCAA-C-...---GAA...T-T----AC--AAR-A-AC--ATG---AGC--G---KGG--GGAG--C--T--G-A-GACA-GG--A-C--AACT.............................. 5346
MUS-1.CM.01.CM1239           ..-ATGGTACAG-GCA-AGC--GTTC--T--T--A------CAG-CT...---GAA...T-C----AC--AAA-A--C-GATG--CT-C--A-----C-CTGA----C-G--G-GTGAG---------................................ 6661
MUS-1.CM.01.SIVmus_01CM1085  ..GAT-GT-CTGG-CA-CCT--GTTC--T-----C------CAG-C-...---GAG...T-C--G-A---AAA-AGAC-GATG---T----G--C-T-GTTGA----C-G--GAG-GACAGCTCA--G-G-GAGAATGAT........................... 6689
MUS-2.CM.01.CM1246           ..-ATGA-ACAC--TTGACC--GT-C--T--T--C------CAG-C-...--CGAG...T-C----A-GTAAAAA-AC-GATG--CT-TC-C--C-T-GT-AAA--CC-C--GAA-GGG--G-----G.................................. 6796
MUS-2.CM.01.CM2500           ..-ATGATACA-TGCA--TC--GT-C--T--T--C------CAG-C-...---GAG...T-C--G-AC--AAA-A--C--ATG---T-TC-G----T-GTGGAC---T----GAGGGAT-GG-GC---................................. 6722
MUS-3.GA.09.09GabOI81        AC-GGTCCACTGG-CA--GAG-GT-C--T-----C--T---CAG-C-...---GAA...T-C--G-A----ACAA-AC--ATG---AGT--A------GTGGA----G-C--GT-CTC-C-A--C-GG-A-GTAACAGAA........................... 6290
MUS-3.GA.11.11GabPts02       ATGGT-CA--AGG-CAGCCTG--T-T--T-----A--T--CCAG-CT...---GAA...T-T----A----AAAA--C-GATG---AGT--G--C---GT-GAA-----T--GT-T-GT-GA-G--GG-ACAACACAGAG........................... 6291
OLC.CI.97.97CI12             ..-AGCCA--A-C--A-A--C-T---C-A--TAGG-AT-----T-CA...GGGTTA...TGT----ACTGTAAAATAG--ATGA--CA-ACA---G--TTGGAA---G--GC-TGT-C-A----C................................. 6373
RCM.CM.00.SIVagi_00CM312     ..GC-G-A---G-T-A-A-TG-T-TGC-G--TA-A--TGCAG-AGCA...GGGTTG...-A-----A-G-AAA---T---T-CA---ATAC--GG---TC-AGA--CC-CTGGTGTCAGA-AGC-A-C..........................AAT 6286
RCM.CM.02.SIVrcm_02CM8081    ..GC-G-TCCAG-T-AG--TG-T-TG--T--TGAG--TGCAG-GGCA...GG-TTA...-A---G-A-G-AAAAT-T----CTA---ACACG-GG--CTC-AGA---C-TTGGTGTGAGA-AG--A-A-A-.........................AATAATAAGACTACT 6364
RCM.GA.x.SIVRCMGAB1          ..GA-G-TC--G-T-C---TG-G-TGTCT--TGAA--TGC-G-AGCA...GG-CTA...-AG----A-G--AAA--C--GT-TA---ATACC-GG---TCTAGA--CC-TTGGTGTGA-A-GG-AACA..........................AATTCTACAAATTCTACA 6502
RCM.NG.x.NG411               ..-GTG-CCCTG-C-C---TG-C-TG--T--TGAA--T-CGG-GGCA...GGGTTA...-A-----A-G-AAAAAT-----T-A---ACACA-GG--CTC-AGA---C-CTGGTGT-AG-CAGG-A--...........................AATAAT 6546
SAB.SN.x.SAB1                ..GAG-----A----A-C--GCC-TG-----T--C--TGC---GGCA...GG-TAT...-G---G-A-GTAAA-A-A--TT-C-CAA--G-A-GGG--G-T-AA--GG-GG-TTGTGA---GGGAAGGGAAAAATCA.....................AACGCAACACAT 7071
SMM.CI.79.SIVsmCI2           ..GG-C-A--ACC--A-CCCC---TTGG----AAA--T-----G-CA...GG-CTT...-A---G-AC--AAA-A-AG--T-TA---AGACA-GG---TC--AG---T--G-CTGTGA-CGA---.............................GGTAATGGCAGT 6356
SMM.LR.89.SIVsmLIB1          ..GG-T----AG---A-CCA--G-T-GGA--TAGG-----C--G-CA...GG-CTA...-AG--G-AC---AAAA--G-GT-CA---AGACA-GG---TCTAGA---T-GG-GTGTGA-C-A---...........................AATAATGAGAAT 6265
SMM.SL.92.SIVsmSL92A         ..GG---A---C---A-CCA--G-T-GGA--TAAG--T-G---G-CT...GGGCTA...-GG----AC---AAAA--C--T-TA---AGACC-GG---GCCAGG-----TG-GTGTGA---AG-C.........................AACCAGACTGGC 6294
SMM.SL.92.SIVsmSL92B         ..GGC--A---C-G-A-CCA--G-T-GGT--TCAA--T--C--G-CA...GG-CTA...-A--A--ACC--AAAAG-C-GT-CA---A-ACC-GG--CTCCAGA---C--G-CTGTGAGC-AGGA...........................GGAAATGAA 6265
SMM.SL.92.SL92B              ..GGC--A---C-G-A-CCA--G-T-GGT--TCAA--T--C--G-CA...GG-CTA...-A-----ACC--AAAAG-C-GT-CA---A-ACC-GG--CTCCAGA---C--G-CTGTGAGC-A.........................GGAGGAAATGAA 6645
SMM.US.04.G078               ..GACT-A---C-G-A-CCA--G-T--G---TAAA--T-----G-CA...GGGTTA...-A-----A---AAA-A-AG--T-TA---A-ACA-GG---TC-AGA---T-GG--TGTGA-C-A--G.........................AGCAATGGAAGT 6490
SMM.US.05.D215               ..GGC--A--AC-G-A-CCA--GGT-TCA---AGA--T--C--GTC-...GGGCTA...-AG----AC---ACAA-AG--T-CA---AGACC-GG--CTCTAGG------G--TGTGA-C-AGGG.......................GCTGAAAATGATACA 6466
SMM.US.06.FTq                ..GGG--A-GA-GT-A---AT-GGT-TCT---AAA--------G-CA...GGGTTA...-AG----AC--AAA-A--G-GT-CA---AGACC-GG--CTC-AGG---T-GG-CTGTGA-CGAG-C.....................GGAAATGAAATAGCAAAT 6491
SMM.US.86.CFU212             ..GG-T----ACC--A-CCA--G-T-GGA--TAAA--------G-CA...GG-TTG...-A------C-G-AAAA--G-GT-CA---A-ACA-GG--CTCTAG----T-GG-CTGTGAGC-AGG-.........................GGTAACAACACAAAT 6461
SMM.US.x.H9                  ..GGCT----AC-G-AGCCC--GGT--GT--TAAA--T--C--G-CA...GGGTTA...-A-----AC--AAA-AGAG--T-TA---A-ACA-GG---TC-AGA---T--G-GTGTGA-C-A---......................AGCAATGGAAATGAAACT 6686
SMM.US.x.PGM53               ..GGCT-A---C-G-A-CCA--GGT--G---TAGA--T--C--G-CA...GGGTTA...-A--AG-AC--AAA-A-AG--T-TA---A-ACA-GG---TC-AGA------G--TGTGA-C-A........................AGTACTACTAGTAGAAAT 7143
SMM.US.x.pE660.CG7G          ..GGCT----AC-G-AGCCC--G-T-GGT--TAAA--T--C--G-CA...GGGTTA...-A---G-AC--AAG-ATAG--T-TA---A-ACA-GG---TC-AGA---T----CTGTGAGC-GTCA...........................GCGAATGAAAGT 7199
STM.US.89.STM_37_16          ..GGCT----AG---A-TCAC--GT-GGT--TAAA--------G-CA...GG-CTA...-A-----A---AAA-AGAG-GT-CA---A-ACG-GG--CTCTAGC---C----CTGTGA-C-A.........................AATGTCACTGGCGAA 6848
SUN.GA.98.L14                ..GAG-ACATAGCG-AGAGC-AT---GTA--TAAA-AT---G-G-CA...GGGCTT...TGT----A-TGTAAAAC-G-G-TGG-ACA-AAC---AGGG-TACA--AG-T-CTTGC-AT-G------C.............................. 7039
SYK.KE.x.KE51                ..GAC-ACTCT-C--A-ACT-GGT-C--T--T--A--T--C--G-CA...GGGG-T...T-T-AG-A----AAA--TC-GT-TA-------C-----C--GTCA--------GAG-GAG--A...........................GGAAAT 6463
SYK.KE.x.SYK173_COMGNM       ..GAC-ACAGT-C--AGCCA-G-TTC--T--------T--CT-G-CA...GGGG-G...T-C-AG-A---AAAA---C--T-TAGA--T--C--C-----GGA---CT-G--GAA-GAG--A...........................GGCAAT 6767
TAL.CM.00.266                ..GAC-ATACA-C--A-CCAG--T-T--T--TA-A-GG--C--G-CA...GG-GCA...T-T--G-A----AG-T-T-CCT-C-T-T-TCAC-GG--CCT-GCA--CC----GAAGGA-CCAGGCA--............................. 6576
TAL.CM.01.8023               ..-AC-AT-CC-C--A-CCAG-GT-C-----TA-A-GG-----G-CA...GG-G-A...T-T--G-A----AA-T-C-CCT-T-T-T-TCA--GG---CT-GCA--CT----GAAGGA-CCAGGCA--GGC..........................AAT 6120
TAN.UG.x.SIVagmTAN1          ..-GC-----TT-T-AGA-T-GTTCT-TA-----C--TGC---GGCA...GG-TAT...-GG----A-GTAAAAA-----T-TA--AGTAC--GG---G-T-AG--GT-GG--TGTGAGA-GG-AA-C-AC.......................ACCACAGGC 6860
VER.DE.x.AGM3                ..GA-T-AA-TG---AGCCTGCTTCT--T--TA-A--TGCA--GGCA...GGGTAT...G------A-C--AAAA----GT-T-CA-TGG-G-GGA--G-TGCA--A--C--GTGT..............................AAGAAGGGTAACAATTCT 6395
VER.KE.x.9063                ..GA-C-----G---A--CAGCTTC---T--TA-A--TGC---GGCA...GGGTA-...--T----A-C--AA-A-A--TT-T-CA-T-G-A-GGA--G-TGCA--G--CTATTGT-AGA-C-GCTC--GCTCTAACTCTAGC.......................... 6897
VER.KE.x.AGM155              ..GAG---A-TG-C-AG-CAGCGTC-------A-C--TGC---GGCT...GGGTA-...--T--G-ACC-AAA-A----TT-C-CA-T-G-A-GGA--G-TGCA--A--CT-TTGT-AGCG--G-ACAT-GCATAATGGG............................ 6877
VER.KE.x.TYO1_patent         ..GAG------G-T-AGCCCGCCTCT--T--TA-A---GCA--GGCT...GGCTAT...G------A-C--AA-A-A--TT-T-C--TGG-G-GGA--G-TGCA--A--CTATTGC-A-A----GAC--A-AGCACT................................ 6876
WRC.CI.97.97CI14             ..G-T-ATACTTC--AGATT-ATC-TGTG--TAGA--T--C-C--CA...GG--TG...TGT-AG-ACTGTAGAATAG--AT-A-A-A--C----AGAGG-GA----G-GG-TTGT-ATA--.................................... 6868
WRC.CI.98.98CI04             ..G-G-ATACCTCT-AGATA-ACC-TGTA--TAGA--T----C--CA...GG--TG...TGC-A--A-TGCAG-ATAG--ATTA-A-A--C----AGAGG-GA----G-GG-TTGC-ACA-A---..............................GGG 6919
WRC.GM.05.Pbt_05GM_X02       ..T-G-ATCTC-C--ACACA-ATC-TGT---TAGA--T---G-A-C-...GGGTTA...TGC----A-TGTAAA--AG--AT-GTA-AG-CC---AGAGG-GA----G---G-TGC-C-C-AC--.........................GGTAATGGT 5812
```

Draft of 30 September 2015, not for public release
plvdna_main.tex r8876–10546

V2 loop end

```
H1B.FR.83.HXB2          ...ACCAGCTATAAGTTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACGTTCAATGGAACAGGACCA...TGTACAAATGTCAG 6952
Env                     ..  T  S  Y  K  L  T  S  C  N  T  S  V  I  T  Q  A  C  P  K  V  S  F  E  P  I  P  I  H  Y  C  A  P  A  G  F  A  I  L  K  C  N  N  K  T  F  N  G  T  G  P  .  C  T  N  V  S
H1O.CM.91.MVP5180       GGA--A-CA----T---A-T-A-------T--A--A-T--C-AG-----------G------AGT--------C----------A-----C--------T--AA-A--A-A---C--CT-T--G------G-C-CAGAC--T----------C-T-----CCAC---A-TTC 6986
H1N.CM.95.YBF30         AAT-AA-CA----G-C-A-T--A-------T---A---CTG-G-----A--T-----T---ACT----------A-----C--------A--AC-A--C-A---C--A-G-------G-AGGA-AC--T-G-------AT---AGC-----------G-- 6545
H1P.CM.06.U14788        ------CA----GA--A-TT-A---C------A-TAC----------A--C-----A--GAA--------GC-A--------G-------A--A--A--G-A---AC-AA-G-----C---C-G-CAGGA--T-----T-------T-------AT--GACAGT 6465
CPZ.CD.06.BF1167        AAT---TCT---T-T--A-T--A---C---T---A---C---A--T----AA---GAG---TCT--A-----A---G----T--T--G--C-----T--AC-A--A-A---T--G-----G-----AG--G-A-AT----GG----A----AAC----------CA--G-- 7080
CPZ.TZ.00.TAN1          AAT-AT--T---T-T--ACAT-A---C---T-------------T--A-AA---GAT--ATCTA-T-----A----------CAGA--C-----T--A--A--C-----CC-GT---------GAG--C---AT----CA--G-A-----A---C---CT-C------TC 6596
GOR.CM.07.SIVgorCP2135con TAT-GTGAG---G---A-TC-A------T-----TACT--AT-G--A--T------C--A--C-----------C-A--T-----G-------A--A--R--C-A---CC-GA-G---------G----ACAA--T-----T------GA-----C-A------GTC 6509

V2 loop end

MAC.US.x.239            GAA-GT--A-G-T-CA---ACCAC-------T--T--T--CCA-G--T-T---GAC--ACAT-AT-GG--TG-T---AGAT-TAGG-------A--TC-A----A---TT-G--T-G-------G-C-CA-AT-ATTCA--CTTTATG--TAAA---T-T--G--GGT 7379
Env                      E  S  R  C  Y  M  N  H  C  N  T  S  V  I  Q  E  S  C  D  K  H  Y  W  D  A  I  R  F  R  Y  C  A  P  P  G  Y  A  L  L  R  C  N  D  T  N  Y  S  G  F  M  P  K  C  S  K  V  V
H2A.SN.86.ST_JSP4_27    GAG-AA-CA-G-T-CA---ACCAC--C-----A-------C---G--T-A---GAC---CAC-AT-GG--TA-T--GAGGT-TAGA-------A--AC-G--------CC-G----G---C----G----CC-AT-ATTCA--CTTT-AG--CAAT---T-T--G--AGT 6888
H2B.CI.x.EHO            GAA-G--AA-G-T-TA-A-A--CC-----T--A---A-T--CCA-G-AT-A---GAC--ACAT-AT-GG--TAGTT-AAGAT-TAGG-------T--CC-A--A-----TT-G----G---C----G---CT-AA-ATTCA--CTTCATG--CAAC----GT--G--AGT 7423
H2G.CI.92.Abt96         GAG-GT--A-GCT-TA---GGCAC-----T--A---------CAGG-AT-A---GAT--ACAT-AT-GG--TG-T--AAGAT-TAGA-------T---C-A--G-----TT-G----G-------G----GT-AC-ATTCA---TTT-C---TAAT--CT-T--A--AGT 6792
H2U.CI.07.07IC_TNP3     AGC-G---A-GCT-TA---ATCA-----T-----T--T---CA-G--T-----GAC--ACAT-AT-GG--TG----AAGAT-TAGG--C---------C-A--G-----TT-G----GG------G----CTGAT-ATTCA--CTTCA-C--TAAG--CT-T--G--AGT 6897
H2U.FR.96.12034         GAA-GT-AA-G-T-CA---ACCAC--C-------T--T--CCA-G--T-T---GAC--ACAT-AT-GG--TG-C--AAGAT-TAGG-------A---C-A--C-A---TC-GT---G-------G----CT-AT-A-TCA---TTTATG--TAAT--CT-T--A--AGT 6921
ASC.UG.10.RT03          ACAGGGGAA---T-TG-A-AGTA---C-----G--T-CT--G-----A------GAT--AT-TCAT---C----TT-C--AG-T-----C---C-A-A--A--G-A---AT-GT-C-----C---G-C-TAC-A-GG-----CA---G-----------A------G-C 6386
ASC.UG.10.RT08          GAAGGGGAA---T-TG-A-AGTA---C-----G--T-CA--A-----A------GAT---T-TCAT---C-A--TT-C--AG-T-----C--CC-A--A--A--G-A---AT-GT-C-----C---G-C-TAC-A-GG-----GCA---G-----------A------G-C 6416
ASC.UG.10.RT11          ACAGGGGAA---T-TG-A-AGTA---C-----G--T-CA--G-----A------GAC--AT-TCAT---C----TT-C--AG-T-----C---C-A--A--A--G-A---AT-GT-C-----C---G-C-TAC-A-GG-----GCA---G-A-------A------GTC 6374
COL.CM.x.CGU1           ....T---GGCA-CAA---A-GA----GCA-AT-ATACAG----C--A-TA---GAT-T-ACTATAG-G--T---G--AGG-C-GGG-TC-----AG-TC-A--A-ACAT-C--T---G----G--G----A-A--GGG---G------G------C-AC-----G-C 6359
COL.UG.10.BWC01         ....TA--GGC--CTA---AGGA------G-AT-ACACAG----T----TA---GATC-TACTATAA-G--T--TG-AAGT-C-GGA-TC-----AG-AC-A--G-ACAT-T--T-G-G---G-G----A-A--GG----------G-A------T-C-----G-C 6185
COL.UG.10.BWC07         ....TA--AGCA-CAA---A-GAC---C-A-AT-ACTCAG-AT--GCTATA---GATG--AGTATAC-G--A--T-C-T-A---GGC-T------AG-AC-A--A-A---TT-AT--GT-------G---CTCA--GG----------GAG-------ACC-G--ATA 6386
DEB.CM.04.SIVdeb04CMPF3061 ....GGG-AGC-T-TA---GGTA---C-----GG--T---A--A--T--A------GAC--AAGTCA----A-A--TT-CAAA---AGG--C-----A--A-----A-A--GTC-AA-G-G----G-G-C--A-AT--------T-------TT---------C-----GTC 6822
DEB.CM.99.CM40          ....GA-CAGGGT-CA---A-CAC-----TGA---T-----C--C---T-T---GA--CTAGCCG----A-A--TT--AAG--CAGG--C-----A--A-----A-A--GTC-A----G----G--G----A-AT--T-----T-----GTT-.....C-AC--C--G-C 6692
DEB.CM.99.CM5           ATA-GACATGGGT-CA---ACCAC--C----AT-----T--A------T-----GA--CCAGTCG----A----TT--AAG--AGA-C-----A--A-----A-A--GTC-GT--G----G--G-C-GA-AC--T----T---C--CTTG---------C-----G-C 6643
DEN.CD.x.CD1_CMO580407  CAA--ATAT---T-TA---GGTA---C--TCAA--CAC------C--A--T---GA------TGT------A---T-CAAA---AGGATG-----A--A-----C-----TC-C--T--G------G--C-TC-T-GG-G---GTAT--C-TT.....A----A-GTC 7003
DRL.DE.11.D3            ACA-GA-CA-G-T-TG----C------C---T-AA----G-----C----AT----AAT----CC---ACA--TATG--GAAGT-CAGA-TG-----T--AC-A--G-A-ATAT-G---G-------G-A--ACT-AAT---AA-T-GTAAG..........A----A-A-C 6622
DRL.x.x.FAO             AAA-GA-CA-G-T-TG----C------C---T-AA----G-----C----AT----AAT----CC---ACA--TATG--GAAAT-CAGA-TG-----T--AC-A--A-A-ATAT-G---GG------G-A--ACTAAAT-TAA-T-GTAAG..........A----A-A-C 6712
GRV.ET.x.GRI_677_gri_1  AGT-AGGAT-G-T-CA---T-CA-----TGAT----G--A-A-G-A--T---GAT--AAC--AT-GG--TA-TT-AAGAG--AGA--C-----A--A--A-G-A---TT-G----------G----GAT-AT-GA--CTTT-CT---AAG--C-AG-----TTC 6938
GSN.CM.99.CN166         CAG-GTG-A---T-TA-T-G--A-------T------TAT--C--T--AAAA---GTC---TCC-----CC-A--GG----------C--C-----T--AC-C--C-A---T--G----------C---GTT-AT-----CA---GT---CA-------CTAC-----GTC 6886
GSN.CM.99.CN71          GAC--TGA----T-TA-T--A---C-----A---TATG-C-G--AAAG---GTC---ACC-----CC-A--GG----------C--C-----T--C-C--C-----T--G----------C---G--GCC-AT-----CA---GTG--CAAG.....T-C-----ATC 6877
LST.CD.88.SIVlhoest447  .......CT-G-T-TA----CCAC----TGATAGTA-T---CAG--A-AT--CAAT----G-GTAA-ATCTAATGC-TATT-TAGGCT------A--A--A--C-A-AT-C--T----G--C--CG-A---ATT-AATTTCTC-G-TAAT..........C---A-A-C 6066
LST.CD.88.SIVlhoest485  .......C--G-T-TA----CCAC----TGATAGTA-T---CAG--A-AT---AAT----G-GTAA-ATCTAATGC-TATT-TAGGCT------A--A--A--C-A-AT-C--T-------C---G-A--ATT-AATTTCTC-G-TAAT..........C---A-A-C 6063
LST.CD.88.SIVlhoest524  .......CT-GCT-CA----CCA---C--TGATAGTA-T---CTG--A-A---CAAT----G-GTTA-GTCAAATGC-TATT-TAGGCT------A--A-----C-ACAT-C--T-------C---G-A---CTAAA-TT-TC-G-TAAT..........T---A-A-C 6069
LST.KE.x.lho7           AAC-GA-C--G-T-CA----TCA-----TGAT---A-A--A-----A-AT---AAT----GGGTAA-GC-AAATGC-TATT-TAGGCT------T--A-A--C-A-AT-C-AT-G-G---C---G-AC-ATTAAA-TT-A-T-A-AA-.........GA----A-A-C 7135
MAC.US.x.251_1A11       GAA-GT--A-G-T-CA----ATCAC-------T--T--T--CCA-G--T--T---GAC--ACAT-AT-GG--TG-T---AGAT-TAGG-------A--TC-A----A---TT-G--T-G-------G-C-CA-AT-ATTCA--CTTTATG--TAAA---T-T--G--GGT 7375
MND-1.GA.x.MNDGB1       ....AT-CT-GCT-TA---ATCA-----TGAG-----T-A----G-A-A-----AG----G-CTT--GATAAG-...TGT--TTTAGG-----TG-TC-A--A-A--TC--GT---G--A----G-G---TTAAAT---AAT-A-TTG.........GC---A-ATC 6463
MND-2.CM.98.CM16        ACATATTCT-G-T-TA-----CAA--C--T--A-----A--A--C----A----AAT--A-C----ACA--TGA---AAAAT-TAGACTG-----A--TC-A--A-A--TTC-CT-G-GG--CCTAG-A-GATT-AATGTCTCT-A-AA-..........C---A---C 7042
MND-2.GA.x.M14          ACTGAAGAA-G-T-TA-----CAC-----T--A-----A--A-------AT---AAT----CC--AACA--CAA---GA-AT-TAGA-TA-------TC-A--G-ACATTC-G----G----CG-G-G--ATT-AAT--AACT-A-TT-.........G-C------TC 6968
MND-2.x.x.5440          GAATATGAG-G-T-CA----TTCA-----G-AA----A--A------A-C--AAC--A-CT---ACA--TAG---GA-AT-TAGG-TG-----A--CC-A--G----TCT-A----------G-A--ATT-AA---GAC--A-TT-.........GGG-----GTC 6669
MNE.US.x.MNE027         GAA-GT--A-G-T-CA---ATCAC-------T--T--T---CA-G-AT-T---GAC--ACAT-AT-GG--TG-T---AGAT-TAGG-------A--TC-A----A---TT-G--T-G-------G-C-CA-A--ATTCA--CTTTATG--TAAC---T-T--G--GGT 6850
MON.CM.99.L1_99CML1     ....GTCA----T-CA-CCTG-A---C-----A--TA---A--C--A--T---GAG---TCCAAT-A------G----TT-G----------A--CC-A--C-A---TT-AT-G-G---G--G-CCCCG-C--T-CG---CA---GAGT-----T-T-----TTC 6897
MON.NG.x.NG1            GATGA-TAT---T-TC--CT--A------T----Y-TA---Y--C--A------GA----TCCAAT-A------T-----T--T----C-----T---C-C--C-A---T--AT-G-G---G--G---CCAK-C----CA--GCAGAACAAT.....CT-R-----TC 5513
MUS-1.CM.01.CM1239      .......CT---T-CA-ATT--AC----------TA---A--T---CAA---GA---AGCAG---CC-A--GG---G-----A--C-----A--T--A-G-A-T-TC-AT-G-------G--GCT-AT--T-CA---CA----AA-------AC-----G-C 6822
MUS-1.CM.01.SIVmus_01CM1085 .......T---T-TA-C----AC--C-----A---TA---C--C-----A---GAG---AGCAG---CC-A--GG---G-----A--C-----A--C--G--G-ACAGCT-G-------C---G--GCC-AT--T-C---GCA----GA-------AT-----G-C 6850
MUS-2.CM.01.CM1246      .....-CA---T-CA-T----AC--C-----T--TTAT--C------CAG---GAG---AGCAG---CC-A--GG-C--A--C-----A--C--G--G-A-A-AGTC-CT-------C---G--GCT-AT----CA---CA---CATT-------AT-----A-C 6957
MUS-2.CM.01.CM2500      ...GA-G----CT-TA-C----A--C-----A---TAT-----C--ACAG---GAG-AAGCAG---CC-A--GG----------C-------A--A--A-A-AGCT-GT----------G--GCT-AT----CA---CA---CAA-------GG-----A-C 6886
MUS-3.GA.09.09GabOI81   .....------T-CA-T----A---C---T---A-T-CT-----C--AT-T---GA-----C-CAT-----C---G-G--A--------C------A--A--A--A-ACA-AT-G--------C---G--GCC-AT----CA--CCA----ATG------CAC-----ATC 6448
MUS-3.GA.11.11GabPts02  .....------T-CA-C----A---C---T---A---CT--A--C---T--GAG----CTCAT-----T---G-G--A--------C------A--A--A--G-ACT-AC-G--C---G--C--CG--GCT-AC--T-CA---CA---GA-C-----TAC-----ATC 6449
OLC.CI.97.97CI12        .......CA-G-T-CA--T--CA----------T--G--A--T----AT----A-T-T-C-GT--CA--T-AGCAAATGG-GA-G-TG-----A--AC-A---AGG-T-C--A----G---------G---AGAA---CC-.................C-G---CA-TTC 6525
RCM.CM.00.SIVagi_00CM312 GTG---CAA-G-TTTA---G-CAC-------AA-TAG---ACAC---TTT---GA-CCAAG--A--GG--A--CT--AGGT--AGA--C-----T--AC-A--C-----CT-G--------C---G--T-T-T-A--CG--TTTT-ATA-T------CAT-----A-C 6453
RCM.CM.02.SIVrcm_02CM8081 AAG-GGCAA-G-TTTA--CG-CA------T--TA-TAGT---CAG--ATT----GAGCCTAA--A--GG--A--TT--AGG--TAGA-----C--T--AC-A--A-----CC-AT-GGT---C--AG----A-AT-AT-CA---TTT-ATA-C------CAT-----G-C 6531
RCM.GA.x.SIVRCMGAB1     AAA-AG-AA-G-TTTG-A-GGCAC--C-T--AA-TTC---ACA---ATTT---GA-CCAAAG-A--GG--A---T--AGGT--AGA--------T--AC-A--G-----CT-A--GGTC-C--AG----A-AT-AT-CA---CTTT-ATA-C------GTT-----T-C 6669
RCM.NG.x.NG411          ACT-GGCAA-G-TTTA-A-GGCA---------AA--AGT--ACA----TT----GA-CCAAA--A--GG--A---T--AGGT--AGG--------A---C-A-----A---C--A--G------CT-G--GCA-AT-AT-CA--CTTT-ATA-C------CA------T-C 6713
SAB.SN.x.SAB1           ACTGTGG-A-G-T-CA---TCCA-----T--------A--C-A-G----T---GAT--AACT-AT-GG--TA-CT--AGGC-TAGG-------T--A--A-----A---TT-G----G---GCAG---CAGAT-ATTCA---CATAAGG-----GG-----A-C 7238
SMM.CI.79.SIVsmCI2      GAA-G---A-G-T-TA---ACCA---------A-----T--ACAGG-AT-A---GAC--ACAT-AT-GG--TG-C---AGAT-TAGG-------A--AC-A----A---TT-G----G-------G---CA-AT-ATTCA--CTTT-A---TAAT--C--T--G--AGT 6526
SMM.LR.89.SIVsmLIB1     GAA-GT--A-GCT-TA---GCCA---C--T--AAGT-----CA-G-AT-----GAC--ACAT-AT-GG--TG-T--AAGAT-TAGA-C-----A--AC-A--A-A---TT-G-T-G-------G---CA-AT-ATTCA--CTTT-C---TAAT--CT-C--G--GGT 6435
SMM.SL.92.SIVsmSL92A    AAC-GTC---G-T-TA---GGCA---------AGTA-----CAGG-AT-T---GAC--ACAT-AT-GG--TG-TG--AGAT-CAGA-------A--TC-A----A---TT-G-T-G-------G---CA-AT-ATTCA--CTTT-C---TAAT--CT-T--G--AGT 6464
SMM.SL.92.SIVsmSL92B    AGC-GT---A-G-T-CA---ATCA---C--T--AAGT---T--ACA-G--T-A---GAC---CAC-AT-GG--TG-C---CAGATATAGG-------A--AC-A----A---TT-G-T-G-------G---CA-AT-ATTCA--CTTT-C---TAAT----GT--G--AGT 6435
SMM.SL.92.SL92B         AGC-GT---A-G-T-CA---ATCA---C--T--AAGT---T--ACA-G--T-A---GAC---CAC-AT-GG--TG-C--CAGATATAGG-------A--AC-A----A---TT-G-T-G-------G---CA-AT-ATTCA--CTTT-C---TAAT----GT--G--AGT 6815
SMM.US.04.G078          GAA-GT-A--G-T-CA--CATCA----------AGT--T---CAGG-AT-----GAC---CAT-AT-GG--TG-T---AGAT-TAGA-------A---C-A----A---TT-G--T-G-------G--TCA-AT-ATTCA--CTTT-CT--TAAC---T-T--G--AGT 6660
SMM.US.05.D215          GAG-GT--A-GCT-TA---ACCA------T--AAGT-----CG-G--T--T---GAC---CAC-AT-GG--TG-TG--AGAT-TAGG-------A--AC-A--A-A---TT-G--C-G-------G---CT-AT-ATTCA---TTT-C---TAAT----GT--A--AGT 6636
SMM.US.06.FTq           GAA-GT--A-GCT-CA---ATCA---C--T--T--T--A--CCAGG-AT-T---GAC---CAT-AT-GG--TG-T--AAGAT-TAGA-------A--AC-----C-A---TT-G--T-G-------G---CA-AT-ATTCA--CTTT-C---TAAT--C-GT--G--AGT 6661
SMM.US.86.CFU212        GAA-G---A-G-T-TA---ATCA-----T--AAGT---A--CCA-G--T-----GAC--ACAT-AT-GG--TG-T---AGAT-TAGA-C-----A--AC-A--C-A---TT-G-T-G-------G---TCA-AT-ATTCA--TTTTATG--TAAT---T-T--G--AGT 6631
SMM.US.x.H9             GAC-GT-AA-GCT-TA---ACCA---------AGT--T---CA-G-AT-----GAC---CAT-AT-GG--TG-T---AGAT-YAGA-------A---C-A----A---TT-G--T-GG------G---CC-AT-ATTCA--CTTT-CT--TAAC-----T--G--AGT 6856
SMM.US.x.PGM53          GAA-GT-AA-G-T-CA---ATCA---------AGT--T---CA-G-AT-----GAC---CAT-AT-GG--TG-T---AGAT-TAGA-C-----A---C-A----A---TT-G--T-GG------G---CA-AT-ATTCA--CTTT-CT--TAAC--T-T--G--AGT 7313
SMM.US.x.pE660.CG7G     GAG-GT-AA-G-T-CA---CATCA---------AGT--T---CAGG-AT-----GAC---CAT-AT-GG--TG-T---AGAT-TAGA-------A---C-A----A---TT-G--T-GG------G---TCA-AT-ATTCA--CTTT-CT--TAAC--CT-T--G--AGT 7369
STM.US.89.STM_37_16     GAG-G---A-GCT-CA----G-CA----C--T---AGT--A---CA-G-AT-----GAC--ACAT-AT-GG--TG-CT-AAGGT-TAGG-------A--AC-A----A---TT-G--C-G-------G---CA-AT-AT-CA--CTTT-C---TAAT--CT-T--G--AGT 7018
SUN.GA.98.L14           .......CA-G-T-TA-----CA---C--TGAT---A-T--------A-A---C-AT---GTATTA-GC-AAATGC-TA-T-CAGGCT------A--T--G--G-A-AT-T-AT---G----G--G-A--ATTAAATGC-ACT-AGAA-..........A----A-A-C 7197
SYK.KE.x.KE51           GAA-GTTAT--CT-T--ACAGCA------T--TAGT--A--GCGCA-T--T---GA---ACAGA-T---CGC---T-------T--A--C-----T--AC-G--A-ACT-AC-GT-G--------G---CAGAA--TG-A---GAT-CTGAG------CA------A-C 6630
SYK.KE.x.SYK173_COMGNM  AGT-G-TAT---T-CC--TT-CA------T--AAGT--T--AT-CGCA--A---GAG--ACAGA-----C-A--CT----G-T--A--C-----T--AC-G----ACT-AT-GT-G-----C---G----CA-AT--TG-A---GAC-ATGTC.....------------A-C 6934
TAL.CM.00.266           GGAT--TT-----T-TGCC---G-A---C--TGTG------T--AG-----A--T---GA---AACTCA--ACC--G--T-----T---G--C-----T--A--C-----T--AC-G---CC-GA----G-------G---GCC-A-GG-CA--T-G----AAG------CAC--CA--TC 6743
TAL.CM.01.8023          GACT-ATA-----T-TGCA---G-A---C--TTTG------A--AG-T-------T---GA---AC-CAT-A-C-AG-GT-----A---T---G--C-----T--A--A--C-----TC-GA----G-------G---CC-A-GG-CA--T-G----CAAG.....CAT--CA--TC 6287
TAN.UG.x.SIVagmTAN1     ACG-GAG-T-G-T-CA---TTCAC--C---GA---T--A--A-A-G-A--T----GAG--AACT-AT-GG--TA-CT-AAGAT--AGA--C-----A--A--A--C-----T--CT---------GG---CT-AT-AT-CA----TTT--TGTT.....G------TTC 7027
VER.DE.x.AGM3           AAC-GAGAA-G-T-TA---TTCA-----TGAT-----T--A-A-G-A------GAT--AAC-AT-GG--TGAGT-AAGAT--AGG--C---------A--A--A-----TC-GT-------CC--C-G--T-TGAT-ATGCA--GTTTAAGA--AAC---T-T-----TTC 6565
VER.KE.x.9063           ACA-AAGAA-GCT-TA---TTCAC-----TGAT-----T--A-A-G-A------GA---AAC--AT-GG--TGAGC-AAGGC--AGA--C---------A--A--A-----TC-GT-------CC--G--C-TGAT-AT-----TTATAAG-A-AAT----GT-----TTC 7067
VER.KE.x.AGM155         ACA-AAGAG-GCT-TA---TCCAC-----TGAT-----T--A-AGG-A--T---GAT---AC--AT-GG--TGA-T-AAGAC--AGA--------T--A--A--A-AC--TT-G--T---------G--TG-GAT-ATGCA---TTTAAG--GAA---T-T----TTC 7047
VER.KE.x.TY01_patent    AGC-AAGAG-G-T-CA---TTCA-----TGA------T--A-A-G-A--A---GAC--AAC--AT-GG--T-AGT-GAGGT--AGG--------T--A--A-----A---TT-G---------G--G-AGAT-AT-----GTATAA--A-AAT--CT------ATC 7046
WRC.CI.97.97CI14        AAT-GTGAA-GCT-TA-A-ATCAC-----TGAG---------A-AG--A-A-----A---A-G--GGCAG--TGA-CAATAT---AGAATG-----T--AC-A--A-CA-TTT-GT-G--------G-G--ATT-AA-TCCTCT-AGCT-.....---------GAG-----TC 7032
WRC.CI.98.98CI04        ACT...GAA-GCT-TG-A-ATCA-----TGAA--T--G--A-A---A-AT----AG--A-GT-GGCAA--TGA-CAATATG--AGAATG-----T--AC-A--A-C--TTT-G--T---------G-GCCCTTAAA-TCCTCT-A-TT-.....-----GAG-----ATC 7080
WRC.GM.05.Pbt_05GM_X02  ACACAA-CT-G-T-CA-A-ATCA-----TGAG--T--A--CCAT----AT----A-----C--GGCA---TGAGCAATA----AGAATG-----A--AC-A----A--TAT-AT---G-------G-A--ATTAAAT-G-A-T-AGCT-.....-----GAG-----TTC 5976
```

Draft of 30 September 2015, not for public release
plvdna_main.tex r8876–10546

```
H1B.FR.83.HXB2              CACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGA.........................TCTGTCAATTTCACG...GACAATGCTAAAACCATAATAGTACAGCTGAACACAT 7095
Env                          T  V  Q  C  T  H  G  I  R  P  V  V  S  T  Q  L  L  L  N  G  S  L  A  E  E  E  V  V  I  R  .  .  .  .  .  .  .  .  S  V  N  F  T  .  D  N  A  K  T  I  I  V  Q  L  N  T
H1O.CM.91.MVP5180           AGTG--TACT----------C--C-A----AC----AGT--------AA-AC-G-----G-CA--CT-TAG----A--.........--A--.....................ATTATGGGAAAA---A-T--A...--ATCA--A--G-AT--C------ACC--A-----TC 7129
H1N.CM.95.YBF30             T--T--------C-----------A-A------GA----C-----GT-AA-CC-------A--CT--AATAC---T-GAA-T--T------.........................-AT...--T-G-CAC-GT-ATC-GT-G--G--ATG---TGAGA 6676
H1P.CM.06.U14788            T-T-AC---C----------------A----AC---G-----A---T-AA-AC-C-----A-C-------A-G-G----CCCT--G-A-TT.......................A--CAG---G-GT-A...--T-CA-GA---GT---C------A-AT-A--TGAGA 6608
CPZ.CD.06.BF1167            G-TT-----T---------A-T--ATTT---A-GA--G-T--AGCTT-A-AT--------GTC-G---A-----G---AAC-AAGGCATAC.......................-A---AG-AGC--AACATA-TGCAC-ATT-TTA----A-T-CA-TAAAGCA--GG 7226
CPZ.TZ.00.TAN1              AGT---T--C----------G---TAT--TA-GA--G-C--AGC-T-A-AC--------GTCC--G-A---------AAC-AA-GC-TAC.........................T-TGT-AAT--C...TCAGT-AA--C-C--T--T-----A-AT-T--TGT-- 6736
GOR.CM.07.SIVgorCP2135con   A-T---T--T-----C-----G--C------AC---C--C--A--GT--A-AC-C-----A-C-T-----A----GC-T-CCAC---A-TT.......................AGCAAA---AGT--T...--T-G--GA---GAT--T-----CA-A--AGC--G-- 6652

MAC.US.x.239                GGTCTCTTC---C---AGGATG--GGA-A--CAGACT--T---TGGT-TGGC--T-----A-C-AG-------A-TAGAACTTAT---TAC..........................TGGC--GGT-G-...--T---AGG-CT-TA--T-GTT--A-TAA-T-TTATA 7519
Env                          V  S  S  C  T  R  M  M  E  T  Q  T  S  T  W  F  G  F  N  G  T  R  A  E  N  R  T  Y  I  Y  .  .  .  .  .  .  .  .  .  W  H  G  R  .  D  N  R  T  I  I  S  L  N  K  Y  Y
H2A.SN.86.ST_JSP4_27        AG-T-CTAC-------AGGATG--GGAAA-GCA-ACC--C---TGGT-TGGC--T-------C-AGG------A-TAGAAC-TAT--CTAT..........................TGGC--GGT-G-...--T---AGA-CT-T---T-GCT--A-CAA-TTTTATA 7028
H2B.CI.x.EHO                AGT-TCT-TG...TACAGAATG--GGAAA--CAGACC--T--ATGGT-TGGC--C-----T-CAAGG------A-CAG-AC-TAT--CTAT..........................TGGC--GGT-AA...------AGG-CT-T-----GCT--A-TTCAT-TTATA 7560
H2G.CI.92.Abt96             GGTCTCTAC-------AGAATG--GGAAA--CA-ACC--T--ATGGT-CGG---T-----A-CAAGG------A-TAGAAC-TAT--CTAT..........................TGGC--GGG-GA...------AGG-CT-TA----GTT--A-TAA-T-TTATA 6932
H2U.CI.07.07IC_TNP3         AGTCTCTAC-------AGGATG--GGAAA--CA-ACC--T---TGGT-TGG---T-----A-CAAG-------A-TAG-AC-TAC---TAT..........................TGGC--GGA-GA...------AGA-CT-TA----GTT--A-TAA-T-TTATA 7037
H2U.FR.96.12034             AGTCTCTAC---C---AGAATG-GGG-AA--CA-ACC--T---TGGT-TGGA--T-------C-AG-T-----A-TAGAACTTAT--ATAC..........................TGGC--GGT-GA...------AGA-C--TA--T-GCT--A-TAA-T-TTATA 7061
ASC.UG.10.RT03              AG-------T---------C-----ATA--T-G-CTAGT--ATGG--C-A----------CATAT-A-A-GA-TACA-AT-AGG-GCAGGTT....................ATGAGA--G-AT-ATTCCA-TT-CAG-GT-G-TT-T--CT-CTCAGA--G-CA-A 6535
ASC.UG.10.RT08              AG-------T---------C-----ATA--T-G-CTAGC--ATGG--T-A-C--------CATAT-A-A-GA-CACA-ATAAGG-GCAGGTC....................ATGAGA--G-AT-ATTCCA-TT-CAG-GT-G-TT-T--CT-CTCAGAA-G-CA-A 6565
ASC.UG.10.RT11              AG-------T---------C-----ATA--T-G-CTAGT--ATGG--C-A----------CATAT-A-A-GA-CACA-AT-AGG-GCAGGTC....................ATGAGA--G-AT-ATTCCA-TT-CAG-GT-G-TT-T--CT-TTCAGA--G-CA-A 6523
COL.CM.x.CGU1               AG----GTCC-----C----AGT---ATAT-AC---CATGT----TG-AT-AG-C----C-TCAAA--AGTTGTCT--TTGG-CG-AAGAT.................AGAGAGGGA---TGGAAG-AT...----TCA-GG-C--TTGAGTATTATTG-T-TCCA-A-G 6508
COL.UG.10.BWC01             AG----CTCC--C-------AGT---ATGTGAC----ATGT----TG--T--G------CTTCAGA--A-TTGA----ATGGTC--AGGAT.................AGAGAAGGA--TTGGAAG-AT...--A-G--GA-CCTTTGA-TATTATTG-T-TCCA-A-G 6334
COL.UG.10.BWC07             TG-CAC-TCC-----C----AGT---ATTT-AC-A-C--CT-G--TA-AG-AG-C-----TTC-AA-A-TTT--TG--CTAT-C--AA-C-.................AGAGAGGGAA--TGGGC--AT...--A-G--G---C-TAGA-TATTACTG-T-TCCA-A-G 6535
DEB.CM.04.SIVdeb04CMPF3061  TG----GGC-------A-CTT---ACACA--A-G-C---C--CTGGT-A-A---C-----GTCAGAT-AG---AG--CA-A--A----CATGTC............ATAAGAAAG......GG-C-TGAA...A-T--GA-A-T---A--T-GGA-T-CT-CCG-GTACA 6971
DEB.CM.99.CM40              TG----GGC-------A--CTC--ACACA-CA-G-C---C--ATGGG-A-AA--T-----GTC-GAT-AG--GAGG-CA-A--AGC--CATATA......ATAAGAAAAGAGGTAAAA-GTG-GG-GCA-...A-TGGAAGC-TT--T----G---G-CCGCT--ATACA 6853
DEB.CM.99.CM5               TG----GGC-------A--CTC--ACATA-CA-G-C------ATGGA---AA--C-----TTCAGAT-A----AGG-CA-A--AC--ACATATC............ATAAGGAAG......GGCC-GGAA...A-T--GA-C-TT--TG-C-G-A-C-CA-CA--GTATA 6792
DEN.CD.x.CD1_CMO580407      AG--AG-GTG------G---AG--ACACA-CA-G-CCAGC--CTGGT--A-CA-G--CA-T-CAG---A-CTTCT---TAATCAT-CACA-ATT..............ATAAGAAACACA-GCAAAGAT...A-TC-GA-A-T-G----TT--T-CAGAAA--GTT-TG 7155
DRL.DE.11.D3                AG----G--G-----TAGACAC--GCCTG-CAC-A-C--T-G-ATGT-TGGC--C-------CAAAGCAT--TC----T-AGT----AGA-ACA..............AGAGGAAGACCAG-A--TGT-...--TC--AAAT-TGTGT-T-GG--TA-TAA---ATGGA 6774
DRL.x.x.FAO                 AG----G--G-----TAGACAT--GCCTG-CAC-A-C--T-GC--TT-TGGC--C-------CAAAGCAT-GTC-G--T-AGT----AGA-ACA..............AGGGGAAAACCAG-C--TCT-...A-TC--AAAT-TGTGT-T-GG--TA-TAA---ATGGA 6864
GRV.ET.x.GRI_677_gri_1      AGT---G--T-----TAGATT---C-ATA-TACTA--A-T--AGGGA-AGGA--------T---AG-T-----A-TAGAAC--AG--AT-G......................CAGAAA-GAGGAAATGAT...A-TG--A-AGTT-TA----AGT-GA-TAA-TTTTACA 7084
GSN.CM.99.CN166             AG-----ACC--C---------C---CAA--CT-----G----ATGG--C-AT--------G-CCTATCAGCC--G-A-TAATACC-GA-T-ATG....................ATGAATGG-A-G-A-...A-TG-AT---T-GTT---GGCT-TGGAGAAG-TTATC 7032
GSN.CM.99.CN71              AG-----AC---C--G------C---CAA--TT-----G----ATGG--T-ATC-------G-C-TATCAGCC--GGA-CAATACC-GAGT-ATG....................ATGAATGGCAAA-A-...A-TG-AT---T-GT----GGCT-TGG-GAAG-TTATC 7023
LST.CD.88.SIVlhoest447      AG--ACC-CT--C---G-CTAT--G-TTAGTTC---GAGT-G-TTCT-TGGC--T-----A-CCAATCATAC---C--T-A-T-----CC-CTA.........ACTCCAAAAAAGATG-GA---C--CTA...--AGCAAAGTTTGTATAT-A----GCAGGA--ATGGG 6224
LST.CD.88.SIVlhoest485      AG--ACT-CT--C---G-CTAT--G-TTAGTAC---GAGT-G-TTCT-TGGC--T-----A-CCAATCATAC---G--TAAGT-----CC-CTA.........ACTCCAAAAAAGATG-GA---C--CT-...----GCGAAGTTTGTATAT-A----TCAGGA--ATGGG 6221
LST.CD.88.SIVlhoest524      AG-TAC--C---C---G--TAC--G-TAAGTTC---TAGT-GCTTCT-TGGC--T-----A-CCAATCATAC------C-AGT-----CC-CTA.........ACTCCAAAAAAGATG-GAG--C-T-AT...-GAGC-AAATTTGTTTAC-AG---GCAGGA--ATGGG 6227
LST.KE.x.lho7               AG-TAC--C------TGGATAT--GCTTAGTAGT---AGT-G-TTCT-TGGC--T-----A-C-AATCATACCAGG--T-AGC-T---CC-CTA.........ACTCCAAATAAGATG-AAG--C-A-AT...-GTGCAAAGTTTGTGTAT-A----GCAGGA--ATGGG 7293
MAC.US.x.251_1A11           GGTCTCTTC---C---AGGATG--GGA-A--CAGACT--T---TGGT-TGGC--T-----A-C-AG-------A-TAGAACTTAT---TAC..........................TGGC--GGT-G-...--T---AGA-CT-TA--T-GTT--A-TAA-T-TTATA 7515
MND-1.GA.x.MNDGB1           AG----G--G-----T--GCACT-AGTAG-CAC----AGT-GCTTTT-TGGC--T-----A-C-A-GCATA-G...--A-G--A-T-G-T-CCCATAGATGATAAATATAGGGGCC-A-AGG-A--TCAT...C-A-GGAAGTTTGT-TAT-AG--G-CAGGA--ATATG 6627
MND-2.CM.98.CM16            AG-------G--------GCC-T-GCCAG-CACCACC------ATGTGTGGA-CT-G---G-CAAA-CAT--TT-TA-T-A-T----ACAGACA......AATACAAAGAAAGGCAAA-AGG-G--TCAT...---C--AAGT-TGTGTAT-G----G-TGA---GTACA 7203
MND-2.GA.x.M14              AG-------G-----TG-CCC---GCCTG-TACTA--------ATGT-TGGC--T-----A-CAAA-CATA-TT-T--T-AGT-----TTGGTG......AACCCTCAA.........AAGG-G--TCAT...---C--AAAT-TGTTTAT-GG--GA-CAAA--ATGGG 7120
MND-2.x.x.5440              TG----G--G--C---GCACCGC-GCCAG-CAC-A-C--T--AATGT-TGGC--C-----A-C-AAGCAT--TT-T--T-AGC-T---CAGACA......AACCCACGAAAAGGAAAA-ATG-G--TCAT...---C--AAAT-TGTATAT-G---GG-TAAA--ATGGG 6830
MNE.US.x.MNE027             GGTCTCTTC---C---AGAATG--GGA-A--CAGACT--T---TGGT-TGGC--T-----A-C-AG-------A-TAGAACTTAT---TAC..........................TGGC--AG--AA...--T---AGG-CT-TA--T-GTT-GA-TAA-T-T-ATA 6990
MON.CM.99.L1_99CML1         TG-----AC------T----CT--CCAA--TA-----G-T---TGGT-C-A---G--CA-T-CGGGTAAT-CTCCCA-TAC-AC-G-A-TGATG..........................AAT--GCAG-AA...A-TG-GT---T-GTAG-T-G-T-GGCTAA-C--CT-C 7040
MON.NG.x.NG1                AG-----AC---C--G-----T--A-T----A-------T---TGGT-T-A-C-T----CT-CCTCT--CAC-A-TR--ACCAC-G---TGATG..........................AATGGGAAA-GA...A-TG-AT-C-T-GTA----GGC-TGM-AA--C-CTC- 5656
MUS-1.CM.01.CM1239          AG-TACG--T--C--G-----C--ATTA--CT-G--TG-T--CTGGT-A-A-C-T-------CATATTT-----G-A-TAATAC-GCAGT-ATG....................ATGAATGGGGA--AA...A-TG-AT---T-GGA----AGT-TGG-GAA---TTTA 6968
MUS-1.CM.01.SIVmus_01CM1085 TG-TAGG--T--C--G-----C--ATTA--CT-G--TG-T---TGG--A-A-C-T-----A-C-TATTT---G-G-A-TAATAC-GCAGT-ATG....................ATGAATGGGGAT-AA...A-TG-AT-A-T-GGA----A-T-TGG-GAA--TTTTA 6996
MUS-2.CM.01.CM1246          -G-TAGG--C--C--T----AC--ATTA--GC--A-TG-T---TGG--A-AA--G-----T-CCTACCT-CC--GCA-T-ACAC-GC-GT-ATG....................ATGAATGG-GAT-AA...A-TG-AT---TTG------A-T-TGGAGAA---TTTA 7103
MUS-2.CM.01.CM2500          -G-TACG--T--C--G-----C--ATT----C----TG-T---TGG--A-A-C-------T-CCTACTT-CCT-GCA-T-ATAC-GCAGT-ATG....................ATGAATGGCAAG-AA...A-TG-GT---T-G-T---CA-T-TGGAGAA--TTTTA 7032
MUS-3.GA.09.09GabOI81       AG----G--T--------------CTT---C-G---G----ATGGT-C-A-C-------G-CCT-GCA-CC--GGA-T-ACACTGCAGTGGTA....................ATGAATGGGAAA-A-...A-TG-AAG-GT-GTG----G-T-GGGAGAA--TTATA 6594
MUS-3.GA.11.11GabPts02      AG----G--T----------G---CTA--TT-G---G----ATGGT-C-AA---------CATACCA-----GCA-TAACAC-GCAGT-ATG....................ATGAATGGGAAA-A-...A-TG-GAG--T-GTAG-C-G-T-GGGAGAA--TCATA 6595
OLC.CI.97.97CI12            AG----TACT--C--C-CGCC-C-GTCAG--AC---GAGT-G-TTCT-TGGC--T-----T-C-AG-CATA-G-C-AGCAG-C-G---CCTATA...ACAGGGGATTTGAATCATAAACCA-GAGAGGGA...ACAGGA-TAT-TGTT--GGC-A--GCTGAT--GTATA 6689
RCM.CM.00.SIVagi_00CM312    AG-TACTTC-------G-CATG--A-ACA--AC---G--T---AGTT-TGGA--------ATCAA--AGT---A-TAGAACTTGG--CTAT..........................CAA-GAAAACA-...AGT--CAGA-C--T----GGGT-GA-TAGTTT-TATA 6593
RCM.CM.02.SIVrcm_02CM8081   AG-CACTTC---C---G--ATG--A-ATA--AC---------CTC-T-TGGAC-------ATCAA-TAGT---A-CAGAACTTGG---TAT..........................CAG-GAAAGCAA...TCG---CGG-C-GTA---GGGT--A-TAGAG-GTATA 6671
RCM.GA.x.SIVRCMGAB1         TG-CACTTC---C------ATG----ATA-TACT--GG-CT-AGGGT-TGGA--------ATCAA-TAAT-T-A-T---ACTTGG--ATAT..........................CAG-GAAGGCAA...AGT---AGG-C-GTT---GGTC-CA-TAGTTTTTATA 6809
RCM.NG.x.NG411              AG--ACGTC---C-----CATG--A-ATA--AC------C--ATC-T-TGGA--------ATC-A-TAGT--GA-TAGAAC-TGG--CTAT..........................CAA-GAAAACAA...AGT---AGG-C-GTT---GGGT--A-TAGAC-TTTTA 6853
SAB.SN.x.SAB1               AGT-TC-GCT--C--TAGGTT---C-ATA-TACT---AGC--AGG-A-AGGAA-T-----T---TAT-T--C-A-TAGAAC--A---CT-G......................CAGAAAAATGGCAATT-C...A-TG--T--GTT-TT--T-GGC-CA-TAGAT-TTTTA 7384
SMM.CI.79.SIVsmCI2          GGTCTCTAC------TAGAATG--GGAAA--CA-ACC--T--ATGGT-TGGA--T-----A-CAAG---T---A-TAGAAC-TAT--ATAT..........................TGGC--GGT-GA...------AGA-CT-TA--T-GTC-TA-TAA-T--TATA 6666
SMM.LR.89.SIVsmLIB1         GGTCTCGTC---C---AGGATG--GGA-A--CAGACC--T---TGGT-TGGC--T-------C-AGG------A-TAGAAC-TAT--CTAT..........................TGGC--GG--G-...--T---AGA-CT-TA--T-GCT--A-TAA-T--TATA 6575
SMM.SL.92.SIVsmSL92A        AGTTTCTTC---C--TAGAATG--GGAAA--CA-ACC--T---TGGT-TGGT--T-----A-C-AG-------A-TAG-ACCTAC---TAC..........................TGGC-CGGT-GA...AGT---AGA-CT-TA--T-GCT--A-TAAA-CTTATG 6604
SMM.SL.92.SIVsmSL92B        AGT-TC-TC-------AGGATG--GGAAA--CAGACT--T--ATGGT-TGGC--T-----T-C-AG------GA-TAGAAC-TAT--ATAT..........................TGGC--GG--GA...AGT---AGA-C--TA--T-GCT--A-TAA-T-TTATA 6575
SMM.SL.92.SL92B             AGT-TC-TC-------AGGATG--GGAAA--CAGACT--T--ATGGT-TGGC--T-----T-C-AG------GA-TAGAAC-TAT--ATAT..........................TGGC--GG-GGA...AGT---AGA-C--TA--T-GCT--A-TAA-T-TTATA 6955
SMM.US.04.G078              GGTTTCTTC---C---AGAATG--GGA-A-GCA-ACT------TGGT-TGGC--T-----A-C-AG-------A-TAG-ACTTAT--ATAT..........................TGGC--GGT-GA...--T---AGA-CT-TA----GCT--A-TAA-TTTTACA 6800
SMM.US.05.D215              AGTTTCTTCT------AGAATG--GGAAA--CA-ACC--T--ATGGT-TGGT--T-----A-C-AG-------A-TAG-ACTTAC--CTAT..........................TGGC--GGT-A-...AG----AGG-CT-TA----GCT-GA-TAA-T-TTATA 6776
SMM.US.06.FTq               GGTCTCTTC---C---AGGATG--GGAAA-TCA-ACC--T--ATGGT-TGGA--T-------C-AG-------A-TAGAAC-TAC--ATAT..........................TGGC--GGG-GA...------AGG-CT-TA--T-GTT--A-TAA-T-TTATA 6801
SMM.US.86.CFU212            AGTCTCTTC-------AGGATG--GGAAA--CA-ACC--T--ATGGT-TGGT--T-------C-AGG------A-TAGAACTTAC---TAT..........................TGGC--GGTCAT...AGG---AG--CT-TA--T-GTT--A-TAGCT-TTACA 6771
SMM.US.x.H9                 AGTTTCTTC---C---AGAATG--GGAAA--CA-ACC--T---TGGT-TGGT--C-----T-C-AG-------A-TAGAAC-TAC---TAT..........................TGGC--RG--GA...AG----AGA-CC-TA--T-GCT--A-TAA-T-TTRTA 6996
SMM.US.x.PGM53              GGTTTCTTC---C---AGAATG--GGA-A-GCA-ACC--T---TGGT-TGGT--C-----T-C-AGG------A-TAGAAC-TAC---TAT..........................TGGC--AGT-GA...AGT---AGA-CC-TA--T-GCT--A-TAA-T-TTATA 7453
SMM.US.x.pE660.CG7G         GGTTTCTTC---C---AGAATG--GGA-A--CA-ACC--T---TGGT-TGGC--C-----T-C-AGG------A-TAGAAC-TAC---TAT..........................TGGC--GG--AA...AGT---AGA-CC-TA--T-GCT--A-TAA-T-TTATA 7509
STM.US.89.STM_37_16         AGTCTCCTCG-----TAGAATG--GGA-A--CA-ACT------TGGT-TGGC--T-----A-C-AG-------A-TAGAAC-TAT--CTAT..........................TGGC--GG--G-...--T---AGG-CT-TT--T-GCT-GA-CAA-T-TTACA 7158
SUN.GA.98.L14               TG-TACC-CT-----TA--TAT--G-CAAGTAC---TAGT-GCTTTT-TGGC--T-----A-C-AG-CATA-G-----T-AGT----ACC-ATC............AATAACAAGGTA-GAC-AGGGG-T...--GGGA-AAT-TGTATGG-A----GCAGCT--GTGGG 7352
SYK.KE.x.KE51               AG----TTC---C------CCTT-C-ACA-TT-G-CTAGT--ATGGT-T-A----------CCTATAA--C-A----CAA---T-GGTTTATT....................AGACAA--GGAT-AA...A-TG-AT-AGTG-TT---C----G-CAGAA--ATT-G 6776
SYK.KE.x.SYK173_COMGNM      TG-----TC---C-----G-A-T-C-ATA--T-G-CTAGC--ATGGT-C-AAC-G-----G-CCTACAA--C-A----TAA---T-GATTTATC....................AAACAA--GGAT-AA...A-TG-GT-AGT--TT--TT----T-CAGAAGCACTTA 7080
TAL.CM.00.266               AG-------C------G-CCCT--C-CCA--A---C-G-T--GTGGT-AA-TC-C-----G-CAAAG...---A-TC-CAC-CAGG-C-T-.......................CA-CAGGC-CCA-A-...A-TG-AAG-GCT-TA--CC-CC--GGAAA-G-ACACA 6883
TAL.CM.01.8023              AG-------C--C---G--CCC--C-CCA-CA---C-G----ATGGT--G-TC-------T-CAAAG...----A-TC--AC-CA---A-T-.......................CA-CAGGCACC--AA...A-TG-AT-AGCC-TA--CC-CT--GG-AA-G-ACACA 6427
TAN.UG.x.SIVagmTAN1         AGT---TAGT-----TGGATTG--G-ATA-TAC---GAGCT-AGC-T-TGGCA------------AG------A-CAGAAC--A---AT-G......................CAAAAGCATGGAG-G-GC...A-T---T--GTG-TA----A-C--A-TAAAC-TTATA 7173
VER.DE.x.AGM3               AGTG--G--T-----TA-CTTG--A-ATA--AC---GA-T---GG----T----G-----G--CTACT---GA-TCGAACCCAG--AT-G......................CAGAAACAT-GAG-A-GC...A-TG-CT-AGTGTTAG-GT--T-TA-TAAAC-TTACA 6711
VER.KE.x.9063               AGT------C-----TGGACTT--G-ATA--AC---T--T--AGG---A-----G-----A--CTATT-T--GA-TCGAAC-CA---AT-G......................CAGAAGCAT-GAG-G-AT...A-TG--T--GT--TAG-TC-TT-TA-CAA-T-TTACA 7213
VER.KE.x.AGM155             AGT---G--T--C---AC-TT---G-ATA--AC----A-C---GGT---T-A--G-----A--CTATT-----A-TCGAACCCAG--CT-G......................CAAAAACATGGAG-G-GC...A-TG-CT-AGTGTTA--CT-GC-CA-TAA-C-TTATA 7193
VER.KE.x.TYO1_patent        AGT---G--T------GGCTT---G-ATA--AC---GA----AGGGT--T--C-G-----A--CTATCAT--GA-TCGAACCCAG--AT-G......................CAGAAACAT-GGG-A-AT...A-T--CA-AGT-TTG--CT-GT-CA-CAA-C--TATA 7192
WRC.CI.97.97CI14            AG-----ATG------G-CATG--G-CAG--ACCA-TAGT-GCTTCT-TGGA--T-----A-C-AG-CATA-G--G-GA-A-T-----GA-GTA...ATAGGAAATATAAATCATAAACCA--GGAAGGA...TCAGGA--ATTTGTGTAT-AG--G-CAGAC--GTATA 7196
WRC.CI.98.98CI04            AG-----TT---C---G-CATG--G-CAG-TACTA-TAGT-G-TTCT-TGGA--T-----T-CCAG-TATA-GA---G--A-T--G-AGAGATA...ATAGGAAATATAAATCATAGGCCTC--GAAGGT...ACAGGA--ATTTGTATAT-A---T-CTAGT-GGTATA 7244
WRC.GM.05.Pbt_05GM_X02      TG----TACT-----TG--TT---G-CAG--ACTA-TAGT-GCTTTT-TGG---T-----A-C-AAGTAT--TA---G--A-T-G--AGA-GTA...ATAGGAAATATAAACCATAAACCA--AGATGGA...ACAGGA--ATTTGTTTAT-A-----CAGCT--ATATA 6140
```

```
                                         ⌐V3 loop start                                                                                              V3 loop end¬
H1B.FR.83.HXB2              CTGTAGAAATTAATTGTACAAGACCCAACAACAAT...ACAAGAAAAAGAATCCGTATCCAGAGAGGACCAGGGAGAGCATTTGTT...........ACAATAGGAAAA...ATAGGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATAACACT 7247
Env                         S  V  E  I  N  C  T  R  P  N  N  N  .  T  R  K  R  I  R  I  Q  R  G  P  G  R  A  F  V  .  .  .  .  T  I  G  K  .  I  G  N  M  R  Q  A  H  C  N  I  S  R  A  K  W  N  N  T
H1O.CM.91.MVP5180           --A--A-C--G-CC--C-T----GAAGGA-TTGCA...GAGGT-C--GAT--ATA--CA........-T---AT----TGGCGCAG-ATGACACTTAAA-G--GTAAC--TACATC-CC--GATCA--GGT---TT------CATA--A--AGA-TGT----G-A--TG-C 7290
H1N.CM.95.YBF30             -A--GCC---A-----------G--AGGA--T---...---G--GG-CAGG-G-AG--A..........-----T-CT-TGA-----TA-............-AC----A----ATAG-----G-C--T---------T-C-----TG-CTC--A--A-CT----G-ACCA-TG 6825
H1P.CM.06.U14788            G---GTC-C-C-CC----T------AGGT--T--C...------GG-CA-G-G-AAC-A.........---GTC...-TGA---GGTA-............-AT--GAA-C-CTACG----CG----C--GGC---T--------TG--TC---G-A---T-----A-GA--C 6754
CPZ.CD.06.BF1167            TCT-TA-C--A-C----GT-------GG------C...------GG-CAGG-A-AG--A........-G-----C-TGA-C--CTA-............-AT---A----TGTT--T---G-C-CT---A-G---TT---CT-TG-A-AC-A-A--CT----G-A--TG-- 7375
CPZ.TZ.00.TAN1              -AA--A-TT-A-CG---GA---A-AGGA------...------GGTCA-G-A-AG--A........-T-----T-TGA-C---TA-............-AT----A---TGTAG----GG-C-CC--GA----TT-----TCAG-C-A-GC-A---C----T-C-GG-AC 6885
GOR.CM.07.SIVgorCP2135con   -A---------CC---GA---GA-A-GT------...------GG-CAG----AAG-A........-G--CAT--C-ATT-A-................-ACTC--A---CATAG-------C-CA---A-G--TTTC--C--GTA--A--AGA----C---C-G--TG-A 6798
                                         ⌐V3 loop start                                                                                              V3 loop end¬
MAC.US.x.239                A-C--AC---G--A----G------AGGA--T--G...---GTTTT-CC-G--ACC--T......ATGT-T--ATTG-TT--C...............CACTC-CA-CC-ATC-AT-AT-GGCCA-AG--G---TGG---TGGT--...G---G---------GG-TG-A 7662
Env                         N  L  T  M  K  C  R  R  P  G  N  K  .  T  V  L  P  V  T  I  .  .  M  S  G  L  V  F  .  .  .  .  .  .  H  S  Q  P  I  N  D  R  P  K  Q  A  W  C  W  F  .  G  G  K  W  K  D  A
H2A.SN.86.ST_JSP4_27        A-C-CACCG-AC------AG--G--AGGA-----G...---GTTGT-CC---AACAC--......ATGT-----TT--TG--.............CACTCCCAGCC-ATC-ATA---GACCC--G------TGG--CTGGT-C...-A--GCG-G-----GG-AG-C 7171
H2B.CI.x.EHO                A-C-GAC---GC-C----A---G--AGGA-----A...-TGGTTGTGCC---AA-A-C-......-TGT----T-TTCTC--C.............CATTC-CAGCCTATC-ATAA--GACCT-A------TTGG--CTGGT--...-A--G---C----CAG-AG-C 7703
H2G.CI.92.Abt96             A-T--AC---A-RA----GG-----AGGG--T--A...---GT-TT-CC---TAC---T......ATGT-----CT--TC--C.............CACTCTCA-CC-ATC-AT-A--GACCC-----G--TTGG---TGGT--...G---GC---------GG-AG-C 7075
H2U.CI.07.07IC_TNP3         A-T-GAC-G-A-GA----G---G--AGGA--T--A...---GT-TTGCC-G-AACC--T......ATGT-----TT--TT--C.............CACTCTCAGCC-ATC-AT-A--GACCC-AG-----CTGG---TGGT--...G-T-GG-GT-----AG-GG-A 7180
H2U.FR.96.12034             A-T--AC---GT-C----G---G--AGGG--T--G...---GTCTT-CC-G-AACC--T......ATGT----ACT--TG--C.............CATTCTCA-GCTATC-ATAAC---CCA-A------TTGG---TGGT--...G-T-G---T-----AGGGG-A 7204
ASC.UG.10.RT03              TA---A-TT-A-CA---GT------AGGG--T--A...T---T--GG-ACT-G-AA--A..........---G----C-TGA-C---TAC...........T--CA-CTC-T-GTAGGG-------CT---AG---TT----C--A--A-A-G-CA--C-G---T-C--TG-- 6684
ASC.UG.10.RT08              TA---A-TT-A-CA---GT------AGGG--T--A...T---T--GG-ACT-G-AA--A..........---G----C-TGA-C---TAC...........T--CA-CTC-T-GTGGG--------CT---AG---TT----C--A--A-A-G-CA--C-G---T-CTCTG-- 6714
ASC.UG.10.RT11              TA---A-TT-G-CA---GT---G--TGGG--T--A...----T--GG-ATT-G-AG--A..........---T----C-TGA-C---TAC...........T--CA-CTC-T-GTAGGG---G---CT---AG---TTT---C-GA--A-A-G-CA--C-G---T-CC-TG-- 6672
COL.CM.x.CGU1               A-A---CTT-AGGG----T----AGAGGA--T-G-...T-TCAC-G--ATC-AAAC-CT......-CCAAT--AGC-AA---CTAC...........TATGAGCTT-TTCCTTATTC---AGGC-TTT-T-G-AGG---C-GT--...GT-C---TGACTGGGC-G-A- 6654
COL.UG.10.BWC01             ACA---CTT-AGGA----TC---AGAGGA--T-G-...T-TCAT-G--ATC--AA--CG......-CCAAT--AGC-AA---CTA-...........TATGA--TT-GGCCAGAG-C-G-C-GAT-GA---G-AGG---C-AT--...GT-C-T-TGAATC-G--G-AC 6480
COL.UG.10.BWC07             A-T--T-CT-AGGG--C-AG---AGAGGA--TGT-...--T-AT-GG-ATT-AAAC-CA......-CTAAT--AGC--T---CTA-...........TATGAGTT--G-GGACAG-ATGCAGATT-GA----GAGG---C-AT--...GT-AA--T-G-A-TG--T-AA 6681
DEB.CM.04.SIVdeb04CMPF3061  ACC-CCGTT-AG-C---GTCC-T--AGGT--T--G...---TAT-G-GC--CT-AC--G......-CCA---ATT-T----TAC...........----CCTTTGT-GAA-GGCT--GG--C-A-AG---------C-GA--G...C---GT--T----G-GG-G-A 7117
DEB.CM.99.CM40              ACT-GAC-T-A-CA---GT---G--AGGA--T--G...-CTACCG-GCC--A-AC--G......-CTA---TCTGT-C---TA-...........--T-CCTTT-T-CAA-GGCTG-GG--A-AGAG----------GAC-A...-AC-G--GC---GCA---G-A 6999
DEB.CM.99.CM5               A-C--TCTT-GG-----GTT--G--AGGG-----G...-GTAT-G-GCTG-T-A---G......-CCA----CT-T-C---TA-...........----CCTTC-T-CCA-GGCTG-GG--A-AGAG----------GG--G......---GT--T----CAGGAG-A 6938
DEN.CD.x.CD1_CMO580407      TAA-CA--T-G--G-----C--G--AGGG--TGTA...--CCT---GG--T-AAAC--G......-C-G----A-T--TCA-GAGA...........G-C-CTATT-T-CCA-G-A-GGC---T--GAG----GCA--C--G--AGTAG---GC--C-----C--GA--A 7304
DRL.DE.11.D3                ACC-CA----C-TG----G----AAAGGA----GG...-GTGT--T-TCT-CA-CCTCA......-CTA-----TTGTTG---TA-...........CATGGGTT-G--CCAGGGAA----C-A-ACAG--G-ATG-----AT--...GAG-G-C-----GGA--GG-- 6920
DRL.x.x.FAO                 ACC-CA------TC----G----AAAGGA--T-GG...-GC-T--T-TCT-CA-CCTCA......-CCT-----TTGTTG---TA-...........CATGGGTT-G--CCAGG-AA----T-A-A-AG--G-ATG-----AT--...GA--G-C-----GGA--GG-- 7010
GRV.ET.x.GRI_677_gri_1      ACT-GAC-G-G-GA--CCG------TGGT--T--A...---GTGTTGCC-G-AACA---......ATGG-----TT--T----...........CACTCTCAG---TAT-ATACC-GGT-A-A------GTGG--CC--T-C...CA--G-G-T-----AGGGG-A 7227
GSN.CM.99.CN166             A-C-TA-TC---CA----TC-----TGGA--T--A...--C-T--G--ATT-G-AA--A..........---G-G--A-TGA-----TA-...........T-CCA--TG-T-GTAGG--------CT--GA----CT----CCGAC--GA-CCT-AT-GG-------GGG-A 7181
GSN.CM.99.CN71              AGC-TACTC---CA----TC-----TGGA--T--A...--CCT-----ATT-A-AG--A..........---G----A-TGA----CTAC...........T-CCA-ATT-T-GTAGGG---G---CT--GA-G--CT----C--AC-C-ACCATA----G---G---TAG-A 7172
LST.CD.88.SIVlhoest447      GAC-CATT----GA----TT--GAAAGGA--T-GG...T--GA-GT-TCT-C-AT--GT......TC-A----ACTTCTG---TAC...........TATGGCTTGG--CATGGTA-C-GGC-----TTG--T--G-----AT--...GA--G-C-G---GGA-GA-TG 6370
LST.CD.88.SIVlhoest485      GAC-CATT--C-GA----TT--GAAAGGA----GG...T--GA-GTCTCT-C-ATC-GT......TC-A----ACTTCTG---TAC...........TATGGCTTGG--CATGGTA-C-GGC-----TTG--T--G-----AT--...GA--G-C-----GGA-GG-TG 6367
LST.CD.88.SIVlhoest524      GAC--ATC----GA----T----AAAGGA--T-GG...T--GA-GTCTCC-CAATC-GT......TC-A----ACTTCTG---TAC...........TATGGGTT-G--CATGGGTCT-GAC-A---TTG-----A--C--AT--...GAG-GTC-----GGA-GA-TG 6373
LST.KE.x.lho7               G-T--ATT----GA--C-T----AAAGGA--T-GA...T--GA-GT-TCC-CAAT--G-......TCTA----ATATTT----TA-...........TATGGCCT-G--CATGGCA-C-GAT-A---TT------G-----AT--...GA--GCC-----GGA-GA-TG 7439
MAC.US.x.251_1A11           A-C--AC---A--A----G------AGGA--T--G...---GTTTT-CC-G--ACC--T......ATGT-TCAATTG-TT--C...............CACTC-CA-CC-ATC-AT-AT-GGCCA-AG--G---TGG---TGGT--...----G---------GG-TG-A 7658
MND-1.GA.x.MNDGB1           GCT--A-G--AG-A---CAC---AAAGGA--T-GG...T--GT-GTG--T-CT-CATCA......-CTA----ATT-TT----TA-...........CATGGGTT-G--CCTGG-AAG---T-A-AGA---GCATG--C-C-T-C...-A--G-CGT---GGGTTAG-A 6773
MND-2.CM.98.CM16            ACT--C--G-AGTA----G----AAAGGA----GA...T-TGTG-TCTC--CA-C--GT......-C-A-G--ATTGATC---TA-...........-GTGGGTT-G--CCAGGGAA----C-T-AGA---GTATG---C-GT-G...-AG-G-C-----GGA-TAG-C 7349
MND-2.GA.x.M14              GGT-GA-GG-GGTA---GAT--GAAGGGA----GG...T-T-T-GTTTC-G---CA-GT......-C-T----ATT-ATT--CTA-...........CATGGGTT-G--CCAGGCA-G---C---A-A---G-ATG--C-GGT-A...-T--G-C-----GGA-GAG-- 7266
MND-2.x.x.5440              GAT--C-GG-A-GG----G---GAAGGGA--T-GA...T-T-TC-TTTCC-CA-CA-GT......TCTA----ATTGCTG--CTA-...........CATGGGTTGG-GCCAGG-AAG---C-T---A-G-GCA-G---C-GT-A...GA--GG------GGAC-GG-- 6976
MNE.US.x.MNE027             A-C--AC---G--A----G------AGGA--T--G...---GTTTT-CC-G--ACC--T......ATGT-T--ATTG-TT--C...............CACTC-CA-CC-ATC-AT-A--GGCCA-A---G---TGG---GGT--...GA--G---T-----GG-GG-A 7133
MON.CM.99.L1_99CML1         A---CA-C----C---C-TT--G--AGGA-----A...--T-TT-GG-ATT-A-AG--A..........---G----C-TGA----CTA-...........T-GCA-CTT-T-GTGGG---T----CTC-TA----CT-C--C--AG--A--AGA-CC-----G-G-CTG-C 7189
MON.NG.x.NG1                -A--GA-C--A-C---C-TT-----AGGA--T--G...--CTT--G--ATT-A-AG--T........--TG----A-TGA----CTA-...........T-GCA-ATC-T-GTGGG------C-CY---A-G---T-C---C-GG-A-A--AGA--C-G---G-CGCTG-C 5805
MUS-1.CM.01.CM1239          GA--TA-CT-A-CC--CTT---G--AGG---T--G...--T-T--GG-ACC---AG--A........--GG----C-TGA-T--CTA-...........T--CAGCTT-T-GTAGG---GG-C-CAC--AG----T-C---C--G--A-AC-A-A--C-G---G--GTAG-A 7117
MUS-1.CM.01.SIVmus_01CM1085 GA---A-TC---C-----T------AGG------G...----T--GG-ACC-T-AA--A..........---G----C-TGA-T--CTA-...........T-GCAGCTT-T-GTAGAT--CG---CAC-CAG---CT----C--A--A-A--AGA--C-G---G--CCAG-A 7145
MUS-2.CM.01.CM1246          GG---A-TC-A-CC----TT--G--AGGA--T--G...----TT-G--ATC-A-AG--A..........---G----C-TGA-C--CTA-...........T-TCAGCTG-T-GTGGAT---G-C-CCC-CAGG---T----C--G--C-A--A-A-TC-----G--CCTG-- 7252
MUS-2.CM.01.CM2500          GG--CA-TT-A-CA--CCTG--G--AGGA--T--G...--T-TT-G--ACC-T-AA--A..........---G----C-TGT-T--CTA-...........T-TCAGCT--T-GTAGAT---G---CCC-CAG---CT----C--A--C-AC-AGA-CC-G---G-CCC-G-A 7181
MUS-3.GA.09.09GabOI81       AG--TA-TC---C-----TT-----AGGG--T--A...--C-TT-GG-ATT-A-AA--A..........---G----A-TGA----CTA-...........T-TCA-TT--T-GTAGGT--GG---CA---AG---TT-------A---GA--A-A--C-----G--GCTG-C 6743
MUS-3.GA.11.11GabPts02      GA---C--T-G-C-----TT--G--AGGA-----G...----T--G--ATT-A-AG--A..........---G------TGT-C---TA-...........T--CAGCTT-T-GTAGAT---G-C-CT---AG---TT----C--A--AGA--AG-AT------G----TG-A 6744
OLC.CI.97.97CI12            AGT-------AG-G----TT--GAAGGG---T-GA...-GT-T-GTTTC-G-TAA-TCA......-CTT----TCTG-T--A-TA-...........G--GG-TT-G--CCTTAT--G-----AGATA---GCTTG----GGT--...-AG-G-C-----GGATGG-TG 6835
RCM.CM.00.SIVagi_00CM312    ACT-GAC-G-A-G-----G------ATCA--T-GG...-----T---GG----TTCCT-A......-CCA----A...-T----..............-T-TCTCT----GTAGA-AA--GGCCA-A-GG---TTGG--C-GAT--...TAT-G---T-----AGGAG-- 6733
RCM.CM.02.SIVrcm_02CM8081   A-C--AC-G-GG-A----G---G--ATCA--T-GGACAGTG-AGGG--TCTCA...T-G......-CTA----A...-TC---..............-TCTCCTT--G-GTAGAGAA-CGACCA-A-GG---TTGG---GAT--...TAT-GC--T----CAGGTG-- 6811
RCM.GA.x.SIVRCMGAB1         A-T-GTC-G-A-CA--C-GG-----TTCA--T-GA...----GTG---G-G--ATCGC-A......-C-A----A...-TC---..............-TCTC-CT--G-GTAGAGAAG-GACCA-A-GG---TTGG---GAT--...GA--GG--T----CGG-TG-A 6949
RCM.NG.x.NG411              A-T--AC-G-A-C-----GG-----A-GT----GAACAGTG-A-GG--TCTCA...T-A......-C-A----A...-T---C..............-TCTCTCT--GGGTGCA-AA--GGCCA-A-GG---TTGG---GAT--...GAT-G-G-T-----AGGTG-G 6993
SAB.SN.x.SAB1               A-T--ACT--C-GA----G------AGGA--T--G...---GT-TTGCCTG-AACC---......ATGG-T--CTTG-TC--C...............CAT-GCCAG---TAT-ACACC-GGT---A------TTGG--CTGGT--...G---GT--T----GAGGAG-C 7527
SMM.CI.79.SIVsmCI2          ACT--AC---G-GA----GG-----AGGA--T--G...---GTGTTGCCTG-GACC---......ATGT-----TT--TT--C...............CACTCCCA-CCTATC-AT-A--GACCA-A------TTGG---TGGT--...G-G-G---G-----AG-GG-A 6809
SMM.LR.89.SIVsmLIB1         A-C--AC---A-GA----G------AGGG--T--G...---GTCTT-CC-G--ACC---......ATGT----ATTG-T----.............CATTC-CAGCC-ATC-AT-AG-GGCCA-A-------TGG---TGGT--...G-T-GGTC-----GA--AG-A 6718
SMM.SL.92.SIVsmSL92A        GCT----GT-AC-C----GG-----AGGG--T--G...--GGTCTT-CCTG-TAC---A......ATGT----ATT--TT---.............CACTCCCA-CC-ATA-AT-AG-GACCA-A------TTGG---TGGT--...-A--GGG-G-----G--GG-- 6747
SMM.SL.92.SIVsmSL92B        A-T-GAC---A--G----G------AGGA--T--G...---GTCCT-CC-G--ACC--T......ATGT----TTTG-TC---.............CATTCGCA-CC-ATA-AT-AG-GGCCA---------TGG--CTGGT--...G---G--------GAG-AG-- 6718
SMM.SL.92.SL92B             A-T-GAC---A--G--C-G------AGGA--T--G...---GTCCT-CC-G--ACC--T......ATGT----TTTG-TC---.............CATTCGCA-CC-ATA-AT-AG-GGCCA---------TGG--CTGGT--...G---G--------GAG-AG-- 7098
SMM.US.04.G078              A-T--AC---A-GA--C-G------AGGA--T--G...---GTCTT-CC-G--ACC--T......ATGT-----TTG-TC--C.............CATTCGCA-CCCATA-AT-AG-GACCA-A---G--CTGG--CTGGT--...G---G--GT-----G--GG-C 6943
SMM.US.05.D215              A-T--AC---A-GA----G---G--AGGA--T--G...---GTCCT-CC----ACC---......ATGT-----CT--TC---.............CATGC-CAGCC-ATA-AT-AG-GACCT--G------TGG--CTGGT--...G---GC--G-----AC-GG-A 6919
SMM.US.06.FTq               A-C--AC---A--A----G------AGGA-----G...---GTT-TGCCTG-GACA---......ATGT----CTTG-TC--C.............CATTCCCA-CC-ATA-AT-AC-GGCCA--------CTGG---TGGT--...G-T-G---------ATCAG-- 6944
SMM.US.86.CFU212            A-T--AC---A-GA----G---G--AGGG--T--A...---GT-CT-CC-G--ACC--A......ATGT-----TT--TT--C.............CATTC-CA-CCCATA-AT-AG-GGCCA-----G---TGG--CTGGT--...G---G---------GTCAG-A 6914
SMM.US.x.H9                 A-C--AC---G-GA----G------AGGA--T--G...---GTCTT-CC-G--ACC--T......ATGT-----TTG-TC--C.............CATTCGCA-CCCATA-AT-AG-GACCA-A---G--CTGG--CTGGT--...G---G-G-------AG-GG-C 7139
SMM.US.x.PGM53              A-C--AC---A-GA----G---G--AGGA--T--G...---GTCTT-CC-G--ACC--T......ATGT-----TTGATC--C.............CATTCGCA-CCCATA-AT-AG-GACCA-A------CTGG--CTGGT--...G---G--GT----GAG-GG-C 7596
SMM.US.x.pE660.CG7G         A-C--AC---G-GA----G------AGGA--T--G...---GTTTT-CC-G--ACC--T......ATGT-----TTG-TC--C.............CATTCGCA-CCCATA-AT-AG-GACCA-A---G--CTGG--CTGGT--...G---G--GC-----AG-GG-C 7652
STM.US.89.STM_37_16         A-T--AC---G-GC----G------AGGG-----A...---GTCTT-CC-G--ACC--T......ATGT----CTTG-TC--C.............CATTCGCA-CCCATT-AT-AG-GACCA-A------TTGG--CTGGT--...G---G-G-----GAGGAG-A 7301
SUN.GA.98.L14               GGT---T----C-G--C-T---GAAAGGG--T-GG...T--CA-GT---T-CTATA-GT......A-TA----ACTGCTG---TA-...........TATGGGCT-G-GCCAGGTTCT--GT-A----TG-----G-----AT--...GT--G-C-----GGA-GAG-C 7498
SYK.KE.x.KE51               GAT--C--C-GGTG---GA------AGGA--TG-G...T-T-T-----AC----AAT-G......-C-G----ATTTTTC--GCCAGTGATT...CAAGGG-G-TTG---ACAGGGAA-GC-GC--A-AG----TT---C--AG-C...-A--G-G-----GGA--GTT- 6931
SYK.KE.x.SYK173_COMGNM      GAT-GC-G----TA---GA------AGGA--TG-G...T-T-TC----AT--T-AAT-G......-CTG----CTATTTCC-CCCAGTAATA...CAAGGG-A-TTG---ACCGGGA-GG-CGCA-A--GT--CTTC---CGAG-A...-C--G---T----CAG-ATT- 7235
TAL.CM.00.266               ACC-CAC---C-CA--C-T------AGGA--T-GG...----T---GGACT-A-AG--A......-C-G----ACT-ATG---CAC...........-GCCAGAT--T-GCAGGGAAG--CC-C-ACAGG--CT-------GC-A...GAG-GC--T----C---AG-C 7029
TAL.CM.01.8023              ACA-CACC--AC-G--C-TT-----AGGA--T-GA...----TC---GATC-T-AG--A......-C-G----TCT-ATG---CAC...........-GTCAGATC-T-GCAGGTAA-G--T-A-AGAG---CT-------G--A...GA--G---T----CG-GGG-G 6573
TAN.UG.x.SIVagmTAN1         AGT-GA-G--AGTG----G------AGGA--T--G...---GTCTT-CC-G-AACC---......ATGG----TCTC-TG--C.............CACTC-CA-C--TAT-ATAC---AT-A---------TGG--CC-TT-C...CAG-GTG-T-----AGGAG-C 7316
VER.DE.x.AGM3               A-C--AC-G---C---C-A------AGGA-----A...---GTCTT-CC-G-AACA---......ATGG-----CT--TG---.............CATTCTCAG-GGTAC-ATAC--GGC---------TTGG---C--T-C...CAG-GC--C----GAGGAG-C 6854
VER.KE.x.9063               A-T-GAC-G-C-C---C-A---G--AGGA--T--G...---GTCTT-CC-G-AACC--A......ATGG----TTT--TT---.............CATTCGCAG---TAC-ATACG--GC-T-----G--CTGG---C-TT-C...CA--GC-GG----GAG-AG-C 7356
VER.KE.x.AGM155             ACC-GAC-G---CA--C-A---G--AGGG--T--G...---GTCTTGCC-G-AACG--A......ATGG---ATT--TC--C.............CACTC-CAG--GTAT-ATAC--GAC-A--G--G--CTGG--CC--T-C...CAG-GC--T-----AGGAG-- 7336
VER.KE.x.TYO1_patent        A-C--TC-G-C-CC----GG-----AGGA-----G...---GTCCT-CCGG-AACG--A......ATGG-G--ACTG-TT--C.............CACTCTCA----TAC-ACATG--GC-T-----G--TTGG---C--T-C...GA--GC--T----GAGGTG-C 7335
WRC.CI.97.97CI14            A---TA-G--AG-A---CAT--GAAAGGT--T-GA...T-TTT-GT-TC-G-AAA-TCA......-CTT----TTTGAT--A-TA-...........G--GG-TT-G--CCTCATA-G-----T---A---GCTTA--C--AT--...-CT-GGCTT---GGA--A-TG 7342
WRC.CI.98.98CI04            AG-----G--AG-A--C-AT--GAAGGGG--T-GA...T-TGT-GTGTC-G-AAACTCA......-CCT----ACTTAT--A-TA-...........G--GGCTT-G--CCACATA----C--A--GA---GCTTA--C--GT--...-A--GGTTG---GGAC---TG 7390
WRC.GM.05.Pbt_05GM_X02      A---G-CT--AG-A---CAT---AAAGGA--T-GA...-GTCT-GTCTC-G-AA--TCA......-CCT-----TTGATT-A-TA-...........G--GG-TTGG--CCTTATAAG-----T--GA-----TTA--C--AT--...-C--G-C-G---GGAC-T-TG 6286
```

Draft of 30 September 2015, not for public release
plvdna_main.tex r8876–10546


```
H1B.FR.83.HXB2             TTAAAACAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATAATC.......................................................................................TTTAAGCAATCCTCAGG 7321
Env                         L  K  Q  I  A  S  K  L  R  E  Q  F  G  N  N  K  T  I  I  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  F  K  Q  S  S  G
H1O.CM.91.MVP5180          C--C---A-C-----TA...AGGTATTT-.....AATCT-GTA--CCA--C-..............GAGAATGTTACCATAATATTC.............AGCAGAACT.......................................AGT-- 7370
H1N.CM.95.YBF30            -GG--TAGA-C-AGAGAGG--A---AGA--ATCC-G--G--A--C--C..............................................................ATAACCTTCAGGGC-CGAG-GAGGAAT-A 6902
H1P.CM.06.U14788           ---G--TG-G-CAG-GAGGCCA-CT-GA-GGCT-A-CCTC-AC--TCC-ACCAT-CATCCAAA..................................................CACAACCATACATTTGTT-----AA-CAGTA---- 6852
CPZ.CD.06.BF1167           ACTG-GGCAGC-AAAGTAGC-A---AC---ACTC--AA---C-CC-CGTTC-A--A-...........................ACCACCTATACATTTAGGACAGATGGGGATC-A---TGGAT------ 7479
CPZ.TZ.00.TAN1             ---G-TTG-GCTATGGCTGCCA---ACAC-ACCA-GA-GGCC-GA--TGA--CGG-A..................................................................CAACAAACG--CC-ATGGCAGAGG-A 6968
GOR.CM.07.SIVgorCP2135con  --G---G-T-C--TAC-GGCTC-C-A--CGA-TCAGATCCGCTTCC-G........................................................................................A-T--T----- 6857

MAC.US.x.239               A-----G--G-GAAGCAG-CCA-TGTCA----TCCCA-GT--.........-CTGGAACTAACAATACT...GATAAAATCAATTTGACGGCTCCTGGAGGA........................................................ 7752
Env                         I  K  E  V  K  Q  T  I  V  K  H  P  R  Y  .  .  .  T  G  T  N  N  T  .  D  K  I  N  L  T  A  P  G  G  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
H2A.SN.86.ST_JSP4_27       A-G--GG--G-GAAGCTA-CCC-TGC-A----TCCCA-GT--.........-A-GGAACCAACGACACA...GAAAAAATTCGTTTTATAGCGCCAGGAGAACGC.................................................... 7264
H2B.CI.x.EHO               A--C-GG--G-GAAAGAG-CCA-T-A-A-T--TCCCA--T--TCAGG-----C--ATATC..................................................TCACAGATAAGGTTAGCAG----CG-GAG-A- 7795
H2G.CI.92.Abt96            A-GC-GG--G--AAGGAA-CCG-GGTCA-G--CCCCA-CT--.........-A-GGAACAAATGACACAAAGAAMATAACCTTTACAACACCAGGAGAAGGT....................................................... 7168
H2U.CI.07.07IC_TNP3        A-CC-GG--G--AAGGAA-CCC-GGT-A----TCCCA--T---C-GG--TC-ATGATACAAAAAAGATTAATTTGACACAG...........................................................C-AGG--- 7269
H2U.FR.96.12034            A-CC-GG--G-GAAGGAG-CCA-TGCCA-C--TCCCA-GT--.........CATGGTACCAAGAATATT...AGTCAGATAAATCTGGCAGAACCAGCAGGA....................................................... 7294
ASC.UG.10.RT03             -----TGCT-CCTA---GGCCA---AGA-GG-G-A-AATCT--CACGGGATCA-CCAATCACCTGGAGATCTCAACCAAGG................................................................ 6765
ASC.UG.10.RT08             -----TGCT-CCTA-C-GGCTA--CAGA-GG-G-A-AATCT--CAG-G-ACCA-TCAATCACCTGGAGATCTCAACCAAGG................................................................ 6795
ASC.UG.10.RT11             -----TGCT-CTTA---GGCTA---AGA-GG-G-A-AATCT--CAC-G-ACCA-CCAATCACCTGGAGATCTCAACCAAGG................................................................ 6753
COL.CM.x.CGU1              AAG---A-C-A-A-ACAATTCGC--T----ATCAAAAAG---TTA-C-G-TTGGC-AGAGCGCATTTCACGCAAAAATATA..........................................ACTATCACTCCAAGG--TGG-AATAG-AC 6764
COL.UG.10.BWC01            A-G-C-A---A-...GAGTTTGC-G-GA--ATTG-C-A----CTG-C--ATTGGC-ACAGAATATTACAGAGTATGGAAAT......................................ATTACCATTAAGCCTAGA--TGGTAA-AA-AC 6590
COL.UG.10.BWC07            ACT-C-TTAGC-CAGGAA-T-AAG-A-A-TTT-AC-TATTGGTTG--G-AC--T-CAAGTTATGATAATATTACCATTAAG..................CCTAGGAATGGTAATAAAACCTCAGA-CCAG--G--A-TTA 6803
DEB.CM.04.SIVdeb04CMPF3061 G-TTCTG-C---AAACAAGT-----AG-GCAT--A-.........--TGTT-CT-A-...............................................................ATAACA-T---T-A-C---- 7185
DEB.CM.99.CM40             AC---GG-A--GCGACAA--GA--TTG---AT---------GGCA--T-GG-C--ATAAT............................................................CTGACA-TA--C-A-C--AA 7079
DEB.CM.99.CM5              G-T--GG-----AGG----G-----AG--GAT--A-.........--TGTG-C--AA...............................................................ATAACAGTA--C-ATC---- 7006
DEN.CD.x.CD1_CMO580407     AGGC-GG-A----AA-AAGC-G-T-A---GG-G---AAC-TC---GGC---G-G-AATGG...................................................................GA-GTA-C-CGGGTG-- 7381
DRL.DE.11.D3               ---G--AGCT----C-AAG------A--C-GT-AA--CCTCC-T-TGG-G--AT-CTACTATTAAAGGGAAAGAATGTAAACGTCTCAAAGAAACCAAAGAC.....ACAAAAACAGTTGGGACAGGGTGCATGTTGAAGCAGCTTAAGGTAGA---TT-CA--A-CAA 7084
DRL.x.x.FAO                ---GG-AGTC----C-AAG------A-TC-GT-AG----TCC-T-TGG-G--AT-CTACT......AATAGCACATGCATAAGTCTCAAAGGCAAAAAGATC.....CAGGACAGAAACGTGACAGGGTGCATGTTAAAGCAGCTTAAGGTGGA---TT-CA--A-CAA 7168
GRV.ET.x.GRI_677_gri_1     -GG---G-AG-CAGAGAAG--G---AGA--GTGAAAAATCT--CAG--GT--GC--AGAA.............................................AATATACATCTGAGA-G-AT--GG..... 7311
GSN.CM.99.CN166            A---G-G-AGCTATG-AAGC-A-G-AT-----T-GG-A------A-CTGGG.....................................AGAAATGATACTCAAATAAGA-GG-CCTC-GAGC-TAA 7270
GSN.CM.99.CN71             A---GG--AGCCATG-AGGCCA-G-AGA-C--T-GG-A---G-TC--C-ATGAT-C-ACC............................CCTAATAAGACTCAAATAAGG-GG-CCTCCGAAC-TAA 7270
LST.CD.88.SIVlhoest447     --TC-TA-TT---GA-A--TGC-T-A-AGG-TTAA--C--CAGCG-TG-ATTAT-C-CATGGAAACTGTATAGATAACACA.....................AAGCCTTGTGGTAGACAACTAAAGGGATTACCTATAGCAAA--T-AC-AGAAG--- 6507
LST.CD.88.SIVlhoest485     --T--TA-TT---GG-AA-TGC-TCA-A-G-TTAA--C--CAGCA-TG-ATTAT-CAAATGGACAATGTCCTGAGGGAAAA.....................CAGCCTTGTGGTAGACATCTAAAGGGATTACCTATAGCAAA--T-AC-AGACGT-- 6504
LST.CD.88.SIVlhoest524     --T--CA-TT---GA--G-TG----AGA-G-TTAA--AG-C-GC--TG-ATTAT-AAGAAGGAAATTGTTCCACACCAGGA.....................AAACCATGTGGTAGACAATTAAAGGGTCTACCCATAGCAAA--T-ACGAGAAG--- 6510
LST.KE.x.lho7              --T--TA-TC---GC-AA-TGC---AG--GTTGAA--C-G-AGCC-TG-ATTAT-CAGAAGGGACAGGGACATGTGATTCTAAGAAA.................ACAACTTGTGGACGGAAATTAAAAGGCTTACCAATAGCTAA--T-ACTAGGCAT-- 7582
MAC.US.x.251_1A11          A-----G--G-GAAGCAG-CCA-TGTCA----TCCCA-GT--.........-CTGGAACTAACAATACT...GATAAAATCAATTTGACGGCTCCTGGAGGA....................................................... 7748
MND-1.GA.x.MNDGB1          C-TTGGAGTC------AAG--C---ATA--TT-AA--ACTCC-TC---GTG-ACCAGACCTGTAAAAATTTTACTAGCACTGGAGAGGAGAACAAACAAAACACGGACAAGCAAAAGGAGTTTGCCAAATGCATAAAGACTCTTAAGATAGATAA-T-TACTA-A----.. 6941
MND-2.CM.98.CM16           A-GC-TG-TC----A-TAG--------A--AT-GACA-CTCC-TATGG-G--ATG-AACA...AAGGGCTATTGCAAGGGGTTAAAAAGAAAG....................GAAAATAGAACGGGTTGTGCCCTTAAAACAATAAAGGTC-C-G-CT-CA--A-GAA 7495
MND-2.GA.x.M14             --G--CGCAT--T-AC-TG------AGA--GT-AA--C--G--T-T-T-G-GGTT-TAATGCAAGTGGCCCTTGCAGGGGGATAAACAATAGA..............GGAGTGGTAAACAAAACAGGGTGTGCCTTGAAAACAATAGAAGTAAG---CT-CA--A-CAT 7421
MND-2.x.x.5440             ---C-TAGTC--T-ATTAG-GC-----A-GAT-AA--AT-GC-T-T-C-A-GAC-A-CAC...AATATGACATGTAAAAGCAGTAACAACAAA....................AAGAACACAACAGGGTGTCATCTTAAAACTATAAGTATAAG-G--TCCA--GTTAA 7122
MNE.US.x.MNE027            A-----G--G-GAAGCAG-CCA-TGTCA----TCCCA-GT--.........-CTGGAACTAACAATACT...GATAAAATCAATTTGACGGCTCCTCGAGGA....................................................... 7223
MON.CM.99.L1_99CML1        C--C-GGCAG-TCACGAGGC-G---A-ACTG-G-GG-A---G--A--C-ATGG--C-AAT...........................................GTGACCACAATATCT-GG-GATTTCAAC-CCA 7281
MON.NG.x.NG1               ---C-TGCAGCCCA--TRGCTA-C-RGR-GG---GG-A--GC--A--T-AT-ATG--ACA...........................................GTGGCCAACATAACT-GG-GAT-CCAGC--AA 5897
MUS-1.CM.01.CM1239         ----GG--AGCTATGTTAGC-A-G---A-G--T-GGCATGGCCGCTTGGAG-A-CAAGGAGAA...AATATTACAGAAGAA...............................AAGATAAAAATCAGG-GG-CCTCCCAGC-CAA 7227
MUS-1.CM.01.SIVmus_01CM1085 --G-G---AGCTATGCTAGCCA-G---A----T-GGCATGG--GATTGCA--A-CAGAACCACAGCTCGATAAACGAGACT...............................AAAATAAAGATAAGA-GG-GATC-GAGC--AA 7258
MUS-2.CM.01.CM1246          --G-G----GCTATG-TAGCTA-G-----C--T-GGCAC----GACTCCA--A-GCTAATAAG...ACAATAAACAAAACG...............................GAAATAAAAATAAGG-GG-CCTC-GAGC--AA 7362
MUS-2.CM.01.CM2500         ----G---AGCTATG-AAGC-A-G----C---T-GG-CC...--A-TGGAC..............AACCGAAGTGAAAGT...............................GATATAAAGATCAGG-GG-CCTCTGAGC--AA 7276
MUS-3.GA.09.09GabOI81      A---G---A-GTATG-TAGCCA-G-----C--T-GGAT-...--A-T--AC-A--ATGATTCAACACATAATTACACAAAA.............................GAAAACATAACAATCAGA-GG-TCTCGCAGC--AA 6856
MUS-3.GA.11.11GabPts02     A-T-GG--AG--A--TTAGC-G----G--T--T-GG-T-...--G-T--AT-ATGGGACTAATGCA..........................................GAAAAAGTAACAATGAGG-GG-CAAGCCAGC-CA- 6842
OLC.CI.97.97CI12           C--T-TGCAC---GA-AAGG-C--GAGA--G--AA-AAC-C-TGGGC--ACTAT-CAGGGGTTAACATAACAGGCATAGGGAGG..............AACTCTACAACCTTTGCCAAAGCAATCAAAAGA.........TTCAAGTTTGATGAG--A-TTAAGAG--- 6981
RCM.CM.00.SIVagi_00CM312   -GG---G--G--AAAGAG--GG-C-T-C--AC-AAG---T----GGG-----ATGA-ACA.............................................ACTAAAATAAACA-A-G-TC-GTT-AT-- 6822
RCM.CM.02.SIVrcm_02CM8081  -GGG--G-AG-TAAA-AAG-GG---T-A--ACGAAG---T----AGGG----ATGATACT.............................................AAGAAAATATACA-A-GATC-CAT-AT-- 6900
RCM.GA.x.SIVRCMGAB1        -GG---G-AG--AAAGAG-G-G-G-A-AC-AC-AAA--GT--CGAGGT--T-GT-A-ACA.............................................GACAAAATAAAGA-A-GAAC-GTA-AT-- 7038
RCM.NG.x.NG411             -GG---G--G--AAAGAGGC-A-T-ACA-GAC-AAA---T----AGGT---CG-CAGCCT.............................................GAGAATATAACTA-A-GATC-GT--AT-- 7082
SAB.SN.x.SAB1              -GG--GG-AG-CAAAGAA-C-A--GTGAG--T-CCACC---A--GT-T-GTGGC-CAAAT...........................................GACACAAACAAGATT---CTA---AGACA-T- 7619
SMM.CI.79.SIVsmCI2         C----GG--G-GAAGGAA-CCC-TGTCA----CCCCA-GT--.......-AGGG-GGAACCAATGATACT...AGGAAGATCAATTTCACAGAGCCAGGAAGAAAG..................................................... 6905
SMM.LR.89.SIVsmLIB1        A--C--G--G-CAAGGAA-CC---GTCA----TCCCA-GT--.........-CTGGAACCAATGATACA...AACAAGATTAATTTGACAGCTCCAGGAGGA....................................................... 6808
SMM.SL.92.SIVsmSL92A       A---G-G--G--AAAGAA-CCA-TGTCC----TCCCA-GT--.........-CTGGAACAAATAGAACA...GAAAAGATTTACCTAACGGCCCCAGGAGGA....................................................... 6837
SMM.SL.92.SIVsmSL92B       A-GC-GG--G-TAAG-AA-CCA--GTCA----CCCCA-GT--.........-CTGGAACCAATGATACT...AGGAAAATTAATCTAACGGCCCCGGGAGGA....................................................... 6808
SMM.SL.92.SL92B            A-GC-GG--G-TAAG-AA-CCA--GTCA----CCCCA-GT--.........-CTGGAACCAATGATACT...AGGAAAATTAATCTAACGGCCCCGGGAGGA....................................................... 7188
SMM.US.04.G078             A-CC-GG--G--AAAGAG-CCA-TGTCA-G--TCCCA-GT--.........-C-GGAACTAATGATACT...GAGAAAATCAATTTGACAGCTCCAAGAGGA....................................................... 7033
SMM.US.05.D215             A----GG--G-CAAGGAA-CC--GGTCA----TCCCA-GT--.........-C-GGAACCAATGATACA...AGCAAGATTAAGCTGACAAAGCCAGGAGGA....................................................... 7009
SMM.US.06.FTq              A-GC-GG--G-GAAACAA-CCC-TGT-A----TCCCA--T--.........-CTGGAACCAATGATACTAAAAACATAAACTTTACTGAGCCAGGAAAAACC....................................................... 7037
SMM.US.86.CFU212           A-----G--G-GAAGGAG-CC--GGTCA----TCCCA-GT--.........-C-GGAACCAATGATACA...AGTAAAATTAATCTGACAGCACCAGGAGGA....................................................... 7004
SMM.US.x.H9                A-CCGGG-AG-GAAGGMAMCC--GGTCA----TCCCA-GT--.........-AGGGAACTAATRAKAYT...RAACAAATTAAGCTAACAGCTCCAGGAGGA....................................................... 7229
SMM.US.x.PGM53             A-CC-GG--G-GAAGGAA-CCA-TGTCA----TCCCA--T--.........-CGGGAACTAATGAGACT...AAGAAAATTAATCTGACAGCTCCAGGAGGA....................................................... 7686
SMM.US.x.pE660.CG7G        A-CC-GG-AG-GAAGGAA-CC--GGTCA----TCCCA-GT--.........-CGGGAACTAATGATACT...AAGAAAATTAATCTAACAGCTCCAGCAGGA....................................................... 7742
STM.US.89.STM_37_16        A-----G--G-TAAGGAA-CC--GGTCA----TCCCA-GT--.........-C-GGAACCAATGACACG...GCAAAAATAAGGATAGTGGCTCCTGGGGGA....................................................... 7391
SUN.GA.98.L14              --TGCCACCT---GG-AACT---G--GC-GGTGGAACCTTTAGC---C-TG-GT-ATGGC.....TGTACATTTGATAATATAACAAAGACA...........TGCAGATTTACCAATGGTACAGATTTTAAAAAGATGATTAAGTTTAAACCA-T-G-GAAGCAT-- 7650
SYK.KE.x.KE51              --TG-G---G-CCA--ATG-GAGT-T-A-GGTG-GGAA----...........................................GTGACCTCAACATCA-GG-G-TCGCAAC--CA 7005
SYK.KE.x.SYK173_COMGNM     --T--G---G--CA-GAAC-GGCC-CGA--ACC-GGAA----...........................................GTAACCAACACCACA-GG-G-TCCCAGC---- 7309
TAL.CM.00.266              A-GG--TC-GCCAA-GAAG--A--GC-A--AGG-ACAAGG-GC---TC-AC-AC-CAGAT.............................................AAAGTGAACATCACA-GG-C-GC-GT-AGGA- 7121
TAL.CM.01.8023             A-GG--G-TG-GAA-GAG--GA-TGC-A--GCC-ACAAGG--C-C-T--AC-GC-GTAGT............................................GGAACAGCACAAATAAGG-GGC-AGG-GT-AG-AA 6668
TAN.UG.x.SIVagmTAN1        -GG-GGG-AG-GAGA-AA-C-A--GTG--G-TTCCAAA-G-G--AT-TCG-GGG-CAAAT.............................................AACACAAGGCAGATT-GGCTAAGTAGACA-T- 7408
VER.DE.x.AGM3              -GG---G-AG--AAA-ATG--A--GT-A--TT-CCAAA-G---GAT-CCA-GG--C-AAT.............................................GATACTGAAGAGATT-A-CT---GAGACT-TT 6946
VER.KE.x.9063              -GG--GG--G-TAAAGAAG--A---TCA--TT-CCAAA-G---GGT-TCGGGG--C-AAT.............................................GATACAGAGAAAATA-A-CT---GAGACAGTT 7448
VER.KE.x.AGM155            -GG--GG-AG--CAAGAGG--A--GT-A--TT-CCAAA-G-ACGGT-CCA-GGC-C-AAT.............................................GATACAAACAAAATC---TT----AGACA-TT 7428
VER.KE.x.TYO1_patent       -GGCGGG-AG--AAACAA---A--GT---GTT-CCAAA-G-C-GGT-T-A-GG--C-AAT.............................................AATACAGAACACATA-ACCT----AGACA-T- 7427
WRC.CI.97.97CI14           C-TT-TG-TT---GG-AG--T---GAGA--T-TAA-A-TTC-TGGGT--ACTAT-CAGAGGGGGACTGCAATAGCAAGAAT..............AATTCAGCAGACTTTGCTAAATGTATTAAACGG.........TTCCAGATAAGAAACTTTAC-AGACA--- 7485
WRC.CI.98.98CI04           ---T-TG-TT---GA-AG--TC-GCAG---T-TAA-AAT-CATGGGT--ATTAC-CTGAA...GATTGCAATCAGAAAAAC..............GATTCTAGAGAGTTTGCTAAATGTATTAAACAG.........TTCCAAATAAGAAACTTCAC-AGAAA--- 7530
WRC.GM.05.Pbt_05GM_X02     ---T-TGG-C---G--AAG-----CA--CTT-TAA-A-CTC-TGGGT--ACTAT-CAGAG...GAGTGTAATAAGAAAGGG..............AATAGCAGTGCATTTGCAACTTGTATTAAACAG.........TTTCAAATAAGAAACTTCAC-AAACA--- 6426
```

Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546


```
                               V4 loop start
H1B.FR.83.HXB2          AGGGGACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGTACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACT................................GAAGGAAGTG 7458
Env                      G  D  P  E  I  V  T  H  S  F  N  C  G  G  E  F  F  Y  C  N  S  T  Q  L  F  N  S  T  W  F  N  S  T  W  S  T  E  G  S  N  N  T  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  E  G  S
H1O.CM.91.MVP5180       T--A--TG-----G-AAGCCATTTACA------C---CAT--A-----C--T--T-----CA--T-TGGGA-------CTA-...ACT----TC-ACTG-ACA-AGT-C-G-.........TG-CAGGAGATCAAAGGGAGCAATGAGACCAATAAAAAT-GTACT.... 7524
H1N.CM.95.YBF30         ---A----T----G-GAC-CACTTA-TG--C------A----A--G--------T-----CA-TT-CA--T-A------GAGGAA-TAC----CGAG--AG-TGAGC--............................................................... 7011
H1P.CM.06.U14788        G--A--------GG--TCTT---TGCAC--C-G----CAT--------C--T--T-----CA-C-GCAGC--------CTT--G--ATAC-TGG-A-GACACC-CCTGGCGG-CTG--GGC--T.....................CATAATACA.........A-TGAAA 6992
CPZ.CD.06.BF1167        ---A----A---G-C-CAG--T--TT-G--C--C---CA---A-----------T---G-CCTC-A-A-TTG-ACA--C-A-GGCATCGAG-TCGA-.............................................................- 7577
CPZ.TZ.00.TAN1          T--A-----T--GG-CACT-GCTT-T-G--C------CA---A-----C--T---------CTC---A-TTG-AC-----CCTGGACAGC-...-A--GAACC-A--A-ACTCATGGT-C-CTT.............................. 7089
GOR.CM.07.SIVgorCP2135con ---A--T------G--ACCTTT-TGCA---------CAT--A--------------G--A-C--TA--A-------CTAC-A---TACA--GGAG-GC--TGA-TGCAT-ACAGGG---TG-.............................. 6981
                               V4 loop start
MAC.US.x.239            .--A--T--G---G--ACCTTCATGT-GACA-----CA---A--G--CC----------AATG-ATTGGT-TC-A---T-GGTAGAAGA--GG-A---AGCT-AC........................................CAGAAGCCAAAG---CAGCA-A 7879
Env                      G  D  P  E  V  T  F  M  W  T  N  C  R  G  E  F  L  Y  C  K  M  N  W  F  L  N  W  V  E  D  R  N  T  A  N  .  .  .  .  .  .  .  .  .  .  .  .  .  Q  K  P  K  E  Q  H
H2A.SN.86.ST_JSP4_27    .TCA---------G-G-C-TACATGT-GAC---C--CA---A------C-------C---ATG--TTGGT-CC-C---T-GGTAGAAAAC-GA-CG-A-CA--CACAG.................................................. 7372
H2B.CI.x.EHO            CTCA--T------G-AAG-TATATGT-GAC---C---A-G--A--G---C-----------ATG--TTTTT-C--A--CT-GGTAGAAAA--GA-C-GGGCTC.........................................................A 7899
H2G.CI.92.Abt96         .TCA--T------G-AAAGTTCATGT-GAC------CA---A-----C--G--T-----AATG--TTGGT-TC-A---T-G-TAGA-GAC-GA-A--TGACTCTCCTA.............................AAGCCCCAG---C--CAAA 7295
H2U.CI.07.07IC_TNP3     ---A---------GCCACCTTCATGT-GAC------CA---A--------G--T-----GATG-ATTGGT-TC-G---T-GGTAGAAGAC-GAG--GAAAACGAC.............ACGAAAGACCAAAACATTAGA--GC--CACA 7406
H2U.FR.96.12034        .--T--TT----GG-AAA-TTTATGT-GAC---C--CA---A-----CC----------GATG---TGGT-CC-C---T-GGTAGAAAAC-GG--C-TGGA-G-ACAA.............................ACGTCAAGG--GA--CA-A 7421
ASC.UG.10.RT03          .-----TAG---G-G-AG-GT---T-G--CC-G---CA------G--C--T---------T-T-TAGGA-C---C-G-AA--CAAT--C-CA-A-GA-ACCTTCTA-CC-C-AAAC-TA..................ACAAAGGAAGGG-C--AGTTGA 6907
ASC.UG.10.RT08          .-----TAG---G-A-AG-GT---T-G---C-G--CCA---------C--T--T-------T-T-TAGGA-C---C-G-AA--CAATC---CA-A--ACACCTT-TACCC-CAAAAC-TA..................ACACAGGAAGGG----CGTTGA 6937
ASC.UG.10.RT11          .------AG---G-A-AG-GT---T-G--CC-A--CCA---------C--T--T-----C-T-T-TAGGA-C---C-GCAA--CAAT----CC-A-GACACCTTCTA-CC-A-AAAC-TA..................ACAGAGGAAGGG---A-GTTGA 6895
COL.CM.x.CGU1           -A-T--T-----GGC-ACTTTCAC-TT-G-G-T----CATA--TT---C--T---------G--T-TAGTT-A-GG--ACA-GA--CC....................................................... 6852
COL.UG.10.BWC01         TA-T--T------GCAACTTTCAC-TTC--A-T---CCATC--TTG--C--T--T---G--G-TT--TC-T---GGGCA-AA-A--CA-CAGTA-TG-A-.......................................... 6690
COL.UG.10.BWC07         TACTTTTAT-ACTT-CCATGGCATGTTCC-CT----CAAT-CTAGT-CC--G......................GG-G-G----CACC..................................................... 6870
DEB.CM.04.SIVdeb04CMPF3061 ---A--TAG---GG-ACA--G-GTAT-G--CC-A--CCAT--A----A---T--T--C---GTC-GTA-GGCC-A-C-ACT-TTCACCAA-G-C..........................ACCATAAC-T-G---A 7295
DEB.CM.99.CM40          ---A--TAGG---G-ACA--GTGT-T-G---C-G---CAT--A-----C-----T-----CATTT-CA--GCCC-AG--CTCCTACTCC-AC-A-A--ACACT...............AGAAATAGCACTT-GTC-- 7201
DEB.CM.99.CM5           ---T--TAGG--GG-ACA--ACACGT-G---C-A--CCAT--A--G--C--T--T-----GTTT-CA--GCA--.......-A-CT-C-C-TG-A--AAACT...............AATAACATAACCT-G---A 7122
DEN.CD.x.CD1_CMO580407  ---A----------C-AGC-GTGCAC-G---C-G---A-G--A-----------T-----CCTC---AG-A----C--CTCCT--CAA-A-CCAGT--ACAAC..................AGTAACAT-ATT-ACA 7500
DRL.DE.11.D3            G-----T-T--C-TCA-AG-AC-T--TGA-GTTG--------A----AC----T-------ATCT-TA--A-A-GG-GA-CATGGAATAA--GATCA-G-AATGTGTGG.............................. 7193
DRL.x.x.FAO             G-----T-TG-C-TCA-AG-AC-T--TGA-GTTG--------A----AC----T-------GTTT-TA--A-A-GG-GA-CATGGAACAA--GA-CA-G-AATGTGTGG.............................. 7277
GRV.ET.x.GRI_677_gri_1  .--A--T------TCA-CG-ATTTTT-G--C------CA---T-----------T-----GATGGACTGGT-TA-C---TA-CTAAACAA-CGA-CAGAAGATGCAGAA.........GG--CT-A-AGGACCTGTGACAAAGGGAAGCCAGGACCAGGACC-T-TGT-C 7471
GSN.CM.99.CN166         ---A--T-T---GG--CA---T--TT-G--CC-A---CA---A-----------T--C----TGT--AT---C---C-ACT--A-AATAAC...-CC-T-AAC--C-GC...AATATTGGG---.........ATCACAAGCAA-TACA 7404
GSN.CM.99.CN71          ---A--T-T---GG-CCA---T--TT-G--CC-A---CAG--A-----------T-------TGT--GT---C---C-ATTC-ACAACAA--T--CGG-AGT-GA-GACAC-AACATT--C--T.........GTCACAAGCAA-TACA 7410
LST.CD.88.SIVlhoest447  G--AA-T-TT-C--CA-AG-T-TTG-TGCA--CA------AA---A-G---T---C---GT----AG-A-A--CC-GGAATGGAATAA---A-ACT-AGAT-AATGG.............................. 6616
LST.CD.88.SIVlhoest485  G-AA-GT-TT-C--CA-AG-T-TTG-TGCA--CA------A----A-G--T-T---C---AT----AG-A-AC-CC-AGAATGGAATAA---A-ACT-AGAT-AATGG.............................. 6613
LST.CD.88.SIVlhoest524  --TA--T-TC-C--CA-AG-T--T--TGCAC-C------T-A----A-G----TT------GTC---AG-A-C--CC-GGAATGGAACAA---A-ACT-AGAT-AATGG.............................. 6619
LST.KE.x.lho7           --CA---TT--C--CA-AG-T-TTA-TGCA--C--------A----A-G----TT------GT---TAGGA-A--CC-AGAATGGAACAA---A-A-T-AGAT-AATGG.............................. 7691
MAC.US.x.251_1A11       .--A--T--G---G--ACCTTCATGT-GACA-----CA---A--G---C----------AATG-ATTGGT-T--A---T-GGTAGAAGA--GG-A---AGCT-AC..................CAGAAGCCAAAG---CA-CA-A 7875
MND-1.GA.x.MNDGB1       .--A--TAG--C-GCA-A--T-ATG-TGA-G-CA---CA---T---A-G----T-------GT----AG-A-CA-G-GGGCATGGAATGA-CC--A-GAGAA--AGTGG.............................. 7049
MND-2.CM.98.CM16        ---C--A----G-GCA-A---TATT-TGC-CCTC--------A----AC----TT-------GG---A--A-T-GG-GA-CATGGAATGA-C-A-A-T-A-CAGTATGG.............................. 7604
MND-2.GA.x.M14          T--C--A----G-GCA-A---AATT-TGA-CCTC--------A----A-----TT-------GG---AG-A-T-GG--A-CCTGGAATAACC-A-ACT-AAATGT-TGG.............................. 7530
MND-2.x.x.5440          ---C--A----G-GC--A---TATT-TGC-CCTC--------A----A-----TT-------GG--TA--A-A-GG--GGCATGGAATAGC--ACAA-G--CAGT-TGG.............................. 7231
MNE.US.x.MNE027         .--A--T--G---G--ACCTTCATGT-GACA-----CA---A--G---C----------AATG-ATTGGT-TC-A---T-GGTAGAAGA---G-A-CTGACTG-A..................ACTACCCAGAAGCCACAG---CAGCA-A 7356
MON.CM.99.L1_99CML1     -------AAG--GG-CCAG--T---T-G-----C--CCAG--A-----C--T--T-----GTCT--GCTT-A--C-T--A--GAA-AACC--C-AA--GGACG-G...........ATATCTCCATTCGACGTGAATAACAAGCCCAACACCACGTA-C 7430
MON.NG.x.NG1            ---A---AAT---G-CCAG--C--TT-G--C------CAG--A--G--C--T--T--C---CTGT-TAG---A--C-TYTCG-G-GACCCG-R--AG-A-AA--A-GAAAC-TTTGA-TCC-C-TTCAATGTTTCTACCAATGCCTCCACTGCCTATAAC-G--AGT-GT 6067
MUS-1.CM.01.CM1239      ------TAAT--GG-CCA---A--TT-G----------C---------C--T-----C---CT-T--GTG------C-ATT--ACAACAA--TC-CACAAATT-ACGAG..........AGCAATATACATAATGTGACAACCA--TACCACC 7370
MUS-1.CM.01.SIVmus_01CM1085 G--A--T-A---GG-CCAG--A--TT-G-----------G--A-----C-----T------CTCT-TGTG-----CC-ATT-GT-AACCAC-CAGAG-TAAATGAA..........AGCAACATCCATACTGTAACAACTA-GTACCACA 7398
MUS-2.CM.01.CM1246      G--A--TAGG--GG--CAG--A---T-G--------CCA---------C--------C--C-TGT--GT-T-A--CC-ATT-GAGAATAAC-CC--AGAAATT-A-GAC..........TCAAATATACATACAGTAAAAACCA--TATC-CC 7505
MUS-2.CM.01.CM2500      ---A---AGG--GG-CCA---T--TT-G-------------A-----C--T--T--C--CCTT-GTGTC--A--CC-ATT-GAGAATGAC-CC--ACAAATTCAC-A-..........TCCAATATTGATACTGTAACAACCA--TACCACA 7419
MUS-3.GA.09.09GabOI81   ---A--T-----GG--CA---A--TT-G---GC-----AT--A-----C-----T------CTCT--GCC--T--CC-ATT-GAGAATAA-GTA-CAGAAATT-A-T-A-GC..........AATATTGATAATGTGACGTCAA-GTAT-AA- 6999
MUS-3.GA.11.11GabPts02  ---A---AG----G-A-AG--A--TT-G---GC-----AT--A--G--C--------C--CCTTT--GT-T-A---C-ATT-GA-A-TG-G-GAGAG-TAACAGACGAA..........AATATACACAATCTTACATCAA--TAT-AGA 6982
OLC.CI.97.97CI12        ---A--TGAT-CTGCA-A--ATATG-TGA-G-TG-----T--CT-CA-G-GG-TT------ATC--TA--A-TC----A-CATGGAATAA-GCGTC-GA-CCT-AGTGG.............................. 7090
RCM.CM.00.SIVagi_00CM312 ---A--TGAT--G-C-AA-TATTTTT-GC-A-----C-AT--A-----CC----T-----ACTC-ATTGGT-T--G---TA-TTGAACAA---G-CAGAGG-T..................AATGAGAAG---A----GC 6944
RCM.CM.02.SIVrcm_02CM8081 ------TGAT---GC-AA-TACTTTT-G--A-------AT--------C--A--T--C--A-TG-ATTGGT-TC-C--CCTATTAAATAA--GA-C-GA-G-A..................AACATAAAT--GC----AC 7022
RCM.GA.x.SIVRCMGAB1     T--A--TGAT--GGCAAG-TATTT-T-GC-A------AAT--A--------A--T--C--G-T--ATTGGT-T--A---TTGTTAAATAA-G-G-CAGTAG--..................ACAACAAAT--GAAG--AA 7160
RCM.NG.x.NG411          ---A--TGAT--GGCAAA-TGTTTTT-G--A-------AT--A------C-T--T--C--AATG-ATTGGT-TC-A--CTA-CTAAATAA---A-CAGAAG-C-A--AA..................AAT----CT--GC 7204
SAB.SN.x.SAB1           G--A--T------TC--A-TTCTT-TT----------CA---T-----C-----T-----AATGGATTGGT-CC-C--CTA-TTGAATAA---GTCAGTAGACCCAGAC......CAT--C---TG-GCAAAAAAATAAT..............ACAAAGCC-T-TT-GC 7768
SMM.CI.79.SIVsmCI2      .TCT--T------G-A-C-TTCATGT-GACA--C--CA---A-----CC-------C--GATG---TGGT-CC-C---T-GGTAGA-GAA-GG-AC..........ATGACAGAACAAAACCCGGAG---AA--AGA 7032
SMM.LR.89.SIVsmLIB1     .--A--T------G-AAC-TTTATGT-GACA-----CA---A--G--CC----T-----GATG-ATTGGT-CC-G---T-GGTAGA-GA--GAGAC--AAATCACG--..........AGATGGACATCACAAACTAGCAAA---C---AGA 6950
SMM.SL.92.SIVsmSL92A    .-----------GG--ACCTTTATGT-GACA--C--CA---A-----CC----T-----AATG-ATTGGT-TC-G---T-GGTAGAAGAC-GA-A-TTGACACTCC-A..........AGATGGAAAAACCAAAAGCCAGAG---AA-GC-A 6979
SMM.SL.92.SIVsmSL92B    .-----T--G---G-CAC-TTCATGT-GACA-----CA---A-----CC-T-------AATG-ATTGGT-CC-C---T-GGTAGA-GA--GG-A---GA-C--CC-A..........AGGTGGACAACTCAAACCAAAAAG---CAGCACA 6950
SMM.SL.92.SL92B         .-----T--G---G-CAC-TTCATGT-GACA-----CA---A-----CC-T-------AATG-ATTGGT-CC-C---T-GGTAGA-GA--GG-A---GA-C--CC-A..........AGGTGGACAACTCAAACCAAAAAG---CAGCACA 7330
SMM.US.04.G078          .-----T------G--ACTTTTATGT-GACA--C---C---A-----C--A--T--C--AATG-ATTGGT-CC-----T-GGTAGA-GA--GA-ACGAAAAAG--T--..........ATATGGGAACAACAAAACAAGAAA--GCA--AGA 7175
SMM.US.05.D215          .--A---------G-CACCTTCATGT-GACA-----CA---A--G---C----T-----AATG-ATTGGT-C--A---T-GGTAGAAGAA-GA-A-G--ACAGAG-TG..........AATGCGAGAAACAAAAACAACCGG---AAG-AAA 7151
SMM.US.06.FTq           .TCT---------G-A-CTTTTATGT-GACA-----CA---A-----CC----T-----AATG--TTGGT-CC-C---T-GGT-GAAGAC-GA-AC-TG-CT....................TCACAAAACAAGGCA---CA--AAA 7164
SMM.US.86.CFU212        .--A-----T--GG-CACCTTCATGT-GACA------A---A-----CC----T-----AATG-ATTGGT-CC-A---T-GGTAGAAGA--GA-AC-GAAC-CAG-TG..........GAGGCGAGAAACAACAACGATAAA--GA-G--AA 7146
SMM.US.x.H9             .--A--T------G--ACTTTCATGT-GACA------C---A-----C--A--T--C--AATG-ATTGGT-TC-----T-GGTAGAAGY--TA-AA-A-G-TTC-..........AGATGGACAAGTCAAAACCAGAAA---C--TACC 7368
SMM.US.x.PGM53          .--A--T------G-CACTTTTATGT-GACA------A---A--G---C----T--C--AATG-ATTGGT-CC-----T-GGTAGAAGA--GA-AC--A-CAG--T--..........CTATGGACACAACAAACAAAGGCA--C-A--AGA 7828
SMM.US.x.pE660.CG7G     .--A--T------G-CACTTTTATGT-GACA------A---A-----C--A-----C--AATG-ATTGGT-TC-----T-GGTAGA-GAC-GAGACCAAAA-----GC..........AGATGGAGACAACAAAATACGAGA--GC--CAGA 7884
STM.US.89.STM_37_16     .--A--T-----GG-CACCTTCATGT-GACA-----CC-T--A------C-T--T--C--AATG-ATTGGT-C--A---T-G-TAGAAAAC-GA-----A-CAGAA.......ATGAGAGATTGGAACAAAAACAAAAAG---CA-CAAA 7533
SUN.GA.98.L14           --CA--TG---C--CA-AG-T-TTG-TGA-G-C--------AA--GA-G----TT-------T----AG-A-T--C--AGTATGGAATGAC-CA-CCT-CAAT-AATGG.............................. 7759
SYK.KE.x.KE51           ---A----TG---G-GAG---A--TT-G---C-A-----C--A--------T-----C---GTTT--A-GA-T---GCA-A-GTG-CAAA-GGCTA-G--AATCCA.............................AGCA-TTATGCAA 7124
SYK.KE.x.SYK173_COMGNM  G--A--T-T---GG-ACG---A--TT-G--CC-A-----G--A--G--C-----------GT-T-TA-G--C---GC--AC-T-AC-AA-GGA-ACG-AA-T-AG-A-...............................A-TTATGCCA 7428
TAL.CM.00.266           ---A----T---GG-ACA-CTTTT-T-GCA-GCA---CAG--A-----C-----T--C--CAT--GTTCTA----CCTG-A-----CC-A-G-GTA--AC-AC-CAT-AA-CCAAGGC--ATTACACCAAGGGGTCATAGGAAAG...AAGCCGAGGTCGAGT-AT--GA 7288
TAL.CM.01.8023          G--A----T----G-ATC-CTCTTTT-GCA-GCG--CCA---A--G--C-----T------ATTT-TAGCA----C-TG-AA--CAAC-A--CATAC-A--AC-CAG-A---AAAGG-GTACTGCACCAGGAGAAGATAGGCCGACAAGAGCACAGGTATAGT-AC--AA 6838
TAN.UG.x.SIVagmTAN1     G--A--T------GCA-CT-ACATTT-GC-C-----CCA---A--G--------T----C-C-TGATTGGT-TG-----T-GCTGAATAA-G-GTC--A-A-TG-A-GA......AATGTAG-TGTAGAAGGTAATAATTGCACCACTGGAAAGGATAAACGCT-CTACA 7572
VER.DE.x.AGM3           T--A--T------GCA-C--ATTTAT-G---------CAG--------C-----T-----AATGGATTGGT-TC-A---TACCTGAATAA-CG--CAGTAGATCCGGAC......CAT---CCGTG-AATGGTACGAAGGGAAAAGGTAAGGCACCAGGACCCT-TGCAC 7110
VER.KE.x.9063           T--A--T------GCA-C--ATTTGT-G--------CCAG--A--G--C--------C--GATGGACTGGT-TC-A---TA-TTGAATAA-C-ATCAGTAGATGCA-A-......CAT---C-ATG-AGTAATGAAACAAAAAAAGGGAATGCCCCAGGACC-T-TGTAC 7612
VER.KE.x.AGM155         T--A---------GCA-C--AT-TAT-G--C--C---CA---------C-----------AATGGACTGGT-T--A---TA-CTGAATAA-TTA-CAGTGGATGC-GA-......CAT---C-TTG-AAAAACAACGCAGGGAAAGGTCGAAGTCCAGGTCCCT-TGTAC 7592
VER.KE.x.TYO1_patent    G--A---------GCATCC-ACTTGT-G---------CA---A-----C--T--T-----GATGGATTGGT-T--A---TACTTAAATAA---A-CATGGGATGCAGAC......CAT---TTTTG-AGCAGCAAAAAG...AAAGGACACGCACCAGGACC-T-TGTAC 7588
WRC.CI.97.97CI14        --CC--TAG--CT-CA-A--ATATA-TGA-G-TC--------A---A-G----T---C---ATC---AG-A-AGCAG--GCATGGAACAA-CGC---GAGACC-AATGG.............................. 7594
WRC.CI.98.98CI04        --CT---AG--CT-CA-A--ATATA-TGA-G-T---------A---A-G----TT--C---ATT--TA-GA-TGC-G--GCATGGAACAAC--G-A-GAGA-C-AATGG.............................. 7639
WRC.GM.05.Pbt_05GM_X02  --CT---AG--CC-CA-A--AT-TT-TGA-G-TA--------A---A-G----TT-----CATT---A--A-AGCAG--T-GTGGAATAA---A-CAGAAAA---ATGG.............................. 6535
```

V4 loop end ⌐ ⌐V5 start

```
H1B.FR.83.HXB2          ACACAATCACCCTCCCATGCAGAATAAAACAAATTATAAACATGTGGCAGAAAGTAGGAAAAGCAATGTATGCCCCTCCCATCAGTGGA.........CAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAAT..... 7604
Env                     D  T  I  T  L  P  C  R  I  K  Q  I  I  N  M  W  Q  K  V  G  K  A  M  Y  A  P  P  I  S  G  _ _ _ _  Q  I  R  C  S  S  N  I  T  G  L  L  L  T  R  D  G  G  N  _ _ _ _ _ 
H1O.CM.91.MVP5180       .........A-A--T---AGT---G---GC-AG---GATCA---AT--G-G--AGTCGAG---C-----A-------CCC--C.........A-CT-A-C---CAT--C--C--A--T--AA--A-TC--CAGTT----CAACCA..... 7656
H1N.CM.95.YBF30         .....-T--G--T--T-------G---G---G----TT-----ACA-GG-------G---T-----A--A--A--TC-G---------GTTC---AC---A-C--C---------T--A---G-TC--GA-TATAG------GCC-..... 7152
H1P.CM.06.U14788        -T-T-._.--A--GT-T--T---C-G--G---G-AG-T--TTCA---AT--G---G---TC--GCT---T---GA-A--T--TGAG--G.........G--T-A-AG---CA---C--C-------TT-AA-G---GA------ACTACC-T-CAAT..... 7138
CPZ.CD.06.BF1167        GA-TGC-G-TAGCA-----TC--C-T-G---G--AG----TCAT---GGA-T----TCC-GG-G---T---CT-------GGGAG-----------GG-C-A---GT---T-G---------AT-CA--A-G---GCT--GACGAA-G-CACTGATGGC..... 7732
CPZ.TZ.00.TAN1          ........GTTGCA---------C-G-GG--G--AG----TCAT---GGT-T---GTC-----GGG-T--CCTT--C--A-GG--G---------AC-G-A-A-----CAC-----C--C------A--TA-CA-G---GC---AAAA-AC--CAAT..... 7230
GOR.CM.07.SIVgorCP2135con ..-----ATTATA----C---CATC-G-----GG-AG-C-TGTCC---AT--G------CTC--GCC-A-T---T--G--T--T--G---------ACTT-A------AA---T----------AA-AA-----CA---G----TACCCCTGAATAAG...ACTAAT 7137
```

V4 loop end ⌐ ⌐V5 start

```
MAC.US.x.239            -A-GG-ATTA-G-G-----TCAT--T-G------A--C----CT-----T--------C---AATG-T---TTG-----A-GAGAG----------G-CC-C-CG---AAC--C-CAG-G--CA-T--CA--GC--AC-T----T-GAT-G--GGA..... 8028
Env                     K  R  N  Y  V  P  C  H  I  R  Q  I  I  N  T  W  H  K  V  G  K  N  V  Y  L  P  P  R  E  G  _ _ _ _  D  L  T  C  N  S  T  V  T  S  L  I  A  N  I  D  W  I  D  G  _ _ _ _ 
H2A.SN.86.ST_JSP4_27    ..CAC-ATTATG-G-----CAT----G-----A--T--T-CC-----C--G----G---AATG-A---TTG-----T-GGGAA----------GT-A-CC--CAAC--T-CAG-G--CA-CA-AA-TGCT-AC-TT--C--A--AG-G..... 7516
H2B.CI.x.EHO            -G-G--ATTATGCGT-----CAC--C-G---G--AG-C----C------C--GA-T----G-AATG-----TTG-----A-GAGAG--T.........G--C-CTCC---AAT--C-C-G----CA-C--CA--GCC-AC-TT--CT-GATAG--AAG..... 8048
H2G.CI.92.Abt96         GACGC-ATTATG-A--C---CAT--T-GG--G--A--T----CA-----C-----------AATG-----TTG-C---GAGAA--T.........G-TC-G-CG---AAT----CAG-C--CA-C---AA--GCG-AT-T---A-AGAA--GGTCT..... 7444
H2U.CI.07.07IC_TNP3     -A-G--A-TATG-A--T--TCAC--T-G---GG-A--T---T-CA---A-T-G---------GAATG-A---TTG-----A-GGGAA----------G-TC-C-CT---AAT--C-CAG-G---A-T----A--GCC-AC-T----T-GAC-G--ACA..... 7555
H2U.FR.96.12034         -A-GG-A-TATG-G--T---CAT--T-GG---GG-A--T----CA-----T-----A-G-AATG-----TTG-----A-GAGAA----------G-TC-C-CC--CAAT--T-CAG-A--TA-CA--A-TGC--AT-TT----C-GAT---CAAA..... 7570
ASC.UG.10.RT03          G--AC--ATGGT--A---T-CC--C-GG---G-AG-G----GA---TCAC-T----A---GTTG-----CTTA--A--A-GA--G----------TG--CAG---A----T---G-C-----AA-CT-GG-GGA-TCT--ATACTA-CCAGGA..... 7056
ASC.UG.10.RT08          G--AC--ATGGT--A----T-CC--C-GG---G-AG-G----GA---TCAC-T----A---GTTG-----CCTA--A--G-GA--G----------TG--CAG---A----C--C--C-----AA-CT-GG-GGAGTC---ATACTA-CCAGGA..... 7086
ASC.UG.10.RT11          G--G--CATGGT--A----T-CC--C-GG---G-AG-G--T-G----TCAC-T----A---GTTG-----CTTA--A--A-GA--G----------TG--CAG---A----T-----C-----AG-CT-GG-GGA-TC---ATACTA-CCAGGG..... 7044
COL.CM.x.CGU1           .....CCAGT-A-GAAC--T-CC--T-G-A-GC-GG-G--T-GT---GTC-CTCAT-CC-GGATCC------GG--G--TCCAG-A--T.........TC-ACA----AAC-GGG-G-AACAGCCAG-TA-TGCGTTC-TG-GCACCATAG-AGGA..... 6996
COL.UG.10.BWC01         ..................-T-CC--T-G-A--T-GG----T-GT---GT-TC-CAT-CC-GGCTGC-A---G------GCCTGA----------TT-GCA----AAC-GGG-A-AGCA-CCAG-AA-TGCTTTT-TG-GCACA-TAG-AAAA..... 6822
COL.UG.10.BWC07         ......-CAGTAA-GAAC----CT----G-A--T-GG-G--T-GT----GT---CCCAT-C---GAT-T------GG--A--TC-A-AA----------ACTC-GCAGAC--AT-GGG-G-AACA-CCAG-TT-GGCCTTT-TG-G-TCA-TAG-AAAA..... 7014
DEB.CM.04.SIVdeb04CMPF3061 ---AC-CTTTGA-G--C----A---T---T---GG-------CC---T-C--GA----GC-GACT--A--CTTG--A--A-GAGAA--G.........G-GT-G--G---GAC---C--G--T-T-CAA-AA-T--TGATGTG-C-CACTA-GGACAC..... 7444
DEB.CM.99.CM40          ---AGTGGTTGA-G--T--T--G-----T--GT-AG------CT---T-T-CG-----GC--CAT--C---CTA-----G-AAGAA----------G-GT-G-AG--CAGC---C-C---T-T-CAT-TG-T--TGATGT----CACTA----GGG..... 7350
DEB.CM.99.CM5           -T-AC-CTTTGA-G--T--T-AG--T---T---TG-------CC---T-C--GA----GC-GATCA--A--TTG--T--A-AGGAA----------G--T-G-AG---CAT--TTC-G--T-T-CA--TG-T--TGAGGT--C-CA-TA-GGGCAA..... 7271
DEN.CD.x.CD1_CMO580407  --CAG-CATATA-GGTG--T-AC---------T-AG-G----CA-----C-C------C-TG-AT--T--CTTA--A--AGCAGAA----------G------C---CA-C--C---G-GT----CA-TA-TG--GAC-CG----TCCA-CCGAAC..... 7649
DRL.DE.11.D3            ..TATCCATATGCAT-C---CAC----------AG--G-TGAT---GCA-G---G-C-GGAA---T---ATG-----TG-ATCA---TTTAATAATG----A------A--CA-G--G-C----AAA--T-C--TGAG-TCC-AAAATTGG-GACAAAT.....GAA 7355
DRL.x.x.FAO             ..TATCCATATGCAT-C---CAC----------A---G-TGAT---GCA-G------C-GGAAG--T---ATG-----TG-ATCA---TTTAATAATG----A------A--CA-G--G-C----AAA--T-C--TGAG-TCC-AAAATTGG-GACAGAG.....GAA 7439
GRV.ET.x.GRI_677_gri_1  -G-G--CTTATG-TG-C----CAT---CG----G-AG----TGAT---T-C-CT---CTCT---AAGG-A-----T---A--A-GGGAA--T.........TT-GGAG---AAC---TCAG-C--G-CA--ATACG-GG---T----TA-AAC--CAAG..... 7620
GSN.CM.99.CN166         -AGG-CAATGG--GG-T----A---T-GG---T--G-C-CAA---GGAT-T--GTCT---T----C--CTTA-----ACGGCAG--T.........T--A-AT--CA-----C--C-------T---T-GA--GATG-C-C-ATGTA-G-AAGC..... 7553
GSN.CM.99.CN71          -AGG-CAATGGA-GG-T----A---T-GG---T--G-C-CACAA---GGAT-T--GTCT-GGT-G--C--CTTA-----ACGGCAA--T.........T--A-AT--A-----T--C--C--T--C--C---A--GACG---CGATGTA-GGACAT..... 7559
LST.CD.88.SIVlhoest447  ..TATCC-TGGGC-AAT---CAC------TC-G-----G-TGAT---GCATCT----G--GAA---T--CT------A-CATCA--CTTTAACAACAG---A------A-TCAT-GAG-A----AAA--TGG--TCA--TG--AAAGTT-G-ACCA.....GAAGAA 6778
LST.CD.88.SIVlhoest485  ..TATCC-TGGGC-AAT---CAC------TC-G-C---G-TGAT---GCA-CT-----G--GAA---C---AT------A-CATCA--CTTTAATAACAG---A------A-TCAT-GAG-A----AAA--TGG--TGA-CTG--AAAGT-GG-GCCA.....GAGGAA 6775
LST.CD.88.SIVlhoest524  ..TATCC-TGGGC-AAT---CAC------TC-------G-TGAT---GCA-CT--T--G--GAA---T--CTT----A-CATCA---TTTAACAATAG---A------A-TCAC-GAG-G----AAA--TGG--TG--AAAAT-GG-ACCA.....CATGAA 6781
LST.KE.x.lho7           ..TATCCTTGGGC-AAC---CAC------AGC--C--TG-TGAT---GCT-C-A------GAAG--T---CTA--A--T-CTTCA--TTTCAACAATAG---C------A-TCAT-GAG-A----AAA--T-C--TGAG-TG--AAAAT-GG-ACCA.....CATGAG 7853
MAC.US.x.251_1A11       -A-GG-ATTA-G-G-----TCAT--T-G------A--C----CT-----T--------C---AATG-T---TTG-----A-GAGAG----------G-CC-C-CG---ATC--C-CAG-G--CA-T--CA--GC--AC-T----T-GAC-G--GGA..... 8024
MND-1.GA.x.MNDGB1       ..TATCCTTATGC-T----TCA---T-GG-----AG--G-TGAC---AT-C--------G-AAG--A---TTA--A--T-CATCA---TTTAATAAT--C--A--G---A--CAT-GGG-A----AAA--TAC--TGA--TGC-AAAGATAG--AGT..... 7205
MND-2.CM.98.CM16        ..TATCCATGGA-GT----T-AT--T-G------A---G-TGAT----T--------G---AATG-T---TTG-----A-GAGAG----------TG-ATC----TTTAATAATG-G--A--------TAATG--G-A----AAA--T-C--TGAGGTTC-AAAGACAG-AGAG..... 7760
MND-2.GA.x.M14          ..TATCCATATA-GT----T-A---C-G---G--AG-GG-TGAT-----T-----G----GGAAG--T---ATG-----TG-GTC----TTTAATAAT--T--A-----CA-TAATG--G-C----A-A--T-T--TGAGGTCC-ACTA-T-G-AGGA..... 7686
MND-2.x.x.5440          ..TACCCTTA-A-GT-TC---AT--T-G------AG--GGTGAT-----T--------C--GAA---T--ATG-----TG-GTC--GTTTAATAATG----A------A-TAATG--G-G----AAA--T-C--TGAGGTCC-AAAAAC-G--GATGAC..... 7390
MNE.US.x.MNE027         -A-GG-ATTA-G-G-----TCAT--T-G------A--C----CT-----T--------C---AATG-T---TTG-----A-GAGAG----------G-CC-C-CG---AAC--C-CAG-A--CA-T--CA--GC--AC-T----T-GAT-G--GGA..... 7505
MON.CM.99.L1_99CML1     -TGGTGGGTGG--AG----T-CC----G------GG-G-CACAA---GGAT-T--GTC---GAGT--C--CCTA-----ACGG-AA--C.........TG-GCAG---A----T------C--T-CC---C--GA-T---G-G--GTTGTA-C-AAAC..... 7579
MON.NG.x.NG1            TA--GTG-CGA---AGGCAACT-G---CT--GTGGGG-T----C-CCA---GTR-CTATCTGC--CCAAGACAAGGG-ATG-GCAGT-T.....................A-C--T-----C----CC--CT-GG--GATG-G--GTTGTA-GGCAGC..... 6204
MUS-1.CM.01.CM1239      -AGATCAATGGA-GGTC--T--G---T-G---GT---G-C-CGCAA---GGCT-T--GTC----T-C--C---CTTG-----ACGAC-G----------TG-GCAG---A-T---T---G-G--T-----C--TA-TGATG---C-ATGTA---AGAC..... 7519
MUS-1.CM.01.SIVmus_01CM1085 -GGATCAATGGA-GGTC--T-----T-G----T--G-C---CA---GGAT-T--GTC----GT-T--C--C---CTTG-----ACG--AG--C.........TTG--C-----GAG---T-----A----------T--CA-TGACG---CCATGTA---GAGAT..... 7547
MUS-2.CM.01.CM1246      -TGACCAATGGA-GGTC----------GG--GT---G---C-CA-----GGTT-C---TC---GT-C--A--CCTG--C--TCGA-AA-T.........TG--CA-----A-----CC--G-A--------AT--A--GATG-G-CCATGTACGGCAAT..... 7654
MUS-2.CM.01.CM2500      -GGACCAGTGGA-GGTT--T--G----G----T--G-T-C-CA----GGTT-TACGTCG--GT-C--C---CTG--C--TCG-CAA--C.........T---CCAG---A-G--C---G-G--T--T--C--CA--GATG---CCATGTA-GGCGAC..... 7568
MUS-3.GA.09.09GabOI81   -A-ATCAATGGA-GA-T----A------G---T--G-C-C-CAA---GGCTTTACCTCT--GT----A---CT---A--A-AAC-A--G.........-T---CA---CA-C--C--------T--TT-AT-GA--GA-G-G-CAATGTA-CGCGAC..... 7148
MUS-3.GA.11.11GabPts02  -A-ACCAATGGA-G-TC--T-----C-G----T--G-T-C-CA----GGCT-CACCTCC--GT-T--A--CCT---A--A-AACAA--G.........-T-C-AC--CA-T--C--C--------AA--T-G--GGA--AT-CGATGTA-GGAAAT..... 7131
OLC.CI.97.97CI12        ..TATCCATGGTCA-AG--T---C-G-G-----A-G--G-TGAT---AT----------G-AG---T---CTT----ACAACCA---TTCAATAATAG---CC-----------------CAAAGAG--ACTGA--TGTT-TTG-AA-TACAA..... 7234
RCM.CM.00.SIVagi_00CM312 -AG-C--GTTTG-T--T---GTC-CG---ATG---G-T--TGAT---T-C-CG--GTC--GGAA-G-T---A-A--G--A-GGCAA-AT.........GCGT-A-C----AATG-T-CAG-A-GCTATT--T--GC-GAC-T---GTA--TG---CAA..... 7093
RCM.CM.02.SIVrcm_02CM8081 -AG-C--GTTTG-G--T--T-TC-C----ATGG-GG-T--TGAT---T-T-C---GTC----AAGG-A--CA-A--A-GACC--AT.........GC-T-G-A---CAGGG-C-CGG-C--TTACT--T-GGC-GAC-T----TA-ATAC--ACC..... 7171
RCM.GA.x.SIVRCMGAB1     -AG--CCTTTTG-A-------TC-C----ATG--AG-C--TGAT---T-T-C----TCG-GGAAGG-A--CA-G--A--G-GGCCA-AT.........GCGT-A-AG--CAGTG--C-GG-AT-CTAT---T-GGC-GAC-T---CTA-AT----GAC..... 7309
RCM.NG.x.NG411          -AG-C--GTTTG-T--T--T-TC-CT---ATGG--G-T--TGAT---T-T-C----GTCT-G-AA-G-A--CA-T--A--A-GGCCA-AT.........GCCT-G-----CAACG-C-C-G-G---TACT--T--GC-GAT-TT--CTA-ACAG--AAT..... 7353
SAB.SN.x.SAB1           -G-G--C-TATG-T--C---CAC----GG---G-AG----TGAT---T-C-CT-GTCT---AAG-CA--C----------GAGAA--C.........-CT-GGAG---AAC--G-CAGCC----C----TATG--GA-TTGA-CTA-AAC-G-AAA..... 7917
SMM.CI.79.SIVsmCI2      GACAT-ATTATGCA--C--TCAT--T-G------A-----T-CT----T--------T---AACG-A---TTG-------GAGAA-T.........G-TC-G-C---CAAT----CAG-A--CA-T--TA--GCC-AT-TCA-CT-GAA-G-C..... 7178
SMM.LR.89.SIVsmLIB1     -A-AG-A-TATG-G--T---CAC--T-GG---G-A--T--T-CC----T-G--------GAATG-----TTG-----A-GGAA----------G-TC-C-CG---AAT--C-C-G-A---A-T--AA-TGC--AT-T----T-GATA---..... 7096
SMM.SL.92.SIVsmSL92A    -A-GG-ATTATG-A--T---TCAC--C-G-----G-G--T--T-CT----T-G--------CAATG-T---TTG-----A-GAGAA----------G-CC-GCA----AAC--C-CGG-A---A-TT-AA--GC-GAT-T----T-GAT---C..... 7125
SMM.SL.92.SIVsmSL92B    -A-G--ATTATG-A--T---CAC--C-GG-----A--T--T-C-----T-GG--------AATG-C---TTG-----A-GGAA----------G-CT-A-CT---AAC--C-CAG-G---CA-C--AA--GC--AT-TT----T-GAT-G-C..... 7096
SMM.SL.92.SL92B         -A-G--ATTATG-A--T---CAC--C-GG-----A--T--T-C-----C-GG--------AATG-C---TTG-----A-GGAA----------G-CT-A-CT---AAC--C-CAG-G---CA-C--AA--GC--AT-TT---T-GAT-G-CAAC..... 7479
SMM.US.04.G078          -A-G--ATTATG-G-----TCAT--T-G---G--A--C--T-C------C--G-----T---AATG-----TTG-----T-GGGAA----------G-CC-G-C----AAT--C-C-G-A--TA-C--CA--GC--AT-T----CT-GATAG--..... 7321
SMM.US.05.D215          -ACAC-ATTATG-G--G---CAC--T-G------A--T--T-CC-----C-G--------GAATG-----TTG-----T-GAGAA----------G-CC-G-C----AAC--T-C-G-G--CA-TA-CA--GC-GA--T----T-GCT-----..... 7297
SMM.US.06.FTq           GACGT-A-TATG-G--T---CAC--TCGG--G--A--T----CT----C--G-----G---AATG-A---TTG-----T-GGGAA----------G--C-G-C----AAT----CAG-A--CA-T--TA-TGCT-AC-T-T--T-GAT-G-..... 7310
SMM.US.86.CFU212        -G-GG-A-TATG-G-----TCAC--T-G---GG-G--C--T-CA-----C-G------T---AATG-A---TTG-G--T-GAGAA----------G-CC-G-C----AAC--C-C-G-G--CA-T--AA--GC-GA--T---CT-GTA-G-..... 7292
SMM.US.x.H9             -G-A--ATTATG-G-----TCAT--TRG---G--A--C----C-----C--------C---AATG---R-TTG-----T-GGGAA----------G-CC-G-C----AAC--C-C-G-A--TA-C--CA--GC-GA--T----T-GATC---GGC..... 7517
SMM.US.x.PGM53          -G-G--A-TATG-G-----TCAT--T-G---G--A--C----C-----C-G------T---AATG-----TTG-----T-GGGAA----------G-CC-G-C----AAT--C-C-G-A--TA-C--CA--GC-GAT-T---CT-GAT-G--GGC..... 7977
SMM.US.x.pE660.CG7G     -G-A--ATTATG-G-----TCAT--T-G------A--C----C-----C--------C---AATG-A---TTG-----T-GGGAA----------G-CC-G-C----AAT--C-C-G-A--TA-T--CA--GC-GAG-T----T-GACC---..... 8030
STM.US.89.STM_37_16     -G-GG-ATTATG-A------CAC--C-G---G-A--T--T-CA-----T-------G---AATG-T---TTG-----A-GACAA----------G-CC---CT---AAT--C-C-G-G-CA-TA-AA--GC--AC-TT--CT-GAC---CAAT..... 7682
SUN.GA.98.L14           ..TATCCTTGGGC-AAT---CAC------AGTG-G---G-TGAT---GC-TC---G--C-GAAG--T--CTTG--C--A-CATCA-GTTTAATAACAG---A------G--AAC-GAG-G----AAGCATGG--C---CTG--AA-G-TGG-AGACTGG...AAGGTA 7924
SYK.KE.x.KE51           -A-ATT-ATAT---T-G--TGCC----CG---G--------TAT---GG-T-T--ACC--GCT-----TTG--A--T-CAGAA-G..........C----AG--CA-T--T--------T-CTG-AT--AC-GAT-T---GTACTA-CCAAAT..... 7273
SYK.KE.x.SYK173_COMGNM  GT-ACC-TCGT--GT----TGCC--TCGC-----A--T--TGAT---AGAT-T---A-----TTG--A--CCTG--C--T-CGGCA-G..........T----A----A-T--C---G-C--T-CAG-AT-GACGGAC-T---ATACTACCCAGGA..... 7577
TAL.CM.00.266           C--AGT---TA-A-GT-----CA--C-G-GG-----G--C----G-----AGC-GG--G-AG-GTT-----CCT---A---GGGAA----------CC-GCAG--CAA-AGT---C--C-----AT--T-GG-GGAT-TCA-CCACCACCGCGGC..... 7437
TAL.CM.01.8023          C--A-T-TTT-A-GT-C----CAGT---GG----G-C--T-G---AGC--------AG-GCT---------CTG--G--A-GAGAG----------CC-G---T--CAA-AGC---CG-C-----TA-C--GG-GGAC-TCA-CCACTA---GGGG..... 6987
TAN.UG.x.SIVagmTAN1     -A-GG-CATATG----C---CAT--T-GGTC----G-C--TGAC---T-C-CGC-CA-T--GAA--CC-----A--A--A-GGGAA----------CT-AGA----CA------CAG-G---TCTA--A-GG--T--CTG----TA-AAC-GCAAA..... 7721
VER.DE.x.AGM3           -A-G--CATATG-TG-T---CAT---CG-TCTG-C--T--TGAT---T-C-C-C--TC--GGAA--CC-----A--G--A-GAGAA--G..........-TT-GCA---CA-----C-CGG-A--G--TA--TC-G-GGAGCT-A---TACAA--G-AAG..... 7259
VER.KE.x.9063           -A-GG-CTTATG-TG-T---CAC--C-G-TCTG-C-----TGAT---T-T-CCA--TC---GAA-G-------T--A--A-GGGAA----------TC-GGA---CA----C-CGG-C--T--TA--AC-G--GA-CT-A--TA-AAC---AAC..... 7761
VER.KE.x.AGM155         -G-G--CTTATG-TG-C---CAT--CCG-TCTG-C-----TGAT---T-T-CTA--TC---GAA--CA----T--A--A-GAGAA----------TT-GCAG--CA-G-C-CAG----T--A--AC-G--GAGCT-A-CTA-AA---CCAG..... 7741
VER.KE.x.TY01_patent    -A-GG-CGTATG-TG-T---CAT--C-GGTCTG-A-----TGAT---T-T-CCC--TC---GAA--CT-----T--G--A-GAGAA----------TT-GCA---CAGG--C-CAG-C--T---A---AC-G-TGAGCTTA--TA-AA--G-AAA..... 7737
WRC.CI.97.97CI14        ..TATCCATGGAGT-------C---T-G-A-T--AG--G-GGAC---GCAC-T--G----T-AA-G-A---TTG-----AG-ATCA--GTTCAATAAT--T--------CA-TCA--GA-----C-CAG----T---A--AC-G--TTGTT-AACTGGGAA..... 7738
WRC.CI.98.98CI04        ..TACCCTTGGAGT-------CG--T-G-A-T--AG-GG-TGAT---GCAC-T------GT-AA-G-A---TTG-----AG-ATCA---TTCAGGAAT--T--A------A--CA--GA.............G-GACTGA--TTCTATTCAA-TGGGAA..... 7783
WRC.GM.05.Pbt_05GM_X02  ..TATCCTTGGGCA--T----CT--C-G-TC---AG-GG-TGAT---GCATCTA----G--GAA---A--CTTG-----GG-GTC----TTTAACAAT--C--------A-CCA--GG.............G---C-GA-GT-CTCTT-AA-TGGGAA..... 6679
```

```
                                                   V5 end                                                                        Rev Responsive Element (RRE) region
H1B.FR.83.HXB2               ...AGCAACAATGAGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATAT...AAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAG.....................AG 7747
Env                           .  S  N  N  E  S  E  I  F  R  P  G  G  G  D  M  R  D  N  W  R  S  E  L  Y  K  Y  .  K  V  V  K  I  E  P  L  G  V  A  P  T  K  A  K  R  R  V  V  Q  .  .  .  .  .  .  .  R
H1O.CM.91.MVP5180            TGG-ATTC--CA-GTGAAA-T-CAC-T-----A-T---G----------AA--T-TA-------C-A---G--C--C--C..---------C-G--AA----T--TA-T--------T--A--AATGTCA---CC-A-AA-AA-CATT......CACACCCCTCAC-- 7817
H1N.CM.95.YBF30              ......G---CCA--GAAACA--AG-ATAT--CTC-------A-C---GTTA-TCTC------CAA--G--G-----G--C..--------T-GC--A-----CA-------------AGGT--A--T--A---C-CACA---AGT.......................-- 7292
H1P.CM.06.U14788             ...GA--G--GCA--AA-ACC-CT--A--T---AT---G-----C----CAA-T-TC-----------G------CCC--C..--------TC--G-AA--G-TC-G-CC--G-----T--A--AATTTCC---CCTACAA-TAT-GCT......CATGATGGACAT-- 7296
CPZ.CD.06.BF1167             .................ACT---ACACCCA-CTTCTCT-C-A-GG--GAA---T-C----A-GTA---A-GAG--G---C..--------GG-G--AC-G------TC---G--------T--A-G---A---CC---A--TGGTGACCCA...GAAGCAAAAAGT-- 7878
CPZ.TZ.00.TAN1               ..............AATAGTTATAC-CCC-AATTTTCT-CT-TAG-AGAA---T------A-GTA------GCA-G----..-----G--GG-----C-G--C--GTC---G--T--A-GGCCA-G---A--GCCT-AAA-TA--GCCAAT.........CATACT-- 7373
GOR.CM.07.SIVgorCP2135con    GAG-AT-GT-C---T-A-A-C-C---A-----AAT------G--C----CAA---T------G-----G--G--CCCC--C..-----G-----GG-CA-GG---GTCT--G-----A--A--A----G----CCT--TA-TAGC...............CATAAC-- 7289
                                       V5 end                                                                                     Rev Responsive Element (RRE) start

MAC.US.x.239                 ...............AA-C-A-CTAAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG-----GGGAG-T---..---T-----G-G--CACT---A-T--CT-G--C-----AG-T-TG-----GTACACTACTGGTGGC............ACCTCA-- 8168
Env                          .  .  .  .  .  .  .  N  Q  T  N  I  T  M  S  A  E  V  A  E  L  Y  R  L  E  L  G  D  Y  .  K  L  V  E  I  T  P  I  G  L  A  P  T  D  V  K  R  Y  T  T  G  G  .  .  .  .  .  T  S  R
H2A.SN.86.ST_JSP4_27         ...............AA-C---CAAAT-TTA-CTTTA-T-C---GG--GCA--ACTA-ACC--TTA-----GGGGG-T---..---T-GA--G--G-AAC----A-T--CT-T-----T--ACCA-T---A---TACTCCTCTGCT..............CCAGTG-- 7653
H2B.CI.x.EHO                 ...............AA-CTT-CTAAT-TTA---TGA-T-C---AG--TCA--ACTG-AT-A-TTG---C-GGGGG-C--C..---T-----G-G--AAC----A-T--CT-T-----T--A-GTAT---A---TATTCCTCAGT-...............ACACCG-- 8185
H2G.CI.92.Abt96              .............A-CAA-C-T-CAAAT-TCAT-TTCA-T-C---GG--GCA--ACTG-ATC--TTA-----GGGGG-T---..---T-----G----AACT---A-T--CT-T-------A--T-T-------TATTCCTCTGT-...............ACACCG-A 7584
H2U.CI.07.07IC_TNP3          ...............AATCGA-CTAAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG---C-GGG-G-T--C..---C-----G-G--AAC----A-T--CT-T--T-----A-GT-TG--A---TATTCCTCAGCA..............ACACCG-- 7692
H2U.FR.96.12034              ...............AATC-A-CAAAT-TCT-CATGA-C-C---AG--GCA--ACTG-AT---TTA--G--GGGGG-T---..---T-G---G----AAG----A-T--CT-T--T-----ACCA-TG--A---TATTCCTCAGT-...............ACACCG-- 7707
ASC.UG.10.RT03               ...............AGTCCAT-TAA--TGA-CCC-TC--CCA----C-A----TTG----A--TA--T---AGA-G-----..--G--G--GG----A--T---A-T--CC----------TCCAAT---A---TAT-AACCA-CACCAAGC......AAAAGCGTTGT 7202
ASC.UG.10.RT08               ...............AG-CCAT-TAA--TGA-CCC-TC--CCA----C-A----CTG----A--TA--T---AGA-G-----..--G--G--GG----A--C---A-T--CC----------TCCACT---A--GTAT-AACCA-CACCAAGC......AAAAGCATTGT 7232
ASC.UG.10.RT11               ...............AGTCCAT-TAAT-TGA-CCC-TC--CCA-C--C-A----CTG----A--TA--T---AGA-G-----..--G--G--GG----A--T---A-T--CT----------TCCAATT--A--GTAT-AACCA-CATCAAACACC...AAAAGTGTTGT 7193
COL.CM.x.CGU1                GAC-ATG-TGG-A-TGGGTGTGCT-ATCC-G-A-C-CC-AATTT--A-CAT-CATTAA-T-CTTTG-----GGGA-G---C..---T-G--------GAGGA--ACTACTTAT-TC--A--AG-TAT---A---TC----AATGTAAAT......TGGCATCATGGG-- 7157
COL.UG.10.BWC01              GGGG-A--TG-GACAGGATGTGCT-ATCCTG-A-CTCC-AATTT--A-CAT-CACTGA-T-C-CT-------GGGC-G-----..---T-G-----G--GAG-A--AC--TGTAT--T--T--AG-T-T---A--GCA-----CTATTAAT......TGGCATCATGGG-- 6983
COL.UG.10.BWC07              GGGGAG--T--AACTGGATGTGCA-A-CCGG-A-C--TGAATTGG-A-CAT-CTGCCA-T-C-TTG--G---GGA--T--C..---T-G--G-----AACT--TAC--C-T-T-TT-----TG-TCAC-GA--GG---GAACAGCT..............GTGGTAC- 7166
DEB.CM.04.SIVdeb04CMPF3061   ..................AT-T-GAATTTGA-CCC-TC-TC------T-A--CAGTG------GCA--TC-G-----G--C..---A-CA--G--G-GC-G---A----CT-T-----T--TGCT-TC-GAC-CTAC-A--GA-CAGAG...........TCTGTG-- 7581
DEB.CM.99.CM40               ..................AGT--AAC-CTTA-ACC-TC--C---C--C---CAGTA------GCA--CC---T---------..---A-CA--G--G-AA----TA-----T-C--T---T--GCA-TG-GA--GTAC-A--GA-C-GAG...........TCGGTA-- 7487
DEB.CM.99.CM5                ..................A-C--GAA--T-A-ACCCTCCTC------T-A--CAGT------GGCA--TC----C--G-----..--GA-CA--G--G--A----CA-C--GA-G--C--T--A-GTAT--GA--GTAC-A--GA-C-GAG...........TCTGCA-- 7408
DEN.CD.x.CD1_CMO580407       .............AGAA-AC-T--AACATTGA-GCTCTC--CT---G-----C-AGTC------GC-------GGCA-G--GG..---T-----G-GT-AAC----A-T--CC-T--T--T--TTCA-TG-GAC--TAT-AA-GA-CAGAATCGGCAGAGTCTGTTTCAC- 7804
DRL.DE.11.D3                 ATC-ATG-TTCCA-C-ATC----TAAATTTGTGCC-CA--AT--AG--CAAA-TC-G-TC-C-GCAAT-GG-GCGC-T--C..---T-G--C--GG-G--G---A-T--CT-T--T--A--TG-T-T-C-T---TACAACC---C-GAA...............CAC-- 7507
DRL.x.x.FAO                  ACAG-AG-TTCCA-C-ATC----TAAATTTATACCGCA--AT--AG--CAAA-TC-G-TC-C-GCA-T-GG-GCAC-T--C..---T-G--C--GG-G--G---A-T--CT-T--T--A--TG-T-T-C-T---TACAACC---C-GGA...............CAC-- 7591
GRV.ET.x.GRI_677_gri_1       .............TCTGG-CCA--AAATGTGA-CCT-A-TCCTC-GG-AC-CAG--TA---GCGTAC---C-GGGAG-C---..---T-----G-G--AAC----A-T--CT-T--T--T--AG-T-T--GA---TATACT-GC-CCACA....................-- 7757
GSN.CM.99.CN166              ..................TCT---AA--TGA-CCCGTCG-CC---G--GCT--TGCG----ACTA------G-CC-GG--C..-----G--GG-G--A--T--GC-CTCGA-G--C--A--TCC---T--A--G-A--AACAC-CC..................GCTGT 7684
GSN.CM.99.CN71               ..................TCC---AAT-TGA-CCC-TCG-CC---G-AGCC--TGCA---AGTA---G----CC-GG--C..--G--G--GG-G--A--T---C-TTCGC-G--C--A---TC---TC-A--G---CCACAT-CAGGG............GTGCATGT 7696
LST.CD.88.SIVlhoest447       AACCTGGGAGGAA-T-TGTCTG--AGATTCTTGCCTCCCTCCTGGGAA-CAA-TC-G-TTGTGGCAAC-GG-GC-C-C---..---T--A---GGC--AG----A-T--CT-T--T--A--AG-T-TGC-TC-GTAC-CACCC..................AGTAGT-A 6927
LST.CD.88.SIVlhoest485       AACCTGGGGGGAA-T-TGTCTG--AGATTCTTGCCTCCCTC-TGGGA--CAA-TC-A-TTGTGGCAACGGG-GC-C-T---..---T--A---GG---AG----A-T--CT-T--T--A--AGCT-AGC-CCAGTAC-CACCC..................AGTAGT-A 6924
LST.CD.88.SIVlhoest524       GACTTAGGAGGAA-TCTGTCT--TAAGTTTTTGCCTCCCTC-TGGGA--CAA-TC-A-TTGTGGCAAC-GG-GC-C-T--C..---T-GA-C-G-C-GAG----A-T--CT-T--T--A--ATCA-AGC-TC-GTAC-CACCC..................AGGGGA-- 6930
LST.KE.x.lho7                GACTTAGGAGGCA-CCTAAGT--AAAGTTCTTGCCTCC-TC-TGGGAA-CTA-TC-A-TTGTGGCA---GGG-C---------..---T-GA-----T-GA-T---A-T--GT-T--T--T--AG-T-AGC-C--GTAC-CACCC..................AGGGGA-- 8002
MAC.US.x.251_1A11            ...............AA-C-A-CTAAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG-----GGGAG-T---..---T-----G-G--CACT---A-T--CT-G--C-----A--T-TG-G---GTACACTACTGGTGGC............ACCTCA-- 8164
MND-1.GA.x.MNDGB1            ...-ATG-A-CAA-AATGC-A--TAAATTCTTGCCTCCCA-T--A-CCTCCA-TC-A-TTGTTGC-T-TGG-GC-C-T---..---T----C-----AATG---A-T--CA-------T--AG-T-TG--A---CACACTT-A-CTGAA............CATCAT-A 7357
MND-2.CM.98.CM16             .............-GA-A-ATC---AAATTTGTGCC-CA--AT--GG-ACAAA--CGA-TT-C-GCG-T-GG-GCAC-C---..C--C---------G-G--C---A-T--CT-T--T--T--AG-A-TGGCA---TATCATC---CAGAA...............GCT-- 7900
MND-2.GA.x.M14               ...........--AAGA-A-CTA--AAAGTTT-TGCC-CA--AT--GG-ACAAA-TC-A-TC-C-GCA-TGGG-GCAC-T---..---T-G------G-G--T--GA-----T-C--C--T--TG-T-T-C-C---TATCATC-A-CAGAT...............GCA-A 7829
MND-2.x.x.5440               ..............----GA-ATATC--AAAATTTATACC-CA--ATTGG--TCAAA-TC-A-AC-C-GCA-TGGGGGC-C-T--C..---T-G--G--GG-G--T--GA----TT-T--C-----AG-CAT-C-C---CATCATT---CTAAT...............ACA-- 7533
MNE.US.x.MNE027              ...............AA-C-A-CTAAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG-----GGGAG-T---..---T-----G-G--CACT---A-T--CT-G--C-----A--T-TG-----GTACACTACTGGTGGC............ACCCCA-- 7645
MON.CM.99.L1_99CML1          ..................A-CG-GAC-CTTGT-CCCTCG-C-C-GG----T---TCA------TCA--GC-CAG-CG----..--G--G--GG-------C--C---TCTA-G-----T--A-CA--GC-------GACT-G-GTA..................CAC-- 7710
MON.NG.x.NG1                 ..................ACAG--ACGTTGAT-CC-TC--C-CGAG--TCT---GCA---C-CTCA--G--GAGC--G--C..--G--G--CG-G--C--C--C--GTCT--G-----T--AGCA--CC-A--G---CAT-G--TA..................ATG-- 6335
MUS-1.CM.01.CM1239           ..................ACAG--AA--TGA--CC-TCC-CGA--G---T----GCC---C---TG---C--A-A--T--C..--G-----GG----A--C--GC-TTCCA-G--T-----AG---TGC----G--GCCT--T-CC..................CATGT 7650
MUS-1.CM.01.SIVmus_01CM1085  ..................TCA---AAT-TGA--CC-TC--C-A--G--TT----GCC---C---TG--GCCTA-G--GGTACAA--G--G--GG-G--C--T---C-TTCCA-G--------GG-A-TGC-A-----GCCT--C-CA..................CATGT 7681
MUS-2.CM.01.CM1246           ..................ACA--AAA--TGA-CCCCTC--CCA--G-C......TGGAC-C-TGCGTCTAG-CTG-G-G--...-C---G---G-G--A--C--TC-TTCC--G-----G--AG-A-TGCGC--G---CCT----CC..................CACGT 7779
MUS-2.CM.01.CM2500           ...............A-T-T--ATCGATTGA-CCC-TC--CCA--G--CTT---GC----C-C-TG------A-G-G---C..-----G--GG----A--T---C-GTCC--G-----A--GG-A-TGC-------GCCT----C-..................CATGT 7702
MUS-3.GA.09.09GabOI81        ..................AGT---AAT-TGA-CCCCTC--C-A--G-A-T---TGC-------GCA-----GAG--G---C..-----G--GG-G--A--C--GC-CTCC--G--C--A--G-GA-TTCCA--GC---A--C--CT..................AGACT 7279
MUS-3.GA.11.11GabPts02       ..................TCT---AAT-TGA-ACC-TCG-CGA--G-C-TA--TGCA-----GGCA-----GAGC-G---C..--G--G--GG----C--C--CC--TC-A-G-----T--T-C--TGC----GC---AAACA-CC..................CGTTT 7262
OLC.CI.97.97CI12             AAAGTAG-G--C-GAA-GC-AG-G-CATTG-T-CC-CC-TC---C-ACGT-A-TC-A-TTGTCTCA-TGGC-AG-G-T---..---T-G---G-GT-G---TTGA-T--CT-T--T--T--AG--A--C-A--GTATCA-CCT.........................-- 7377
RCM.CM.00.SIVagi_00CM312     ...............AAT----CTAATGTGA-ACTGTCT-C---AG-AG-A---CTA---GC-GCA-----GGGA-G---C..---------GG-----A-----A-T---TAT-----A--AGCTAT---A--GTAC-A-ACA..........................-A 7221
RCM.CM.02.SIVrcm_02CM8081    ...............AAT--A-CAAATGTTA--CTGACT-C---GG-AG-A--TCTA---GC-GCA--G---GGA-G---C..---------G-----A-----A-T--CTAT-----T--AG-T-TT--A--GTAC-A-ACA..........................-- 7299
RCM.GA.x.SIVRCMGAB1          ...............AGT----CAAA--TCA-CCTCTC--CG--AG--G-T--TT-----GC-GCA-----GGGG-G-----..--G-C-A--G----CAG----A-T--CT-T-----A--AG-AAT---A--GTACCA-AC-..........................-A 7437
RCM.NG.x.NG411               ...............ATGACA-CAAATGTCA-CTTGTCT-CT--AG-AG-A--T-TA---GC-GCA-----GGGC-G---C..---------GG-G---ATG---A-T---TAT-----A--A--T-TC-GA--GTAC-A-ACA..........................-A 7481
SAB.SN.x.SAB1                ...............AA-AG--C-AATGTGA-GTT-TC-CCTC-G--TGAAAG--T----GCT-A------GGGAG-T--C..---T----GG----CA-G---A-T---T-T--T-A-----A-T--GA---TACACT-GC.................CCCGAA-- 8051
SMM.CI.79.SIVsmCI2           ..............-GAAAT-GC-CAAAT-TCA-CTTTA-T-CC--GG--GCA--GTTG-AC---TTG--G--GGGGG-C--C..---T-G---G-----AC----A-T--CT-T---------A--T-TG--A---TACTCCTC-GCAACG............CCTGCC-- 7321
SMM.LR.89.SIVsmLIB1          .............A-TAAT----CTAAT-TCA-CATGA-T-C---G--TGCA--ACTG-ATC--TTG-----GGGAG-T--C..---T----TG-G---ACT---A-T---T-G--C--A--A--T-TG--A--GTACTCTACC.....................ATG-- 7230
SMM.SL.92.SIVsmSL92A         .............--AAAT----C-AAT-TCA-CATGA-T-C---AG--GCA--ACTG-ATC--TTG-----GGGAG-T--C..---T-G---G---CACT---A-T--TT-G--C--T--A--T-TG--A--GTACACCACCAGTACC...............TCA-- 7265
SMM.SL.92.SIVsmSL92B         .............A-CAAT----C-AAT-TTA-CATGA-T-C---GG--GCA--ATTG-ATC--TTG-----GGGGG-T---..---T-G---G-G--CACT---A-T--CA-G--T-----AC-T-TG--A--GTACACCACAAGT..............ACCTCA-A 7236
SMM.SL.92.SL92B              ...............AAT----C-AAT-TTA-CATGA-T-C---AG--GCA--ATTG-ATC--TTG-----GGGGG-T---..---T-G---G-G--CACT---A-T--CA-G--T-----AC-T-TG--A--GTACACCACAAGT..............ACCTCA-A 7616
SMM.US.04.G078               .............A-TAA-C---CTAA--TTA-CATGA-T-C---GG--GCA--ACTG-AT---TTG-----GGGGG-C--C..---T-----G-----ACT---A-T--CA-G--C--A--A-GT-TG--A--GTACACTACTGGTACC...............TCA-- 7461
SMM.US.05.D215               .............A-CAAT----CTAAT-TCA-CATGA-T-CG--CG--GCA--ACTG-ACC--TTG-----GGGGG-T--C..---T-----G----CACT---A-T--CT-G--T--------GT-TG-----GTATACCACA---.............ACCTCA-- 7437
SMM.US.06.FTq                .............-GAAA-AGA-CAAAT-TCA-CTTCA-T-C---GG--GCA--ACTG-AC---TTG--G--GGGGG-T--C..---T-----G-----ACT---A-T--CT-T-----T---G-T-T---A---TATTCCTCTGTAACG............ACACCG-- 7453
SMM.US.86.CFU212             .............A-AAA-C-A-CTAAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG-----GGGGG-T--C..---T-----G----CACT---A-T--CA----C--A--A--T-T------GTACACCACTACTACT............ACCTCA-- 7435
SMM.US.x.H9                  ...............AATA---C-AAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG--G--GGGAG-T--C..---T-----G-G---ACT---A-T--CT-C--C-----A-GT-T---A--GTACACCACAACTGGT............GCCTCA-- 7657
SMM.US.x.PGM53               ...............AA-C---CTAA--TCA-CATGA-T-C---GG--GCA--ACTG-AC---TTG--G--GGGAG-T---..---T-----G-G---ACT---A-T--CT-G--C-----A-GT-T--G---GTACACAACTGGTGCC...............TCA-- 8114
SMM.US.x.pE660.CG7G          .............A-CAAT----C-AAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG--G--GGGAG-T--C..---T-----G-G---ACT---A-T--CT-G--C-----A-GT-T--GA--GTACACCACAACT............GGTGCCTCA-- 8173
STM.US.89.STM_37_16          ...............AAT----CTAA--TCA---C-A-T-C---GG--GCA--ACTG-ATC--TTG-----GGGAG-T--C..---T-----G----AACT---A-T--CT-G--C--A--G--T-TG------TACACAACTAGT..............ACCTCA-- 7819
SUN.GA.98.L14                AATG-ATCA--CATC--A-TTG-AGC-TTC-AGCC-CCTACCA-C-CCTTAA-TC-A-TTGTG--CAC-GG-GCAC-T--C..---C------G----AGG---A-T--GT-T--T--A--AG-T-AGC-CC-GTAC-CACCC....................AGG-A 8070
SYK.KE.x.KE51                ...............--AC--T-GAATTTTG-ACCTACT-CTA--G--GAA--TGTG------GC---T----TC--T-------..---T-GA-TCGG---A----TA-T--CT-T--G--T--A-GTCA--G-C-CTAT-A-C---CT...................ACG-A 7407
SYK.KE.x.SYK173_COMGNM       ...............--TAC-T-GAA-TTTA--CC-AC--C-A--G--GA---TGT-------GCA--TC-G-TC--T-------..---T-GA-TC-----A-----A-T--CT-T--G-----TG-TCA-CGTC-GTAT-A-T---CT...............AATACC-- 7714
TAL.CM.00.266                ............TCTAGT--C--AACAGT-GAGCCCTCGTC-C-G--AC-T---TCG---C--GCA-C-G-T-CACGG--C..--G--GA--G-G--C-TG--T--T--GTAT--G--G--A--TATC-GA--GTAC-AT-G--CC...............GCTAAC-A 7577
TAL.CM.01.8023               ............A-TGA-ATC--AAGTGTGGAACCCTCTTC-C-A--C--A--TGC----C--GCA-CGG--GCACGG--C..--G--GA--G-G--CAC----A-T--CT-T-----A--G--TATC-GA--GTAT-AT-G--C-..............GCAACA-A 7127
TAN.UG.x.SIVagmTAN1          ...............GAAAG--CTAATGTGA-ATTGAC--CTA--T--GA-A---TA---GCTTA------GGGA-G-----..--GC-CA--G-----------A-C--TT-C--G--A--AG--AT--GA---TAT--T-G--CAACT..................C- 7855
VER.DE.x.AGM3                ...............AA-AG--CTAATGT-A-ATT-A-TCCCC-G--AGAAAC--TC---GC-GCA-----GGGC-GG--C..---T-----G-----AC----A-T--CT-C--------AG-A-T--GA--GTATAC--GAGGT..............CATGAC-- 7396
VER.KE.x.9063                ...............-ATAGA-C-AATGTGA-ATT-A-TCCCC-G--AGAA-GA-TC---GC-GCA-----GGGACG-----..---T-G---G---CAC----A-C---T-T--------G--T-T--GAC-TTATAC--GAGGA..............CAAGAC-- 7898
VER.KE.x.AGM155              ...............AA-AG--CAAATGT-A-ATTGA-TCCCC-G--AGAAAC--TC---GCGGCA-----GGGC-G---C..---T-G---G-G---AC----A-T---T-T--------AG-A-TC-G-C--TACAC--GAGGC..............CAAGAG-- 7878
VER.KE.x.TYO1_patent         ...............AA-AGA-C-AA-GTGA-ACT-A-TCCCC-G--AGAATCT-TC---GCGGC------GGGC-G---C..---T----GG---CAC----A-T--CT-T--------AG-A-T--G-C-TTATAC--GAGGA..............CATGAG-- 7874
WRC.CI.97.97CI14             ............AGAAATC-AG-GA-GTTC-AA-CTCCTTCCA-C--ACA-A--TC--TTGTGGCA-TGGG-C-A-------..---T-G------T-AA--ATGA----GA-G-----AGGGG-T-G------GGTT--ACCC-CTGAA............CATCGC-- 7881
WRC.CI.98.98CI04             ............A-TAAGC-AG-GGAATTC-AA-CTCCTTCCA-C--ACA-A-TTCA-TTGTGGCA-TGGG-C-A--G---..---C-G--G-G-C-AA--ATGA----GC-G--T--TGGTG-T-GG-GA---GTGACACCAA-AGAA............CATCAT-- 7926
WRC.GM.05.Pbt_05GM_X02       ............A-CGGGC--G-GGAGTTC-AA-CTCCTTCCC-C---CAAA--TCC-TTGTGGCA-T-GG-C-A------..---C-G------C-AA--TTGA-----------T-----TG-T-G------GGTGACACCT-CTGAC............CATCAT-- 6822
```

```
                              Env gp120 end Env gp41 start
H1B.FR.83.HXB2                AGAAAAAAGAGCAGTG......GGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTG...AGGGCTATTGAGGCGC 7908
Env                            E K R A V . . G I G A L F L G F L G A A G S T M G A A S M T L T V Q A R Q L L S G I V Q Q Q N N L L . R A I E A
H1O.CM.91.MVP5180             G-------------A.......--T-G---ATGC-A---T-G---G-GC-AA-T--------T----------------GG---CAG-----------G-A--CACAGTG--C--AAG-----------A------G---CC-----...--A--G--AC----C- 7978
H1N.CM.95.YBF30               --------------CC...TTT---C----T--GC----T--------TC-T-----------G----------------G-----A------------------C-GAC------A-----G---------A--------T-T-C--T-...--A--A--A------- 7456
H1P.CM.06.U14788              GA-G-----G----CA......--CC-----ATGC-A-----------TC--A-C----------------------------GGG---C-------------------AG----------CA-------------------------A----A...--A-----A-----A- 7457
CPZ.CD.06.BF1167              -AGT--------A-......ATA-----G---ATGA-C--TA-G--A--TC-CA-T-----------T--------------G-----AG------------A------GG---GC-C-----------A--------AGCA-----------...CA---A--A-----T- 8039
CPZ.TZ.00.TAN1                GTC--G-----AT---.......--C------CTG-----T-----A--TC-TA-T-----------T--A-----------G-----AG--------C---------GG----C-C--------T--A--------AC-A--CC----T...CA---C--A--A--- 7534
GOR.CM.07.SIVgorCP2135con     G------C-T----CA......--CC-G---ATGC----T----------A--A-T--------------G-----------GG--G-A-------------------G---G-C---CA------C-----------AM----CA-----...--A-----A-CA--T- 7450
                              Env gp120 end Env gp41 start

MAC.US.x.239                  -A-T-------GG--C...TTT-TGC----G......---T-G--T--TC-C-C-A-G-----TTCTG-A--------G--G--GT----------C-CT---T--C--ACT------G----G-----------A---C-AC-GC--T--...GAC-TGG-CA--AGA- 8326
Env                            N K R G V . F V L G . . F L G F L A T A G S A M G A A S L T L T A Q S R T L L A G I V Q Q Q Q Q L L . D V V K R
H2A.SN.86.ST_JSP4_27          GA-T-------GT--A...TTC-TGC----G.......---T-A--T--TC-CACGA--------GCTG-A--------G--G--CT--------T---CT---T-TC-GACT------G-C--G-----------A---C-AC-GC--T--...GAC-TGG-CA--AGA- 7811
H2B.CI.x.EHO                  GA-T-------GT--A...CTT-TGC----G.......--TT-G--A---C-T-CGA-G-----TTCTG-A--------G--G--CT-----T--T---CT---T--C-GACT---C--G----G---------------C-AC-GC-CG--...GAC-TGG-CA-AAGA- 8343
H2G.CI.92.Abt96               -A-T-------GT--A...TTT-TGC----G.......--TT-G--A--TC-T-C-ATG-----TTCTG-A------A----G--GT--------T---CT---T--C-GACT--G---G----G-----------A---C-AC-GC--T--...GAT-TGG-CA--AGA- 7742
H2U.CI.07.07IC_TNP3           GA-T-----G-G-----...TTT-TGC-T--G.......---T-G--T--TC-C-CTA-G-----T-CTG-A------A-G--G--TT--------T-T-CT---T-TC-GACT------G----G-----------A--AC-AC-G---T--...GAC-TGG-CA-AAGA- 7850
H2U.FR.96.12034               GA-T-----G-G---A...TTT-TGC-T--G.......---T-G--T--TC-C-CCA-G-----TTCTG-A-----T--G--G--GT--------T-A-CT---T--C-GACT---C--G----G-----------A---C-AC-GC--T--...GAC-TGG-CA-AAGA- 7865
ASC.UG.10.RT03                CA---G-GC---C-CA...TTGACTC-G---.......---T-G--A--TC-CA-TA-G-----GGCG--A------T-G-TGG---CCG-CT----A--C---T-GC-TTCCC-C--AGAA--C-----AGCC-----A--GGCGC-A--T...GA---AG-----CAC- 7360
ASC.UG.10.RT08                CA---G-GC---C-CA...TTGACCC-G---.......---T-G--A--TC-CA-TA-G-----GGCG--A------T-G-TAG-G-CCG-CT----A--C---T-GC-TTCCC-C--AGAA--C-----AGCC-----A--GGCGC-A--T...GA---AG-----CAC- 7390
ASC.UG.10.RT11                CA---G-GC---C-CA...TTGACCC-G---.......---T-G--A--TC-CA-TA-G-----GGCG--A------T-G-TGG---CCG-CT----A--C---T-GC-TTCCC-C--AGAA--C-----AGCC-----A--GGCGC-A--T...GA---AG-----CAC- 7351
COL.CM.x.CGU1                 -C----G----GCA-C...TTT-C-T-CTCC......A--T-A-CCC-----A-T-GG--C--TGCTG-A------T----G--TG--G--------A-C-----TCA-TCCC-TAATGGAA-GGCTAGTGCCAGCAGC---CGAA--T--CTG-AACTGG-C---A--- 7318
COL.UG.10.BWC01               GC---------GCA-T...TTT-C-T-CTCC......A--T-A-CTC----AA-T-GG--C--TGCTG-A------T----G---G--G--T-----A-C-----TCA-TCC------GAA--GC-----------A---CGG-TGC--T--...-AACTGG-C--AA--- 7141
COL.UG.10.BWC07               ---G---C-G-G-A-C...TTT-CTT-CTCC......--GA-A-CTC-----A-T-GG--C--CGCTG-A------T----G--TG--G--------A-C-----TCAGTCC------GA---GC------G-------CGGC-GC-------...-A-CTGG----AA-A- 7324
DEB.CM.04.SIVdeb04CMPF3061    -C--------------GGAATG-CTT----T......C-AG-G-CA--TC-CA-CA-------TGCTG----------G--G--C-CAG--------------T-TC-CTCT--GC--AGC--G---------------AC-AG-G-------...-A----G-------T- 7742
DEB.CM.99.CM40                -C-C--G--G----CAGGAATT-CTT-T--T......C-GG-A-CA--TC-CA-TA-------TGCTG------------G----CAG--------------T-TC-CTCT--GC--AG---G-----------A---C-AG-GC--T--...-A----G-G-----T- 7648
DEB.CM.99.CM5                 -C-------G--G-C-GGAATT-CTT-T--T......C-AG-G-CT--TC--A-TA-------TGCTG------------G----CAG--------------T--C-CTCT--GC--AG---G-----------A---C-AG-GC--T--...-A----G-G-----T- 7569
DEN.CD.x.CD1_CMO580407        CCG--GGC-G-A----TCTTTA-TGC-T--G......C--A-A--A--TC-CA-C-----------CTGTC--------G-T-GG--CAGGCT-AG--------T--C-GG-T---C-TG-C--G-----------A---C-AG-GC--T--...--A----G-----A-GT- 7965
DRL.DE.11.D3                  GC-G--G----GT-CA...GTC-TGC-T--G......A-------CC-----A-CTTG-----TTC---A------T-G-TGG-G-CAG-CA----T--C---T-TCAGGCT--G---------G------G-------A--GC-AC------...C-CCTGG-G--ACA-- 7665
DRL.x.x.FAO                   GC-G--G----GT-CA...GTC-TGC-T--G......A-------CC-----A-CTTG-----TTC-G-A---T--T---TGG-G-C-G-CA----T--C---T--CAGGCT--GC--A----G--G---G-------A--GC-AC------...C-CCTGG-G--ACA-- 7749
GRV.ET.x.GRI_677_gri_1        ---------G-TGCCA...TTC-TGC----G......--T--A--C------------T--T-----CTG-A------------GG---CA---------A--C---T-TC-G--T--GC-TG----G---T-------------G--C-------...GC----G-G--ACA-- 7915
GSN.CM.99.CN166               G--G--------C......CTTTC-T-G---ATA-CC---T-GAC---TC-CA-T--G-----T-C---G------------GG---CAG---------A----T--C-GTCT--G---G----G---A-----------G-G--CC------...-------G-A-----C- 7845
GSN.CM.99.CN71                GA----G--G--C......ATTTC-T-G--CATA-CC---T-GAC---TC-CA-T--G-----TGC---G------------GG---CAG---------------AT--C-GTC---G---G----G---------------G-G--CC------...-------G-A-----C- 7857
LST.CD.88.SIVlhoest447        GC-G--G-------CACCGCTG-CTC----T......GC-T-G--AC-TC-CA-T--T-----T-CTG-A------TT--TA--G-C--TA--A--T--C--A---CA-GT-G--A-ACAA--G---T-------------GC-G-------...GT-CT-G-----AAA- 7088
LST.CD.88.SIVlhoest485        GC-G--G-------CACCGCTG-CTC----T......GC-T-G--AC-TC-CA-T--T-----T-CTG-A------TT--TA--G-C--TA--A--T--C------CA-GT-G--A-ACAA--G---T-------------GC-G-------...GT-CT-G-----AAA- 7085
LST.CD.88.SIVlhoest524        GC---C--------CTCCACTA-CTC----T......GC-T-G--AC-TC-CA-T--T-----T-CTG-A------TT--TA--G-C--TA--A--T--C------CA-GT-G--A-ACAA--G---T-------------GC-G-------...GT-CT-G-----AAA- 7091
LST.KE.x.lho7                 GC---CC-------C-CCTCTG-CTT-----......GC-T-G--AC-TC-CA-T--T-----C-CTG-A------TT--TA--G-C--TA--A--T--C------CA-GC-G-C---CAA--G---T-------------GC-G-------...GT-CT-G-----AAA- 8163
MAC.US.x.251_1A11             -A-T-------GG--C...TTT-TGC----G......---T-G--T--TC-C-C-A-G-----TTCTG-A--------G--G--GT----------C-CT---T--C-GACT------G----G-----------A---C-ACTGC--T--...GAC-TGG-CA--AGA- 8322
MND-1.GA.x.MNDGB1             ---G--G----G--CA...GTAAT-C-T--T......A-------TC-GC-CTCGCTG------TC-G-G------T---TG--GG--G-A-----T--C--AT-TCAGTCT--GG--A-----G-----G-A--A--A--AC-G---T--...-A-CTC--A---CAA- 7515
MND-2.CM.98.CM16              GC-G--G--G-G--C-...GTCTTGC-T--G......A-GT-C--CC--C-A--TTTG-----TTC---G------T---TGG-GG--G-C-----T--C---T-TCAGGCT--G---AA---G--T---G----------GGT-C------...--CCTG--A--TCA-- 8058
MND-2.GA.x.M14                GC-G--G----G--CA...GTCTTGC-T---......A-G--C--CC-------CTTG-----TTC-G-G------T---TGG-GG--G-C-----T--C---T--CAGGCT--G---AA---G--T---G----------G-T-C------...--CCTG--A--TCA-- 7987
MND-2.x.x.5440                GC---------G--C-...GTCTT-C-T---......A-G--C--CC--C-A--TTTG-----TTC-G-G------T---TGG-GG--G-C-----T--C---T--CAGACT---C--AA---G--T---G----------GGT-C------...--CCTG--A--TCA-- 7691
MNE.US.x.MNE027               -A-T-------GG--C...TTT-TGC----G......---T-A--T--TC-C-C-A-G-----TTCTG-A--------G--G--GT----------C-CT---T--C-GACT------G----G-----------A---C-AC-GC--T--...GAC-TGG-CA--AGA- 7803
MON.CM.99.L1_99CML1           G-----G------A-T......ACTT-G--TATGGCC--T--G--C---C-CA-CA-G------G----A-----------GG-G-CAG-C------------T-TC-GTC----C--G----G---------------G-G--CC------...------CG-AACA--C- 7871
MON.NG.x.NG1                  -----G-C----TA-C......AC-T-G---ATGGCC--T--G--C--TC-CA-CA-G------G----G-----------GG-G-CAG-C------------T--C-GTCT--G---G----G---------------G-G--CC------...------CG-AACT--C- 6496
MUS-1.CM.01.CM1239            -A--GC-------A-A......TCTC----CATAACC--T--G--A--TC-CA-C--G------G----G-----------GG-G-CAG-C------------T-TC-GTC----C--G-C--G---------------GCA---C------...------CG-A--A--C- 7811
MUS-1.CM.01.SIVmus_01CM1085   GC-TGC------CA-C......TC-T-G--CATAACC--TT-A--A--TC-CA-C----------G----A-----------GG-G-CAG-C------------T-TC-GTCT--G---G-C--G---------------GCA---C------...--A--CG-A--A--C- 7842
MUS-2.CM.01.CM1246            CA-GGCT--G---A-C......A-CT-G---ATAACC---T-A--A---C-CA-T--T--G---G----A-----------GG-G-CAG-C------------T-TC-GTCT--G---G----G-----------A---GCT---C--T--...--A--CG-A--A--C- 7940
MUS-2.CM.01.CM2500            CC--GCT--G---A-A......ACCC-----ATAACC--TT-G--A--TC-CA-T--T------G----A-----------GG-G-CAG--------------T--C-GTCT--G---G----G-----------A---GCT---C--T--...-----CG-A--AA-C- 7863
MUS-3.GA.09.09GabOI81         GC-T-CCC-T--GA-A......TC-C-G--GATAAC---T--CAC---TC-CTC---T--T---G----G------T----GG-G--AG-C------------T-TC-GTC-------G----G---------------C-A--CC--T--...C-A--AG-AC--A--- 7440
MUS-3.GA.11.11GabPts02        GC-C-C------CA-A......TC-C----GATAACC--T--CAC---TC-CTC---G------G----G------T----GG----AG--------------T-TC-GTCT--G---G----G-------A------C-G---C------...-----CG-GC--A-C- 7423
OLC.CI.97.97CI12              -A-T-------G--CA...TTG-TGC-G--G.......--AT-G---G-TC-C--CCT-------TC-G-A------A---TA----C--TA-----T--C---T-TTCG--T--GC-TG-G--G-------G---------GC-CA------...GA-ATG------AA-- 7535
RCM.CM.00.SIVagi_00CM312      -C-G-------T-CCT...TTG-TGC-G--T.......--T--A--T---C-CTC---------T-CTG-A-----------GG-G-CAG-C-----A--C---T--C-G--T--GC-TG----G---T-------------GC-G-------...GACATAG-----CA-- 7379
RCM.CM.02.SIVrcm_02CM8081     -C-G-------T-CCT...TTG-TGC-G--T.......--T--A--T---C-CTCG--------T-CTG-A-----------GG-G-CAG-C-----A--C---T--C-G--T--GC-TG----G---T-------------G--G-------...GACATAG-----CA-- 7457
RCM.GA.x.SIVRCMGAB1           -C-G-----G-T-CCT...TTG-TGC-G--T.......--T--A--T---C-CTC---------T-CTG-A-----------GG-G-CAG-C-----T--C---T--C-G--T--GC-TG-A--G---T----------A--G--CC------...GACATAG--A--CG-- 7595
RCM.NG.x.NG411                -C----G----T-CCT...TTG-TGC-G--T.......--T--A--T---C-CTCG--------T-CTG-A-----------GG-G-CAG-C-----A--C---T--C-G--T--GC-TG----G---T-------------G--G-------...GACATAG----ACA-- 7639
SAB.SN.x.SAB1                 -C-------G-TGCCA...TTC-TGC----G......-----A--T---------------T---GCTG-A-----T-----GG-G-CAG-C---------C---T-TCAG--G---C-CG----G---C-------------A---C----T...GC----G-A---CA-- 8209
SMM.CI.79.SIVsmCI2            GA-T-------GT--A...TTT-TGC----G......---T-G--A--TC-T-CGA-G-----TTCTG-A-----------G--CT--------T-T-CT---T-TC-GACT------G----G-----------A---C-AC-GC--T--...GAC-TGG-CAC-AGA- 7479
SMM.LR.89.SIVsmLIB1           -CC--------GG--C...TTT-TGC----G......---T-G--A--TC-C-C-A-------TTCTG-A-----------G--CT------------CT---T--C-GACT--G---G----G-----------A---C-AC-GC--T--...GAC-TGG-CA--AGA- 7388
SMM.SL.92.SIVsmSL92A          -A-T--G----GG--C...TTT-TGC-G--G......---T-G--T--TC-T-CGA-G-----TTCTG-A-----------G--GT----------T-CT---T--C-GACT------G----G-----------A---CGAC-GC--T--...GAC-TGG-CA--AGA- 7423
SMM.SL.92.SIVsmSL92B          GA-T-------G---C...TTT-TGC----G......---T-G--T--TC-C-CGA-G-----TTCTG-A-----------G--GC-A---------T-CT---T--C-GACT------G----G---------A--A---C-AC-GC--T--...GAC-TGG-CA--AGA- 7394
SMM.SL.92.SL92B               GA-T-------G---C...TTT-TGC----G......---T-G--T--TC-C-CGA-G-----TTCTG-A-----------G--GC-A---------T-CT---T--C-GACT------G----G---------A--A---C-AC-GC--T--...GAC-TGG-CA--AGA- 7774
SMM.US.04.G078                -A-T-------GG--C...TTT-TGC----G......---T-G--T--TC-T-CGA-G-----TTCTG-A-----------G--CT--------T---CT---T--C-GACT--G---G----G-----------A---C-AC-GC--T--...GAT-TGG-CA--AGA- 7619
SMM.US.05.D215                -C---G-----G---C...TTT-TGC----G......---T-G--A--TC-C-C-A-------TTCTG-A--------G--G--CT--------T---CT---T--C-GACT------G----G-----------A---C-AC-GC--T--...GAC-TGG-CA--AGA- 7595
SMM.US.06.FTq                 GA-T-------GT----...TTC-TGC-T--G......---T-G--A--TC-C-C-A-G-----TTCTG-A-----------G--GT--------T-C-CT---T--C-GACT--G---G----G-----------A---C-AC-GC--T--...GAC-TGG-CA-AAGA- 7611
SMM.US.86.CFU212              --CT-------GG--C...TTT-TGC----G......---T-G--A--TC-C-C-A-G-----TTCTG-A--------G--G--GT----------T-CT---T-TC-GACT---C--G----G-----------A---C-AC-GC--T--...GAC-TGG-CA--AGA- 7593
SMM.US.x.H9                   -A-T-------GG--C...TTT-TGC----G......---T-G--T--TC-C-CGA-------TTCTG-A--------G--G--GT--------T---CT---T--C-GACT--G---G----G-----------A---C-AC-GC--T--...GAT-TGG-CA--AGA- 7815
SMM.US.x.PGM53                -A-T--G----GG--C...TTT-TGC----G......---T-G--T--TC-C-CGA-------TTCTG-A--------G--G--CG----T-G-T-C-CT---T--C-GACT--G---G----G-----------A---C-AC-GC--T--...GAT-TGG-CA-AAGA- 8272
SMM.US.x.pE660.CG7G           -A-T--G----GG--C...TTT-TGC----G......---T-G--T--T--C-CGA-------TTCTG-A--------G--G--GC--------T---CT---T--C-GACT--G---G----G-----------A---C-AC-GC--T--...GAT-TGG-CA--AGA- 8331
STM.US.89.STM_37_16           -ACT--G----GG--C...TTC-TGC----G......---T-G--T--TC-C-CGA-------TTCTG-A-----------G--GC------------CT---T--C-GACT------A----G-----------A---C-AC-GC--T--...GAC-TGG-CA--AGA- 7977
SUN.GA.98.L14                 G-----G--G----CACCTGTT-CCC-----......GCTT-G-CAC-TC-CA-T--T--G--T-CTG-A------TT--TA--G-C--TA--A--T--C------CA-GC-G-C---CAA--G---T-------------GC-G-------...GT-CT-G-----AA-- 8231
SYK.KE.x.KE51                 -C-G-----------CACCATTG-C-C-T--G......-----G--A--AC-CTCG-----G--T-CTG-A-----GAGC--TG-G-CAG----T--AC-C---T--CAGACT--G---G----G---------------C-A--G-------...GA----G-A-----A- 7568
SYK.KE.x.SYK173_COMGNM        G---------G----CACCATTG-C-C-T--G......-----G--A--AC-CTCG--T--G--C-CTG-A-----G-GC--TG-G-CAG----A--AC-C---T-TCAGACT--G---G----G---------------C-A--G-------...GA---AG-------A- 7875
TAL.CM.00.266                 GC-----C----G-CCCCGCTA-CCT----C......---A-A--A--TC-CA-CA-------C-CTG-C----------TGG-G-CAG-C---------G---T-TC-CTCT--GC--AG---G---------------G-AC-CC------...-----A------CAT- 7738
TAL.CM.01.8023                ---G--------C-C-CCTCTG-CTT-G---......---A-A--A--TC-CA-CA-------C-CTG-C----------TGG-G-CAG-C---------G---T-TC-TTCT--GC--AGC--G---------------G-GC-CC------...-----A------CAT- 7288
TAN.UG.x.SIVagmTAN1           ------G--G-TGCCC...TTC-TGT-G--G......-----A--C--T------G-----T---GCTG-A-----T---A-AG-G-CAG--T--------C---T--CAG------C-TG-A--G---T-------------G---C------...GC---AG-----CA-- 8013
VER.DE.x.AGM3                 -AC---GC---TCCC-...TTC-TGC----G......-----A--C-----A-----T--T--G-CTG-A-----A-----GG-G-CAG-C---------C---T-TCAG--T---C-TG----G---C-------------G---C------...GC----G-G-----T- 7554
VER.KE.x.9063                 -TC----C---TCCC-...TTT-TGC----G......-----A--C-----A-----T--T----CTG-A-----A-----GG-G-CAG-C---------C---T-TCAG--T---C-TG----G---A-------------A---C------...GC----G-G-----T- 8056
VER.KE.x.AGM155               GC-----C---TCCC-...TTC-TGC----G......-----A--C-----------T--T--G-CTG-A-----A-----GG-G-CAG-C---------C---T-TCAG--T---C-TG----G---T-------------G---C------...GC----G-G-GA--T- 8036
VER.KE.x.TYO1_patent          GC----G----TCCC-...TTT-TGC----G......---T-A--C--------G--T--T----CTG-A-----A-----GG-G-CAG-C---------C---T-TCAG--T--GC-TG----G---C-------------G---C------...GC----G-G-----T- 8032
WRC.CI.97.97CI14              ---G-------GT-CC...ATTATTC-----......A-T--A--A---C-C--GCTG------TCTG-A------A---TA----CAGTA------A-C---T-TCAG--T--GC-TG----G-------------A--G--CC------...GA-CT-G-A---AA-- 8039
WRC.CI.98.98CI04              -----------G--CT...ATT-TCT-----......A-T--A--A---C-C--GCTG------TCTG-A------A---TA----CAGTA------A-C---T-TCAG--T--GC-TG----G---T----------A--G--CC------...GA-CT-G-----AA-- 8084
WRC.GM.05.Pbt_05GM_X02        -------C---G--CA...ATT-TTT-----......C-T--A--C---C-C--CCTG------TCTG-A------A-G-TA----CAGTA------A-C---T-TCAG--T--GC-TG----G-----------------G--C-------...GA-CT-G-A---AA-- 6980
```

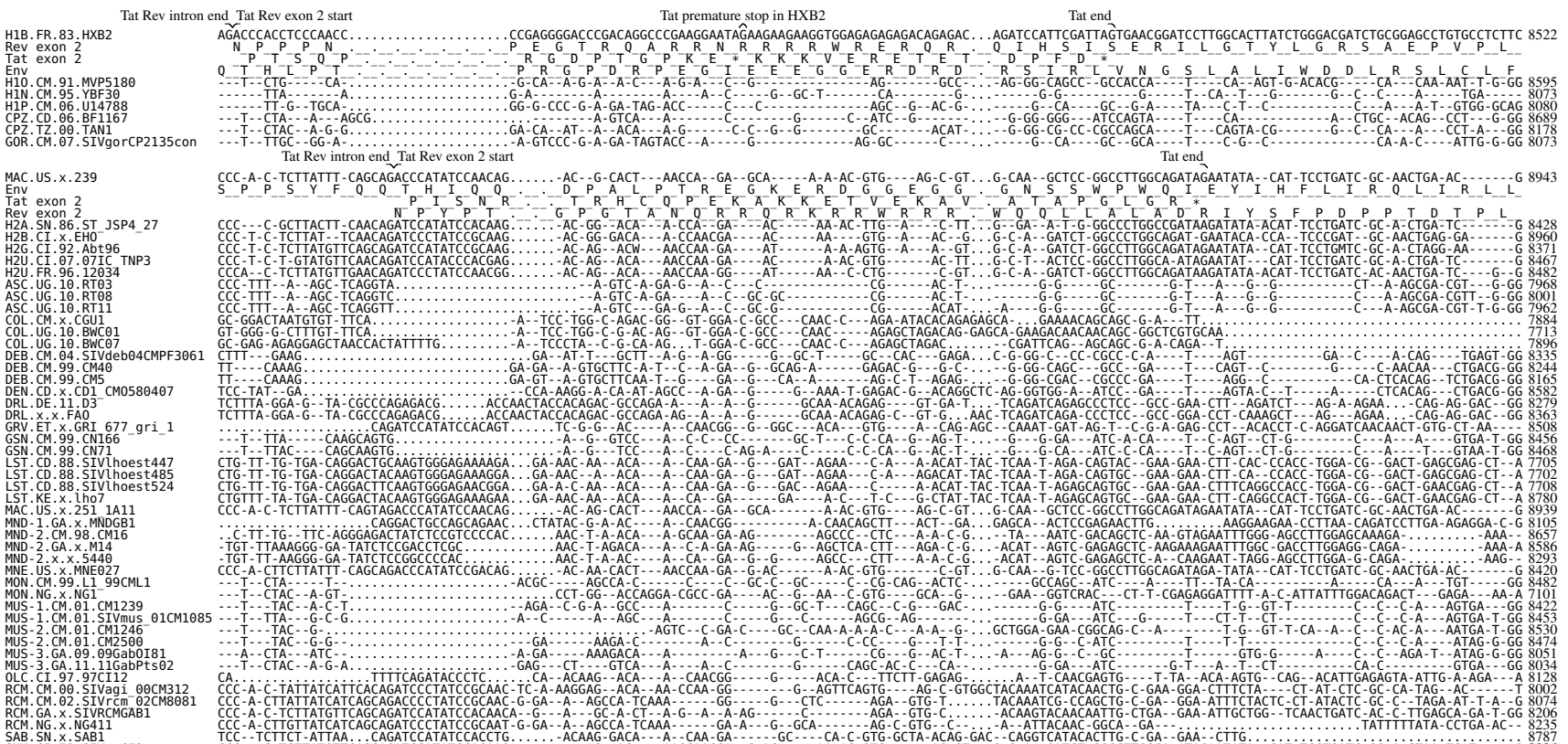
```
                                                                                                                                                                                          RRE end
H1B.FR.83.HXB2              AACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCT 8078
Env                         Q  Q  H  L  L  Q  L  T  V  W  G  I  K  Q  L  Q  A  R  I  L  A  V  E  R  Y  L  K  D  Q  Q  L  L  G  I  W  G  C  S  G  K  L  I  C  T  T  A  V  P  W  N  A  S  W  S  N  K  S
H1O.CM.91.MVP5180           -G--A--CT--C--AGGT-AT-T--A-----T--T-GA--A---GA--TC-CC-G-AA--CT-A----CCCTTA--C--A---G--A-G---AAACC-A-----C--TAAA--------A--C--TTA---AT-A--AAAA-----CA-ATCA---TCAGGA-G-... 8145
H1N.CM.95.YBF30             ----A---T-----------T--A---------T--A-------------A-G----T---A-A-----------T-G------G--AA----AA-TC-A-----C-----A------ACA--A---TAT--CA---------------AG-C------C--C--TA-C 7626
H1P.CM.06.U14788            -G--AG-AT--C--AG----T-T--G--------A-GA-------GA--TC-CC-G--T--CA-T----CC--TT---G------G--A-----A--AC-G-----A-----A---C--A-AR-A--TTAT---AA------A------AG--CC----CC--C---AA- 7627
CPZ.CD.06.BF1167            ----A---T-------G---T-G--G--------A--A--------A--C---C-G--A--A-----G-A--------GA-----------------TC-A-----A--TG--AAC---T-G--C--TCAT-G-T-AT-AGAA------AGCTC----TTA--C---A-C 8209
CPZ.TZ.00.TAN1              ----A--CT-------G---T-T--A--------T--------------C-----G--T--A--A--G-A---A---GA--C-----------AA-CC-C-----A--TG--AAC---T-GG-G--TCA--G-AG------A-----CCTC-CC---GC-G-AG-T--- 7704
GOR.CM.07.SIVgorCP2135con   -G--AG-AT--C--AG----T-T--T--------A-GA-------GA--TC-CC-G--A--CA-A----CT-------GA--C--G--------A--AC-G-----A---A-A--G-G-T-A-----TTA----AA------A-----CAAG-CC----CCGGG---... 7617
                                                                                                                                                                                          RRE end
MAC.US.x.239                ----AG-AT------G---G--C---------A-CA---A-C------A-T--GG--ACT--CA-C--G-AG---T-------C--GGC--AG---AATGC------A--TG-GTTT-G--AAG-C---CA----A----A--A---CCAAA-GCAA-T............ 8484
Env                         Q  Q  E  L  L  R  L  T  V  W  G  T  K  N  L  Q  T  R  V  T  A  I  E  K  Y  L  K  D  Q  A  Q  L  N  A  W  G  C  A  F  R  Q  V  C  H  T  T  V  P  W  P  N  A  S  .  .  .  .  .
H2A.SN.86.ST_JSP4_27        ----AG-AA------G---G--C---------A-CA--AA-T------------G--ACT---A-C--G-A----T-------C--GGC--AA--AAATTCA-----A--TG-GTTT-G--AAG-C---CA----A----A--A---GTAAA-GACACC........... 7969
H2B.CI.x.EHO                ----AG-A-------GG--G--C---------G-CG--AA-C------------G--ACT--CA-C--G-A---T--C--A--C--GGCA-AA--AAATTCA-----A--TG--TTC-G--AGG-C--TCA---GA----A--A---GTAAA-GAA-CC............ 8501
H2G.CI.92.Abt96             -----G-AA------G---G--C---------G-CG--AA-C------A-T---G--ACT--CA-C--G-A---T---------C--GGCA-GG--AAATTCA-----A--TG-GTTT-G--AGG-C---CA----A----A------G----CCT-G-TGC-------A-A 7912
H2U.CI.07.07IC_TNP3         ----AG-G-------GG--G--C---------G-CT--AA-C------A-T---G--ACT--CA-C--G-A---------------GGCAAGT--AAATGCA-----A--TG-GTTT-GG-AGG-C---CAT---A----A--A---GTAAA--A--CT............ 8008
H2U.FR.96.12034             ----AG-AA------GG--G--C---------A-CT--AA-C------A-T---G--ACT--CA-C--G-A---------------GGCAAGT--AAATGCA-----A--TG-CTTT-G--AGG-C---CAT---A-------A----TAAACGACACA........... 8023
ASC.UG.10.RT03              ----A--G--TC-A--G--GT-T--G-----G----GAA-T---A-T--CC-CC-TACC---A-T--G-A---TG-C-----C---GG-A--T---CTTCA-AT--G---CAATTC----AA-----TTAT--CT-A--T---------AAATCC-TTGC-......... 7521
ASC.UG.10.RT08              ----A--G--TC-A--G--GT-T--G-----G----GAA-T---A-T--CC-CC-TACC---A-T--G-A---TG-C-----C---GG-A--T---CTTCA-AT--G---CAATTC----AA-----TTA---CT-A--T---------AAATC--TTGC-......... 7551
ASC.UG.10.RT11              ----A--G--TC-A--G--GT-T--G-----G----GAA-T---A-T--CC-CC-TACC---A-T--G-A---TG-C-----C---GG-A--T---CTTCA-AT--G---CACTTC----AA-----TTA---CT-A--T---------ATATC--TTGC-......... 7512
COL.CM.x.CGU1               -GTCTGCA---C-A--G-----G--------GG-T--AA-C-------T-C--G-TGCCA--A-A---G-T--TT-GG----A---GCAAAG--T-CC-GCATT--G--TG-AAAC-TG-AA-----T-GA---AT------C-----CAAG-CC---............ 7476
COL.UG.10.BWC01             -GTCAGCA---C-C---------G--A-----GG-T--AA-C-------T-C--G-TGCCA--A-A-----T--TT-GG----A---GCAAAAT---CC-G-ATT--A--T--AAAT-TG-AA-------GA---AT--------------AAG-CC---............ 7299
COL.UG.10.BWC07             -GTCTGCA-----A---------G--A-----GG-T--AA-C-------T-C--G-TGCCA-AA-A--G--G--T---G----A---GCAAAG--T-CAG--ATA--A--TG-AAATTT--AAG-C----GA--A--A--T--A------T-C-CG---............ 7482
DEB.CM.04.SIVdeb04CMPF3061  -CGG---C-----AAC-T-G----CG-----AG-G--AA-C---A-TA-CC-CC-GACC--GC-T--G-A-------------C---AGCAAGT--AATGAG-----A--TG--TTT----AG--C---CA---CA-A-------------GATCAATA........... 7900
DEB.CM.99.CM40              -CGG---G-----AAC-T-A----C------AG-G-GAA-C---A-CA-CC-CC-GACT---A-A--G-AG---T---A--C---GCAAAAT-AAATGAA-----C---G-CTTC----AG--C--TCAT--CA-G--C---------AACTC-CT-............ 7806
DEB.CM.99.CM5               -CGG-----------TC---A----CA-----AG-G-GAA-C---A-CA-CC-CC-GACT---A-A----AG--------A-----GAGCAAGT-AAATGAA-----A--TG-CTTT----AA--C--TCA----A-G--C--------CCA--CA---........... 7727
DEN.CD.x.CD1_CMO580407      -CAGTGCA--AC----GT-GT----A---------T--AA-C---A-C--TC-CC-GACT--AA-A--G-A---T--T--A--C---GCA--G--AAATGAA-----A--TG-CT-G--G-AA-----TCA---CT-A--AGAA----GCAA-CAA-CA........... 8123
DRL.DE.11.D3                -----G-GT-AC-CA-G--A--TA---------G-A---A-C---------CC-CC--ACA---C-T--GGA----G--GGA--C---GCAAGGT-AA-CC-C-----A-----CTTTGC--AAG-G--TCA---AAA----GTG---CC-AA-GAG-CA........... 7823
DRL.x.x.FAO                 -----G-GT-AC-CA-GT-A--TA---------G-A---A-C---------CC-CC--ACA---C-T--GGA----A--GGA--C---GCAA-GT--A-CC-C-----A-----CTTTGC--AAG-A---CAT--AAA----GTA---CCCAA-GAG-CA........... 7907
GRV.ET.x.GRI_677_gri_1      ----A--GT-----A-G--G--CA-T-----TG-G--AA-C---A-T--CC-CG--ACA---C-C--G-AG------G-------GGCA-GG--AAATTCA-----A--TG-GT-G---AAG-A--TCA---CA-A-----A-----GTA--A-AAC........... 8073
GSN.CM.99.CN166             -----AGCT-A--A--G-CAT-T--A---------A--A-----A--AA-CC-CC-GTCAT-AC-A--G-A---TT-G-G-------ACCA-T--CCAAGCA-----C--TG-GAAC-GG-CA--C--TCA---CATA--A---------CA-CTCA.............. 8000
GSN.CM.99.CN71              -----AGCT-A--A--G--AT-T--A---------A--A-----A--A--CC-CC-GTCTT-AC-A--G-A------G-G-------ACAA-T--CCAAGC------C--TG-GAATC-G-CA--C--TCA---CATA--A-----------ACTCA.............. 8012
LST.CD.88.SIVlhoest447      -G--AG-G--C--AAG------CA--------AG-G---A-T-----A---C-CC-GACA---A-T--GGAG------------C---ACT--G--T-CTTCC------G--TCAAT-G----AAG-A--TCAT---AA---AGAA-----CTA--A-.............. 7243
LST.CD.88.SIVlhoest485      -G--AG-G--C--AAG------CA-A------AG-G---A-T-----A---C-CC-GACA---A-T--GGAG------------C---GCT--G--T-CTTCC------G--TCAAT-G----AA--A--TCAT---AA----GAA-----TA--A-.............. 7240
LST.CD.88.SIVlhoest524      -G--AG-G---C-AAG------CA--------AG-G---A-T-----A---C-CC-GACA---A-T--GGAG---T-G--A--C---GCA--A--T-CCTC------G--TCAGT-G----AAG-G--TCAT---AA---A--A-----TA--A-.............. 7246
LST.KE.x.lho7               -G--AG-A-----AAG------CA-A------AG-G---A-T-----A---C-CC-GACA--CC-T--GGAG--TG-G--AC-----GC-------T-CCTC------G--TCAAT-G----AGG----TCAT---AA----GAG----CATA--A-.............. 8318
MAC.US.x.251_1A11           ----AG-AT------G---G--C---------A-CA---A-C------A-T--GG--ACT--CA-C--G-A----T-------C--GGC--AG---AATGC------A--TG-GTTT-G--AAG-C---CA----A----A--A---CCAAA-GCAA-T............ 8480
MND-1.GA.x.MNDGB1           -GTCTG-A--C--AA-------CA-A------AG-A---A-TT-A---A-TC-CC-GACCAG-T----G-AT--TA-C------C---GCTT-G---TCTCAA-----G--T--AT-GGC--AGG-G--TCAT---T----AGAG----C-AA--CAA-C............ 7673
MND-2.CM.98.CM16            -CTCCG-GT-A--AA----A--TA-------TG-A---A-T--T-----CC-CC-TACA--CT----GGAC--TG--GCT------GCAAGA--ATCC--G-----C-----ATTCGCC-AGG-A---CA---ACA--A--A---CCAAA-GAC-CT........... 8216
MND-2.GA.x.M14              -CTCCG-GT-A--AA----A--TA-------TG-A--AA-T--T-----CC-CC--ACA--CT----GGAC--TG--GC---C---TCAAGA--A-CAG-A-----A-----ATTCTCC-AAG-G--TCAT---AA---T--G---CCAAA-GAAA-T........... 8145
MND-2.x.x.5440              -CTCCG-GT-A--AA----A--TA-------TG-A---A-T--T----TCC-CC--ACA--CT----GGA----G--GC---C---TCAAGG--CTC-G-A-----A--T--ATTCTCC-AAG---TCAT---AG---AAAA---CCCAA--A-A-T........... 7849
MNE.US.x.MNE027             ----AG-AT------G---G--C---------A-CA---A-C------A-T---G--ACT--CA-C--G-AG---T-------C--GGC--AG--AAATGC------A--TG-ATTT-G--AAG-C---CAT---A----A--A---CCAAA-GCAA-T........... 7961
MON.CM.99.L1_99CML1         -----AGCT-AC-A--G--T--T--T-----AG-----------A-----CC-CC-GAC-------A--G-AG-T-A-------C---ACAT-G--AAATGCA-----A--TG-CAAT---GC-G-G--TCA---AA-A-----G-----CAA-TC-.............. 8026
MON.NG.x.NG1                -G---AGCT-A--A--G---T-T---------------A--AT-A-----CC-CC-GAC-------A---AG-T-A-------C---ACAT-GT-AAATTCA-----A--TG-CAATCGCGCAG-R---CA---ACAA----TA-----CAA--CG.............. 6651
MUS-1.CM.01.CM1239          -G--A---------A--GT-GT-T---------T--A--A-----A-----CC-CC-GACA--GC-T--G-AG-T-A-----------GCCT-GT--AACC-------A--TG-GAAC-GG-AG--C--TCA----AGG-----A------A-ATCA.............. 7966
MUS-1.CM.01.SIVmus_01CM1085 ----A-----A--A--GT-GT-T--T--------T--A-----A-----TC-CC-GACA--GC-T--G-AG-T-A-------C---GCC--T---AATC-C-----A--TG-AAAC-GG-AG-----TCA----AGA-----A-------A-TCA.............. 7997
MUS-2.CM.01.CM1246          ----A-----AC-A--GT-GT-T--T--------A--A--A--A-----TC-CC-GACA--AC----G-A--T-G-G-----C---GCCA-A--AAATC-G-----G---G-CAAT-GG-AA--C---CAT--CT-G-----A-------A-TC-.............. 8095
MUS-2.CM.01.CM2500          ----A---------A--GT-GT-T--T--------A--A-----A-----CC-CC-GACA--AC----G-A--T-A-----------GCCT-GT-AAATC-G-----A--TG--AAC-GG-AG--C--TCA---CAGG-----A------AACTC-.............. 8018
MUS-3.GA.09.09GabOI81       -G--A--C---C----GT-AT-T--A-----T--T--A--A-----A--TC-C--GAC----A-T--G-A----T-G-GA--C---ACC--G---AATTCA-----A---G-AAAC-G-GCA------CA---CAA-C-T--A-----CAAATCA.............. 7595
MUS-3.GA.11.11GabPts02      ----A--------A---A-GT-T--T-----G--------A--G-----CC-CC-GACC--CA-A--G-A---TT----A------ACC--AT--AAT--A-----C--TG-AAAC-GGGCG------CA---CAA---------------ATC-.............. 7578
OLC.CI.97.97CI12            -G--AG-AT--C--A-G-----TA-------AG-G---A-C--------TC-CC-G----GA--A--G-A------C--A-----GGCA--A--AAAACAA-----A--TGAATTT-TG-AAG-G---CAT---A-A-----A-----CTAC-C-.............. 7690
RCM.CM.00.SIVagi_00CM312    -G--A--CT--C-AA-G-----C--------G-CT--AA-C--------CC-TG--ACT---A-T--G-A-------GCA--C---TCT--AT--AAT-CA-TT-----TG-CT-G-G--AAG-G--TCA---CT-A---------GCTT-.................. 7531
RCM.CM.02.SIVrcm_02CM8081   -G-----CT--C--AGG-----C--------G-CT--AA-C--A-----GC-TG--ACT---A-T--G-A-------GCA--C---TCT--AT--AAT-CA-TT--C--TG-AT-G-GG-AAG-C--TCAT--AA-A----AG----C-TTC.................. 7609
RCM.GA.x.SIVRCMGAB1         -G---A-----C-AA-G-----C--------A-CT--AA-T--------GC-TG--ACT---A-T--G-A-------GCA--C---TCT--AT--AAT-CA-TT--G--TG-CT-G-G--AAG-C---CAT--A-TG-----G----CATTC.................. 7747
RCM.NG.x.NG411              -G---G-A---C-AA-G-----C--------G-CT--AA-C--------GC-TG--ACT---A-T--G-A-------GC---C---TCT--AT-AAAT-CA-TT--C--TG-AT-G-G--AAG-G--TCA---AA-G---GAG----TATA----CAA............ 7797
SAB.SN.x.SAB1               -G-----AA--C-AA-G-----TA-T-----TG-A--AA-C---A-T--CC-CG--ACT---C-T--G-A-------G-------GGCA-GGT--AAC--------A--TG--TTC-GG-AGG-G--TCA---GA-G--C-T------GTA--A-AAC............ 8367
SMM.CI.79.SIVsmCI2          ----AG-A-------G---GG-C---------A-CT--AA-C------A-T---G--ACT--CA-C--G-A---------------GGCA-GG--AAATGC------A--TG-ATTT-G--AAG-C---CA----A-G--A------G-AAA--ACACT............ 7637
SMM.LR.89.SIVsmLIB1         ----AG-AT------G---G--C---------A-CT--AA-C------------G--ACT---A-C--G-AG---T-------C--GGCA-AG--AAATTCA-----A--TG--TTT-GG-AAG-C--TCA----A----AGAG---CCAAA--AAACA............ 7546
SMM.SL.92.SIVsmSL92A        ----AG-A-------G---G--C---------A-CT--AA-C------A-T---G--ACT---A-C--G-AG---------------GGCA-AG--AAATTC------A---G-ATTT-GG-AGG-C---CAT--CA----A--A---CCAAA-GA--CT............ 7581
SMM.SL.92.SIVsmSL92B        ----AG-AT------G---G--C---------A-CT--AA-T--T---A-T---G--ACC--CA-C-........................................................................................................ 7460
SMM.SL.92.SL92B             ----AG-AT------G---G--C---------A-CT--AA-T--T---A-T---G--ACC--CA-C--G-A---------------GGCA-AG--AAATTCA-----A--TG-ATTT-GG-AAG-C---CA----A----A-TA---CCAAA-GACA-T........... 7932
SMM.US.04.G078              ----AG-AT------G-T-G--C---------A-CT--AA-C------A-T---G--ACT---A-C--G-AG---------------GGCA-AG--AAATTC------A--TG-ATTT-G--AAG-C---CA---CA----A--A---CCAAA-GA-A-C........... 7777
SMM.US.05.D215              ----AG-AT------G---G--C---------A-CT--AA-C------A-T---G--ACT---A-C--G-AG---------------GGCA-AG--AAATTCA-----A--TG--TTT-GG-AAG-C---CA----A-C-----A---CCAAA-GA-ACA........... 7753
SMM.US.06.FTq               ----AG-A-------GG--G--C---------G-CT--AA-C--T---A-C---G--ACT--CA-T--G-A-------C-------GGCAAAG--AAATTCA-----A--TG-ATTC-GG-AAG-C---CA----A----A--G---GTCAA-GA--CA........... 7769
SMM.US.86.CFU212            ----AG-AT------G---G--C---------A-CT--AA-C------A-----G--ACT--CA-C--G-AG---T-------C--GGCA-AG--AAATTCA-----A--TG-GTTT-GG-AGG-C---CA----A----CGAA---CCA--CGCAAAT........... 7751
SMM.US.x.H9                 ----AG-AT------G---G--C---------A-CT---A-C------A-T---G--ACT---A-C--G-AG---------------GGC--CRN-AAATTCA-----A--TG--TTT-GG-AGG-C---CA----A-K--A--A---CCAAA-GACACA........... 7973
SMM.US.x.PGM53              ----AG-AT------G-T-G--C---------A-CT---A-C------A-T---G--ACT---A-C--G-AG----G---------GGC--AG--AAATTCA-----A--TG--TTT-GG-AGG-C---CA----A----A--A---CCAAA-GCC-CA........... 8430
SMM.US.x.pE660.CG7G         ----AG-AT------G---G--C---------A-CT---A-C------A-T---G--ACT---A-C--G-AG-----G--------GGC--A---AAATTCA-----A--TG--TTT-GG-AAG-C--TCA----A----A--A---CCAAA-GAAACA........... 8489
STM.US.89.STM_37_16         ----AG-AT------G---G--C---------A-C----A-C------A-T---G--ACT---A-C--G-A---------------GGCA-AG--AAATTC------A--TG-GTTT-G--AGG-C---CA---CA----A--A---CCAAA-GA--CT........... 8135
SUN.GA.98.L14               -G--AG-G--AC-AAG------CA-A------AG-G---A-T-----A---C-CC-GACA--CC-A--GGAG--TG-GC-A--C---AGC--TT-A-CTTC------G--TCAAT-G----AAG-A---CAT---AA------C------TA--A-.............. 8386
SYK.KE.x.KE51               -------C--C--AGGCT-A--G---------TG-G--AA-C--TA-C--CC--C--ACT---C-A--G-CG--T--C-GA--C---GCCA-TA--TCA-A------A---G-ATTT--G-AA-----TCA---A--A--AA-G---C-GCAAGCA--TG-C-----C--- 7738
SYK.KE.x.SYK173_COMGNM      -G--A--C--C--AGGCT-A--T---------TG----AA-C--TA-C--CC-GC--ACA---T-A--G-CG--T--C-GA--C--GGCCA--T--TCA-A------A---G-ATTC--G-AG--C---CAT--A--A---A-G---G-GAAAGCA--TG-A--C--C--C 8045
TAL.CM.00.266               -G--A--C-----A---T-A-----G------------A-T---A-T--CC-CC-GACC--GC-T--G-A----T--G----C--GGCTAAG--AAATTCA-----G---G-AT-G----AG--C---TA---CT-G--A--G-----CAAA-CA.............. 7893
TAL.CM.01.8023              -G--A--C-----A---T-A-----G------------A-T---A-T--CC-CC-GACT--GC-T--G-A----T--G----C---GCCAGGT-AAATTCA-----A--TG--T-G----AG--C---TA---CT----A--G-----AAG-CA.............. 7443
TAN.UG.x.SIVagmTAN1         -G-----GA--C-AA-G-----CA-T-----TG-G--AA-C---A-T--CC-CG--ACT--CC-T----A-------G-------GACA-GG--AAATT-G-----A--TG-ATTC----AGG-G--TCA----A-G-----G----C-TTC-ACAAT............ 8171
VER.DE.x.AGM3               -------GA-----A-G--G--CA-T-----TG-G--AA-C---A-T--CC-CG--ACA---C-T--G-AG------G----C--GGC--GGT--AATGC------G---G-AT-G--G-AAG-C--TCAT--AA-C--A--G---C-GTGG-A-AAT............ 7712
VER.KE.x.9063               -------AA-----A-G--G--CA-T-----TG-G--AA-C---A-T--CC-CG--ACA---C-T--G-AG------G-------GGCA-GAT--AACG-C-----G---G-AT-G----AGG-C--TCAT--CA-A--A--A---C-ATGGCAAAAC............ 8214
VER.KE.x.AGM155             -------GA-----A-G--G--CA-T-----TG-G--AA-C---A-T--CC-CG--ACA---C-T--G-AG-----GGC------GGCA-GGT-AAACGC------G---G-GT-G----AAG-A--TCAT--AA-A--A--C----CGTGG-A-AAT............ 8194
VER.KE.x.TYO1_patent        -------GA-----A-G--G--CA-T-----TG-T--AA-C---A-T--CC-CG--ACA--CC-T--G-AG------G-------GGCA-GA--AAACTCC-----G---G-AT-G----AAG-A--TCAT--CA-A---GAG---CCCTGG-CAAAT............ 8190
WRC.CI.97.97CI14            ----AG-G-------G------CA-------GG-G--AA-C--------TC-CC-GTCA--CA-T---G---T-T--G-A--C--GTT-AAAT--AAACAA-----A--TGAATTG-C--AAG-A---CA---C-----T---------CAGTC-.............. 8194
WRC.CI.98.98CI04            ----AG-AT------G------CA-------GG-G--AA-C--------TC-CC-GTCA--CA-T---G---T-T--G-A--C--GTTAAAAT--AAACAA--------TGAATTG-C--AGG-A---CAT--CAA---A----------AGTC-.............. 8239
WRC.GM.05.Pbt_05GM_X02      ----AG-G-------G------CA-------GG-T--AA-C--------TC-CC-GTCA--A--C---G-----T--G-A--C--GGT-AAAT--AAACAA-----A--TGAATTG-C--AAG-C--TCA---A--A--C------G-G-AGGCA.............. 7135
```

```
H1B.FR.83.HXB2               .......................................CTGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGA 8209
Env                          _.__.__.__.__.__.__.__.__.__.__.__.__._ L  E  Q  I  W  N  H  T  T  W  M  E  W  D  R  E  I  N  N  Y  T  S  L  I  H  S  L  I  E  E  S  Q  N  Q  Q  E  K  N  E  Q  E  L  L  E
H1O.CM.91.MVP5180            ..............................TATAAT...GAT--CAGT------G-CA--CTT--A---CA-C-A------CA-C-C--A------GTA-GCTC-A-T---T-TGATGA---AC---C-G-A---G----A--G----------T-A---C---G----- 8282
H1N.CM.95.YBF30              TCT....................................TAT--TACA--C------A-TTTA------CAAC-A-----TGAGA--G-A-GA--C--TT--G-TG-C--TTTTGGAC-T--A---C-GG-A---G-A--A--GA-C-CA-----GA--TC-C-C----- 7760
H1P.CM.06.U14788             .............................GAAACAGAAT-A--TGGC------GG-A-TCTA--A---CA-----------A-CTGG-GG-T--------TGA-AC---TT---TGGA---AC-GA--G-A---G-G-----GA-GG-A------AG-A-GC----A-- 7767
CPZ.CD.06.BF1167             CTGTGCGGTAATAAGACAAACAATTATGAAAACAATTATGA-TGC...--A---TC---ATAC------CAAC-------TGAG-----A-GG--CAT-T--GATG-T--CT--CTTAGTT-A---AC-G-T--GGTG--A-------GA---C-CA-GC--C-CC-A-- 8376
CPZ.TZ.00.TAN1               ACAAAGTGCAATCACAGTGAT.........GCAAAGTACTAT--CTGT--A-----CA-TTT---T---CA---A-----TC--TT-G-AG-A--C-CT---G-AACC---T-----C-GT-A--GA--G-A----CA--A--G--G--A--CA-------G--G--A-- 7865
GOR.CM.07.SIVgorCP2135con    ...........................AATGAAACTGAA--A--TG-C------GGAA---T---T---CAAC-A-------A-TTGG-AG----C-----TGA-AC----TTT-TAGA---AC--A-GG-A--GG-A-----G--GGTT------A-G-CC---C-A-- 7760

MAC.US.x.239                 .......................................--AAC--CA-AG-----CA-TGA---T---CAA--------GC--A-GG--G--TTC-TGGA-GAAAAT---ACAG--C-CC-A--G--GG-A----TT--A-----G-----CATGT-T------CAAA- 8615
Env                          _.__.__.__.__.__.__.__.__.__.__.__.__._ L  T  P  K  W  N  N  E  T  W  Q  E  W  E  R  K  V  D  F  L  E  E  N  I  T  A  L  L  E  E  A  Q  I  Q  Q  E  K  N  M  Y  E  L  Q  K
H2A.SN.86.ST_JSP4_27         .......................................T-AACG-CTGA------CA---T---A---CA---A----GCA-CG---CCG---CCTAGAGGCAAAT--CAGTGAAAGTT-A---C-GG-A----T---------------CATGT-T---C--CAAA- 8100
H2B.CI.x.EHO                 .......................................--TA-G-CAGAC------A---T---A---CAAC-A-----G--GC--G-CCG-TTC-TGGATGCAAAT---ACAAAA---C-A-----GG-A--G-TA--A-----------CATGT-T--G---CA-A- 8632
H2G.CI.92.Abt96              .......................................T----G-CACAA------A-T-T---A---CAA--A-----A-AGC-G--C---TT--TGGAGGA-AAC--CACACGGC--T-A------G-A----TT-----------A---ATGT-T--G---CA-A- 8043
H2U.CI.07.07IC_TNP3          .......................................T-AC---CTGAC-----CA---T---G---CAA--------AGT-A--G-A--AG-C-TGGAGGAAAAT--TACAGT----T-G------G-T----TT--A-----------ATGT-TC-G---CAAA- 8139
H2U.FR.96.12034              .......................................T-AAC--CT-A----G--A-T-T---G---CAA--A-----GGA-A-GG-G---T-C-TGGA-GAAAAT--CACACAGC--T-A----C-G-T----TA--A-----G--A--CATGT-T------CA-A- 8154
ASC.UG.10.RT03               ...............................GATGGAGCA--TCCTGCTTGGCAC--C...-T---------A--------A---A--G-G-CT---C-T---G-TA-C--CA-TA-T--GT-G-T---GG-C----GGAG-------G-A--CAC---TA-G---CA-A- 7658
ASC.UG.10.RT08               ...............................GATGGTGAAGC-CTT-CTGAA-----CA---T---------A--------A---A--G-G-CT---C-T---G-TA-C--CA-TGGG--GT-G-T---GG-C----GGAG-------G-A--CAC---TA-G---CA-A- 7691
ASC.UG.10.RT11               ...............................GATGGAGGAG--CTT-CTGAC------A-T-T---------A--A------A---A--G-G-C----C-T---G--A-C--TA-TAAT--GT-G-T---GG-T----GGAG-------G-A--CAC---TA-G---CA-A- 7652
COL.CM.x.CGU1                ....................................GGAGA---GG-TCCA---C-GA---T---------AAC-----C-TGAGCG-G-A-GA---------GATA-----G-GG-AGACT-A-T---GG-AT-TG--TTA-----GG-A------A-GA----GGCA-- 7610
COL.UG.10.BWC01              ....................................GGCA----GG--CCA---C-GA--TAT--A----A-C-----T-TGA-C--G-GGCA---------ATA-T---G-GG-AGATT-A-T---GG-TT-TG-TTT---G---GCA--C--GA-GA-----GCA-- 7433
COL.UG.10.BWC07              ....................................GGAGAT--G--ACC-----TGA---T---T----AAC-----C-TGA-AG-G-A--A-----T---GATG-------AGAGGA-T-GC-----G-CT-TG-GTTA-----GG-------GA--A-G---GCA-- 7616
DEB.CM.04.SIVdeb04CMPF3061   .......................................GGAC---CAGA-------A---T---A---CA---A-----ACAGA-GG-AGCA--C------GATA--TGGG-AGAAGC-T-AC-GA--G-A--GG-AA-A--G--GC----C-T---T-CCC--CA-TC 8031
DEB.CM.99.CM40               .......................................GA---T-CAGAC---G-CA-T-T-------CA---------A-TGA-GG-AGCA-----T---GATGAGTGGG-AGGAGCTT-AC-GAG-G-A---G-G-----G--G-G-----T---T-CCC--CAATC 7937
DEB.CM.99.CM5                .......................................GGA----CTGA-------A-T-T-------CAA--------A--GA--G-AGCA--C--T---GATGA-TGGG-AGGAGCCT-AC-GAG-G-A--GG-A--A-----G-GA----TG--T-CC---CA-TC 7858
DEN.CD.x.CD1_CMO580407       .....................ATACCTGGATTGGCATTATACC-G-CTGAC-TCC-GA---T---A---CAAC-A-----G-C---C--AGCAGC-CTATTTG--AATG-GTCAGA-GCCT-GA-GA--G-A--GGCA-----G----GA----TG--C---G--CAATC 8272
DRL.DE.11.D3                 .......................................G--ACT-CA-A-----CCTCAGA---A-----------C-A-A-AG-G-AG-T-G-ATAT---A-AAC--TACTCTTGATT-AC-GA-GG-AT-TG-A--AG-GC---------ATTTTT--G---CAAA- 7954
DRL.x.x.FAO                  .......................................G--ACT-CA-A-----CCTCAGA---A-----------C-A-A-AG-G-AG-T-G-ATAT---A-AAC--TACTCTTGATT-AC-GA-GG-AT-TG-A--AG--C---------ATCTTT--G---CAAA- 8038
GRV.ET.x.GRI_677_gri_1       ..........................................ACT-CT-AG---G-CA-T-T---T---T---------G---C-------TGCC-TGGA-G--AAC---ACTCAAC--T-G------G-A-----T---G--TC-------CTGG-TCTG-ACCA-A- 8201
GSN.CM.99.CN166              ...........................TGGGCTAATGGCTC-CTT-CAGAC---G-GA---T-------CAAA-----AGT-TGTT-G--G-G---G-----TATAC------GCAG--GT-A--GC--G-AA-CC-G--A----C-TCA--CCTGA-T--GC-GA--A- 8143
GSN.CM.99.CN71               ...........................TGGGCTAAGAACTC-ACG-CAGA----G-G----T-------CA-------AGT-AGTT----G-A---G-----TATACC-----GCAA---T-A---A-TG-AA-CC-T--A---TC-------CATGA-T--C------AA- 8155
LST.CD.88.SIVlhoest447       .......................................A-AACT-CT-A-----CAAGGGAT--------T------------C--G--GGAGTA-TAGA-GCAAAT--CTC-A--C-GT-GC-----G-CT-T-C-ACAG-GTT-G-A---AG-A-T-CC--CAA-A- 7374
LST.CD.88.SIVlhoest485       .......................................A-AACT-CT-A-----CAAGGGAT--------T------------CG-G--GGAGTA-TAGA-GCAAAT--CTC-A--C-GT-GC-----G-CT-T-C-ACAG-GTT-G-A---AG-A-T-CC--CAA-A- 7371
LST.CD.88.SIVlhoest524       .......................................G-TACT-CT-A-----CAAGGGAT--------A--A-----A---C--G--GGATC--TAGA-GCAAAC--CAC-A-AC-GT-GC----GG-TT-C-CAACAG-GTT-G-----AGGA-TA-C--TAAAA- 7377
LST.KE.x.lho7                .......................................A-CAC--CC-A-----CCA-GGAT-------GA--A-----ATC-A-GG-AGCA-T---TGAC-AGAAC--TACT--TC-GT-AC-----G-AT-C-C-ACAG-GCT-G-----C-GA-TA----TAA-A- 8449
MAC.US.x.251_1A11            .......................................--AAC--CA-AG-----CA-TGA---T---CAA--------GC--A-GG--G--TTC-TGGAGGAAAAT---ACAG--C-CC-A--G--GG-A----TT--A-----G-----CATGT-T------CAAA- 8611
MND-1.GA.x.MNDGB1            .......................................A-CACT-CA-A-----CATCAGAA--T----A---A-----G-C-AG--C-G-TT--CTGCA-CAAAAC--TACAGAAA-GT-AA--C-GG-AT-TG-T-GAG-GC---GA--CAC-T-T------CA-A- 7804
MND-2.CM.98.CM16             .......................................A-AAC--CA-A-----CCTCAGA---A---C----A-----T-A-AG-G-A-CAGC-CTA--TGA-AAT--GACAGTAAACT-AC-GA-GG-AT-TG-ATT-G-GC---------AT-T-T--GC--GAAA- 8347
MND-2.GA.x.M14               .......................................A-AACT-CT-A-----CCTCAGAA--A---C----A-----T---AG-G-G-CAGC-AT---T-A-AAT--GACAAT-GATT-AC-GAGGG-AT-TG-ATT-G--C--------ATGT-T------CAAA- 8276
MND-2.x.x.5440               .......................................A-A-TG-CC-A-----CCTCAGA---A---C----A---------AG-G-A----GCATAGTG-CAAAT--GAC-ATAGACT-AC-GAGGG-AT-TG-ATTAG--C-G-G---CATCTTT--G--GCAAA- 7980
MNE.US.x.MNE027              .......................................--AAC--CA-A------CA-TGA---T---CAA--------GC--A-GG--G--TTC-TGGAGGAAAAT---ACGG--C-TC-A-----GG-A----TT--A-----------CATGT-T------CAAA- 8092
MON.CM.99.L1_99CML1          ...........................TGGGCAAAAGGT-ACTTC-CTGAG---G-CA-T-T---A---CAAC-----AGTGAGTT-G-AG-T---G------TGACC--T--GCAGC-CT-G----CTG------G-G--------GC--A--CC----T--G---A--A- 8169
MON.NG.x.NG1                 ...........................TGGGCTAAAGGC-ACTTC-CCGAG---G-CA---T---A---CAAC-A---AGT-TGCT-G--G-----G-T--GGCAC-T--C--GCADC-CT-G--G--GG-T--GG-G--A----G---A----C---T--GC-CA--A- 6794
MUS-1.CM.01.CM1239           ...........................TGGGCCAATCATACCTT--CAGGC---G-AA---T-------CA-C-----AG--AGTTGG-GG-T---G-T--T-ATACC-----GGG-C-CC-A----C-G-A-----G--A--G---G-A--CC-----CA----GCAAA- 8109
MUS-1.CM.01.SIVmus_01CM1085  ...........................TGGGCAAATCACACCC-G-CAGG----G-AA-T-T---A---CAAC-----AG---GTTGG-AG-----G-T--T-CAACT-----GGAG---T-G---TTGG-A-----GA-----G---G------C-G--TA-----CAAA- 8140
MUS-2.CM.01.CM1246           ...........................TGGGCAAACCACACCC---CAGGA---G-CA---T-------CAAC-----AG----TT-G-GG-----G-----GCAG-C--C--TGGTC--T-G---ATGG-T---C-A--A--G---G-A--CC-----CA-G--GCA-A- 8238
MUS-2.CM.01.CM2500           ...........................TGGGCTAATCATACCC-G-CAGAA---G--A---T-------CA-------AG---GTTGG-GG-----G----TGCAACC--C--AGGTC--T-A--GTTGG-T--G--A-----G--GG-A--CC-G--TA-GC-GCAAA- 8161
MUS-3.GA.09.09GabOI81        ...........................TGGGCAAATGAAACAATG-CAGGG---G--A-T-T---T---CA---A---TC-C--TTG---G-----G----TGTT--G---TCAAA---GT-G--G--GG-TG-G--G--A-----G-G----A-T--T--CC--A--A- 7738
MUS-3.GA.11.11GabPts02       ...........................TGGGCAAATCATACAC---CAGGA---G--A-T-T---T---CA-------AG-C-CTTGG-AG-----G-T---CT-GA---TG-AAGG--GT-G---TTGG--A-C--A-----G---G-A---A-T--CA-GC-TA--A- 7721
OLC.CI.97.97CI12             .......................................-A-TC--CAGA---GGCA-TGAT--T---T--A-----C-A--GA--G-AG-T----TA---CAGGA----G-TGA-AG-T-GAC-AT-G-AT-TC-AG----GC---G----C--T-T--G--GCA-AG 7821
RCM.CM.00.SIVagi_00CM312     ....................................AATAA-AC--C-GAC---C-GA-ACA---A---C-T--A-----A--GA-TG-CTCAT---TAGA-GCAAAC---ACAG---C-T-AC-----G-A---G-T-----T--G--A----T---T---C--GAAA- 7665
RCM.CM.02.SIVrcm_02CM8081    ....................................AATGCAAC--CTGAA------A-ACA---A---C-T--A-----A--GA-T--ATCA-TAATAGA-G-GAAT--CTCT-TAGC-T-AC-G---G-C--GG-T--A--T--G-GA----TG--T---C-GGA-A- 7743
RCM.GA.x.SIVRCMGAB1          ....................................AACAAAACG-CTGAG---C-GA-AGAAT-A---T--C-------A---A-T--CTCTT---TAGAGGCTAAC--TACAATAGC-T-AC-G--GG-C--GG-T--A--T--G--A----TG--T-----GGA-A- 7881
RCM.NG.x.NG411               ..........................................AC--CAGAG-----CA-ACA---A---C-T--------A---A-T---TCA-GA-TAGA-G-AAAC--CTCAGTAGCTT-AC-G--TG-A---G-A-----T---CGT----TG--T--C---GAAA- 7925
SAB.SN.x.SAB1                ..........................................ACT-CCGAC---G-AA---T---G---CAA--A-----G--GC----AG-A--G--TGA-GCTAAC--CAGTAGGA-TT-A--GC--G-A--TG-A---G--C---------TTGG-TAGT-ATCAAA- 8495
SMM.CI.79.SIVsmCI2           .......................................T--C---CTGA-------A-T-T---T---CAA--A-----A-T-A-G--C-GAG--CTGGAGGCAAAT--CTCAGAA-C-T-G--G--GG-A----TA--A------------ATGT-T--G--GCAAA- 7768
SMM.LR.89.SIVsmLIB1          .......................................T-AC---CT-A----G-GA-T-T---T---CAA--A-----A--GA-GG--G--TTCCTTGAGGCAAAT---ACACAAC--T-A------G-A----TT--A-----G--A--CATGT-T------CAAA- 7677
SMM.SL.92.SIVsmSL92A         .......................................T--AC--CAGA-------A-T-T---T---CAA--------A--GA-GG-GG--TTCCTTGAGGCAAAT---ACACTGA-GT-A------G-C----TT--A------------ATGT-T------CA-A- 7712
SMM.SL.92.SIVsmSL92B         ......................................................................................................................................................................... 7460
SMM.SL.92.SL92B              .......................................--T-TC-CAGAC-----CA-T-T---A---CAA--A-----G-A-A-GG--G-ATTCCTAGAGGCAAAT---ACTCAAA-GT-G--G---G-A-G-CTA--------G--A--CATGT-T------CA-A- 8063
SMM.US.04.G078               .......................................T--ACG-CT-A-------A-T-T-------CAA--------AC--A-GG-CG--T--CTAGAGGCAAAT---ACTGAGA-GT-G------G-A--G-T---A-----G--A---ATGT-T--G--GCAAA- 7908
SMM.US.05.D215               .......................................T--AC--CA-A------CA-T-T---A---CAA--------A---C-GG--G-TTT-CTAGAGGCAAAT--TACACAAA-GC-A------G-A----TT--A-----G-----CATGT-T--G--GCAAA- 7884
SMM.US.06.FTq                .......................................--CC---CAGAC-----CA-T-T---T---CAA--A-----G-A-C-GG-AGCTT-CCTAGAGGCAAAT--TACACAGCA-T-G------G-A----TA-TA------------ATGT-T--T---CAAA- 7900
SMM.US.86.CFU212             .......................................G--AC--CAGA-------A---T-------CAA--------G---A-GG--G--T--CTAGAGGCAAAT---ACGCTTC--T-A------G-A----TT--A-----------CATGT-T--G--GCAAA- 7882
SMM.US.x.H9                  .......................................T--AC--CT-RC-----CA-T-T--YT---CAA--------A-A-C-GG-G---TTCCTAGAGGCAAAT---ACTSAA-Y-T-G------G-A----TT---------M-----ATGT-T-----GCAAA- 8104
SMM.US.x.PGM53               .......................................T---TG-CT-A------CA-T-T---T---CAA--A-----A---C-GG--G--G-CCTAGAGGCAAAT---ACTCAGGC-T-G------G-A----TT---------------ATGT-T-----GCAAA- 8561
SMM.US.x.pE660.CG7G          .......................................T---TG-CT-A-----GCA-T-T---T---CAA--------A---C-GG--G--TTCCTAGAGGCAAAT---ACTCAA---T-A------G-A----TT---------------ATGT-T-----GCAAA- 8620
STM.US.89.STM_37_16          .......................................T---T--C-GA------CA-T-T---A---CA---A-----G---A-GG--G--TT-CTTGAGGCAAAT---ACACAAC-GT-A------G-T---GTT--A--G----------ATGT-T--G--GCAAA- 8266
SUN.GA.98.L14                .......................................A-CACG-CC-A-----C-A-GGAC-----------A-----T--GC--G-A--A-TG--TGATGATAAT---ACAG-TC-TC-CC-G--GG-AT-TGTAACAG-GTT-G------C--A-CA-G--CAAAC- 8517
SYK.KE.x.KE51                .................................CGTTGC-CAAC--CCCAA---G-AA---T---A---CATACA-----AC-GC-GG-AG-T----TG---GA-CAC---G-TAAT---C-GAG----G-T--GG-G---------G-------TG--T--TC-TACTA- 7875
SYK.KE.x.SYK173_COMGNM       AAC..............................TTCTGT-CCA---CACAG-----AA-T-T---A---CACAGA-----ACA---GG-AG------TG--TGA-CAT--TG-TGGA--GC-CAG----G-A---G-A--A--G--G-GA----TG--T--TC--ACAA- 8185
TAL.CM.00.266                ...........................TGGTCGAATTACACA--T-CTCAA---C-GA-T-T---A---CAA--A-----G-TGA--G-AG-----C-----G-GC-C---TCACAG--GT-AC-----G-A---G-G--A--G----GA----TC--T--G--GCAAA- 8036
TAL.CM.01.8023               ...........................TGGACAAACTCCACAA-C-CTGA----C-GA---T---A---CAA--------G-AGTTGG-AG-----GC-T--GA-ACT--CACAGTA--GT-AC-----G-A---G-A--A--G----GA----TC--T--G--GCAAA- 7586
TAN.UG.x.SIVagmTAN1          ..........................................AC--CAGAC---G-CA-T-T-------CA---A-----G--CC----A-CTGCC-TGGA-G-AAAT--TAGTA-TACTC---TCA--G-AT-TG-G--AG-GC-G--A---ATGG-TACT-ATCAAA- 8299
VER.DE.x.AGM3                .......................................AG-ACC-CTGA-------A-T-T---T---C----A-----A---C-G--ATCGT---TGGA-G-TAAC---ACAA-ACA-T-A--G---G-CAG-GCA---G-G--G------TTGG-T-C--ACCAAA- 7843
VER.KE.x.9063                .......................................A--ACG-CC-A----C-AA---T-------T----------A---C-G---GGAG-G-TGGA-G--AAC---ACAGAACAGT-A-T-A--G-AAG-G-A--AG----G--A---TT-G-T-CC-ATCAAAG 8345
VER.KE.x.AGM155              ..........................................AC--CAGAG------A-T-T-------T----------A-A-C-G--AG-AGGA-TGGAGG--AAC---ACAAAACA-T-G---C-GG-AAGGG-A--AG-G---------TTGG-T-CT-ATCAAA- 8322
VER.KE.x.TYO1_patent         .......................................-G-ACT-C-GA----C-AA-T-T---T---T----------A---C----AGCTG---TGGA----AAC--TACGAGACA-T-A-TGA-GG-TAG-G-A--AG-G---------CT-G-T-CC-ATCA-A- 8321
WRC.CI.97.97CI14             .................................TGGAATGCAACT-CA-A----G-AA-TGAA--T---CAAC-T---T-T-A-AG--CAG-TG-ACTGCAGGCAAAT---AGTG-AA--T-A-T-ATTG-C--TG-T--AG-GC--GCA---ATGTTT--G--GCA-A- 8331
WRC.CI.98.98CI04             .................................TGGAATGCTACT-CA-A-------A-TGA-------CAAC-T---T-T--GAG--CAG-TG-A-TGCAGGCAAAT---AGTGTAC-CC-A-T---GG-T--T---ACAG--C-------CATGTTT------CAAA- 8376
WRC.GM.05.Pbt_05GM_X02       .................................TGGAATA--AC--CTGC-------A-TGAA--A---CAAC-A---C-TTT-AG--CAG---CC-TG-TG--TAAT--CTCAGATA-GT-A-TGC--G-T--TGCT--AG--C--TCA---ATGTTT--G---CA-A- 7272
```

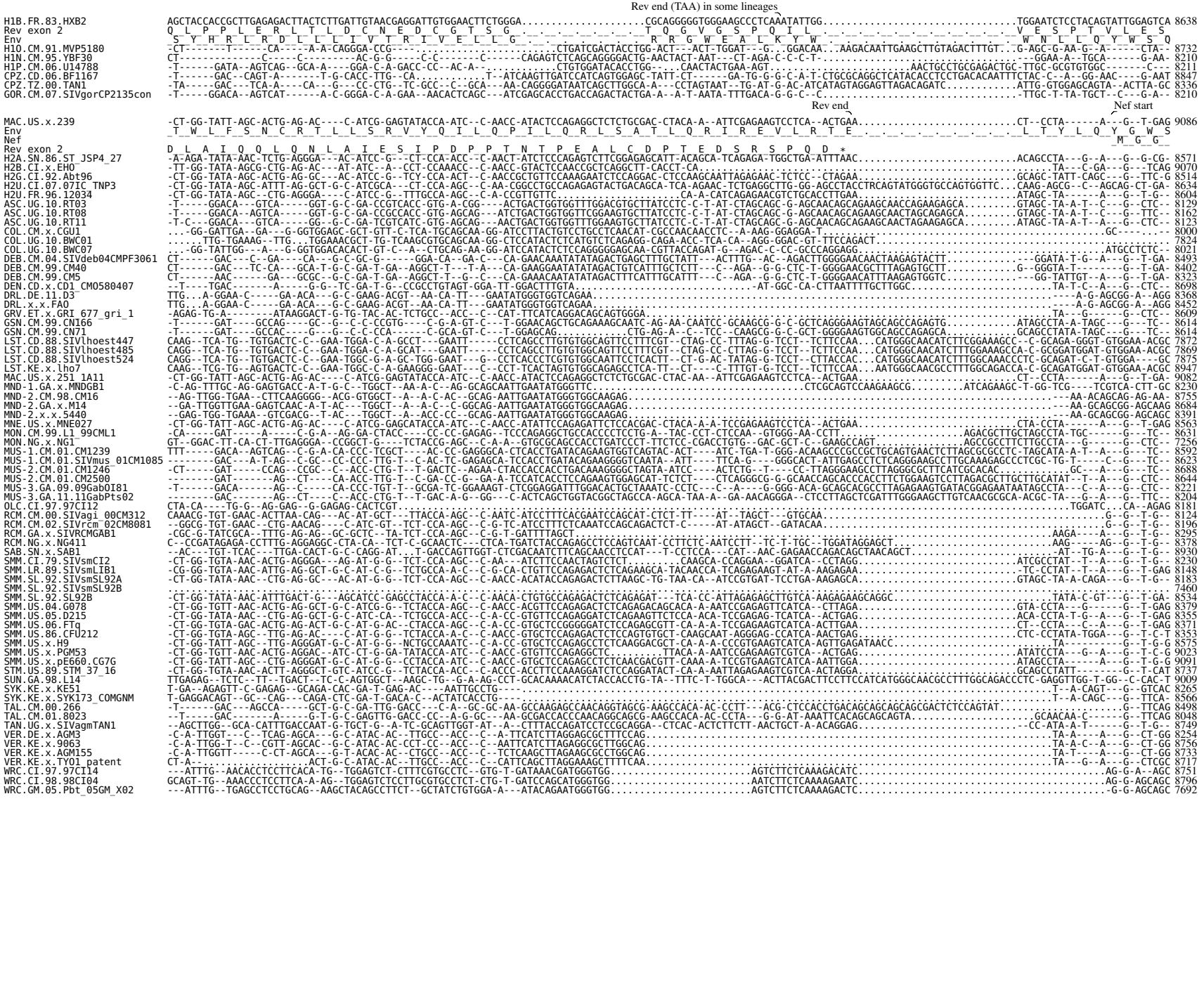
Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546



Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546


Rev end (TAA) in some lineages
Rev end
Nef start

```
                                                  Rev end
H1B.FR.83.HXB2              GGAACTAAAGAATAGTGCTGTTAGCTTGCTCAATGCCACAGCCATAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAGTAGTA.........CAAGGAGCTTGTAGAGCTATTCGCCACATACCTAGAAGAATAAGACAGGGCTTGGAAAGG......ATTTTGCTAT 8793
Rev exon 2                   G  T  K  E  *
Env                           E  L  K  N  S  A  V  S  L  L  N  A  T  A  I  A  V  A  E  G  T  D  R  V  I  E  V  V  .  .  .  Q  G  A  C  R  A  I  R  H  I  P  R  R  I  R  Q  G  L  E  R  .  .  I  L  L
H1O.CM.91.MVP5180           A---T-G--A---------ACA-A-C----TG--A-T-TT--AG-GT----T--CA-TT----T--CG-CA-C--CTT--GTC--.........---A--ATAG-ACA--GAT-C-TT-----C--A--------T-----A--TGCA-----A.........-C--AG-G- 8887
H1N.CM.95.YBF30             A--GT---G-G--------A-C--------T---A-A-----T--T-T------A--A--A---------A--------T---C-.........---A--ATAG-A--G-GA--ATTA-----------------C-----A---C-A-----A......GCAC--A--- 8365
H1P.CM.06.U14788            A---T----AC-R-----AACC---Y--T-G--CA-TGTT--T---T----T--AA-TT----T---CAA--A----C-----GG.........---CA-ATAG-A----GCT-CTTGA-------A-----GT-------A--GC-A-----A......-GC--A--G- 8366
CPZ.CD.06.BF1167            TCT-A-T-GAC-GCA-TGCA--GTT-GCA-AG--A-TTT---AGAGTTCACT-GCTG-T-------CG-----------C-C--.........AG--T---AGTAGACATA--AA-A--T-----A-CT-----C--G-----A--A----TT......GC-C-AAAC- 9002
CPZ.TZ.00.TAN1              TAGCT---GAGT-TTGCTAA-ACAAGCCA-AG-CAGACTT--T-ACTTTAC--GGTG-T--------TTAA-C-------G---G.........GTTTACATAGCC--G-GA--CA-AA-T--T-----------T--------TC-----CTA......GCC--AAAT- 8491
GOR.CM.07.SIVgorCP2135con   A-----T---C-A------A----------AG-CTGTGTT--TG--TGGACT-----TT----T---CAA--A----C-A---CT.........---A--ATAG----G-GC--CTTAA-------C-----G--T-----A--A--A-----A......-GCC-AT--- 8365

MAC.US.x.239                CT-TT-CC-TG-GGCG-TCCAGGC-G-CTGG-GAT-TG-GA-AGAGA-TC-T--G-GC-C-TGG-GAGACT-ATGG--GACTC-T.........AGGA---G-G-A---TGG--A-T-GCA--C--C--G--G--T-----A--GC-T--GCTC......-C-C-CT-G- 9241
Env                           Y  F  H  E  A  V  Q  A  V  W  R  S  A  T  E  T  L  A  G  A  W  G  D  L  W  E  T  L  .  .  .  R  R  G  G  R  W  I  L  A  I  P  R  R  I  R  Q  G  L  E  L  .  .  T  L  L
Nef                         A  I  S  M  R  R  S  R  P  S  G  D  L  R  Q  R  L  L  R  A  R  G  E  T  Y  G  R  L  L  .  .  .  G  E  V  E  D  G  Y  S  Q  S  P  G  G  L  D  K  G  L  S  S  .  .  L  S  C
H2A.SN.86.ST_JSP4_27        -TGGA-CC-AG-AGCGTTCCGAGC---CGCG-GG--T--G-GAGAGA-TC-TA-AA-C-CCTGGAGAG-CT-CTGG-GGAC-C-G.........GG-CA-AT-G-G--G-GA--A-TTGCAG-C--A-----G--C--G-----GGCA----TC......GCCC-C--G- 8726
H2B.CI.x.EHO                CTGGT-CC-AG-AGCAATCCAAGCAGCAGC--GG--TG-G-GAGAGA-TC-T--GAGC-C-G-GAGG-CCTCGTGG-G---CC-C.........AGGA----GGCAG---AG--CATTGCA-----C--G--G--C-----A--GGCT---CTC......GCCC-CT-G- 9225
H2G.CI.92.Abt96             CTGG--CC-AG-AGCGTGCACCGCAGCAACT-GG--TG-GCAAGAGA-TC-TA-AAGCAC-TGGAGAGCCT-ATGGA--ACTC-G.........GG-A-G-T-G-A----GG--A-T-GCA-----G--------C--G--A--GC-T--GCTC......-C-C-CT-G- 8669
H2U.CI.07.07IC_TNP3         --CTRCG-GAG-G-C-CT--CA---GC-TGG-GA-A-TT-TGGGAGA-TC-G-GACG--TTGG-AGAG--A-AC-C-C-A-CCCG...................................--G--G--C--G-----A--T---CTT......GCCC-C---- 8756
H2U.FR.96.12034             -T-TT-C--AG-AGCATTCCAAGCG--CGGG--A--TG-GCGAGAGA-TC-TAGCCGCAC-GGGAGAGAAC-ATGG--GACTC-G.........GGGC-G-T-G-A---TGGC----AGCT-----A--G--G--C--G--A--A--T--GCTC......GCGC-TT-G- 8759
ASC.UG.10.RT03              A---T-C--AGCAGCA-TCA-CTCAGCCTGGG-----CTC------TTCACCTGGA-CT-------AGCT--AC-G-CGTATC-T.........-GGC--CT-GCCCAG-GG---ATTGC---C--G--G-----C--G-----AGCC--G-TA......C-AC-CAAT- 8284
ASC.UG.10.RT08              A---T-C--AGCAGCA-TCA-CTCAGCCTGGG-C---CTC------TTCACCTGGA-CT-------AGCT--GT-G-CG-CTT-G.........-GGCAGCT-GCCCAG-GA---ATTGC---C--A--G--G--C--G-----AGCC--G-TA......C-CC-CAAT- 8317
ASC.UG.10.RT11              A---T-C--AGCAGCA-TCA-CTCAGGCTGGG-CA--CTT------TTCACCTGGA-TT-------GGCT--AT-G-TG-CTC-G.........-GGC--CT-GCCCAG-AGT--ATTGC---C--G--G-----C--G-----GGCC--G-TA......C-CC-CAAT- 8278
COL.CM.x.CGU1               -ACT--TCG-T--CTA......TTGG-A-C--GG-GTC--AG-C-TCAGA-TCT----ACACTTC-G---TGGC--CG-TCT-CG.........GCGC-T-GC--GC-TCGAGCG-CTG-ATAT-TGC--G--TGG-T-T-T-A-CGTCCCCAA......GG-CCTGCC- 8149
COL.UG.10.BWC01             ....--TTG---CTTA......CTGG-A-CG-GGTGTC---A-C--T-GA-TC-----ACACTTTCGC--TGGC--CG-TCT-CG.........ACGC-TCGC--GC-TAGAGCATCTG-ATAT-T----G--TGGCT-TTT-A-CGTCCCGA-.......-GCCCTGCC- 7969
COL.UG.10.BWC07             --CT--TTG-CG-CTA......TTGG-A-C--GGAGTG--AA-C-G--GA-TC--C--ACACTTC-G---TGGC--CG-TCT-CG.........ACGA-T--AG-GC-TAGAGCA-TTG-ATCTTG-G-C---TGG-T--GA--AGATCCCCA-......GGACCTGCC- 8170
DEB.CM.04.SIVdeb04CMPF3061  ----T-C--AG-C-CAATC--CT-GGCAGGAGGA-AATTG-GGCA-A--T-TCAAAG-T--G----GGT--CGC--C-G---C-T.........GGGA---GAATC----AA----TTGC------AGCC-----C-----A--AGCT---TT-......T-CC-TAAC- 8648
DEB.CM.99.CM40              A---T-C--AG---CCATCACGT-GGCAGGTGGG-AGCTTTGGCAGTGGC-CCAG-G-ATTG-T--G-T--CGT--C-G-G-C-T.........GTCA----AG-AG----A----T-AGAG-G---GCT-----C-------AGCA--GCTC......--CC-CAAC- 8557
DEB.CM.99.CM5               A---T-C--AG-C-CCATCA--T-GGCAGGGC-A-AA-TCTGGGAGTGGC-CCAA-G-ATTG-TC-GGT--CG---C---G-C-T.........GTCT-G-GAG-C---AACC---T-GC---T---GC------C-----A--AGCA--GTT-......T-AC-CAAC- 8478
DEN.CD.x.CD1_CMO580407      A-----CC-A-CAG-CCTCAGGGAT---GGG-C-AG-G-CAT-CA-CA--G-AGG-C-ACAG-G--AGT---AT-G-CG-CTC-T.........ACCC----AGCA---AAG-GGTGGC---C--C---C-G--C---------C-T----TT......G-CC-CAAC- 8853
DRL.DE.11.D3                CAT-TGC...-CGCT-CTACGGGAGGCTGCGC-GAGGCT-TGGCGC-G--GGAGGC-ATTTGG-CTATCA-CAGCTAG--CCC-C.........-G---CCT-GC-CAG-GAG-GGCGA----T--C-----G--T-----A---GCA----TA......C-CC-CAAC- 8520
DRL.x.x.FAO                 CAT-TGC...-CGCT-CTACAAGAGG-TGCGC-GAGG-T-TGGCGC-A--GGAGGC-ACTTGG-CTATCATCAGCTAG--CCC-C.........-G--CCCT-GC-CAG-AAG-GGCAGC---T--C--G--G--T-----A---GCA--GGTC......C-C--CAAC- 8604
GRV.ET.x.GRI_677_gri_1      -C----CC-AG-GGCA--AACAG-GC-TGCTC-A--TCTG--G-GG--T-CGAGG--A-CCTGG-GC--AC-GGGT-CTA-T--C.........-G-TCC----A-C-G--AG-CAT-A---GT--A------G-GC-G--A---C-T----AA......G-CC--GGG- 8764
GSN.CM.99.CN166             A-----CC-AGCAGCA--AACAG-AA-C---G-CAG-CTT--T--CTTCACGTGGA-CT-------GGCC--AC-TC-C-CTTGT.........AG-C---TA--GC---AAT---TGGCT--C--CC--C-C--C-----A--GGC---G-TA......C-A--TAAC- 8769
GSN.CM.99.CN71              A-----CC-AGCAGCA-TCTCAG-AA-C---G-CAG-CTT--TG-CTTCAC-TGGA-CT-------GCCCC-CG-TC-GACT--T.........GG-C---TA--GC---AAT-CTTGGCT--C--CC--C-C--C-------GGC---G-TA......C-AC-TAAC- 8769
LST.CD.88.SIVlhoest447      TC-GAAG-TTGCG--CTGGAGCT-GAACAAG-TCAGA-GGAA--G-AG--ACATATCA-CAG-TC--GCAAGA-ACCTCAG-C-G.........GG-CA-AAGAAG---TGG-GATT-AGAT-CAGAG------GTG-GTTTCC-AGT--G-CCACAGAG-CCGCT--G- 8033
LST.CD.88.SIVlhoest485      TC-GAAG-TTGCGC-CTGGAGCT-GAACAAG-TCAGA-GGAA--G-AG--ACATATCA-CAG-TC--GC-AGAGACATCAG-C-G.........GG-CA-AAGAAA---TGG-GATT-AGAT-CAGAG-G----GTG-GTTTCC-AGT--G-CCACAGAG-CCGCT--G- 8030
LST.CD.88.SIVlhoest524      TC-GAA--CTGCG--CTGGA-C--GAACAAG-GCAGA-GGAA--G--G--AGCTATCAAAA--TA--G-AACA-ACATTAGGT-G.........GG-CA-AACA-A---AGATGG...AGAT-CAGATTC----GTG-G-TTCC-AGT--G-CCACAGAG-CCACT--G- 8033
LST.KE.x.lho7               TC-GAAG-TTC-G--CTGGC-C--GGA-AAATGCAGA-GGAA--G--G-CAGTTATC-A-C--T-----AAAA-ACATCCA-T-G.........GGCAA-AAGAAG---TGG-GA-T-AGAT-CGGAG------GTG-GATTTC-AGT---GCCACAGAG-CCGCTT-G- 9108
MAC.US.x.251_1A11           CT-TT-CC-TG-GGCG-TCCAGGT-GACTGG-GAT-TG-GA-AGAGA-TC-T--G-GC-C-TGG-GAGACT-ATG---G-CTC-T.........AGGA---G-G-A---TGG--C-T-GCA--C-----G-----T-----A--GC-T--GCTC......-C-C-CT-G- 9237
MND-1.GA.x.MNDGB1           --C-A-T-GTTGG--G--C-G-TA-ACCGGATGGCTACA---A-ATAG-A-CTTCACA---TGCAGAGAA-CA--CATT-CT-CG..................G-G-CGTGC--AT-G-------TC-----G--T-----AA-TGCA---C-A......CCCC-TAAC- 8376
MND-2.CM.98.CM16            AAT-T-T--AGTACTACAA-GCTATGC--AGGGGCTATGGAG--G--GGCACCAATG---ACT-TCA-CA-C-GC-TGCT-CAG-.............-CCAT-GC---G-GC---AT-A----C--A--------TC-------AGCA---GTC......C-CC-CAAC- 8907
MND-2.GA.x.M14              -AT-TGCC-AGTATTACAAAGCTTTGC--AG-GGCTATGGAG-GG--GGCACCAACTT--ACT-TCAGTTA---G--G--C--CG.............-CTAT-G-A--G--C---G-GA----C---C-GC-C--C-----A--AGCT---GTC......C-GC-CAAC- 8836
MND-2.x.x.5440              ATT-TATC-AGTATTACAA-GCTTTGC--AG-GGCTATGGAG--G--GGTACCAACTT--ACT-TCAG-T----G--G--C--C-.............-CCTT-G-C--G-GC--CT-GA----C---C-CC-C--C-----A--TGCT--GGCC......C-AC-CAAC- 8543
MNE.US.x.MNE027             CT-CT-CC-AG-AGCG-TCCAAGT-GCCTGG-GAT-TG-GA-AGAGA-TC-T--G-GC-C-TGG-GAGACT-ATGG--GACTC-G.........GG-A-G-T-G-A---TGG--A-T-GCA--C-----G--G--C-----A-AGC-T--GCTT......-C-C-CT-G- 8718
MON.CM.99.L1_99CML1         A-----CC-AGCAGCA-TCACCTCTC-C---G-C---CTT--T-GGTTCAC-ATC-TCT----G---GC-C-AC-TC-T-CT-G-.........GG-C--CTA--G----A--CGTGGCT--C--C---C-C--CC-------AGC-----TA......T-CC-TAAC- 8786
MON.NG.x.NG1                A--GY-CC-ARCAGCCTGCACAG-AGCCA-TG-----CTC--ACGGTTCAC--TCATCT-------CGC---CC-TCGCC-G-G-.........GG-C--TTA--GC---GAC-CGTGGCTG-C--C---C-C--CC------GGC---G-TT......C-CC-TAAC- 7411
MUS-1.CM.01.CM1239          A-----CC-AGCAGCA-GAACAGCAGCC--TG-CCAGCTC-----CTTCACCTGGA-CT-------GCCCA-AC-TC-GA-T-GT.........AG-C---TA--G----AAT-C-TGGCT--C--C---C-C--CC-------GGC-----TA......C-CC-TAAC- 8747
MUS-1.CM.01.SIVmus_01CM1085 A--G--CC-A-CAGCCAGCACAGCAG-C--TG--CAGCTC----CCTTCACCTGGA-CT----G--A-CA--AC-TC---CT-GT.........-G-A---TA--GC---AAT-CTTGGCT--C--C---C-CG-CC-G-----GGC---G-TA......C-CC-TAAC- 8778
MUS-2.CM.01.CM1246          AA----CC-AGCCGCA-TCTCAG-AGC----G-C---CTT--G--CTTCAC-TGGA-CT----C--ACC-A-AC-TC-GT-TAGT............--AA-ATA--G---AAAT-C-TGGCT--C--C---C-C--CC-C-----GGC-----TA......C-CC-TAAC- 8843
MUS-2.CM.01.CM2500          A---T-CC-AG-AG-A-TCACAGCAGCCA-TG---GGCTC-----CTGGACC--AA-CT-------CGCC--AC-TC----TAGT.........A--CA-CTA-AA-A---AT-CTTGGCT--C--C-A-C-C--CC-C-----GGC-----TA......C-CC-TAAC- 8799
MUS-3.GA.09.09GabOI81       A-----CC-AG-AGCA--AACAG-A--C---G-----CTT--T--CTTCACCTGGA-CT-------GGCC--GC-TC----TT-G.........-G-C---TGGCC----AAT-C-TGGCT--C--C---C-C--C--G-----AGC-----TT......C-C--TAAC- 8376
MUS-3.GA.11.11GabPts02      A--G--CC-AGCAGCA-TCTCAG-AA-C-----CCAGCTT--T--CTTCAC-TGGA-CT-------ACC-C-AT-TC-G-CTS-G.........-G-C---TA--G----AAT-C-TGGCT--C--C---C-C--CC-C-----AGC---G-TA......C-CC-TAAC- 8359
OLC.CI.97.97CI12            CTTG-A-C-CGTCT-GAGAACCTTAAAA-AG--GAGACA-..................CATGGGATC-ATT-GCAGC-GAAGAA-..........--GA--AGGG-A--TC--C-A-AGA-GGCT-AGGAGCA-T-GG-GAGA-AAGGA-C-GAA...ACA-C-GGAG-C- 8321
RCM.CM.00.SIVagi_00CM312    A--GT-C--AGCATTC--A--CT-GC-TGGTG-GATGG-G--AGA-AAT-CTTACT-CACCTGGAGAG--T-ACAGTCGC-G-C-.........-G-CA-CT-GCAG--TGGCCAGC-ACGTGTGGC-----G--C--G--A---C-C--G--A......C--C-CAAT- 8279
RCM.CM.02.SIVrcm_02CM8081   A--G--C--AG-GTTCAG--GCT-GC-TGCAG-AATGG-GCAACAGAAT-CTTACTTC-TCTGGAGAG-CT-ATGC-TTACT-C-.........----ACAT-GCA---TGGCCCG--ACGG-GTGC-----G--C--G-------C--G--A......C--C-TA-C- 8351
RCM.GA.x.SIVRCMGAB1         A--GT-C--AG-ATTCAGCAGCT-G--TGCTG-GATGG-GCTTCA-AAT-CTTACT-CACCTGGAGAG--T-ATGT-CG--G-C-.........-G--ACTT-GC-G--TGGCCAGC-ATGG-GTGC-----G--C--G-----C-C--G--A......C---GTAAT- 8450
RCM.NG.x.NG411              A--GT-C--AGGCTTCAGCAGCT-GC-TGCGG-GATGG-GAGG-CTAAT-CTTACT-CACCTGGAGAG-AC-ATGC-CG--TTGC.........AG--ACC--GCAG--TGGCCAGC-ACGT-GTGC-----G--C--G-------C--G--A......T--C-TAAC- 8533
SAB.SN.x.SAB1               A--G--C--AGCAGCA--A-CGTCAC-CACTGTG--TGTG-T-CA---T-CGA-AAGC-CC-GT---TCA--ACG-C-T-CTTGC.........AG-TCCAT-GT---G--AG-CAT-GCACAT--A--------GC-C--A-AGC-T--G--A......TGG--TAAT- 9085
SMM.CI.79.SIVsmCI2          AT-TT-CC-AG-AGCGTGCCAGGAAGCCGCACGG---G-GCGAGAGA-TC-T--AAGC-C-GGGAGAGACT-ATGG--GACTC-G.........GG-A-GAT-G-A----GGT-A-TTGCAGCC--G--G-----T--G-----AC----GCTC......GC---CT-G- 8385
SMM.LR.89.SIVsmLIB1         CTGGT-CG-AG--GCGCTCCAAGCGGCATGG-GG--TG-GAGAGAGA-TC-T--GAGC-C-TGGAGAGACT-ATGGC-GACTC-G.........GG-A-GAT-G-A----GGT-AT-GGCA--C-----GC-C--C--G--A--GC-T---CTT......-CC---T-G- 8303
SMM.SL.92.SIVsmSL92A        AT-CT-C-TCG-A--CCTGCAAGAAGCCGCA-TC--TG-GAGACA-A-TC-T-TCAGCAC-GGGAGA-CAA-ATGGC-GACTC-G.........GG-CACAT-G-A----GG--CGCAGCG--C--A--G--G--C--G-----A------CTT......-CAC-CAAT- 8338
SMM.SL.92.SIVsmSL92B        ................................................................................................................................................................................ 7460
SMM.SL.92.SL92B             TT-CT-C-TCG-AGCCTGCCAGGAAGCATGG-GAT-TG-GCAAGA---TA-T-TC-GC-C-TGG-GACTCA-ATGG--GACTC-G.........GG-C-G-T-G-A----GG--CGCAGCA--C--A--G--G--C---------C-T---CTC.......--GC-TAAC- 8689
SMM.US.04.G078              CTGGT-C--AG-AGCA-TCCAAGCGGC-TGG--AT-TG-GAGAGAGA-TC-T--AAGC-C-TGGAGAGACA-ATGG--GACTC-G.........GG-A-G-T-G-A----GGT-A-T-GCA--C--G--GC-C--C--G-----G--T--GCTC......-C-C-CT-G- 8534
SMM.US.05.D215              TTGGT-CC-AG-AGCGATCCAAGCG--CT----ATTTG--AGAGAGA-TC-T--AAGC-C-TGGAGAGACT-ATGG--GACTC-G.........GG-A-G-T-G-A----GG--CGCAGCA--C--G--G--G--C--G--A--G--T---CTC......-C-A--T-G- 8510
SMM.US.06.FTq               CTGGT-CC-AG-AGCA--ACAAGCGGCAT-T--A--TG-GAGAGAGA-TC-T--AAGC-C-TGGAGAGACC-ATGG--GA-TC--.........-G-C-G-T-G-AG--TGG--G-TTGCA--C--G--GC-C--C--G------C-T--GCTC......-C-C--T-G- 8526
SMM.US.86.CFU212            TTGGT-C--AG-AGCGCTGCAGGCAGCAGGA-GG--TG-G-GACAGA-TC-T--GAGC-C-TGGAGAGACT-ATGG--GACTC-G.........GG-A-G-T-G-A----GG--A-TTA----C--G--G-----C--G--A--GC-T--GCTC......-C-C-CT-G- 8508
SMM.US.x.H9                 TT-C--CC-AG-AGCA--GCARR---G-TGG--ATTTGYRCGAGAGA-TC-T--AAGC-C-TGGAGAGACT-ATGG--GACTC-G.........GG-A-G-T-G-A--G-GA--A-T-GCA--C--MR--C-C--C--G--A--GC-T--GCTC......-C-C-CT-G- 8730
SMM.US.x.PGM53              CTGGT-CC-AG-AGCA-TCCAGGCAGC-TGG-GG--TCGG-GAGA-A-TC-T--AAGC-C-TGGAGAGACTCATGG---ACTC-G.........GG-A-G-T-G-A---TGGG-A-T-GCA--C--G--G-----T--G-----AC-----CTT......GCAC-C--T- 9178
SMM.US.x.pE660.CG7G         TT-TT-CC-AG-AGCA-TGCAGGCG-G-TGG--ATTTG-GCGAGAGA-TC-T--AAGC-C-TGGAGAGACA-ATGG--GACTC-G.........GG-A-G-T-G-A----GG--A-T-GCA--C--G--GC-C--C--G--A--G--T--GCTC......GC-C-CT-G- 9246
STM.US.89.STM_37_16         CTGGA-CC-AG-AGCA--GCAAGCAGCATGG-GA--TG-G-GAGAGA-TC-T--GAGC-C-GGGAGAGACT-ATGG---ACTC-T.........GG-A-G-T-G-G---AGG--CG-GGC---C--A--GC-C--C-----A--GC-T---CTC......-CAC-CT-G- 8892
SUN.GA.98.L14               TATT-AGGCT-CG--CAGGA-C--GAACGCGTGCAGA--CAG--GG-AGAG--TATCATC-T-TC-G-A--CA-G-AGTAG-ACCTTTAGTCTCGGCA--AAA--G--GC-C-AAT-GA-T-G-A-A---G---GT---ATTCC-AGT--G-CCACAGAG-CCACTT-G- 9179
SYK.KE.x.KE51               A--G--C--AG-AGCA--CA-CTCGC--GGACGA-AAGTG--GCACCAGA-G-TG-CCATTTGGC-AGCCC-AC-C-CCTAT-CC.........AG-C---T-GC-GAGAA-G-CGCAGC-T----AC-----C-TC-------GC----G-TC......TACC-TAAC- 8420
SYK.KE.x.SYK173_COMGNM      C--G--C--A-TCGCA--A-C-TCGC-CGGA-GG-AA-T-CG-GATTGG--G--G-CTATTTGGC-AGCAA-ATAC-CC-CTACC.........AG-C---T-GTCGAG--CG--GCAGC-T-G--GC-G--GC-TC----A--------G-TC......TACC-TAAC- 8721
TAL.CM.00.266               T---A-CC-AGC-GC-TGGCGAGA-GGCA---CA--GCTT--A-GGTTCAC-TGGTG-T----GA--G-T--AT-G-CG-CT-CG.........-G-T-G--AG--GA---G--CGCAGCT--C--C--GC-C-CG--G--A--AC-T---TAT......TGC--TACT- 8653
TAL.CM.01.8023              T---A-CC-AGC-GC-TGG-GAGA--GCG---CA--GCTT--A-GGTTCACCTGGCGCT----GTC-G-T--AT-G-CG-CT-CG.........-G-T-G--AGTGCAG--A---GCAGCG--C--GC-G----CG--G--A--GC-T---TAT......TGCC-TGCT- 8203
TAN.UG.x.SIVagmTAN1         -C--T-C--AG-GG-A--A-CA--G-CTT-TG-G--TTTG--G-GC--T-CGCAAAGC-CC-GTCG--C-C-ATGGA-T-CTTGC.........AG-TCC----A-C-G--A--C-T-G-ACAT--A--------GC----A-AAC-----C--......TGG--CAAC- 8904
VER.DE.x.AGM3               ------C--AGCCGCA--ACAAGAAGCAG-TGTC--TTTG--ACGCCTT-C-CAAA-C-C-GGCT-CCA-A-ATGGCTT-CTTGC.........AG-TCC----A---G--A--CAT-A--TCT--A------G-GC----A---C-T---G-A.......--CC-TAAT- 8409
VER.KE.x.9063               ------C--A-CCGCA--ACAAGAAGCA-GTCTC--TGTG--ACGCTTT-C-CAAA-C-CAGGTC-CCAAA-ATGGCTT-CTTGC.........AG-TCCA---A---GCAC--CAT-AG--GT--A------G--C----A--GC-T--GGAA.......--CC-TAAT- 8911
VER.KE.x.AGM155             ---G--C--AGCCGCA--ACAAGAAGCAG-T--CCATCTG--G-GCTTT-C-CGCA-C-C-G--C-CCA-A-ATGGCTT-CTTGC.........AG-TCC----A-C-G--A--CAT-A--TCT--A------G-GC----A--GC-T--GGAA......G-CC-TAAT- 8888
VER.KE.x.TYO1_patent        A-----C--A-CCG-C--ACAAGAAA-A---C-AA-TCTG--AGGC-TT-C-CAAA-C-CATGTC-CCA-A-ATGGCTT-CTTGC.........AG-TCC----A---GAAC--CGT-A---GT--A------G-GC----A---C-T--GGAA.......--CC-TAAT- 8872
WRC.CI.97.97CI14            AT-TGG-GC--TAC-CAAGAAG-TTGCAAAG-GCAGG-AG-AG-ATA-GCAG--A----TAT---CA-CT-GC-GCAG-AAGAC-.........GC--C-AGCA-----ATC-GGATG-TTGCTA-C---CTCTGCGC-AGCTTTGATA--T-CAGAAGA-GC--CAGTG 8912
WRC.CI.98.98CI04            -C-TGC-TGC-TAC-CAAGAAG-TTACAAAG-GCAGG-AG-AA-C-A-G-AGAAA---ATAT----A-ATAGC-ACAG--AGAC-.........AC---TAGCA------TCTGGAAA-TT--TA-C---CTCTGCTC-AGCCTTAATA-G--CAGAAGA-GAC-CT-C- 8957
WRC.GM.05.Pbt_05GM_X02      -T-TGCTGC-CTACAG-AGAAG-TTGCTGCGCG-CGGCAGAAA-G-AAG-AGAAG---ATATACACGCCT-GC-GCAG--AGAC-.........AGCA--AGCA-C--CATC-GGG-AATT-CTG-C---CTCTGC-C-A-CCTTAATCCGT-CAGAAGAGGACA--AG- 7853
```

```
                             Env gp41, gp160 end        Nef start
H1B.FR.83.HXB2               AAG.........ATGGGTGGCAAGTGGTCAAAAGTAGTGTGATTGGATGGCCTACTGTAAGGGAAAGAATGAGACGA.....................................GCTGAGCCAGCAGCAGATAGGGTGGGAGCAGCATCTCGAGACCTGGAAAAACATGG 8918
Nef                                      _M__G__G__K__W__S__K__S__S__V__I__G__W__P__T__V__R__E__R__M__R__R__.__.__.__.__.__.__.__.__.__.__.__.__A__E__P__A__A__D__R__V__G__A__A__S__R__D__L__E__K__H__G
Env                          *_
H1O.CM.91.MVP5180            --..C......-----GAATGCA---AGC-----C-AAT-TGCA------T-AGAA-----A--T-----......AGA........CGATCCTCCTCTGATCCTCAA......C-A---TGT---CC-G-A--A-----T-TC--CA-G--GT-A-C--CTAGA-- 9027
H1N.CM.95.YBF30              ---.........-----AAAG-TT--------G--C--CC-AG-A--------AGAAA-C--A---------...A--AGA..................CAAACGCAAGAACCA--A-TAGAGC------TAG-A-CA--------T----A---T--A-CT--T-GA-- 8505
H1P.CM.06.U14788             --A.........-----GAATGCA---AAG--------CT-AG-G--C-----AG-A--C-------A---A...AAGCAGACTACCCCGACTACCCCTGACCCG...ACTACCC-A-TAA--C-T---CCCG----T--G-A-ATT--CAA---AT-A-C-C--GGAAA 8521
CPZ.CD.06.BF1167             --CA........-----A-CTGTACTAGG--CC--ATTCCCAGGGAT-GACAAC-TGT-G-ACAG--TT.........AGA..................................AACACA------------GCACCA--GT--CA-GA-......T--C-T--CAGA-- 9116
CPZ.TZ.00.TAN1               --A.........-----AAA--TA-TTGGT-G-T-GCC-......--GGCC-GG-AA-CC-TC---GATC-T.......................................CA-A-CA-CT--AGT--GCCT--A---CAG--C--A-A------CC-G--TA-A-- 8604
GOR.CM.07.SIVgorCP2135con    --..........A-----AAA--CA---AA---G-----CT-C--A--------AG-A------------------A-ACC..........................AGCCCT-AGCCAAAGC-G--TCCGG-------T-A--TG---A---TGT-A-C-C--AGG-- 8499
                             Env gp41, gp160 end
MAC.US.x.239                 G-.........................................................................................................GGGACAGAAAT-CAATCAG-G-C-GTATA--AATA-TC---GGA--A---CA-CTG--G-GA- 9306
Env                          *
Nef                          _E__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__. G  Q  K  Y  N  Q  G  Q  Y  M  N  T  P  W  R  N  P  A  E  E  R
H2A.SN.86.ST_JSP4_27         G-.........................................................................................................GGGACGGCGGT-T-A-CAG-G----TTTA--AATA-CC---GGA---C--CA-C-...G-A-- 8788
H2B.CI.x.EHO                 G-............................................................................................................................................................T-G-C--A-GAC 9239
H2G.CI.92.Abt96              G-.........................................................................................................AGGGCAACAAT-CA-GCAG-GTC-ATTTA--AATA-CC---GGA--A---CA-C--C-G-AA- 8734
H2U.CI.07.07IC_TNP3          G-..............................................................................................................................G-GACA--CATATAGT-A-GGG-AGTT-A--A-T-CC-CGT- 8800
H2U.FR.96.12034              G-...................................................................................................AGCTCAGGGCCAA-GCT-TAGT-A--G---GTTCA--AATA--C-T-GGA--AC--CA-C--C-GGGAA 8830
ASC.UG.10.RT03               -G...............................................................................................................................................................-G--G-AAC 8295
ASC.UG.10.RT08               --...............................................................................................................................................................GG--G-A-C 8328
ASC.UG.10.RT11               G-...............................................................................................................................................................-G--G-A-C 8289
COL.CM.x.CGU1                --................................................................................................................................................T-GC-CTTG--C-CTTTGCT--A- 8175
COL.UG.10.BWC01              --................................................................................................................................................T-G--C-TGT-CTTTGTGCT--C- 7995
COL.UG.10.BWC07              --................................................................................................................................................T-G--CGTGT-C-CTTTGCT--C- 8196
DEB.CM.04.SIVdeb04CMPF3061   G-......................................................................................................................................................G--AACCTTT-T-GGAAA 8668
DEB.CM.99.CM40               G-............................................................................................................................................................-TTCTGGGGAA- 8571
DEB.CM.99.CM5                G-......................................................................................................................................................G--A-CACTCT-TGCA-A 8498
DEN.CD.x.CD1_CMO580407       -G.........................................................................................................GTTACTG-AGC-TGAT-AG-AC--AGAAACCTTT-T--ATG-AGA-...GAAATC-TC--ATC 8915
DRL.DE.11.D3                 -G.............................................................................................................................................CAGC-TAGATACTACTCCT-T---A-C 8549
DRL.x.x.FAO                  -G.............................................................................................................................................CAGC-TAGATACTACTC-T-TG--A-C 8633
GRV.ET.x.GRI_677_gri_1       --................................................................................................................................AG-TAAAATGA-TC--...................G-C-- 8787
GSN.CM.99.CN166              G-........................................................................................................................................................................ 8771
GSN.CM.99.CN71               G-........................................................................................................................................................................ 8771
LST.CD.88.SIVlhoest447       -G........................................................................................................................................................................ 8035
LST.CD.88.SIVlhoest485       -G........................................................................................................................................................................ 8032
LST.CD.88.SIVlhoest524       -G........................................................................................................................................................................ 8035
LST.KE.x.lho7                -G........................................................................................................................................................................ 9110
MAC.US.x.251_1A11            G-.........................................................................................................GGGACAGAAAT-CAATCAG-A-C-GTACA--AATA-TC---GGA--A---CA-CTG--G-GA- 9302
MND-1.GA.x.MNDGB1            --........................................................................................................................................................................ 8378
MND-2.CM.98.CM16             G-........................................................................................................................................................................ 8909
MND-2.GA.x.M14               --........................................................................................................................................................................ 8838
MND-2.x.x.5440               G-............................................................................................................................................................-CGGG-T-TAA- 8557
MNE.US.x.MNE027              G-.........................................................................................................GGGACAGAAAT-CAATCAG-G-C-GTTTA--AATA-TC---GGAA-A---CA-CTGG-G-GA- 8783
MON.CM.99.L1_99CML1          G-........................................................................................................................................................................ 8788
MON.NG.x.NG1                 G-........................................................................................................................................................................ 7413
MUS-1.CM.01.CM1239           G-........................................................................................................................................................................ 8749
MUS-1.CM.01.SIVmus_01CM1085  G-........................................................................................................................................................................ 8780
MUS-2.CM.01.CM1246           G-........................................................................................................................................................................ 8845
MUS-2.CM.01.CM2500           --........................................................................................................................................................................ 8801
MUS-3.GA.09.09GabOI81        G-........................................................................................................................................................................ 8378
MUS-3.GA.11.11GabPts02       G-........................................................................................................................................................................ 8361
OLC.CI.97.97CI12             G-.........................................................................................................................................G-ACCAGCA-.......-AAAC---T-TG-C 8347
RCM.CM.00.SIVagi_00CM312     -G........................................................................................................................................................................ 8281
RCM.CM.02.SIVrcm_02CM8081    -G........................................................................................................................................................................ 8353
RCM.GA.x.SIVRCMGAB1          -G............................................................................................................................................................A-G-GGCAT-AC 8464
RCM.NG.x.NG411               -G........................................................................................................................................................................ 8535
SAB.SN.x.SAB1                -G........................................................................................................................GCA--G---CTCCTCTA--A---A-..................G-A-- 9117
SMM.CI.79.SIVsmCI2           G-..............................................................................................................................--GGCAAACTTATT---AGGGG-A-TTTA--A-T-CT-C-T- 8429
SMM.LR.89.SIVsmLIB1          G-..............................................................................................................................GCCACAAAGATACAATCAGGGAGACTT-G--A-T-CT-C-T- 8347
SMM.SL.92.SIVsmSL92A         G-.................................................................................................................................GCA-AAATACAGCC--GGGGA-T--A--A-T-CC-CAT- 8379
SMM.SL.92.SIVsmSL92B         .......................................................................................................................................................................... 7460
SMM.SL.92.SL92B              -G......................................................................................................GGATGGATATAAAT-CAAT-T--GG--ATTTA--AACA-CC---GGA--A---CT-CT-C-G-AA- 8757
SMM.US.04.G078               G-..............................................................................................................................--C-CA-AAATATAGT-AGGGA-A-T-TA--A-T-CC-CCT- 8578
SMM.US.05.D215               G-.................................................................................................................................GCAA--TTACAGT-AGGGAGATT-TA--A-T-CC-C-T- 8551
SMM.US.06.FTq                G-..............................................................................................................................G-GAGA-AGTTACAGTCAGGG--AGTTTA--A-T-CC-CCT- 8570
SMM.US.86.CFU212             G-..............................................................................................................................GCC-CA-A--TATAAT-A-GGC-AGT--A--A-T-CC-CCT- 8552
SMM.US.x.H9                  G-.........................................................................................................GCCTCAGAAAT-TTGT-AG-GTC-ATTCA--AATA-CC-T-GGA-GA---CA-C--C-G-A-- 8795
SMM.US.x.PGM53               G-.........................................................................................................GGGACAG-GAT-CAGT-AG-G-C-ATTCA--AACA-CC---GGA--A-T-CA-CT-G-G-GA- 9243
SMM.US.x.pE660.CG7G          G-..............................................................................................................................GCC-CA-AAATATAGT-AGGGA-A-T-TA--A-T-CC-CCT- 9290
STM.US.89.STM_37_16          G-.........................................................................................................GCCTCAGAGAT-TAAT-A--GTC-ATTTA--AATA-CC-T-GGAA-A---C--C-GC-G-AA- 8957
SUN.GA.98.L14                --........................................................................................................................................................................ 9181
SYK.KE.x.KE51                G-........................................................................................................................-C-CATCTTGGATCCCAA-AG--AC.........T-T--CTCTGGG-- 8461
SYK.KE.x.SYK173_COMGNM       G-..................................................................................................................GCCCATT-ACC--C-CG--CC--ACAG--AC-GG......GGAC-TTCGGGA-- 8771
TAL.CM.00.266                G-..............................................................................................................T-AA--ATGGCA----ACAG-A-C--ACTACATG......................-- 8691
TAL.CM.01.8023               G-..............................................................................................................T--A-GATGGCA----TCCGAC-CAAATTGGATG......................-- 8241
TAN.UG.x.SIVagmTAN1          -G.............................................................................................................................GAGGG-AAA-A-CA-CC-T....................G-A-- 8930
VER.DE.x.AGM3                -G........................................................................................................................GAAGAG-A--G-TAAAATGA-TC--..................G-A-- 8441
VER.KE.x.9063                -G........................................................................................................................GAAGAGGA--G-CAACATGA-TC--..................G---- 8943
VER.KE.x.AGM155              -G........................................................................................................................GAAGAG-A--G-CAACATGA-TC--..................G-A-- 8920
VER.KE.x.TYO1_patent         -G........................................................................................................................GAA-CA-A--G-CAACATGA-G-A-..................G-G-- 8904
WRC.CI.97.97CI14             --AGAAGACAG............................................................................................................................................................... 8923
WRC.CI.98.98CI04             --........................................................................................................................................................................ 8959
WRC.GM.05.Pbt_05GM_X02       --........................................................................................................................................................................ 7855
```

```
H1B.FR.83.HXB2            AGCAATCACAAGTAGCAATACAGCAGCTACCAATGCTGCTTGTGCCTGGCTAGAAGCACAA.........GAGGAGGAGGAGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGG 9079
Nef                        A  I  T  S  S  N  T  A  A  T  N  A  A  C  A  W  L  E  A  Q  .  .  .  E  E  E  E  V  G  F  P  V  T  P  Q  V  P  L  R  P  M  T  Y  K  A  A  V  D  L  S  H  F  L  K  E  K
H1O.CM.91.MVP5180         G-G---AT-----TC-C-C--TC-TCAA-A------A--CCT---A-TC-----CAGC--CAAAGAT.........-----T--A--C--C----A-G------A--G---C----G--------C-TT--A-----CT-T--C--C---TT---------------- 9188
H1N.CM.95.YBF30           G--C-----C-TA--A-----TAG--AC-AT----AAAG-ATA--T------------...CAAGAA...--A--A-----A--A--C--------ACGC-----------A-----G-----A--C--T--ACAG---T-T------TC-TT------------T---- 8669
H1P.CM.06.U14788          --G---AC-C----AATT--GTT--AAG-A------A--A-TG----TCT-G--T--T--TGAGGAA.........--A--A--A--R--C--------GG-----A-----C------TGC-----A-----------AT-T--C--C---TT---------------- 8682
CPZ.CD.06.BF1167          --GCC-T......-CA-CA-ACA-CATAGGA-CA-AACA-GA--TA-TA-A-C-TT--GG-GACCAT...ACA--A--A-CA--A--C--C-----A-G------A--G--AA-G-----T-----AG-A--ATTG--A-----C--CTCATGG--------------- 9277
CPZ.TZ.00.TAN1            --GTC----T-C--A--CCCT--GTA-CT-AGCA-A--TG-TA-AA-ACTCT-C--AC--TACTGAA.........--A--A--A-------------G---AGCA-----CA-G-----C-----AG-G---CT---AA-------GTCATGG--C-------------- 8765
GOR.CM.07.SIVgorCP2135con --GG--AC-TGCA-AAT----TC--CGA-A----CAA----TA--G-TC--------T--TGAAGAT.........-------A--A---------GY------A--G--------G--------C-TT--A-GG--AT-T--C--C---TT---------------- 8660
                                                                   Premature stop in SMM239
MAC.US.x.239              --A--AATT-GCATA--GA-A-CA-AA--TGG---A-ATA...-ATGA-TA-..................--T--T--CTT---A--GG-AT----G-GG--AA-A--T--CC-----A------G------ATTG--AA----CA-GTCT--T---A------------ 9455
Nef                        E  K  L  A  Y  R  K  Q  N  M  D  D  I  .  D  E  *  .  .  .  .  .  .  .  D  D  D  L  V  G  V  S  V  R  P  K  V  P  L  R  T  M  S  Y  K  L  A  I  D  M  S  H  F  I  K  E  K
H2A.SN.86.ST_JSP4_27      G-AG-AAGGCTCGTA---GCA-CA-AA--TGG---A--TA...-AT-CAGAT..................--T--T--CCTA--A--GG-C--T--------AAGA-----A------GA------A--T-G-TTG--AAG----A-GTCA--T--GA----------- 8937
H2B.CI.x.EHO              TTTGGGAG--GAAG-GGGAGG-AA-CAAGATTCA-A-.......-AGGATGAT.................-----CA-T--A--A--GG-C-GT--A-G---CGG---C--AC-G--G--------A-T---ACT---A-----CA-GTCT--T---------------- 9385
H2G.CI.92.Abt96           -----AGTT-GCATA--GGYAGCA-AA--ATG---A--TA...-ATA-TGAT..................--TA-T--CTTA--A--AG-A------TAC--CA-A--C--GC-----GT-----G------ATTG--AA----CA-GTCA--T---A------------ 8883
H2U.CI.07.07IC_TNP3       GAG--A-C--GCAG-AG-AGGG-A-AAGCTA--ATA-AGACAACAGAACA-G--T-ATGT-GATAGTGAT--T--T--CTTA--AAAAG-AT----A-GG--AAG---T--A--G-----------------ATGGC-A-----CA-GTCA--T---------------- 8970
H2U.FR.96.12034           --AT-AATT-CAATAT--GCA-CA-AAC-TGG---A--TA...-ATGAAGA-..................--TA-T--TTTA--A--AG-AG----A-G---AAGA---G-AC-G---A-------------ATTG--A------A-GTCA--T---A----------A- 8979
ASC.UG.10.RT03            ---CGAG---GCC-AG-GAGGG-A--T-GTACC--TAAGA-ACAA-GAAGGT--CAGCACCCGAGATGAT--T--A--A-GA--------C-----ATGC--A--A-----CC-C----A------A-T----A-G-----G---T-CTC-TGG--------------- 8465
ASC.UG.10.RT08            ---CCAAC--G-AGAAGTAGT-C-CTTA-GGT-CAAC-AGG---A--ACACCCG-A-T.........GAT--T--A--A--A--------C-----ACGC--A--A-----CA-C---A-------A-T---AA-------G---T-CTC-TGG--------------- 8489
ASC.UG.10.RT11            ----GCAG-C-AG--AGGAGA--TT-TAC---TGAGGTACAA--AAG-TGACA-CA-C-G-GATGAT...--T--A--A--A--------C-----ACGC--A--A-----CA-C---AACC----A-T---AAT------G---T-CTC-TGG--------------- 8456
COL.CM.x.CGU1             -CATCCAGATGTATATT-CCACAAC.........C-A-AACCA-AGGACG--...................--A--T-TA--A-CA--A-G-TAT--G-TGT---GAC-G--AG--C---AGGC---CATT---CT-CTG--G--CA-CTCTTG------G----------- 8318
COL.UG.10.BWC01           CCATCCAT-TGTGTA-T--CATT--.........C-A-AGCCA-AAGATGC-..................--T--A--A--A-GA----GCTTT--GCA-G---GA--T---G------A-GGA--AAT----CTTTTA-----CA-CTCTTG------G----------- 8138
COL.UG.10.BWC07           CCATCCACGGGTATA-TT-CACT-T.........C-A-AACCA-AAGAT---..................-GA--AA-T...AACA-C-G-TTC--GTGTG-C-GA-----AG-C----G-CCC--CAT---ATTGTTGT-T--CA-CTC-TT------G----------- 8336
DEB.CM.04.SIVdeb04CMPF3061 TC---GA-G-GAA-T---GTGTA--CAG-GAC-ACAACAAGA-TTTGAC...................C-AC-A-----A-----A-----T----A---CAG---C--AA-T---GACCCC--C--T--ATT-ATGA----CTACTCT--T---------------- 8817
DEB.CM.99.CM40            TC-T-GA-G-GAA-TA--GTGCA-CAA--AGC-ACAACAAGACTTAGAT...................C--ACA--A-CA--A--C--C----ACGC--C------T-CA-CC--GACCCC--A--T---TT-ATGA----CTACTCT--T-----G----------- 8720
DEB.CM.99.CM5             CC-CCA--G-GAG-TA--ATGCA-TCAG-GAGCCCAG-AG---TATGAT...................-----AC-T-G-A-A--C---G-T----G---AAG---C--AA-T--GGA-CCC---------ATTGATGA----CTACTCT--T-A------------- 8647
DEN.CD.x.CD1_CMO580407    -T-TCCAG--CTGGA-CC-CA-CATC-A-GA.........ATACA-CAATAT-TG-T-G-C...CATGAT--T--A--A--A--A--C--------ACGG--A-GACG---AC-GCACG--CC---C--T--A-ATATGA-----A-GTCT--T-----G----------- 9073
DRL.DE.11.D3              ----GCAG--GCA...G-G-A--AT-TC---CT-C-CA-AGAA-AAGAA-A-CCCT--......GAAGAA--A--A--A--A--A--C--------ATAT-----ATGC--G--G---GA-CCA---------A-AGTTA-----CT-CTC--------------------- 8710
DRL.x.x.FAO               ----GCAG--......G-G-A--AT-TC---CT-C-CA-AGAA-AAGAA-A-CCCT--......GAAGAA--A--A--A--A--A--A-----------ATGT-----ACGC---------GAGCCA--C-----A-ATCTA-----CT--TC------------G-------- 8791
GRV.ET.x.GRI_677_gri_1    CCGCCG-CTGCAAGAAGGAGACA-CTT-...G---AGTGG...-ATGATGAT..................--A--A--A...--A--C--C--T--GCA-----GA--C--C-------AG-----C--T--ATT---A--G--CT--TC-------------TC----- 8930
GSN.CM.99.CN166           ........................--AGC-A--AC-CAAGA-A-GAA-C--C...................-----T-T--ACCATCA-GC-----A-G---A-G------AC-T---GACCCC--C-----A-TCATG--G-----CTCT--------G----------- 8899
GSN.CM.99.CN71            ........................--AAT-ACCCT-CAGGCAAAAT--CA-C...................--TT-T--ACCATCA-GC-----G-G---A-----G--GC-T---GACCCC--C-----ACTTATG--G-----CTCG--------G----------- 8896
LST.CD.88.SIVlhoest447    .....................................................................................................................CCCT--C--T---A-TCTGA----C--CTCA--T---A------------ 8085
LST.CD.88.SIVlhoest485    .....................................................................................................................CCCT-----T---A-TCTGA----C--CTCA--T---A------------ 8082
LST.CD.88.SIVlhoest524    .....................................................................................................................CTCT--C--T--AA-CTTGA----C--CTCA--T---A------------ 8085
LST.KE.x.lho7             .....................................................................................................................CCCT--C--T---A-TCT-A----C--CTCA--T---A------------ 9160
MAC.US.x.251_1A11         --A--AATT-GCATA--GA-A-CA-AA--TGG---A-ATA...-ATGA-GA-...-ATG-T...............--TTT---A--GG-AT----G-GG--AA-A--T--CC-----A------G------ATTG--AA----CA-GTCT--T---A------------ 9451
MND-1.GA.x.MNDGB1         ............AGATG-G-A-CTT-ACTTA-CACAGCAGGA-C-AGAAGAG...................-----A--A...--T--A-----T--GTGT-GC--A--TT-C------GTGCCAT-A-----A-ATCTGA----CT-CTCT--T---A------------ 8515
MND-2.CM.98.CM16          ....................................-GAAGAG..................--A--AAC-.......--C--------ATAT--A---TGC--AG-----GA-CCC--C-------ACTTG--T---A-GCC-------------------- 9010
MND-2.GA.x.M14            ...................................AGTA-ATGATGAGC--..................A-A--A--A...AG---C--------TTAT--CACACAG---G-GC-TGA-GCA--C-----AAACCTGA-----A-GTC-------------------- 8951
MND-2.x.x.5440            GCACCAGCAGGA-GTA-C-CTCC-CT-AGAA--A-AGCAACCAAGTGATGA-..................-----A-----...--A--C---------TAC--AAC-CAG--AG-GCCTGA-GCA--C-------ACTTGA----CA-GTC-------------------- 8706
MNE.US.x.MNE027           G-A--AATT-GCATA--GA-A-CA-AA--TGG---A-ATA...-ATGA-GA-..................--T--T-GCTT---A--GG-A-----G-GG--A-GA--T--C------G-------------ATTG--A------A-GTCT--T---A------------ 8932
MON.CM.99.L1_99CML1       .........GG-C--A---GGCAA-CAACG-CGGCAAC--..........-G--C-AC......GACGAC--C--C--C-GT--C----G------A-GG--C-GA--T--GC-TC-CGACCCT----GG---CTCATGA-G--C--CTCT--T-A---G---------- 8934
MON.NG.x.NG1              .........GG-GC-G--CCACCGCT-A-A-CGCCGCC-AACCA-TGTCG-T--T-AC...GCCGACGAG--CC-A--TTT-CCT--C-G---G--TCGC---AT---G--AC-T---GACCCT--A-GG---ATCATGA-G--C---TCT--T---------------- 7571
MUS-1.CM.01.CM1239        .........-G-A------GAGCGGC-ACT--TC..................................GAA--AC-T---...CGTTCA-GC-----GCGG--G-GA--G--CC---C-GGATCCC--A------ATGATGA-T--C--CTCT--T-A-------------- 8874
MUS-1.CM.01.SIVmus_01CM1085 .........-G-AGC--G--A-AG-C--GTG-TA..................................GAA---C-T---...CGTACC-GC--G--GCGT--GAG---G--AA-CC-TGACCCC--C------CTGATG--G---A-GTCG--T--C------------- 8905
MUS-2.CM.01.CM1246        ..........--AC-AGGGCCGCAGGCAA-T--TA.................................GAA--AC-T---...TGCTCC-GC-----ACGC--GA-A--C--GA-TC-TGATG-C-AC-----ACTCATGA----C--CTCT----A---G---A-----A 8970
MUS-2.CM.01.CM2500        .........GG-A--T--C-A-AA---A-T--TA..................................AAAA-AC-TA-A...CG-TCG-G------ACGC--G--ACAC--AA-CCAGAACCCC--C--T---TT-ATAA--A-C--CTCT--T-----G---------- 8926
MUS-3.GA.09.09GabOI81     .........-G-G-AAG-CCGCAG-C-GGT--T-..................................GAC--A-----T...CGTTCC-G--AT--G--T--AA-ATGTGGCC-GC-TGACCCC--C-T----CTTATGA-T--C--CTCT--T-----G--------A- 8503
MUS-3.GA.11.11GabPts02    .........-G-A-AAG-C-GGAG-C-AGT--TC..................................GAG--ATGT----...AGAA-C-G------G-GG--AAGATGTGG-A-CCA-G--G------TT---AG-TTGA-T-----CTCT--T-----G---------- 8486
OLC.CI.97.97CI12          -AT-GAGT-TTTA-TG-C-GGG-AC....................ACTGAC..................AC---T--TCT--AA--A--------G-AG--------G--AG-GC---A-GCC-----T---A-TCTAA-T--G--CTCG--T---------------- 8478
RCM.CM.00.SIVagi_00CM312  --G-GGGTTT-C--TAG-A---CAGCAGT-AGTA-A----ATG-AT---TAT--G-AG.........ACA--T--CACCTT---A--A----------AG--A--A-----A------G------G--T---ATTG--AA----CA-GTC-------------------- 8442
RCM.CM.02.SIVrcm_02CM8081 G-G-GGA-T--CA-TAG-A---CAGCAA-GTGTA-A----ATA-AT---TAT--G-AG.........ACA--T--TACCTT---A-----C-----G-A---A--A-----A--------T-------TT--ATTG--AA-------CTCA--T---------------- 8514
RCM.GA.x.SIVRCMGAB1       ----GAA--GCAAGCTTCA-T--AT-AC...-T--ACTGG-A--AAGATACT..................--T--CACCTT---A--A--------G-A------A-----AC-T----------G-------CT---AA----CA-GTCT------------------- 8613
RCM.NG.x.NG411            ......AGG-G-GTTT-CA-T--A-A--CAGC-GT-A-TGGA---AATAGATTGGTATG--...GATACA--T--TACCTT---A--G--------G-G---A--A--C--CA----G-------T-------ATT---AA-------ATCT------------------- 8696
SAB.SN.x.SAB1             T-T--AA-TG-AACATC-AGG-AG-CAG...CCCT-GTGG-A--ATGA-GAT..................--A--A--A...--A--C--------A-GG---TGCT----AC--C--G-------A--T---TTG--AA----CT--G-------------------- 9263
SMM.CI.79.SIVsmCI2        GAA--A-C--GT--AAG-A-G--G-AAACT---GTA--GGA-ACAA--CA-GA-TAATGT-GGTGAAGAT--TA----CCTA--A--AG-AG-T--G-G---AA-AA----AC-G--GGA------------ATTG--AA----CA-GTCT--T---A------.---A- 8598
SMM.LR.89.SIVsmLIB1       GAG--A-C--GC---GG-A-A--A-AAA......TA-AGGCAACAAAATA-G--T-ATGT-GATGAGGAA--T--T---CTAA-A--AG-GG------TGAA-AGAA-C---C-----G-----T----T--ATTG--AA----CA-GTCA--T---A------------ 8511
SMM.SL.92.SIVsmSL92A      GAG--A-C--GC---AG-A-G--A-AAG.............G-AGAATA-G--T-ATAT-GATGAGGAA--T.........--T--GA-AG----A-GG---A--A----A-----GG-------A-----ATTG--AA----CA-GTCA--T---A------------ 8526
SMM.SL.92.SIVsmSL92B      ............................................................................................................................................................................ 7460
SMM.SL.92.SL92B           -CA--AGCA-CAGTAT-GGCAGCAGTG--TGG---A-ATA...-ATGA-GAT..................--T--T---CT--CT--AG-AGTT---TGG--AA-A-----CC-----G-----T-A-----ATTG--AA----CA-GTCA--T---A------------ 8906
SMM.US.04.G078            GAGG-A-C--GCA-CAG-A-G----AAATTAGG-TA-AGACAACAAAACA-G--T-ATGTGGATGATGAA--T--T--CTTA--A---G-CT----ATAC--A--T-CC--C-----GG-C---T-A-----ATTG-TGA----CA-GTCT--T---A------------ 8748
SMM.US.05.D215            GAG--C-C-TGCA--AG-G-G--AGAAACG--TGTA-AAGCAACAGCAAA-G--T-ATGT-GATGAGGAA--TA-T--CCTAA-A--AG-GG-T--ACGC--AA----C---C-G---G-----G-------A-----ATTG--GA----CA-GTCT--T---A------------ 8721
SMM.US.06.FTq             GAG--A-C--GC---AG-G-G--A-AAATTGGCATA-AGGCAACAAAATA----T-ATAT-GATGATGATA-T--T--TTTAA-A--AG-A-----ACACT-AA-A--C--C------TGC-----------ATT---A-----CA-GTCA--T---A------------ 8740
SMM.US.86.CFU212          GAG--A-C-TGCA--AG-A-G---CAAATTAGCATA-AGGAAACAAAATA-G--T-ATGT-GATGAGGAG--T--T--ATTA--A--GG-AG----ACAT--AA-A--T--A------G-----T-A-----ATTG--AA----CA-GTCA--T---A------------ 8722
SMM.US.x.H9               -----AATT-G--TAT-GACA-CA-AAC-TGG---A--TG...-ATAATGA-..................--T......TTA--A----G------G-GC--AAGA--C--CG-----ATC-----A-----ATTG--AA----CA-GTCT--T---A------------ 8938
SMM.US.x.PGM53            --A--AGTT--AATAT-GACA-CA-AAC-TGG---A--TA...-ATGATGAT..................--T--T---CT-A-A--AG-AT----ACAC--AA----G--A------G-C---T-A-----ATTG--GA----CA-GTCT--T---A------------ 9392
SMM.US.x.pE660.CG7G       GAGG-A-C--GCA-CAG-A-G--A-AAATTAGG-TA-AGACAACAAAACA-G--T-ATGTGGATGATGAA--T--T--CTTA--A---G-CT----GCAC--AAGA--C--C-----GG-C-----A-----ATTG--AA----CA-GTCT--T---A------------ 9460
STM.US.89.STM_37_16       C--T-AGCT-GAATAT-GACAGCA-AAC-TGG---A--TA...-ATGA-GA-..................--T--TA-TCTA--A--AG-AG----ACAT--AAGA--C--AC-T--GGA-------------TTG--AA----C--GTCA--T---A-----AGC---- 9106
SUN.GA.98.L14             .....................................................................................................................GCCA--A--T---CAGCTAA-------CTCA--T---A------------ 9231
SYK.KE.x.KE51             GAA--AATGG-A-GCGGGAGACATCTG-...C--..........-ATGA-GG-..................--T-----TCTA--A--C-----T--TTT---A--A-----C--G-----CT-A--------ACTT--CA-------CTC---T---A-----AG----- 8601
SYK.KE.x.SYK173_COMGNM    -CA--AATTCTCACCTGGAGATAT--TG...C-G..........-ATGAAGG-..................--CACA-GTCTC--A-----C--T--GTGT-----AAC---AC-C---A--T-A--C-----ATTG--AA----C--CTCT--T---A-----A-C---- 8911
TAL.CM.00.266             CAAGGCATT--CACAAGGAGAG-TTTGGTATG-AAAA-AACAA-AAGAAAAT..................--CAT---ACTA-----C--C--T--G-GG---A-A--T--G--G---A-CTGT-----T--ACT--G-A----CT-CTCGTT------G----------- 8843
TAL.CM.01.8023            GAAGGCACTCTCGGAAGGAGA--TCTGGTATG-AAAG-AGCAA-AAGAAAAT..................--CAT----CTT--A--C---------A-GG-----A--T--A--G--G--CTGT-A------ACTG-G-A----CT-CTCGTT------G--------A- 8393
TAN.UG.x.SIVagmTAN1       -AG--AATTGGCAGCTG--GACA--TGG...G---ACTGG...-AACCAGAG..................--A--A--A...------C---------A-G---A-GA--G----------A-------A--T---TTG--A------T-CTC-------------------- 9073
VER.DE.x.AGM3             -AG--AATT-CAAGAAGGAGATAA-TGG...G---AATGG...T-TGATGA-..................--A--T--A...--A--A---------A-G---AAGA--G--GC------A-------C--T--ATT---G--G--CT--TCG------------------ 8584
VER.KE.x.9063             -AG-CG-CTGCAAGAAGGAGAC-A-TGG...G---AATGG...T-AGATGAT..................--A--T--A...------A-----T--A-A---AAGA-----A--G--G-A-------------ATT---A--G--CT--TCG------------------ 9086
VER.KE.x.AGM155           -AG-CGTCT-CAGGA-GGGGACCA-TGG...G---AATGG...T-AGATGA-..................--A--T--A...------A--------A-G---AAGA--G--AC------A---A--A-----ACTT--A------T--TCG------------------ 9063
VER.KE.x.TYO1_patent      -AGG-GGCTTCAAGAAGGAGACA-CTGG...G-A-AGTGG...T-GGATGAT..................----A--A...------A--------G-G---AAGA-----C------G-A------------T--ACTT--A--G---T--TCG------------------ 9047
WRC.CI.97.97CI14          ....................................-AGTGATGA-..................--T--A-GA...--A-----------A--T-----A--T--CA------A-CCA---TT--ACTGCTA-----C--CTCG--T--A--T---------- 9030
WRC.CI.98.98CI04          ................................A-AAGA-AGT-CTGAT..................--T--A--A...--A--C--------TCAT--------T--AA------A-GCT--A-TT--ACTGCTA--------CTCG--T--A--T---------- 9072
WRC.GM.05.Pbt_05GM_X02    ................................A-AAGACT-A-CTGAT..................--T--T-CA...A-A--A--C--T----G---A--A--T----------AGGCA--A-T----ATCTTA-----C--CTCT--T--A--T---------- 7968
```

```
                                   3' LTR U3 start                                  Nef premature end in HXB2
H1B.FR.83.HXB2             GGGGACTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGG....................GTCAGATATCCACTGACCTTTGGATGGTGCTACAAG 9228
Nef                        G_G_L_E_G_L_I_H_S_Q_R_R_Q_D_I_L_D_L_W_I_Y_H_T_Q_G_Y_F_P_D_*_Q_N_Y_T_P_G_P_G_._._._._._._._._._V_R_Y_P_L_T_F_G_W_C_Y_K
H1O.CM.91.MVP5180          -A--------T---T-----T-------T-AG---GC---A-----G-----C-----A--T-----T--G--A-T-----------G----TGT--------G--A-----A....................CCT----TC--------A----------TG-TT--A 9337
H1N.CM.95.YBF30            -----------------G--TGG---AG--A---G--------T--A--C--C-----G--T------------AT-C-------C-G---T--------------------A....................A-T-----C--CG-A----------------T---A 8818
H1P.CM.06.U14788           -A--------T---T--G--T----A-CTGAG---GC---G-----A-----C--------T-----T--G--A-T---------C-G------T-----T-----------A....................-AA--------C-----------G---CTG-TT--A 8831
CPZ.CD.06.BF1167           -----------------TA-TGG--T-C--AG---GC--CA--A--A---ACC-----G-TTA----T-----AATA-----------G------T------------------A....................--T--------T---TGTAGA--T----TA-TT--- 9426
CPZ.TZ.00.TAN1             ----------T------T--TT---T-C--A----GC--CC-----A--CACC-----G--TA-T-----G--TGT---T--A--C-G-----------C--T--A-----A....................A-------C------TGTAGG-------TA-TT--- 8914
GOR.CM.07.SIVgorCP2135con  --------------T--G--TT---T--GGCC---AGT--G---T-A-----C---G-T--T--T--T--G--A-----------C-G------T--T---------------....................A-A--------C---------------CTG-TT--- 8809
                                   3' LTR U3 start

MAC.US.x.239               --------------A-TTA-T--AGTGC---------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-A-A--A---G----G-T-----CT----A-----A....................A-T-----C---AA---A-----C---CTA-GG--A 9604
Nef                        G_G_L_E_G_I_Y_Y_S_A_R_R_H_R_I_L_D_I_Y_L_E_K_E_E_G_I_I_P_D_W_Q_D_Y_T_S_G_P_G_._._._._._._._._._I_R_Y_P_K_T_F_G_W_L_W_K
H2A.SN.86.ST_JSP4_27       ---------------GTA-T--AGTG-T--G----GTAGAG----A--CA-A-ACT-AG-AA-GGA-G-G--AATAA-TGGA--C-G---------T--T-AT--A-----A....................--A--G------AA-TT------G----TA-GG--- 9086
H2B.CI.x.EHO               ---A-----------A-TT-CT-TAGTG-G-----G--TA-A--A--A--CACA-ACT-AG-AA-TGA-G-----ATTG-GT---GA-G--------------AT---------A....................--A--------CAA-TT------C---CTA-GG--- 9534
H2G.CI.92.Abt96            --------------A-TTA-T--AGTG----------TAGA----A--CA-A-AC--GG-AA-GGAG-----TATAA-T--A----G------T------G----A-----T....................A-A----------R----G-----C---CTA-GG--A 9032
H2U.CI.07.07IC_TNP3        --------------A-TTA-T--AAT-----------TAGA--AT-A---A---ACT-AG-AA-TGAGG----GATAG-TG-A----G------T-----TGA---------A....................--T--G-T----AA---A-----C---CTA-GG--A 9119
H2U.FR.96.12034            ----------T---A--T--T-T--T--G--------TAGA---T-A---A-C-ACT-AG-AA-TGAGG-G--AATAA----A----G------T-----CT----A-----A....................ACT----GG---A--TG-----------CTT-GG--- 9128
ASC.UG.10.RT03             -A------------T-GT--T-T----CTGAG-----CA-G-AAT-G--C--A-ATGCT--TA-TGT-TGG--ACTAG-T--A-GG-G---AGGA-----C-AG--------C....................ACA--G-TC---ACATG------CATCCT--GGG-A 8614
ASC.UG.10.RT08             -A---------------T--T-T----CTGAG-----CC-G-AAT-G--C--A-ATGCT--T--TGT-TGG--ACTGACT--A-GG-G----GGA-----C-A---------C....................CCG--G-TC---ACATGT-----CATCCT--GGG-A 8638
ASC.UG.10.RT11             -A------------T-GTA-T-T----CTGAG-----CA-G-AAT-G--C--A-ATGCT--T--TGT-TGG--AGTAG-T--A-GG-G---AGGA-----C--T--C-----C....................CCG--G-TC---ACATGT-----CATCCT--GGG-A 8605
COL.CM.x.CGU1              -T--------G---A-T---AGAA---CTGAG---G-T--CC-GA-A--G-A--ATGC----ATTGAGTGG----G---AAAA-G--G-TT-G-G--TGA-GAT-A--TG--AGAGGATGGAGCATTAAAAGAG-A----A-G--T---GTTGC-------CTT-GG--A 8488
COL.UG.10.BWC01            -T--------G---A-T---AGAA---CTGAG---G-T--CC-GA-A--A----ACGC----ATAGA-TGG--AGTA--GAAA-GA-G--TTG-A--TGA-GAT-A-A-T-ATGAAAACATGCAAATAAGAGAA-AA---A-G--TT-AGTAGC-------CTT-GG--- 8308
COL.UG.10.BWC07            -A--------G---A-T---AGAA---CTGA---GG-CAGAC-G--G--GAGA-ATTG---TATTGA-TGG--AGTGC-AAG-----G--T-G-A--TGAGGAT-A-A-CAATGAGGACGGGACACCAAAGGAA-AA--GCTC--G--TG-T-G-------CTA-GG--- 8506
DEB.CM.04.SIVdeb04CMPF3061 -T--------G-ATA-TT--T----AGCC--G--G--T-CA---T-G--A---CATGC-C-AA-TGA-TGG--AATAA-A----G--G--TCC-G-----TGAG--C------....................A-T---------TAA-TA------G-TCCT--TT--- 8966
DEB.CM.99.CM40             -T--------G-ATG-TT--T----TGCC-----G--T-C---TT-A--G--TCATGC-C-GA-TGA-TGG--AATTA-A--A-GC-G--TAC-G-----AAG--A--T--A....................--T--G------A-TTA---C--G-TCCT--TT--A 8869
DEB.CM.99.CM5              -T--------G-ATA-TT--T----TGCT--------T-C---AT-A--AA-CCATGC-C----TGA-TGG--AATAA-A--A-G--G-TT---G-----GA---C-----A....................CCA--G------ACATA------G-TCCTA-TT--- 8796
DEN.CD.x.CD1_CMO580407     -A------C-G-ATA-TT--T--AGT-C-------G--T-C---TT-A--C--C-ATGC-C-GA-TGA---T----T-A-AA-A-GA-G-----C-------GAT--------C....................A-TC-G--C--G--AGAG-----G-TCCT--GG--A 9222
DRL.DE.11.D3               ---------A-----A-TTGGTGG--T-TC------G--TCA---T-G-----T-ATGCTC-AA-TGA-TGG----T-A-AAAA-GC-G---A-G------CAAG--A--T--C....................A----G--C---AA-GTA-----G-TCCTG-T---- 8859
DRL.x.x.FAO                ---------A-----A-TTGGTGG--T-TC------G--ACA---T-G-----C-ATGCTC-AA-TGA-TGG----T-A-AAAA-GC-G---A-GT-----CAAG--A--T--C....................A----G--C---AA--TG-----G-TCCTG-T---- 8940
GRV.ET.x.GRI_677_gri_1     ----------T---A--TA-T----TG--------G--A-G-----GA--T---ATGC--TGA--GAGTGG--AATAA-AGA-----G---AGCT---T-------C--G---....................A-A--G--C--GAGAGT------C-TC----TT--- 9079
GSN.CM.99.CN166            -A--------G--CA-GT--T---GTG-GGAC--G--CC-G-AGT-G--A-AA-ATGC---T-TGGAGTGG---CTGG-G----GA-G--TCTCA-T---T------C-----A....................AC-C-C-----TAC--TACCA----TC---AT-TGT 9048
GSN.CM.99.CN71             -A--------G-CCA-GT--T-T-GTG--GAT-----CC-A-AG--A--GAGC-ATTG---TT-TGAGTGG---ATTG-G----GA-G--TCC-G-GG--T-----A-----A....................A--C-C--C---ACT-TGCCA--C-TC----GGTGC 9045
LST.CD.88.SIVlhoest447     -T---------------TA-TGG-----G--------G--C---T-ATG---T-ACTGTG-GA-TGA-TGG--TCT-A-TGGA--C-TCAT--CT------GAT---------....................ACGC-T-TC---T-AG-A----------TG-GGC-- 8234
LST.CD.88.SIVlhoest485     -T---------------TA-TGG-----G--------G--C---T-GTG---T-ACTGTG-GA-TGA-TGG--TCT-A-TGGA--C-TCAT--C-------GAT--A------....................ACAC-T--C---T-A--A----------TG-GGC-- 8231
LST.CD.88.SIVlhoest524     -T---------------TA-TGG----G-----GA-T--C---T-AACA--C-ACTGTG-GA-TGA-TGG--TCT-A-TGGA--C-TCAT----------GAT--A-----A....................ACTC-T--C-----CG-----------CTG-GGC-- 8234
LST.KE.x.lho7              -T---------------TA-TGG-----G--------G--C---T-AAT---C-ACTGTG-GA-TGAGTGG--TCT-A-TGGA--C-TCAT----------GAT--A-----A....................ACA-----C--T-----A----------TG-GGC-- 9309
MAC.US.x.251_1A11          --------------A-TTA-T-TAGTGC--------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-A-A--A---G----G-T-----CT----A-----A....................A-T-----C---AA---A-----C---CTA-GG--A 9600
MND-1.GA.x.MNDGB1          ---------G----A--TA-T-TAG-AGG------G----A-----A-----C-ATGCAG-GA-TGAGTGG--A-TTGAA----GA-G---AC-G--T--GA----T-----A....................AC----------TAA---A-------TCCTG-TT--- 8664
MND-2.CM.98.CM16           --------------A-TTGG--T--TA--------G----G---T-G--CT-A-ATGCTC-AA-TGA-TGG--A-T-A-T---ACC-G---ATCT--T--TGAT--C------....................A-A-----C--GAA---A---A-G-TC-TA-T---- 9159
MND-2.GA.x.M14             --------------A-TTGGTT---TAG-------G----G---T-AA-C----ATGCTC-GA-TGA-TGG--A-T-A-A--A----G---AG-A-----TT-----------....................A-A-----C--GAA--GA-----G-TC-TA-T---- 9100
MND-2.x.x.5440             --------------A-TTGGTGG--TA--------G----G---T-G--C--T-ATGCAC-AA--GAGTGG--ACTAA-A--A---G-------------AAG--A--G---....................A-A--G-----GAA---T--C--G-TC-TG-T---- 8855
MNE.US.x.MNE027            --------------A-TTA-T--AGTG----------TA-A---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-G-G--A----G----G-T----CT-G--A-----A....................CCT-----C---AA---A-----C---CTA-GG--A 9081
MON.CM.99.L1_99CML1        -A-------G--A-A-GT--T---GTG--GAC-----TAGA-AGA-A--G-AA-ATGC---T-TGGA-TGG---TAA----A-GA-G--TAC-A--T--GGA----------....................-----G--C--GACT-TGCCC--G-TCCTG-GGTGC 9083
MON.NG.x.NG1               -A--------TAA---GT--T---GTGTWGACC-Y--TC-A-AGT-A--G-AA-ATGC---TGT--TGTGG--ACTTG-G----GA-G--TAC-G-TT--CY-C--A-----C....................--GC-G--C---ACC-T-CC---G-TCCTG-GGTGC 7720
MUS-1.CM.01.CM1239         -A------------A-GT--T-T-GTG-GGA------CC-A-AAT-G--GACT-AT-C---TGTGGACTGG--AATTG-A--A-GA-G--TAC-G-TT---GAG--C--G---....................AC-C-G--C---ACACTGCCA--G-TCCTG-GGTGC 9023
MUS-1.CM.01.SIVmus_01CM1085 -A------------A-GT-CGT-AGTAC-GA---G-TCC-G-AGT-G--GACC---GC---TGTGGAGTGG--AATAA-A--A-GA-G--T--TG--T---GAT--C------....................ACTC-G-----CAA--TACCA--C-TCCTA-GGTGT 9054
MUS-2.CM.01.CM1246         -A-----A-----AA-GT--GTG-G-AT-GAC----TGC-A-AA--A--GACT-ATGC---TGTA-A-TC-AAAATAG----AAGA-CA-TA-TA-----GGAT--C-----C....................AC-C-T--C----G-CAACCA--G-TCCTG-GGTGC 9119
MUS-2.CM.01.CM2500         -A------------A-GT--T--AGTG-GGA------CC-A-AAT-G--GACC-ATGCA--TGTGGAGTGG--TCTTA------GA-G--T--TG------GA---C--T--T....................--TC-C--C--G--T-TGCCA--TGTC-TG-T-TGC 9075
MUS-3.GA.09.09GabOI81       ----------TA--A-GT--T---GTAC-GAT--G--TC-G-AAT-G--GACT-ATGC---TATGGAGTGG--AATAG-G--A-GA-G--TACGA-TT---------C--G---....................AC-C-G--C---ACA-TGCCA--G-TCCTA-GGTGC 8652
MUS-3.GA.11.11GabPts02      -A------------A-GT--T-T-GTG--GA---G--CC-A-AGT-A--GACC-ATGC--TTGTTGA-TGG--AATAG------GA-G--T--G------------A------....................--G--G------AG--TACCC----TC-TG-GGTGC 8635
OLC.CI.97.97CI12           ----------G---A-GTA-TGG--A-CT-----GG-GA--T-GA--T------ATGCT-TGA-TTGTTGG--AATAA-AGG-A-C-TC---GCAACAGAGGAG-TAGA--A-.................GGCACGT-A--GAGA--TAA---GCCA--G----TA-GGG-T 8633
RCM.CM.00.SIVagi_00CM312   --------------A-TTA-T-T--AAT---------TAGA--T--A---A-A-ACT-AG-AA-TGAG--T--AATAA-T-------G------T-----T-----A--T--T....................CC---G--C--GACCTT---C-------TG-GG-T- 8591
RCM.CM.02.SIVrcm_02CM8081  --------------A-TTA-T-----ACC--G-----TAGA---T-A---A-A-ACT-AG-AA-TGAG--T--AATTA---------G------T------AA---A-----T....................A-T-----C---A---TG---------CTG-GG--A 8663
RCM.GA.x.SIVRCMGAB1        --------------A-TTA-TGGAGTATC---------AGA--AT-G---A---ACC-GG-AA-TGAG--T---ATAA-A-------G---A--------T-----------A....................A-A---------ACACTG---------CT--GGC-A 8762
RCM.NG.x.NG411             --------------A-TTA-TT---AG----G---AGGCGC--TT-A--C--C-AC--GG-GA-TGA---T--AATAA-A--A----G------T-----T------------....................A-A---------ACA-TG-----C---CTA-GGC-A 8845
SAB.SN.x.SAB1              --------------A-TTA-T--AGTG----G---A-GA-G---T-G-----C-ATGCTCTTA-TGAGTGG---ATAG-AGA--GA-G------T--T---GAT--------T....................AC---G--C---AA-TG--------------TT--A 9412
SMM.CI.79.SIVsmCI2         --------------A--TA-T--AGTG-G--G-----TA-A-----G---A---ACT-GG-AA-TGAGG---AATAA-A--A----G------T-----CT----------A....................A-------C---A---TG---------CTA-GG--A 8747
SMM.LR.89.SIVsmLIB1        --------------A-CTA-T--AG---T--G-----TAGA--T--G---A-A-ACT-AG-AA-GGAGG----GATAG-A--A----G------T--T---T----------A....................A-A---------A--TT-----------TA-GG--A 8660
SMM.SL.92.SIVsmSL92A       --------------A-TTA-T-TAGTG----------TA-A--AT-G---TCA-ACT-A--TA-TGAGG-----AT-A-A--A----G------T-------G---A-----A....................A-T-----C---A--TT------C----TA-GGC-- 8675
SMM.SL.92.SIVsmSL92B       ............................................................................................................................................................... 7460
SMM.SL.92.SL92B            ----------T---A-TTA-T--AGTG-G--G-----TAGA-----A---A-A-ACC-AG-AA-AGAGG----TAT-A----A----G------T--T---AGT--------A....................A-A--G--C---AA-TT------G---CT--GGC-- 9055
SMM.US.04.G078             --------------A-TTA-T--AGT--G--------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-G----A----G------T------T----------A....................--T-----C--TA--TT------G---CT--GG--A 8897
SMM.US.05.D215             --------------A-TTA-T---GT-CC--------TAGA--AT-A--CA-C-ACC-AG-AA-TGA-G-G--GATTG-A-------G------T-----TT----------A....................A----G--C---A----A----------TA-GG--A 8870
SMM.US.06.FTq              --------------A-TTA-T--AAT--G--------TAGA--A--A--CA-T-ACT-AG-AA-GGAGG----AAT-A-------C-G------T------T---A-----A....................A-A-----C---A--TG--C------CTG-GG--A 8889
SMM.US.86.CFU212           ----------T---A-TTA-T--AGTG-G--------TAGG--T--A--CA---ACC-AG-AA-GGA-G----AATAA-A--A----G------T-----CT-T---------....................A-A-----C---A--TT------C----TA-GG--A 8871
SMM.US.x.H9                --------------G-TTA-T--AGTATT--------TA-A--AT-A-----R-ACT-AG-AA-RGA-G----AATAA-A--A----G---A--T------G----------A....................A-------C--TA--TT----------CT--GG--A 9087
SMM.US.x.PGM53             --------------A-TTA-T--AGTG-G--------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-AC---A----G------T------T----A-----A....................A-------C---A--TT------C---CTA-GG--A 9541
SMM.US.x.pE660.CG7G        --------------A-TTA-T--AATG-G-A------TAGA--AT-A---A---AC--GG-AA-GGA-G----AATAA-A--A----G---A--T------T----------A....................ACT-----C--TA--TA--A---G---CT--GG--A 9609
STM.US.89.STM_37_16        --------------A-TTA-T--AGTG-G--------TAGA--AT-G--CA---ACT-AG-AA-AGA-G----GATAG-T--A----G------T------G----A------....................A-A-----C---AA-CAG---------CTA-GG--A 9255
SUN.GA.98.L14              -T--------------CTGGT-----AG--CC---G----A---T-G---T---ATGCAG-AA-TGAGTGG----T-A-AA---GC-G---AG-T-----CAA------T--A....................--G---------AA-G-------C---CTA-GG--- 9380
SYK.KE.x.KE51              -A------C-G--CA-GT--TT--GT----AG-----T--G--T--AC-G----ATT-AA-GA-TGAG--T--TGTGA-AGA---C...ATC-C---T--GAGT--------A....................ACA-----C--G----T---C--G----TA-GGG-A 8747
SYK.KE.x.SYK173_COMGNM     -A------C-G--CA-G-A-T-T-GTG---AG---G-T--A--T---C----C-ACC-AC-GA-TGAG--T--AAT-A-A...--CAG-ATC--T-----CT----------A....................AC-C-G--C--GT-A-T-----------TA-GGG-A 9057
TAL.CM.00.266              -A------A-G---A--T--T--AA-A--C-T--G--T-CA---T-GA------ATGC-C--A-TGA-TGG---ATAC-G--A--C-G------------GAG--A-----A....................ACA-----C---T--TG------TATCCTG-GG--A 8992
TAL.CM.01.8023             --------A-G---A--T--T--AA-A-C-----G--T-CA---T-AA-C----ATGC-C-TA-TGAGTGG---ATAC-G--A----G----TCA------GA---------A....................ACA--G--C---T-ATG---C--TATC-TG-GG--- 8542
TAN.UG.x.SIVagmTAN1        -A--------G--CA--TA-TGG--A-C--AG---G-GC-A-----GA----C-ATGC-CTAA-TGA-TGG---ATTA-AGA-----G---AGCA--TT-C--T--------C....................AC---GA-G--C-GATG---C--C-TT----TTG-- 9222
VER.DE.x.AGM3              ----------T---A-TTA-T-----G-C--G--GA-TC-G-----AA-C----ACGC-CT-A-TGAGTGG--AAT-A-TGA-----G-A-TGCT-GGT--GA---A-----A....................A-------C---AGATG---C--C-TC----TT--A 8733
VER.KE.x.9063              -A--------T---A-TTA-T-----G----G--GA-TAGA-----CA-C----ATGCTCTTA-TGA-TGG--AAT-A-TGA-----G-A-CGC--GGT--A----A-----C....................--A--------C-GCTG------C-TT----T---- 9235
VER.KE.x.AGM155            -A--------T---A-TTA-T-----G-T--G---A-TA-G------A------ATGC-CT-A-TGA-TGG--AAT-A-TGA-----G-A-CGCA-GGT--AA---A--T---....................A-A-----C--GAG-TG------C-TC----T---- 9212
VER.KE.x.TYO1_patent       -A--------T---A-TTA-T-----G-T--G---A-TA-G-----GA------ATGCTCTTA-TGA-TGG--GATAA-TGA-----G-A-TGC--GGT-GAAG--A-----A....................A-A----TC--TAAATG-----G-TC----TT--- 9196
WRC.CI.97.97CI14           -A--------G-ATA--T--TGG--AGCG------G----A--AG-GA-C--C-ATGCAC--A-TGA-TGG--GCTGA-AGA--GC-G---AGCT------GA---A------....................A-A----GG---AA---A-----G-------GG--A 9179
WRC.CI.98.98CI04           -A--------GTATA--T--TGG--T-CT------G----A--AG-GA-C--T-ATGC--T-A-TGA-TGG--ACTGA-AGAA-G--G----GCA------GAG--A-----C....................A-AC-C-GG---AA---A-----C-----T-GG--A 9221
WRC.GM.05.Pbt_05GM_X02     -A--------G-ATA--T--TGGAGT-CT------G----A--AG-GA-CT-A-ATGC---T--TGA-TGG--TATAG-GGAG-GA-G---AGC------CGAT--A------....................---C-C--C---AA---A-----C-------GG--A 8117
```

```
H1B.FR.83.HXB2          CTAGTACCAGTTGAGCCAGATAAGATAGAAGAGGCC...AATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGAC.........CCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCAC...GTGGCCCG 9383
Nef                     _L_V_P_V_E_P_D_K_I_E_E_A_____N_K_G_E_N_T_S_L_L_H_P_V_S_L_H_G_M_D_D_____P_E_R_E_V_L_E_W_R_F_D_S_R_L_A_F_H_H___V_A_R
H1O.CM.91.MVP5180       --G-------GTCAG----AG--GC---GAGACTGGGT----C-AAT--AG-TG-T--TC-TC----T--A-CTT-TAAT-----AGCT--G--T.........G-AC-CG-G--GA-AC--A-A----CA------T--ATCAT---GC--AACA--T---A-A-----T 9495
H1N.CM.95.YBF30         ----------T-GTCAG-T--AG-AG-----------T.....G-------C----ATGC-C-C-------CA-AT-T-AA-----AGCA-----T.........GATC-T-A-------G-T-----C-A---------TC-------AGAAGA--T.....A--AA- 8973
H1P.CM.06.U14788        ----------CTCTGAG-TAG-AGCT--G--AATGGGAG----GCAG------AG-T-AGC--C----T--A-CCT--ACTT-----T-TTCA--T.........--TC-T-AG--GA-CC---T-----A---------TCA--T-G-AGAG-A--T....--T--- 8989
CPZ.CD.06.BF1167        T---------CA--C---CCAG-AGAG--T--AAGGAAT.....................CTAC-CA-G-------CTT--TCCT----A-GA----------.--TA-TG-----C----G-TA----A-------TCTGAA--G--TAGAAGA------A-A---A- 9566
CPZ.TZ.00.TAN1          T-G-----G--A--C---CCAG-AGAT.....................G-T.....-----G-A--T----C----T--A-CCT-TAGC-----A-CTACC--T.........--A--TG----GACTC-GATC---C-C---------A---------AGAAGG------A-A---A- 9054
GOR.CM.07.SIVgorCP2135con   T----G-----CTCAGAG-CAG-AGCT--G---ATGGGA-CAGC-A-T----GAG-T-TAC--C-G--T--A-C-T--AGC-------GA--C.............--AC-TG----GA-CC-T-TC-----A--C--T--AGCAT-G-G-ACAAC---T....-T--ATT 8967

MAC.US.x.239            T----C--T--AA-T...-TATCAGAT--G-CACAG...G-GG-T...---G-GCATTAT--AA-G--T--A-CTCAAACTTCCCA-TG---------------.--TTG-G----G--TC---CA----A------TCCAACT--G--C-ACAC-T-T...-A---ATA 9753
Nef                     _L_V_P_V_N_._V_S_D_E_A_Q___E_D_._E_E_H_Y_L_M_H_P_A_Q_T_S_Q_W_D_D___.___P_W_G_E_V_L_A_W_K_F_D_P_T_L_A_Y_T_Y___E_A_Y
H2A.SN.86.ST_JSP4_27    ----------A--T.--TCCCACA---G-GA-AT...G-C-GT.------CTCA-T----AG-G--T--A-CACAAACAAGCA--T-T---------------.--C-TG-----ACA----TT-----------CC-ACG-----T---AGCT-----A----TT 9235
H2B.CI.x.EHO            -----G---A-AA-T...ATG-TAGC----CCA-AG...G-CG-G...--A-C-CATT-TC--G-G--T--A-CACAGACCTCC-CATG---------------.--CC-CGA---GACCC-T-TC---CA------TC--T-------A-G-CT-T...-------TT 9683
H2G.CI.92.Abt96         ----------CAC----TCTCTGAT--G-CACAG...G-GG-C.------CTCATT----G-----T--A-CACAAACTAGCCCATG---------------.-------CC-------G-C-----A-------CCTACAT----C-A-G-CT-----AG---ATT 9181
H2U.CI.07.07IC_TNP3     --G--G-----AC-T...-T-TCAGAT--G-CAAG-.....G-T.....-CTCATT--C--G-G--T--A-CACAGACCAG-CA-GA---G-------------.GATG-G-------A--G-CA----A------TGCAACG--G--T-A-G-CT-T...AAA--GTT 9268
H2U.FR.96.12034         ----A-----A--T...-CA--TGATA---CACAG...G-GG-T.....-GGCTGTATC--G-GGGTT-A-CACAGACCAGCT-CGA---G-------------.ACTG-G----G-CC----TA----A------TTC-TCT--G----A-AGCT-T...CA---ATT 9277
ASC.UG.10.RT03          T-G--C--G--G----..ATAGG-CC-C-G-GCCTGGGAG-AGG--AT--A-GAG-ATCG---C-C--TG-G-G-CAA-AAGTCTACCA---C-------------.--CC-TT-T--GACCC-G-TA---CAC--CA-TCC-AC---T--C----G-G-CAGGCA-ACTGAA 8769
ASC.UG.10.RT08          T-G--C--G--G----..---TAGG-CC-C-G-GCCTGGGAG-AGG--AT--A-GGG-A-AG---C-C--TG-A-G-CAA-AATCATTCCA---C-------------.-CT-TG-T--G--CC-G-C----CAC--CA-TCC-AC---T--C----G-G-CAGGCA-ACTAAA 8793
ASC.UG.10.RT11          T-G--C--G--G--T...-TGGG-CC-C-G-GCCTGGGAG-AGG--AT--A-GGG-ATTG---C-C--TG-G-G-CAA-AAGCATACCA---C-------------.TCC-TG-T--GACCC-G-T----CAC--CA-TCC-AC---T--C----G-G-CAGGCA-ACTGAA 8760
COL.CM.x.CGU1           T----CTACA-A----...C-ACTT.....................................................................................G-AGAGTAT----ACAGCT-T..... 8530
COL.UG.10.BWC01         T-G--CTTCA-A--A...C-GCTA.....................................................................................G-A-AGTGG--T-ACAGCT-T..... 8350
COL.UG.10.BWC07         ----G--C-TTC-G--C...TC--TA.....................................................................................G-AAAGTGG--T-A-TGCTTT..... 8548
DEB.CM.04.SIVdeb04CMPF3061  -----G-----A----...ATAGCAGACCC---CTAT...G-G--T-AT--A-GG-A--T---------TGA--CCCA--A-GG-CA------A-------------.-TT-T-AG--GAG-C-G-TA----A------TTCA-AA-----C-ACTGCT-T...AAA--AG- 9118
DEB.CM.99.CM40          --G--G-----A--A...ATAGCAGACCC---TTAT...G-G--T-AT----GG-A--T---------TGA--CCCA--A-GGGAT-----A-------------.--CC-T-----GAG---G-TT----A---------ACT--G--C-A-TG-T-T...AA---AG- 9021
DEB.CM.99.CM5           -----G-----A----...ATAGCAGATCCT--CTAT...G-G--T-AT----GA-AT-T-C-C-----TGA--CCCA--AAGGACAAGCA--A-------------.--CC-T-AG--GAG-C---TA----A------TTC-TC---T--C-A-TGCT-T...AAA--AG- 8948
DEN.CD.x.CD1_CM0580407  T-G-C--G--GACT...ATAG--GAG---T-T-A-.....-G.....G-A-ATT-TC--C-G--TGACAGATATGA-GG-CAACAA-C-------------.-TTG-----GACCC-G-TA------A------CCTGAAT----C-GGTG-T-T...AAA...GC 9368
DRL.DE.11.D3            T-G--C--T--G--T...-TAG-TGAT--CCT-ATA...G--C-G...CC-TGT-A---GC----GA--T-CAGCCAGA--GG-CCT--T-------------.T-A--AG-----AGAC-CAT----CA------TGCAA----G----ACACCTTT...CA---TT- 9008
DRL.x.x.FAO             --G--C--T--G--...-TAG-TGA---TCT--TA.....-C-G...C--TGT-A---GC----GA--T-CAGCCAGA--GG-CCT--T-------------.A--AG-----AGAC-CAT----CA------TCCAG----G----A-ACATTT...CA---TT- 9089
GRV.ET.x.GRI_677_gri_1  -----C-----G--C...CTGC-TGAG--G-CACG-.....-CTGT.....----GACA-T-TC--A-G--T---A-CACAGA--GGG-AAGATCC---TGGAATA..GATC-TG-------C--G-T----A-------CCGAAGT-G---GG-GG-GT-....CGCC-GGA 9234
GSN.CM.99.CN166         --GCGG-----A-CC...ACA-CTGAG--CTCT-AG...CC-GGG-AT--TG-ATATCTGC--ACC---------CCTAT-A-GGAC---GT--A-------------.A-C-T-AG--GT-C---G-TC-TCTCC--CTG-TC-AAG--G--CA-AA-GTCAGGCA-ACAG-T 9203
GSN.CM.99.CN71          T--CGG-----G-CC...ATG-CTGAG--CTCT-AG...CCCGGG-AT--TC-ATATCTG---AAC--T--C-CCTAT-A-GGACA-CAA--G-------------.ACC-T--G--GA-C--G-TC-TCTCC--CTG-TCTA-G-----C--GA-ATCAGGCTG-CAGAT 9200
LST.CD.88.SIVlhoest447  --G-A-----G-CC...TG-G-TGAGT-CACA-AT...CCCTC--AT-T-C-GACAGT--C--C-G---T-AAGTCAG---GGG-TCCA---G-------------.--CTG-G-----AG-C-TATC---CAC---A-TCCTATG-----G-GG-CT-T.....A----T 8386
LST.CD.88.SIVlhoest485  --G-A-----G-CC...TG-G-TGAGT-CACA-AT...CCCTC--AT-CC-G-CAAT--C--C-G---T-AAGCCAG---GGG-TCCA---G-------------.--CTG-G-----GAG-C-TATC---CAC--CAGTCCTATG-----G-GG-CTTT....-A----T 8383
LST.CD.88.SIVlhoest524  T-G-AG-----A-CA...TG-G-TGAGTTTA-A-AT...CCCTC--AT----GACAAT-T---C----TT-AAGCCAGT--GGG-TCC-A--A-------------.--CTG-G-----AG-C-CATC---CAC---A-TCC-ATGT-----G-GG-CT-T.....A----T 8386
LST.KE.x.lho7           T---A-----C-CT...TG-G-TGAGT-TA-A-AT...CCCTC--AT----CACAAT--C-T--G--TT-AAGCCAG---GGG-TCC----A-------------.CTG-G----GAG-C-TATC---CAC--CA-TCCAATGT-G---G-GG-CTTT...A-A----T 9461
MAC.US.x.251_1A11       T----C--T--AA-T...-TATCAGAT--G-CACAG...G-GG-T.....-G-GCATTAT--AA-G--T--A-CTCAAACTTCCCA-TG---------------.--TTG-G----G--TCC--CA----A------CCAACT--G--C-ACAC-T-T...-A---ATA 9749
MND-1.GA.x.MNDGB1       --G-A-----GAGCAG---C--TAGG---T---TAT...GCAGCT.........----ATCATC---------T-CTCCAGT-ATG-CCTCA---A-T.........A--AG-----GACCC-CAT----CTGGGAC-CT-AT-T-GC-TA-GAC-TTG...CA-CAT.. 8811
MND-2.CM.98.CM16        --TTGC--G--G-CA...-TGCCTCC---CC---AG.....C--T.....-TGT-A--AA---C-C--GT-CTCTCAG---GGAATCCA---A--A.......T--TTG-G----GAG-C---T-----A-------TCTG----T--T-ACACCTTT...TAT----C 9309
MND-2.GA.x.M14          T--A-C-----AC-A...-TACCACC---CC---AG.....CC-G.....-TGT-A---GC--C-TA--T-CTCTCAG--AGGAATCCA---A--A.......T--ATG-G-G-GAG-C-GAT-----A-------CCTGAAT-G--T-A-ACATTT...TAT----C 9250
MND-2.x.x.5440          --G--C--G--G-CA...-TGTC-CCCCCCCT--AG...G-GG-T.....-ATG--A---GC--C--A--T--TCTCAAA--GGAATCCA---A--A.........A---G----GAG---GAT-----A-----TTCAG-G--T--T-A-ACGTT---TAT--T-C 9004
MNE.US.x.MNE027         T----C--T--AA-T...-TATCAGAT--G-CACAG...G-GGGT.....-G-G-ATTAT--AC-G--T--A-CTCAAACTTCCCA-TG---------------.--TTG-G----G--TC---TA----A------TCCAACT-----C-ACAC-T-T...-A---ATA 9230
MON.CM.99.L1_99CML1     T-GCGC--G--A-CC...ACC-C-GAG--TTCC-AG...G-GGG--AT---G--TT-CT-C-AACC--T--G-CCTAT-AAGGCC-------G-------------.--CC-T---C-GT-C--G-TC-TCTCA-CTG-TC-AAG--G---G-CA-GAG-GGCAGACAG-T 9238
MON.NG.x.NG1            --GCGT--G--A-CC...ATCC--GA---CTC-AG...G-CGG--AT--YG-GTT-CT-C--ACC-------G-CCTA--A-GG-C--GAT--A-------------.--CC-T-AGC-GT-C--G-TC-TYTCC--CTGTTCTAAG--G-G-G-CA-GAG-GGCCCTCAG-T 7875
MUS-1.CM.01.CM1239      ---AGGGA---C-CC...ATA-CTGAG--CTCT-AG...G-GGGG-AT---G-ATATCT-C-CACC--TG-A-CCTA--AAGGCAA-GA---G-------------.--CC-T-AGC-GT-C--G-TC-TCTCA--CTG-TCT-----C--CA-GA-GTCAGGCAG-CAATT 9178
MUS-1.CM.01.SIVmus_01CM1085  --GAG-GA---G-CC...ATG-C-GAG-GCTCT-AG...G-GGG--AT--CC-ATA-CT-C-AGAC------CCTTT-A-GGGA-AGAA--A-------------.ACC-T-AGC-GT-CC-G-TC-TCTCC--CTG-TC----T-G--CC-GA-GTCAGGCT-ACAA-T 9209
MUS-2.CM.01.CM1246      T-GAG---G--A-CC...ATG-CAGAG--GTC--AG...CCAGG--AT--CC-GTA-CT-C--AC------G-CCTATG--GG-C--GAT--A-------------.ACC-C-A--GT-CC-G-TC-TCTCC---TG-TCTA-G--GC--CC-AA-GTCAGGCCC-CAG-T 9274
MUS-2.CM.01.CM2500      T--AGG--T--GC-C...ATA-CTGAGA-CTCA-AA...CTGGGG-AT--TG-ACATCT-C---GC--T--C-CCTT-A-TGGGA--GAA--A-------------.--CC-T-AG--GT-CC-G-TC-TCTCC--CTG-TC------G--TC-CA-GTCAGGCCGCCAGTT 9230
MUS-3.GA.09.09GabOI81   --CAG---G-A--C...ATGC-AGAG--TTCA-AA...GCAGG--AT--TG-ACA-CT-C--ACC-----G-CCTT--A-GGAA-AGAC--G-----------.ATCC-T-AG---T-C----TC-TCTCA--CTGTTCTAAG--T--CA-AA-GTCTGGCAA-CAG-T 8807
MUS-3.GA.11.11GabPts02  --CCGT--G--G-----..ATAGCAGA---CCCA-GG...G-AG-G-AT--C-CTTT-CT---AGCC--T--A-CCTA--A-GGGA--GA---A-------------.ACC-T-AG--GT-CC-G-TC-TCTCC---TG-TCT--------CC-GA-ATCAGGCCA-CAATT 8790
OLC.CI.97.97CI12        --TAGGAG---G--A---TTGCCTGAGACCCTT-AA...G-G-G---G.....................................................................--TTG-G----G--TC---TA----A------TCCAACT-----C-ACAC-T-T...-A---ATA 8678
RCM.CM.00.SIVagi_00CM312  --G--------G--T...-TG-GTGAT--G-CAAAA...G-GG-T.....-G-ACAT--TC-CC-G--T--A-CAGAAAGTAG-----T--G-------------.--TTG-G----ACA--G-CT----A----A-TCCTATGT-----G-AG--T-....A-A-G-TA 8740
RCM.CM.02.SIVrcm_02CM8081  T-----------A--T...-TG-TTGAT--G-CTAAG...G-AG-T.....-AG-GCAT----C--C-------A-CTGAGACTTG---------A-------------.--CTG-G----G--C----CT----A-AG--T-T.....-G-TA 8812
RCM.GA.x.SIVRCMGAB1     T-G--G-----A--T...-TATCTGAT----CAAGA...G-AG-T.....-AG-GCAT--T---C----T--A-CAGAAACAAG---------A-------------.--ATG-G-G--G--C--G-CC----A----A-TCCTATG--G---G-AG--T-....A-A-G-TA 8911
RCM.NG.x.NG411          T----------A--T...-TC-GTGAT----CAA-A...G-G--T.....-GGCATT-TC-CC-G--T--A-CAGAAACTAG---------C-------------.-CTG-G-G--G--A----CA----A----A-ACC--T-T-----G-AG--T-....-CA-GGTA 8994
SAB.SN.x.SAB1           --G-----T--G--C...CTGTCAGAG--G-CTAAG.....--GT.....-------CA-T--C--C-G--T----CTCAAG--GCCTATGA-------------.G-CTG--A---GACCC-G-T-----A-------CCA-TA------G-GG-CT-T...--T--TT- 9561
SMM.CI.79.SIVsmCI2      T---------GA-C...-TCTC-GAT--G-CACAA...G-GG-T.....-GACA-T--C-AG-C--T----CACAGA-CTC-CATT-A-------------.ATG-G-G--G---C--G-CA----A-------CCA-AAT-G--T-A-A-CT-T...-A---ATT 8896
SMM.LR.89.SIVsmLIB1     --------A-G--T...-TCTCAGAT---CTCAA...G-GG-T.....-CACATT-TC--G-C-----A-CACAAACT---CCA-----T-------------.-ATG-G----G---C-G-CA-----C-----TCCTAAA----T-ACACCT-T...-A---ATA 8809
SMM.SL.92.SIVsmSL92A    T---------G--T...-TCTCAGAT--G-CTAAT...G-AG-T.....-TCATT--C--C-T-.-----CTCAGACTT--AA----AG---T-------------.ATG-G-----GACCC----CC-----A----A-TCCTATGT----T-ACACCT-T...-A---ATT 8823
SMM.SL.92.SIVsmSL92B    ..........................................................................................................--------------------------------------- 7460
SMM.SL.92.SL92B         --G--G-----A--C...-TCTC-GAT--G-CAAA-.....-G-T.....-C-CATT-----G----T--A-CCCAGACAT--CA-CATTCA-------------.--TTG-G--A-G--CC---CC----A----A-TCCT-A------C-ACACAT-T...-AA--ATT 9204
SMM.US.04.G078          T----C-----G--T...-TCTCAGAT----CCCAG...G-AG-T.....-CACATT-TC--G-G-----A-CACAGACC---CA-TG-------------.-CTG-G-G--G--TC---C-----A-------TCCA-AA-----C-ACACAT-T...-A---TTT 9046
SMM.US.05.D215          T----G-----G--T...-TGTCAGAC--G-CACAA...G-AG-T.....-TTACA-T-TC--G-G--T--A-CACAGACAGCACA-TG-------------.--ATG-G-----G--AC---CA----AA-----TGCAAC---G--T-A-A-CT-T...-A---ATT 9019
SMM.US.06.FTq           -----C-----A---...-TCTCAGAT--G-CACAG.....-G-T.....-CACATT-TC--G-G--T--A-CACAGAC-GGCCAATG-------------.---TG-G--A-G--C--G-CA----A-------CC-ACT--G--T-A-ACCT-T...-A---AATG 9038
SMM.US.86.CFU212        --G--------A--T...-TATCAGAT----CTCAA...G-AG-T.....-C-CA-T-TC--G-G--T--A-CACAGACAT--CA-TG-------------.--TTG-G-------CC---CA----A------TCCAACAT----T-A-ACCT-T...CA---ATT 9020
SMM.US.x.H9             T---WC--G--AA-T...-TCTCAGAT----CTCAG...G-AG-C.....-CACATT-----A-G--T--A-CACAAACTTC-CA-TG-------------.-CTG-G----G--AC-G-CA----A-------TCCARAGT----T-A-R-CT-T...AA---ATT 9236
SMM.US.x.PGM53          T-G--C-----AAGT...-CCTCAGAT----CTCAA...G-AG-C.....-CACATTA--AG-G--T--A-CTCAGA-ATCCCA-TG-----T-------------.-CTG-G----G--TC-G-CA----A-------TTCA-AA-----C-ACAGAT-T...-A---TTT 9690
SMM.US.x.pE660.CG7G     T----C-----A--T...-TCTCAGAT----CCCAG...G-AG-C.....-CACATT---C--G-G--T--G-CACAGACT---CA-TG-------------.-CTG-G----G--AC-G-CA----A-------TCCAGAAT----T-A-AGCT-T...AA---ATT 9758
STM.US.89.STM_37_16     ----------A--T...-ATGTCA-AT--G-CACAG...G-GG-T-AT-G--CACATTATC--G-G--T--A-CACAGACA---CA-TG-------------.--TTG-G-------AC-G-TT----AA-----TCCA-TAT----CCA-AC-T-T...-A----TT 9407
SUN.GA.98.L14           ----C------CACC...AT-G-TGA---CAG--AT...CC---TCATCCTTGTCAAGC-C-AC-G--TT-CAGCCAG-AAGGG-TT-AT--A-------------.TTG-G---GAGAC-TATC----CA-----TCC-ACT--G--C-A-G--TTT...AG----AT 9532
SYK.KE.x.KE51           T-G-CC--TAA---A...ATAG-AGGGT-CTT--TA...G--G-G.....-AG----ACTGA--A-G-----A-CAGCAGGAGTG---GCTTCA--AGAC.........--AC-T--G--GAATC-CAT-----AC--CA-TCCA-A-T----C-A-ACA-CAGGCTG--AAAT 8902
SYK.KE.x.SYK173_COMGNM  ----T---AA-----..ATAG--GG-TGCTT--AG...T--G-G.....-C-T--ACT-C-CC-C---T--A-CCTCTGGA--G----GTTCGAG-ATGGGAGAA--TC-TGTG---C-TCA-CCCCCACCTGGCT-T-CA-CAGGTTGGGAGATGGC----CGCTTG-A 9215
TAL.CM.00.266           --CTGC----A-----.ATCC-TGAG--T--CA-G...G-GG-C....GAC-T...CT-C-CCG-G---T--A-CATATGATGGGCA-GGA--A-------------.-ATG-G-----TCCC-G-TC---GTC-----TGAGAAG--G---C-ACACC-C-GGA-C-AAGAT 9141
TAL.CM.01.8023          --GTGC--C--G-----..ATCC-CGAG--C--CA-G...G-GG-C....-GAC-----CT-C-CA-G------A-CATATGACGG-CA-CA---A-------------.CTG-G-C-----CCC-G-TC---GTC-----TGAGAAG--G---C-ACACC-CAGGA-C-AAAAT 8691
TAN.UG.x.SIVagmTAN1     T----G--G--A--T...-TCTCCCAG----CCCAG...G-----.....----GACA-T-T---C-C---------CCCAGA-TG-AT--GAAAGC-------------.---TTG--A---GACA--G-T-----AA---A-TCCA-T------G----GT------.....---- 9360
VER.DE.x.AGM3           T-G--------A--C....CTGC-TGAG----CA-AG....-C-TGT.....----GACATT--C--G-G--T--A-CACAAG--AGG-AAGACCC---TGGAATC...AACC-TG-------C--G-TC----A------TCC-ATGT-----G-C--AT-T--ACC-TAA 8888
VER.KE.x.9063           --------G--A-----..ATGC-TGA---G-CA-AG....-CCTGT.....----GACATTA----GCG--T----CACAAA-AA-A-A-GACCCC--TGGCATC...AGTC-TG-----GACC--G-TC----A-------CCTATGT-----G-G--GT-T....-ACC-AAA 9390
VER.KE.x.AGM155         --------G----CC...CTGC-TGAG----CA-AA....-CATGT.....-A-GGCATT-----G-------A-CACAA-------AAGACCC---TGGTATA...AATC-TG-----A-A--G-CA----A------TCCAATGT-G--TG----GT-----.ACC--TC 9367
VER.KE.x.TYO1_patent    -----G-----G--C...TTAC-TGAG----CACAA....-CATGT.....-A-GACATT--C-AG-C--T--A-C-CAGA--GGA-AAGATCCA--TGGTATC...AGCC-TG----GAC--G-TC-----A------TCCTATGT-G---A-A--GT-----.ACC--AA 9351
WRC.CI.97.97CI14        T-G--T-----AC-C...TTC-GTGA----TCA-AT.....-C-----.CCCATA-TATGA-AG-A-TAAGT-ACTG-AT-CA-CATGT-C---GCCAGAGGAT--CTG-G-----CATC-G-TC----A-------CCTAAG--G--C-G-G--TTTGTA-CA-GAA- 9340
WRC.CI.98.98CI04        T-G--T-----GC-C...TTC-GTGA---GTCTAAT...G-C-GC...TCTC-GTA-TA-GA-AAGA--AAGC--TTA-AT-CA--CTGCC-A--ATCAGAGGAT--TTG-G-----CAC--G-TC----A-------CCT--A-----C-G-G-CTTTGTG-CAAGTA- 9382
WRC.GM.05.Pbt_05GM_X02  T-G--T-----GC-C...TTC-GTGA---GTCTC-T.....CTC-....-CC-AGG-TTTGA-AAGA---TGC--CTA-AT-CA-CAG-A-C---GTCAGAAGAT--ATG--AG---TAC--G-TC----A-------TCCTA-G--G--T-G-G--TTTGTG-CT-GGA- 8278
```

Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

```
                                                     TCF-1 alpha binding
                                                                                                    Nef end
H1B.FR.83.HXB2             AGAGCTGCATCCGGAGTACTTCAAG...................................................................AACTGCTGA................................................ 9417
Nef                         E  L  H  P  E  Y  F  K  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  N  C  *
H1O.CM.91.MVP5180          GC-AAA---C--A---CT----CCCAAG...........................................................-A-................................................ 9526
H1N.CM.95.YBF30            ----------------TT-A----.....................................................---------.........................................CAA.... 9010
H1P.CM.06.U14788           -C-AAA---C-----ACTG--T-TTAAA.................................................GA--A-...........................ATTGCTGACGCCACGTAG 9041
CPZ.CD.06.BF1167           ----AGA----------------A.........................................................-A-...GGACTGCTTCCGGCAACCACACAAGAACTGCAGACGAAGGACTTTT 9640
CPZ.TZ.00.TAN1             ---AAGAT---------------A.........................................................-A-.GGACTTCCGGGTGCCATGACTCAGAACTGCTGACAGAGGACTTTTGGA 9130
GOR.CM.07.SIVgorCP2135con  -C--AA---C--A--ACTG------AGAGAA..................................................-A-....AGACTAAAAAACTTGCTGACAGCATGAAAAA.........GCTGC 9037
                                       TCF-1 alpha binding                                           Nef end

MAC.US.x.239               T-TTAGAT-C--A--AG-G--TGGAAGCAAGTCAGGCCTGTCAGAGGAAGAGGTTAGAAGAAGGCTAACCGCAAGAGGC......CTTCTTAACATGGCTGACAAGAAGGAA-CTC-----................................................ 9868
Nef                         V  R  Y  P  E  E  F  G  S  K  S  G  L  S  E  E  E  V  R  R  R  L  T  A  R  G  .  .  L  L  N  M  A  D  K  K  E  T  R  *
H2A.SN.86.ST_JSP4_27       TATT-GAT-C--A---G-G--TGG-TACAAGTCAGGCCTGCCAGAGGATGAATGGAAGGCAAGACTGAAAGCAAGAGGG.....ATACCGTTTAGC...................-A-...............AAACAGGAACAGCTATACTTGGTCAGGGCAGGA 9362
H2B.CI.x.EHO               CAGCAG-TTC--A---G-G--TGG-TATCAGTCAGGAATGCCAGAGAAAGAGTGGAAGGCTAAACTGAGAGCAAGAGGA.....ATACCTACA...................GAG-AG...........ACAGGAAATAACAGCACAAAGAGGAACTAGCTAACACT 9815
H2G.CI.92.Abt96            TATAT-------T---G-G----G-TGGAAGTCAGGGCTACCAGAAGCAGTATGGAAAGAGAAACTAAAACAGAGAGGA.....CTGCCTATAGAA...................-A-.............................GCAGAAAGAGGAC..... 9288
H2U.CI.07.07IC_TNP3        CATA---T----A--AG-G--TGGTAGTAAGTCAGGACTGCCAGAGAAAGAGGTTAAGAGAAGGCTAGCCGCAAGAGGC......CTTTTGAACATGGCTGACAAGAAGGAA-CTA-----.....................GCTGACAGA.AGGAAAC..... 9399
H2U.FR.96.12034            TATTAA-TGC--A--AG-G--TGG-AGTAAGTCAGGTCTGTCAGAGGAAGAGGTTAAGAGGAGGCTAACCGCAAGAGGC......CTCCCAGTAAAG...............--T------.........................CAGAAAGGAAAC..... 9389
ASC.UG.10.RT03             G-CTGACA-GTT--GAC-GAAGCCTTTGCCTGGGCAAAAGACCTTTCTAACAGGAAGAGAC...................................................-AG.....CTTAACCGCAAACCGCATCCTATGCAAGGTCCC............ 8866
ASC.UG.10.RT08             G-CTGACA-GTT--GAC-GCAGCCTTTGCCTGGGCAGGAGACCTTTCTAACAGGAAGAGAC...................................................-AG.....CTTAACCGCAAACCGCATCCTATGCAAGGTCCC............ 8890
ASC.UG.10.RT11             G-CTGACA-GTT--GCC-GCAGCCATTGCCTGGGATGAAGAGCTGCATAACAGGAAGAGAC...................................................-AG.....CTTAACCGCAAACCGCATCCTATGCAAGGTCCC............ 8857
COL.CM.x.CGU1              ......................G-CTTGAGTCTGTTGAGCGTTACAAGCAGGAAGAAGAAGAAACCACAACAG....................GTTGCCATAGAGATGGTGG--...----.........................................G..... 8606
COL.UG.10.BWC01            ......................G-CATGTCTCTGCTCAGCATTAACAACAGGAAGAAGAAACCGCAGCAGCAGCAG...GTTGCCGTGGACACGGTTGCTATAGAAACACCAG--....-A-................................................ 8443
COL.UG.10.BWC07            ......................G-CCCAGCTTTGCTTAACAGGCCTGTGACTAAGAAGAAGAAACAGGATGCGCCAAACCGCAGCACGGTTGCCATGGAGATAGAGACTGTGG--....-A-................................................ 8644
DEB.CM.04.SIVdeb04CMPF3061 CC-TGAAGC--ATACTA-AGAG-CACATACTAGGAGATGTATGTTCCCTAAGAGGAAG.......................................................-AG..............................CTAGCTTAACCGCAAAAC 9197
DEB.CM.99.CM40             CC-T...GC-GAAC-AAGA...G--CATACTAGAAGATGTATGTTCCCTAAGAGGAAG.......................................................-AG..............................CTAGCTTAACCGCAAACC 9094
DEB.CM.99.CM5              -C-TGAAGCC-ATA-ACCA...G--CATAACAGAAGATGCATGTTCCCTAAGAGGAAG.......................................................-AG..............................CTAGCTTAACCGCAAACC 9024
DEN.CD.x.CD1_CMO580407     --GC-ATA-GAA-CTTC-AACTTCACAGCATAATAACTACTGCCAGAGAGCTAAGCCCAGGAAG.................................................-AGCCTAACCGCAAACCACATC.............CTACTGCAGAAC...... 9466
DRL.DE.11.D3               TATT-AC-----A---G-G---GGACATGTTACAAGTATAAGACATGATAAAAGAGAAGAGCCAGATTGTTGCAAGAGGAAGTGGTGGCAG.........................-A.................................................. 9101
DRL.x.x.FAO                TATT-AC-----A---G-G--TGGACATGTTAAAAGTATAAAGTCA...GTA...GAAGAGCCAGGTTGTTGCAAGAGGAAGTGGTGGCAGTTCAAGCCAACCACAGAAGGTTGTCATG--GACAAC...........................CTTCAGAAGT......... 9220
GRV.ET.x.GRI_677_gri_1     CAT-T-TA-GGACAT-C--GAAC-TGCAAAGCGC...............................................................................-AG....................................TGTCAGCACTTTG 9284
GSN.CM.99.CN166            ---C-A--TG-A-C--G-GGAGCG-AAGATGCGCCTAACCGCAAACCGCTTCCTA..........................................................-AG...................................AACTCTGCCC.... 9271
GSN.CM.99.CN71             GA-C-A--TG-A-C--G-GGAGCG-AAGAAGCGCCTAACCGCAAACCGCTTCCTA..........................................................-AG...................................AACTCTGCTC.... 9268
LST.CD.88.SIVlhoest447     CAGAAAA--A--TA-AAG-A-AC-AGCATCAGCTATTGGCTTG......CAGAGGAAG.......................................................-AGCCAAACCGCAGACCGCAACACATCCTCTTGCAGCCTGGTTGCCAGGGCA 8490
LST.CD.88.SIVlhoest485     CAGAAAA--A--TA--AG-A-AC-AGCATCATCTCTTAGCTTGAACTGTAAGAGGAAG.......................................................-AGCCAAACCGCAGACCGCAGCACATCCTCTTGCAGCCTGGTTGCCAGGGCA 8493
LST.CD.88.SIVlhoest524     GAGAAA---A--A-CAAGTG-TC-AGCAACTGCTATGAGATTTAACTGCGAGAGGAAG.......................................................-AGCCAAACCGCAAACCGCAGAACATCCTCTTGCAGCCCGGTTGCCCTAGCA 8496
LST.KE.x.lho7              CA-AAA---G--T-C-A-GA-AC-AAATACTGCTTTCGCTTTTGACTGCAAGAGGAAG.......................................................-AGCCAAACCGCAAACCGCAGCACATCCTCTTGCAGCCCGGTTGCTAAGGCA 9571
MAC.US.x.251_1A11          T-TTAGAT-C--A--AG-G--TGGAAGCAAGTCAGGCCTGTCAGAGGAAGAGGTTAGAAGAAGGCTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAAGGAAACTAGA.........---.......................................... 9864
MND-1.GA.x.MNDGB1          .................................................................................................................-A-CATATCACCCTGAGTGTTTCAATAAGGCTAAGAGTATTGAACATCTGCC 8863
MND-2.CM.98.CM16           CATAAA-AGA--A-GTG-A------CATGTG...CAAAGCATCAGCCATATTGCTTATGCTAAG.....GAGCATAAGCCAGAGTGCTGCAAGAGGAAGTGGTGGCAGTTC......-AG.........CCGACCACAGAAGGTTGCCATGGAGATGACCTGCAGAAAT 9455
MND-2.GA.x.M14             CATAAAATGC--AA--G-A-A--G-CATGTT..................ACAAGCCTCTCCTAT...ACTGCCTACCAGGAGAAATCAGACTGCTGCAAGAGGAAGTGGTGGC-G-T--AG.........CTAACCACAGAAGGTTGCTATGGAGATGCTCTTTAATTCA 9390
MND-2.x.x.5440             TATAA-TAGG--A---G-A-A----TGTGTGACAAGTCTGAGCTATGAAGCCTACAAGAAGGAGGAAAAGCCTGACTGCTGCAAGAGGAAGTGGTGGCAGTTC...............-AG.........CTAACCACAAAGGGTTGCTATAAGCAACTGCTAACTCTGG 9150
MNE.US.x.MNE027            TATTAGAT-C--A--AG-G--TGGAAGCAAGTCAGGCCTGTCAGAGGAAGAGGTTAGAAGAAGGCTAACCGCAAGAGGC......CTCTTAAAAATGGCTGACAAGAAGGAA-CTA-----................................................ 9345
MON.CM.99.L1_99CML1        --CA-A--T--A-C-AG-GGAGCG-AAGAAGCGCTTAGCCGCAAACCGCATCCTC..........................................................-AG.....................................AACTCTGCCC... 9306
MON.NG.x.NG1               ---TGA-TTA-A-C-AG-AGAACG-AAGAGGCGCTTAACCGCAAACCGCATCCTA..........................................................-AG.....................................AACTCTGCCC... 7943
MUS-1.CM.01.CM1239         ---T-AAATA-A-C-AG-AGAACG-AAGAGGCGCTTAACCGCAAACCGCATCCTA..........................................................-AG.....................................AACTTTGCCC... 9246
MUS-1.CM.01.SIVmus_01CM1085 ---C-AAATG-A-C--G-GGAGCG-AAGAAGCGCTTAGCCACAAACCACATCCTA.........................................................-AG.....................................AACATTGCCC... 9277
MUS-2.CM.01.CM1246         TA-T-AAATA-AAC-AA-AGAGCG-AAGAGACGCTTAACCGCAAACCGCATCCTC..........................................................-AG...................................AACTCTGCTAGAGA 9346
MUS-2.CM.01.CM2500         -A-C-A-ATG-A-C--G-GGAGCGAAAACGGCGCTTAGCCGCAAACCGCATCCTCTCGCAGTCT......CCA........................................-A-................................CTACGGGACTATAAG.. 9315
MUS-3.GA.09.09GabOI81      T--T-A--T--A-C-AG-GGAGCG-AAGAAGCGCCTAACCGCAAACCGCTTCCTA..........................................................-AG.....................................AACTCTGCTC... 8875
MUS-3.GA.11.11GabPts02     ---C-A--T--A-C-AGCAGAACGCAGGAAGCGCTTAACCGCAAACCGCATCCTA..........................................................-AG.....................................AACTCTACCC... 8858
OLC.CI.97.97CI12           ...................................................................................................................................................................... 8678
RCM.CM.00.SIVagi_00CM312   TAGA-----C--T--C-T---TGGAGAGAGA...AAGAACAAGACCCAG................................................................-A-........CCACTTCCTATGGAGTTGCCTTGGAGACCGCGCATGCACA. 8829
RCM.CM.02.SIVrcm_02CM8081  TA-A-----C--A---G----TGGAGAGAAG...AAGAACAAGACCCAA................................................................-A-........CCACATCCTATGGGGTTGCCAAGGAGACAGCACCTGCGCAC 8902
RCM.GA.x.SIVRCMGAB1        TAGA--------A----T---TGG-GAAAGG...AAGAACAAGACCCAG................................................................-A-........CCACATCCTCTGGGGTTGCCTTGGTAACCAGGCAGAAGAA. 9000
RCM.NG.x.NG411             TAGA--------T----T---TGGAGAGAGA...AAGAACAAGACCAAG................................................................-A-........CCACTTCCTCTGGAGTTGCCTTGGAAACAGCGCATGCCTAC 9084
SAB.SN.x.SAB1              GAG------C--A---C-GG-GCC-TCAGCA...CAAGGC.........................................................................-A-CCACAAAACCACATCCTATGGAGTTGTCATGGTGATGACATTAA..... 9645
SMM.CI.79.SIVsmCI2         CATAAGAT-C--T--AG-G--TGGTTCCAAGTCAGGGCTGTCAGAGGATGAGGTTAAGAAAAGGCTAACCGCAAGAGGCCTTTTGAAAATGGCTGACAAGAAGGAAAATAGC......---.................................GACAGC.... 9017
SMM.LR.89.SIVsmLIB1        T-TTAGAT-C--A--AG-A--TGGTAGCAAGTCAGGCTTGTCAGAGGATGAGGTTAAGAGAAGGCTAACCGCAAGAGGCCTTTTAAAAATGGCTGACAAGAAGGAAACACGC......---.................................GACAGC.... 8930
SMM.SL.92.SIVsmSL92A       T-TTAGAT-C--A--AG-G--TGGATGGAAGTCAGGACTGACAGAGGAAGAGGTTAAGCGTAGGCTGGCTAGCAGGAAGAAGCCTAACCACAAAAAGCAACATGGCGCAGAAG-AG-AA-CTAGCTAAGACAGC.............................. 8957
SMM.SL.92.SIVsmSL92B       ...................................................................................................................................................................... 7460
SMM.SL.92.SL92B            T-TAAG---C--A--AG-G--TGG-TGGAAGTCAGGCTTGACAGAGGAAGAGGTTGAGAGAAGGTTAGCTAACAAGCCAAAACCGCAGAAGAAGATGGCGGATAAGAAGGAA-CAA-----...........................................GA.... 9327
SMM.US.04.G078             TATTAGAT-C--A--AG-G--TGGTATGCAGTCAGGCTTGTCAGAGGATGAGGTTCAGAGAAGGCTAACCGCAAGAGGCCTTTTAAAAATGGCTGACAAGGGAGGAAACAG................................................. 9157
SMM.US.05.D215             TATTAAAT-C--A--AG-A--TGGTAGCAAGTCAGGCCTGTCAGAGGAAGAGGTTAGGCGCAGGCTAACCGCAAGAGGCCTTTTAAAAATGGCTGACAAG.GGGAAACAG................................................. 9128
SMM.US.06.FTq              TATTAA-T-C--A--AG-T--TGG-GACAAGTCAGGCTTGTCAGAGGAAGAGGTTAAGAGAAGGCTAACCGCAAGAGGCCTTTTAAAAATGGCTGACAAGAAGGAAACTAG................................................. 9149
SMM.US.86.CFU212           T-TTAAAT-C--A--AG-G--TGGATATAAGTCAGGCCTGTCAGAGGAAGAGGTTAAGAGAAGGCTAACCGCAAGAGGCCTT.TTAAAATGGCTGACAAGAAGGAAACCAG................................................. 9130
SMM.US.x.H9                T-TTGA---C--A--AG-G--TGGTAGTCAGTCAGGCTTGTCAAAGGAAGAGGTGCAGAGAAGGCTAACCGCAAGAGGC......CTTTTAAAAATGGCTGACAAGAAGAAA-CAA---AG..........................................GA.... 9353
SMM.US.x.PGM53             CATTAGA--C--A--AG-G--TGGCAGTAAGTCAGGCTTGTCAGAGGAAGAGGTAAAGAGAAGGCTAACCGCAAGAGGC......CTTTTAAAAATGGCTGACAAGAAGGAA-CAA-----...........................................GA.... 9807
SMM.US.x.pE660.CG7G        TATTAA-T-C--A--AG-G--TGGTAGTAAGTCAGACTTGTCAGAGGAAGAGGTAAAGAGAAGGCTAACCGCAAGAGGCCTTATTAAAATGGCTGACAAGAAGGAAACAAGC......---.................................GACAGC.... 9879
STM.US.89.STM_37_16        T-TTAGA--C--A--AG-G--TGGTTCTAAGTCAGGCTTGCCAAAGGAAGAGGTTGAGAGAAGG......CTAACCGCAAGAGGCCTTCTAAAGATGGCTGACAAGAAGGAA-CAA------.........................................GA.... 9524
SUN.GA.98.L14              -C-AAA---C--T--AG-G------CATGTTACTAGTCTGCAGTGGGAGGTG.............................................................-AGCTTAACCGCAAACCGCATCCTCTTGCATCGCCTAGGCAACGGGGCTAGC 9636
SYK.KE.x.KE51              G-CT-GT--G-AATTAG-GAGAC-AACAGGAAAGCGC............................................................................-AGCTCACCGCAGAAACCACAGGGTACATCTCCTAGGAGACACCTAGGAGAC 8991
SYK.KE.x.SYK173_COMGNM     GTTAGA-AGA-AAACAGGAAAACCACAAGAG.........CTGCAGTCAGCACTTAGTAAGAAC......ATCTCC......................................-AG..............................GAGACTCCATGGTGACAA 9297
TAL.CM.00.266              G-CCGA-TGGGACAG--TGGAG-GAGAGAAGCGGATGCTGCTAGCTCCACCGCAAACCGCATCCTCT..............................................---....................................AGGACTC....... 9218
TAL.CM.01.8023             G-CTGA-T--GACAG--TGGAG-G-AAGAAGAAGGAGCTGCTAGCTCCACCGCAAACCGCATCC......TCT........................................---............................................AA.... 8763
TAN.UG.x.SIVagmTAN1        .......A-C-----C.......GCTTTAAGGACATGCATGGTCTTGTTAAGAGGAAG.......................................................-AGCCTAACCGCAGGCTTGCGGTTAAGCTAACTAGGCCGTTGCCTAGGAGAT 9457
VER.DE.x.AGM3              CAGAAAAT---TCACTG--A-GC-TGATCTT...GGCAAGAGGAAG...................................................................-AG.....CT.AACCGCA.GGCTTGTGGTTAAGCACATCACCATGGTGATGA 8976
VER.KE.x.9063              CAGA-A-T--TTT--AG--A-GC-TGCAATAGTGAAGAGGAAG......................................................................-AG.CTAACCGCAGGCTTGTGGTTAAGCCGTTGCCAGGGAGATGACATTTG. 9483
VER.KE.x.AGM155            -AG-GA-T-CTTTACAG----AT-TTCAACAGTTGGTACAGGAAAC...................................................................-AG.CCGACCACAGGCTTGCGGTTTCCTGGTTGCCTAGGAGATGACATTAA. 9463
VER.KE.x.TYO1_patent       TCG-GA-T-CTTTACTG--A-GC-TGGGCTGGTGAAGAGGAAG......................................................................-AGCCAGACCGCAAGCCTGCGGTTAGAACATCACCATGGAGATGACATTAA. 9445
WRC.CI.97.97CI14           -CTC-CTGGA-A-TCAAT-GC--CTGGGAAAGAGGCCATTGAGAGTCTCTATAAG...AACAAG......AAGCAGTGCAAG................................-AG.......................................AGACATG... 9423
WRC.CI.98.98CI04           -CTC-CTGGA-ATATAATAGC-C-AGGAGAAGATGCTATTGAGACCCTCCATAAAGAGAAGCTG......AAGCGGCACAAG................................-AG.................................................. 9461
WRC.GM.05.Pbt_05GM_X02     -CT-GCCG-CGGTC--ATAGC--CTGGTATTGAGACCCTC......TATGAGATG...AAGAAG......AGAAACAACAAG................................-AG.......................................AGACTTG... 8355
```



```
                                                                                                                              NF-κ-B-II                                            NF-κ-B-I
H1B.FR.83.HXB2               ......................................................CATCGAGCTTGCTACAA............................GGGACTTTCCGCTG.............................................GGGAC 9453

H1O.CM.91.MVP5180            ........................................................CT--CAC---.................................----------AG........ACTGCTGACAC.........................TGCG----- 9568
H1N.CM.95.YBF30              ...................................................................................................--------A-TGCTGACAA....GGGACTTTATA.....CTTGGGGACTTTCCGCCA..........------ 9063
H1P.CM.06.U14788             CTGCT......................................................-AAGCTGCTG-CACTGCA......................-----------GG-A...CGGAAAGTCCCGAGGGCGGAACAAGGGGAGG.....AGCAGGGGA-TGG 9123
CPZ.CD.06.BF1167             ACTAA.................................................................-............................----------A..................................................... 9657
CPZ.TZ.00.TAN1               C..................................................................TC..............................----------AA.................................................... 9145
GOR.CM.07.SIVgorCP2135con    ............................................................TGACACTGC-.............................----------AT.................................................... 9059
                                                                                                                              NF-κ-B-II                                            NF-κ-B-I
MAC.US.x.239                 ...........................................................-A-AGCA.................................----------A-AAG.............................................---.A 9894

H2A.SN.86.ST_JSP4_27         AGTAACTAACAGAAAACAGCTG.........................................A-ACTGC-.............................----------AGAA..............................................---G- 9411
H2B.CI.x.EHO                 GCATAGAGAAGGAAACTAGC..........................................TGATACTGC-............................----------AGAA..............................................---G- 9863
H2G.CI.92.Abt96              ...............................................................AAGCTG-CACAGCA.......................----------AG-A..............................................T--G- 9321
H2U.CI.07.07IC_TNP3          ...............................................................-AGCTG-GCTGACA.......................----------AGA..............................................AT--G- 9432
H2U.FR.96.12034              ...............................................................AAGCTG-GACAGCA.......................----------A-CAG..............................................A--GA 9423
ASC.UG.10.RT03               ...................................................................................................----------AG-TG.................................................... 8881
ASC.UG.10.RT08               ...................................................................................................----------AG-TG.................................................... 8905
ASC.UG.10.RT11               ...................................................................................................----------AG-TGCC.................................................. 8874
COL.CM.x.CGU1                ...............................................................-TGCTG-GACAGCA.......................-----------T-C..............................................----- 8639
COL.UG.10.BWC01              ...............................................................A-ACTGCTGACACGGC..................G------------CC..............................................----- 8479
COL.UG.10.BWC07              ...............................................................GAACTGCTGACACAGC....................---.---C-GCTGTTA..........................................----- 8680
DEB.CM.04.SIVdeb04CMPF3061   CACATCCTATGCAAGGTCTA...............................................................................--T-------A..................................................... 9228
DEB.CM.99.CM40               ACATCCTACAGCAGCGCCCG...............................................................................----------A..................................................... 9125
DEB.CM.99.CM5                GCATCCTCTTGCAGAGAGCCC..............................................................................--T-------A..................................................... 9056
DEN.CD.x.CD1_CMO580407       ...............................................................-TGCTG-GAGGCTCA......................----------A..-...........GTTGCCT....A........................-CA-- 9506
DRL.DE.11.D3                 ........................................................................................................................................................................ 9101
DRL.x.x.FAO                  .................................................................GCTG-CTGAGAA.......................----------A-A-GAGCATGCGCTCTGTGGC.....A........................----- 9270
GRV.ET.x.GRI_677_gri_1       CGGTT..............................................................................................------------CA..............................................----- 9308
GSN.CM.99.CN166              ...................................................................................................----------A...................................................-T- 9285
GSN.CM.99.CN71               ...................................................................................................----------A...................................................-TT 9282
LST.CD.88.SIVlhoest447       AC............................................................CA-GCTAGCGCATGCGC....................A-------G-TGATGAGC..................A........................----- 8533
LST.CD.88.SIVlhoest485       AC............................................................CA-GCTAGCGCATGCGC....................A-------G-TGATAAGC..................A........................----- 8536
LST.CD.88.SIVlhoest524       ACC............................................................G-GCTAGCGCATGCGC....................ACA-GC--G-TGAACAGC..................A........................----- 8539
LST.KE.x.lho7                ACCGGGCTAGCGCATGCGCAATG......................................GC-TGCTG-CGAGCA.........................----------AGGACGGGCGG.........................................----G 9635
MAC.US.x.251_1A11            ...........................................................GA-AGCA.................................----------A-AAG.............................................----. 9889
MND-1.GA.x.MNDGB1            ATTTTGGAAGAGGAAGTAG...CCTAACCGCAAAACCACATCCTACTGCAGAACTGT-GTTGC...TTGG---CCTGCTTAGCAACCTGGACTGGCGCTTGCGC-CTAGGAA..............................................----- 8974
MND-2.CM.98.CM16             ...........................................................GCTGACTCTGG-............................----------A--AGTGCATGCGCACTAGACT....G........................----- 9505
MND-2.GA.x.M14               ...........................................................GCTGACTCTGG-............................----------A--AGTGCATGCGCACTGGACT....G........................----- 9440
MND-2.x.x.5440               ......................................................................-............................----------A---GTGCATGCGCACTGGACT....G........................----- 9189
MNE.US.x.MNE027              ...........................................................GA-AGCA.................................----------ATAA..............................................---GA 9371
MON.CM.99.L1_99CML1          ...................................................................................................----------A...................................................-T- 9320
MON.NG.x.NG1                 ...................................................................................................----------A...................................................-TT 7957
MUS-1.CM.01.CM1239           ...................................................................................................----------A...................................................-T- 9260
MUS-1.CM.01.SIVmus_01CM1085  ...................................................................................................----------A...................................................-T- 9291
MUS-2.CM.01.CM1246           CCTTCCA............................................................................................................................................................-T- 9356
MUS-2.CM.01.CM2500           ........................................................................................................................................................................ 9315
MUS-3.GA.09.09GabOI81        ...................................................................................................----------A...................................................-T- 8889
MUS-3.GA.11.11GabPts02       ...................................................................................................----------A...................................................-TT 8872
OLC.CI.97.97CI12             ........................................................................................................................................................................ 8678
RCM.CM.00.SIVagi_00CM312     .....................................................GA-GAA-CTGCTAACAC-GAA..........................----------A---GCGCCTGCGCGCCAGTGCA....A........................----- 8889
RCM.CM.02.SIVrcm_02CM8081    .....................................................T--AGA-CTGCTGAGAC-GAA..........................----------A---GTGCATGCGCACTGGTGAA....A........................----- 8962
RCM.GA.x.SIVRCMGAB1          ...........................................................TCTGCTGATGC-AAA..........................----------A---GTGCATGCGCACTGGGGAA.............................----- 9053
RCM.NG.x.NG411               .....................................................TG-GAAGATGCTGA..C-CAGAA........................----------A---GCGCCTGCGCGCTGGTGTA....A........................----- 9144
SAB.SN.x.SAB1                ...............................................................GAACTGCTGACTGAGA.....................----------AG.................................................... 9673
SMM.CI.79.SIVsmCI2           ..............................................................A....................................----------........................................................ 9029
SMM.LR.89.SIVsmLIB1          ..............................................................A....................................----------A..................................................... 8942
SMM.SL.92.SIVsmSL92A         ..............................................................A....................................----------A..................................................... 8969
SMM.SL.92.SIVsmSL92B         ........................................................................................................................................................................ 7460
SMM.SL.92.SL92B              .............................................................-AGCA..................................----------A-AA.....................G........................----G 9352
SMM.US.04.G078               ........................................................................................................................................................................ 9157
SMM.US.05.D215               ........................................................................................................................................................................ 9128
SMM.US.06.FTq                ........................................................................................................................................................................ 9149
SMM.US.86.CFU212             ........................................................................................................................................................................ 9130
SMM.US.x.H9                  .............................................................-AGCA..................................----------A-AAA.............................................---GA 9378
SMM.US.x.PGM53               .............................................................-AGCA..................................----------A-AAA.............................................---GA 9832
SMM.US.x.pE660.CG7G          ...............................................................A....................................----------A..................................................... 9891
STM.US.89.STM_37_16          .............................................................-AGCA..................................----------A-AA..............................................---G- 9548
SUN.GA.98.L14                GCAT..........................................................-C--GCTAG-GT-.........................-----------GAC................................................... 9667
SYK.KE.x.KE51                AGAGACCGCCCAGGGGAGGAGTCGGGGCGGGGGAGGCTCACCCAACCTGCATAAAAGCC...-A-GCTCTGT-GCTATCC....................................................................................... 9068
SYK.KE.x.SYK173_COMGNM       GCTCGGCCCAGGGG...............................................AGGA--CTGGGCGGGGGAAGGCTCA............................................................................CCCT 9340
TAL.CM.00.266                ...................................................................................................----------A..................................................... 9229
TAL.CM.01.8023               ............................................................-GACT-.................................----------........................................................ 8780
TAN.UG.x.SIVagmTAN1          GACATAAAGAAGTG..................................................................CTGATCAGCA..........----------AG.................................................... 9493
VER.DE.x.AGM3                CATTAA........................................................GAACTGCTG-C...........................----------AGCAA.............................................----- 9013
VER.KE.x.9063                ...............................................................AAACTGCTGACAG........................----------AA................................................----- 9513
VER.KE.x.AGM155              ...............................................................GAACTGCTGAC..........................----------AGCAC.............................................----- 9494
VER.KE.x.TYO1_patent         ...............................................................AAACTGCTGAC..........................----------AGC-AA............................................----- 9477
WRC.CI.97.97CI14             ......................................................TTG-AGTAGGAA.................................-T-G---G--G-TAAGCTACA............................................ 9457
WRC.CI.98.98CI04             ........................................................................................................................................................................ 9461
WRC.GM.05.Pbt_05GM_X02       ...........................................................TC--AGGAA...............................-CTGG---G--G-TAGGCTAGTA........................................... 8387
```


Draft of 30 September 2015, not for public release

plvdna_main.tex r8876–10546

```
                              NF-κ-B-I            Sp1-III                     Sp1-II      Sp1-I                          TATA box                        TAR element start
                                                                                                                                            3' LTR U3 end  3' LTR repeat start
H1B.FR.83.HXB2                TTTCCAGGGAGGCGTG..................................GCCTGGGCGGGACT.GGGG.AGTGGC.GAGCCCTCAG.ATCCT....GCATATAAGCAG..CTGC....TTTTTGC.CTGT..ACTGGG..TCTCTC...TGGTTAGA.CCAGAT.CTGA 9565
H1O.CM.91.MVP5180             -----...............................AGCGTGGGAGGGATAAGG--C....GG-TC---G------.T-A-------.--G--...-------------.---------.---CC--.T---.---C---.----TAG...-TAGAG--.----G-.---- 9681
H1N.CM.95.YBF30               ------------T---GT.................................-G---.------.TT--------.-.G--...-----A-------.-C--.....--C-C--.T---.------.----------.-T-C-G--.-------.TA-- 9164
H1P.CM.06.U14788              ----A.......................................................................-------.-.G--...-------------.---------.---CAC--.T---.------TC---G-G.....AC---.-------.TA-- 9195
CPZ.CD.06.BF1167              ...............-TATGGGAGTGGTTAGGGGGGTGGCTT..........................--GT-----TTTC---G-T-.-GA-.....A-------------.---------.---GCA-T----AG-.G-CTCT-G---GAGC-ACCAG.............. 9760
CPZ.TZ.00.TAN1                ......................................TGTGGGTGGTTA-TG--C...-GGACA---G-----TTTT.G--.-GCTGAG--...-------------.---------.---GC--T-----..AAA--.C---TG-....CTAATCTG------.---- 9252
GOR.CM.07.SIVgorCP2135con     .......--G--AAA-T.....................CCCCGAGGGC-GAACT-G-A-GAGCA---G-----TTTC.A--CTC-G.AG--...-------------.---------.---GC--.-----.------TC---G-G.....AC---.-------.-A-- 9176

                                         NF-κ-B-I                                                                            TATA box                       TAR element start
                                                                                                                                                    3' LTR U3 end  3' LTR repeat start
MAC.US.x.239                  -G-TA......................................CGGGGAG-TACT--G-A-G.AGCC--TC-G-AAC-CC-A--TTC.T-G......ATG-----AT-TCA------.A---CGCTC.----.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAG10010
H2A.SN.86.ST_JSP4_27          -G-TACC....................................AGGGGAG-GACAT-G-A-G.AGCC--TG-G-AAC-CC-T-AT-CTT--T.....-T---A-T-T-C..-C--...TAC-CGCAT.----.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAG 9529
H2B.CI.x.EHO                  -G-AACA....................................GTGGGAG-GACAT-G-A-G.AGCA--G.AG-AAC-CC-ATCA--AC--T.....-T---A-T-T-C..-C--...T--G-GCAT.----.ATTCA-TC.G----GC.GGA-AGGCTGG-----.-GAG 9980
H2G.CI.92.Abt96               -G-AACA--G-AG..................................GTA--AT-G-A-GGACT--TT..G--AACG.---C.-AA-.T--..CT-T---A-TAT--.------.A---CGCTC.----.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAG 9438
H2U.CI.07.07IC_TNP3           -G-AACAA-G--A..................................G-TACAT-G-T-G..A-CT-G.-A--AACG.-----TAA-A-T...CT-T---A-TA--A..-C--...TC---GCAT.----.ATTCA-TC.G----GC.GGA-AGGCTGG.......... 9540
H2U.FR.96.12034               GG-AACCA-G-AG...................................-AGCAA-G-A-GAGCCA.TG..G--AACG.----.TCAA--GA-..CT-T---A-T-T-C..-C--...TC---GCAT.----.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAG 9539
ASC.UG.10.RT03                .................................................................................................................................................................... 8881
ASC.UG.10.RT08                .................................................................................................................................................................... 8905
ASC.UG.10.RT11                .................................................................................................................................................................... 8874
COL.CM.x.CGU1                 ------A-TT--GAG-...................................-A-CT--G.C-GAGG-AA-GGCGT-..........-GCTA-A.--GCTT-T---A---TCAG.AGC-.....-CGAGAGCT-CGAGG-................................ 8728
COL.UG.10.BWC01               -------A-T--GAG-...................................G-CT--G-C-GAGG-A-AGGCGT--CTGCT-AAACTGC......TT-T---A---ACAG.GGCTTGAGAAGCCTTCGA-GCTGGA-TTAGAACTCTTGCCT-GCT-CA--T--GC-CTG 8607
COL.UG.10.BWC07               ------CA-T--GCG-..................................AGG-AAG-C-...............T-CC--GCTTCTCG-A...GT-T---A---..................................................... 8733
DEB.CM.04.SIVdeb04CMPF3061    ................GTTGCTAAGGAGACACCATGGCAACTGAGGGAG-G-C................................................................................................................ 9265
DEB.CM.99.CM40                ................GTTGCCTAGGAAATATCATGGCAACCAGGGGAG-G-CT--GT--TAC.----AG.----CA..G--CGCTG-A-----.A----A----C-..A-----------.T-----.-GCTATCCA............................ 9227
DEB.CM.99.CM5                 ................GTAGTCAAGGAGACTCCATGGT.TACAGGGGAG-GACT--GC--TAC..---G-----TTTG.----GCTGT-A--...----------T-..A--------.T----..-GCTATCCA............................ 9158
DEN.CD.x.CD1_CMO580407        CC-GGGC---C.............................TCAGGGGAG-G-CT--GA--T--.CT--G-.GT-GCTT.A--CGCTTC-A--...-----A----T-..A-----------.T----.A--A-TCGG-T............................ 9598
DRL.DE.11.D3                  .................................................................................................................................................................... 9101
DRL.x.x.FAO                   -----......................................AGGGTGAC-T-G-TTGGGG-A....----TCTG.-------CT.G-A...-----A-------.--TG....C--ACTGCA---..-AAC--.------CCTG--AGGCTA-------.--- 9377
GRV.ET.x.GRI_677_gri_1        ------CA-T--G........................TGGATCGGAGGCG-TACA--G-C-GTAC.T--G------TTTC-----.AG-G-TG...-----A-------.A--.....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA..----C-..A-- 9433
GSN.CM.99.CN166               ACCATGACT-C-G-GCCCGGTTGCTGAGGCAACCGGGGCGGACTCGTGGGAGGGACTG--AGG-.C---G------TCTA---CGCTAT.--A...-----A-----A..AC---...-C..........AG-GACTCT--GGAGTC.................... 9414
GSN.CM.99.CN71                ACCATGACT-C-G-GCCCGGTTGCTGAGGCAACCGGGGCGGACTCGTGGGTGGGACTG--AGGC.C.--G------TCTA---CGCTAT.--A...-----A-------.AC---...-C..........AG-GACTCT--GGAGTC.................... 9410
LST.CD.88.SIVlhoest447        ------...................................--A----CGGG--A-G.T-T-........--A-T...-G--...-----------------.---------.-C-............................................... 8588
LST.CD.88.SIVlhoest485        ------...................................--A----CGGG--A-G.T-CA.......--A-..TT.G--...-----A.--------------.-C-............................................... 8589
LST.CD.88.SIVlhoest524        ------...................................--A----CGGG--A-G.T-T-.......--A-..TT.G--...-----------------.---------.-C-............................................... 8594
LST.KE.x.lho7                 G--TG...............................................................---A-T...-G--...--------------.-------.-C-CC--.T---..T-G--AGTC--T.GCTC.T-GGGA..----CAGTC-C 9710
MAC.US.x.251_1A11             -G-TA...........................CGGGGAGGTACTGG--A...-G.AGCC--TC-G-AACACC-A--TTC.T-G.....ATG-----AT-TCA-----.A----GCTA.----.AGTCA-TC.G----GC.GGA-AGGCTGG-----.TGAG10005
MND-1.GA.x.MNDGB1             -----..........................AAA...................CAGG-A--GG-A---TCG.---CA..TG.-TG...CT------------------.A---CGCT.-----.T-G--AG.-----ACTACA-AGGCTAA.......... 9063
MND-2.CM.98.CM16              ---T-CA-ACT-A....................................................................................................................................................... 9518
MND-2.GA.x.M14                ---T---ACT-A---CGGG...........................T---G-----TCAG.---CATCTGA.--...---------------.C-C--..T-GCCGCT.----.ATT.-A-TC----C-T........................ 9529
MND-2.x.x.5440                -----.A-ACT-ACGTCGGG.............................T---G-----CAG.---AA.CCTGG--...---------------.CT---..T-GCCGC-.----.A-G--A-TC----C-T..GA-AGGCTG------.TGAG 9298
MNE.US.x.MNE027               -G--..........................ATGGGGGG-TACT--G-A-G.AGCT--TC-G-AACACC-A--TTC.T-G.....ATG-----AT-TCA-----.A---CGCTC.----.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAG 9487
MON.CM.99.L1_99CML1           ACCATGACT-C-G-GCCCGGTTGCTGCGGCAACC.GGGCGGTCCAAAGGACTTG-T-G-T-GACCC---G------TTG.-----TCGT-T....CA----A-------.AC--...CC..........AG--ATCCT--GGAGTC.................... 9448
MON.NG.x.NG1                  GCC-TGACT-C-G-GCCCGGTTGCCGAGGCAACCGGGGCGGACTCGTGGGAGGGACTG--AGG-AC---G------TT-.-------ATC-....T----A-------.AC--...-C..........AGM-ATCCGC-GGAGTCAGCTC-C-TCT-CACA.GC-G 8106
MUS-1.CM.01.CM1239            ACCATGACT-C-G-GCCCGGTTGCTGAGGCAACCGGGGCGGACTCGTGGGTGGGACTG--AGG-ACT--G-----.CTT--------AT-T-...-----A----CA..GC--...-C..........AG--ATCCC--GGAGTCAGTTC---.CT-C-CA.GC-G 9408
MUS-1.CM.01.SIVmus_01CM1085   ACCATGACT-C-G-GCCCGGTTGCTGAGGCAATCGGGGCGGACTCGTGGGTGGGACTG--CGG-ACT--G-----.CTTA-------AT---...-----A-------.AC--....-C.........AG--ATCCC--GGAGTC.................... 9420
MUS-2.CM.01.CM1246            ACCATGACT-C-G-GCCCGGTTGCCGCAGCAACCGGGGCGGACTCA.GGGCGGACT-G-A-GGACCT-AG------TTG.-------ATC-A...-----A-------.AC--...-C..........AG--ATCCAA-GGAGTC.................... 9484
MUS-2.CM.01.CM2500            ..............GCCAGGTTGCCTGGGCAACCTGGGCGAACTCGTGGGTAAAACTG--AGGAGC---A------TT-.-------A--T-...-----A-------.-------.-C..........G--ATCCAGACGGAGTC................... 9430
MUS-3.GA.09.09GabOI81         AC-ATGACT-C-G-C.CTGGTTGCTGAGGCAACCRGGGCGGACTCATGGGTGGAACTG--CGGACC---G-----.CTTA-------AT-T-...-----A-------.AC--...-C..........AG--ATCCC-TGGAGTCAGCTC-T-TCT-C-CA.GC-G 9037
MUS-3.GA.11.11GabPts02        ACCATGACT-C-G-G.CTGGTTGCTGAGGCAACCAGGGCGGACTCGTGGGTGGGACTG--AGGAAC---G-----.CTTA-------AT-TG...-----A-------.AC--...-C..........AG--ATCCA--GGAGTCAGCTC-T-TCT-C-CA.GC-G 9020
OLC.CI.97.97CI12              .................................................................................................................................................................... 8678
RCM.CM.00.SIVagi_00CM312      -------A......................................---ACGT--GAG.----GT-----TTG.-----TCCT.G--...--------------.---------.---GC--T----.A-TCA-..------CCTG--AG.CTG-.M---.--- 8994
RCM.CM.02.SIVrcm_02CM8081     -------A......................................---ACGT--GAGG----GT-----TTG.-----TCCT.G--...--------------.---------.---GC--T----.ATTCA-..------CCTG--AGGCTA------.--- 9070
RCM.GA.x.SIVRCMGAB1           -----.................................G-GA--AC..-TGGGA----G-----TCAG.-----TCCT.G--...--------------.---------.-C-GC-.CT---.A-AAC--.G------CCTG--AGGCTA--G---.TGA. 9160
RCM.NG.x.NG411                -----.................................A-A---AC.-T-GGAG----GT-----TTG.-----TCCT.G--...--------------.--------.A---C--T----.A--CA-..------CCTG--AGGCTA------.--- 9251
SAB.SN.x.SAB1                 ..........................GGTGGAGA-TG--C...-GTAC.T--G------TTG.---CGT-CT.G-----.CA-----------.---------.---.GCGC----AA-ACA-..-----GC.T---AGGCTG------.--- 9777
SMM.CI.79.SIVsmCI2            ........................CCGGGGGAGGTGACAGTGGGA-GA-CAAG--AGGGGCT--T.-G-AAC.GC--A-ATAATCT--.....-T---A-TAA--..-CTG..CT--GCACTG.-A-.TCTCA-TC.G----GC.GGA-AGGCTGG-......... 9147
SMM.LR.89.SIVsmLIB1           ........................CAAGGGGCTGTAAC.ATGGGA-GTACTTG--AGGAGTC--TG-G-AAC.GC-TA-TTCTT-T--....-T---A-TAA-...-CCG..CT-AC-GCTC.----.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAG 9067
SMM.SL.92.SIVsmSL92A          ........................AAAGGGGCTGTAAC.ATGGGA-GAACATG--AGGGGCT--TG-G-AAC.GC--A-A-AATCT--.....-T---A-TAT-...-CAGC..-----GCAC.---ATTTCA-TC.G----GC.GGA-AGGCTGG-......... 9086
SMM.SL.92.SIVsmSL92B          .................................................................................................................................................................... 7460
SMM.SL.92.SL92B               -G-A--A--G..........................AGGAGTCT--GC--GGGGAA..C..G--AACG.---CT-AAC--TG...TAG-----AT-TAC-C--..TT-GCCGCTC.----.ATTCA-TC.G----AC.GGA-AGGCTG------.AGAG 9469
SMM.US.04.G078                .................................................................................................................................................................... 9157
SMM.US.05.D215                .................................................................................................................................................................... 9128
SMM.US.06.FTq                 .................................................................................................................................................................... 9149
SMM.US.86.CFU212              .................................................................................................................................................................... 9130
SMM.US.x.H9                   -G-TAC.........................GGGGAG-TTCT--G-A-G.AGCT--C.TG-AAC-CC-A--T-TTC--T.....-T---A-TA--A.-------.A---CGCTC.----.ATTCA-TC.A----GC.GGA-AGGCTGG-----.TGAG 9493
SMM.US.x.PGM53                -G-TA...........................CGGGGAG-TACT--G-A-G.AGCT--C.TG-AAC-CC-A--T-TTT--T.....-T---A-TA--A.-------.A---CGCTC.----.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAG 9947
SMM.US.x.pE660.CG7G           ...........................C.AAGGGGCTGTCA.TGGGGAG-TACT--G-A-G.AGCT--CT.G-AAC-CC-A--T-TTC--T.....-T---A-TAT-A.-------.A---CGCTC.----.ATTCA-TC.G----GC.GGA-AGGCTGG-----.TGAG10014
STM.US.89.STM_37_16           -G-AACA.........................GGGGAGGTA-TG--A...-G.AACT--TG-G-AAC-CC-TT-T-CTT--T.....-T---A-TA--A..-C--...T-A-AGCTC.----.ATTCA-TC.A----GC.GGA-AGGCTGG-----.TGAG 9664
SUN.GA.98.L14                 ...................................AGG-A--GG-A---.TTA---A-T...-A--...--------------.---------.-C--C--C.----..T-G--AG.------..GC-CAGAG-..AT--GCGAC-- 9753
SYK.KE.x.KE51                 .................................................................................................................................................................... 9068
SYK.KE.x.SYK173_COMGNM        ......ACT....-CATAAAAGCCG..ATGC....TC.......TGTAGCTATCC-CG-A-TTGCAACTA.........................GCT-TCTCC---G-CCT...GGG.AGCCCTGAGA-ACTGG-A-AGCGGC-C.GGGA...................... 9438
TAL.CM.00.266                 .................CGTTGCTAAGGCAACGGGGGACGGACTGGGGCG-GGA-C-G-A-GAGTT.--G------TTG.-------A-.--A...-------------.A--.....-C-G-AG...........CTATC--T-GCTC.................... 9338
TAL.CM.01.8023                .................CGCTGCTAAGGCAACGGGGGACGGACTGGGGCG-GGA-C-G-A-GAGCT.--G------TA..-----.............................................................................. 8845
TAN.UG.x.SIVagmTAN1           .........................................GGTG-TGC-A-G-C-GTACTT--G------TT.C-------.-.G--...-----A-------.A---.....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA..----C-..--- 9596
VER.DE.x.AGM3                 ---------.................................CGGGTCATG--C...-GTAC.---G------TTTA-------.-.G--...-----A-------.A---.....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA..----C-..--- 9123
VER.KE.x.9063                 ------.............................GGGCGGGCCATG--C...-GTAC.---G-----TTTTA-------.-.G--...-----A-------.A---.....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA..----C-..--- 9623
VER.KE.x.AGM155               ------A--C--......................................GACATG--C...-GTAC.---G------TTTA-------.-.G--...-----A-------.A---.....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA..----C-..--- 9604
VER.KE.x.TYO1_patent          ------A--C--G.........................................ACAT-G-C-GTCC.---G------TTTA-------.-.A--...-----A-------.A---.....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA..----C-..--- 9587
WRC.CI.97.97CI14              ....................TCCTGAGACTGTGTACCAACGGACTTA..................................................................................................................... 9484
WRC.CI.98.98CI04              .................................................................................................................................................................... 9461
WRC.GM.05.Pbt_05GM_X02        ....................TCCGCTAGCTGTGTACCAACGGCCTTAGCAA--AT-CTAATGAACT---ACTCT-GGAG-AGACTT-TGGTT-TGCCCT---A---T--ATGCT-TGTAGC-A-CTGG-TGCAGTCTTCTGC--GAGGCTA---GCCTGAGT-G-GG-CC 8537
```

TAR element end

```
H1B.FR.83.HXB2              GCCTGGGAG.CTCTCTGGC.....TAA.............................CTAG.....................................................................GGAACCC.....ACTGCTTAA..... 9606
H1O.CM.91.MVP5180           ---C------.---C---------.CT..............................----------.....................................................CT----------.G---------..... 9722
H1N.CM.95.YBF30             -T--------.-ATAT---G.....................................----------.....................................................----------..... 9182
H1P.CM.06.U14788            ---------.---------.....-CTAA..........................----------.....................................................GCA----------.G----AAA..... 9238
CPZ.CD.06.BF1167            ---------.---------.....................................ATCT-AACCTGGG.AGTTCTCTGGTAGTCA.....................GGGCATGAGACCACTGCTT.AACGCTC..... 9815
CPZ.TZ.00.TAN1              ---------.---------.....G...............................-G-CTGGC.....................................................TA-AG---------.G--------..... 9297
GOR.CM.07.SIVgorCP2135con   ---------.--------T.....CTTA..........................----------.....................................................GCA----------.---------..... 9216

MAC.US.x.239                C--------GT-----CCA.....GC-CTAGCAGGTAGAGCCTGGGTGTTCCCTG----ACTCTC.ACCAGCACTTGGCCGGTGCTGGGCAGAG.....................................................10098
H2A.SN.86.ST_JSP4_27        C--------GT-----CCA.....GC-CTAGCAGGTAGAGCCTGGGTGTTCCCTG----ACTCTC.ACCAGTGCTTGGCCGGCACTGGGCAGA.....................................................9616
H2B.CI.x.EHO                C--------GT-----CCA.....GC-CTAGCAGGTAGAGCCTGGGTGTTCCCTG----ACTCTC.ACCAGCACTTGGCCGGTGCTGGGCAGA.....................................................10067
H2G.CI.92.Abt96             C--C-----GT-----CCA.....GC-CTAGCAGGTAGAGCCTGGGTGTTCCCTG----ACTCTC.ACCAGCACTTGGCCGGTGCTGGGCAGAG....................................................9526
H2U.CI.07.07IC_TNP3         ..........................................................................................................................................9540
H2U.FR.96.12034             C--------GT-----CCA.....GC-CTAGCAGGTAGAGCCTGGGTGTTCCCTG----ACTCTC.ACCAAAAG.....................................................9607
ASC.UG.10.RT03              ..........................................................................................................................................8881
ASC.UG.10.RT08              ..........................................................................................................................................8905
ASC.UG.10.RT11              ..........................................................................................................................................8874
COL.CM.x.CGU1               ..........................................................................................................................................8728
COL.UG.10.BWC01             AGAG-CTT-T-AGC-ATC-AACTAGCCTGAGCCCTGAGTGGCTTGCTAGAGTGGAAGG-CAAGCGGACTGCTCCAGTCCCC.....................................................8688
COL.UG.10.BWC07             ..........................................................................................................................................8733
DEB.CM.04.SIVdeb04CMPF3061  ..........................................................................................................................................9265
DEB.CM.99.CM40              ..........................................................................................................................................9227
DEB.CM.99.CM5               ..........................................................................................................................................9158
DEN.CD.x.CD1_CMO580407      ....C---.-------CCAGCGGCCTGG.....................................................................................AGC--TGAG..GGC----GCAGCGGCCTGGGAGCCC 9656
DRL.DE.11.D3                ..........................................................................................................................................9101
DRL.x.x.FAO                 -------T-.T-------T.....G-GTCTGCTGAAGGAGGCTACCTGCTCGAGC-CC-GGTGGTTCGCAGGGCCTTGGCCTTGGGCTCTGGGTAGCTCGGTCAGGAGGCTTCTGGAGGTAGATCACTAGGG-CTG-T-TGGTCTTG---TCT.GGTAGACCTTCGCTA 9537
GRV.ET.x.GRI_677_gri_1      -------T-.T--G----T.......GCCTA...ACCCGGT.TGGCCA.CCGGGGG--A.....................................................---CT-------.TTG----C-TAT..... 9500
GSN.CM.99.CN166             ..........................................................................................................................................9414
GSN.CM.99.CN71              ..........................................................................................................................................9410
LST.CD.88.SIVlhoest447      ..........................................................................................................................................8588
LST.CD.88.SIVlhoest485      ..........................................................................................................................................8589
LST.CD.88.SIVlhoest524      ..........................................................................................................................................8594
LST.KE.x.lho7               T------T-AGCACACT-GT...CAGCTAAGC.AGTCAGCCCTGGGAGGCTACA.....CTTAGC.....................................ACCCACCTGGGCCTG-T-AGCCTAGGGAGC---G..... 9800
MAC.US.x.251_1A11           C--------GT-----CCA.....GC-CTAGCAGGTAGAGCCTGGGTGTTCCCTG----ACTCTC.ACCAGCACTTGGCCGGTGCTGGGCAGAG.....................TGACTC-A-GCTTGCT-------AG.....10118
MND-1.GA.x.MNDGB1           .......G-GT-G-A-CT.....................CTGAGCAGATCCC--.....................................TAGAGCAAGGACCAGAGTCCTGAGTGACTGGGTCTGAGCACCTCACTCGGGGCTGATCACCTC. 9152
MND-2.CM.98.CM16            ..........................................................................................................................................9518
MND-2.GA.x.M14              ..........................................................................................................................................9529
MND-2.x.x.5440              C.-----..TTGT---CT.GG.TGAGTTCTTG.AAGGAG..TGCCTGCTTGTAGC-CT-GGCGGT.....................TCGCAGGCCCCTGGCTTGTGGCTCTGGGTAGCTCGTCA--TGTT-TGGAA-GGCT-ACGGGAGGACACTCTTGCTT 9431
MNE.US.x.MNE027            C--------GT-----CCAGCACTA........GCAGGTAGAGCCTGGGTGTTCC--.....................................GCTA--CT-TC.ACC-GCA---GGCCGGTGCTGTGCAGAG. 9575
MON.CM.99.L1_99CML1         ..........................................................................................................................................9448
MON.NG.x.NG1                --------C-CT-GGA--T....AGGTAAAGCGGCCT...GGGAGCCCTTGCCTAG--AGGTGCCGGCTCTCGGCTCGCTTGCTGAATAGGCTTTACCAGCCCAATAAAGTGCTTGAAAGCA.....................8221
MUS-1.CM.01.CM1239          --------C-CTCGGAA-T....AGGTAAAGCGGCCT...GAGAGCCCTTGCCTAGGAGAACT..GGTTCTCGGCTTGCTTTGCTTCGAAGCTTATTGAGCCCACAATA.....................9508
MUS-1.CM.01.SIVmus_01CM1085 ..........................................................................................................................................9420
MUS-2.CM.01.CM1246          ..........................................................................................................................................9484
MUS-2.CM.01.CM2500          ..........................................................................................................................................9430
MUS-3.GA.09.09GabOI81       --------C-CTCGGA--T....AGGTAGAGCGGCCT...GGGAGCCCTTGCCTAG--AG.TGCTGATTCTCGGCTTGCTCTTGCTTCGGCTTTATCAGCCCAAATAAAGTAGTTGCTTGAGTCAAGTCTAA-CC-T-ATTA..T-CGCGGTGTGGTGTGCCTTCGGGA 9196
MUS-3.GA.11.11GabPts02      --------C-CTCGGA--T....AGGTAAAGCGGCCT...GGGAGCCCTTGCCTAG--AGTTGCTGATTCTCGGCTTGCTCTAGCTTAGGGTCTTTATAGCCCAATAAAGCGCTTTAACTAGAGTCAAGTCTAGC--TCATCTGT-CGCGGCGAGCGGGGCCTTCGGGA 9182
OLC.CI.97.97CI12            ..........................................................................................................................................8678
RCM.CM.00.SIVagi_00CM312    -------T-.T-------T.....A-GTCTCTAGGAGACT.AGCTTGAGCCTGGGTGTTCGCT.....................................AGT.TCTCTGAA.GGGCTTGCTG--G-G-GCTCTGG---GGGTCGTGGTAGACCTTCAGCT 9108
RCM.CM.02.SIVrcm_02CM8081   -------T-.T-------T.....A-GTCTCTAGGAGACT.AGCTTGAGCCTGGGTGTTCGCT.....................................AGT.TCTCTGAA.GGGCTTGCTG--G-GTACTCTTG--TGG.TCTTGGTAGACCTTCAGCT 9184
RCM.GA.x.SIVRCMGAB1         -------T-.T-------T.....A-GTCTCTAGGAACTCCAGCTTGAGCCTGGGTGTTCGCT.....................................GGTGTCTCTGAACAGGCTTGCTG--GTG--TCTCGCTCTTCGGGTAGACCGCCA.....GT 9273
RCM.NG.x.NG411              -------TTGT-------T.....A-GTCTCTAGGAGACT.AGCTTGAGCCTGGGTGTTCGCT.....................................AGT.TCTCTGAA.GGGCTTGCTG---.G-GCCTTGG--.GGGTCTTGGTAGACCTTCAGCT 9365
SAB.SN.x.SAB1               -------T-.T-------T.....A-GTCTG..AACCAGCT.TGAG...CCTGGGTGTTCGCT.....................................GGGGAAGGCAAGCCTAAGAGCACTCTGGT-AGG.T-.....TTG-TAG-CCC..........TCGCCT 9880
SMM.CI.79.SIVsmCI2          ..........................................................................................................................................9147
SMM.LR.89.SIVsmLIB1         C----A-T-GT-----CCA.....GCGCTAGCAGGTAGAGCCTGGGTGTTCCCTG----ACTCTC.ACCAGCACTTGGCCGGTGCTGGGCAGGG....................................................9155
SMM.SL.92.SIVsmSL92A        ..........................................................................................................................................9086
SMM.SL.92.SIVsmSL92B        ..........................................................................................................................................7460
SMM.SL.92.SL92B             C--C-A---GT------CA.....GC-CTTAGTAGGTGAGCCTGGGTGTTCTCTA---AACTCTCAACCAGCAACAGGCCATTGCTGGGTAGAC.....................................9558
SMM.US.04.G078              ..........................................................................................................................................9157
SMM.US.05.D215              ..........................................................................................................................................9128
SMM.US.06.FTq               ..........................................................................................................................................9149
SMM.US.86.CFU212            ..........................................................................................................................................9130
SMM.US.x.H9                 C--------GT-----CCA.....GC-CTAGCAGGTAGAGCCTGGGTGTTCCCTG--G-ACTCTC.ACCAGCACTTGGCCGGTGCTGGGCAGAG.....................................................9581
SMM.US.x.PGM53              C--------GT-----CCA.....GC-CTAGCAGGTAGAGCCTGGGTGTTCCCTG----ACTCTC.ACCAGCACTTGGCCGGTGCTGGGCAGAG.....................................................10035
SMM.US.x.pE660.CG7G         C--------GT-----CCA.....GC-CTAGCAGGTAGAGCCTGGGTGTTCCCTG----ACTCTC.ACCAGCACTTGGCCGGTGCTGGGCAGAG.....................................................10102
STM.US.89.STM_37_16         C--------GT-----CCA.....GC-CTAGCAGGTAGAGCCTGGGTGTTCCCTG----ACTCTC.ACCAGCACTTGGCCAGTGCTGGGTAGAG.....................................................9752
SUN.GA.98.L14               C----A----G----ACCT.....AG-GCTGGCTTGTCGCTCCTGAGTGAGCACTTGCTAAG.....................................CACCTCACCCGAGCCTGAT-A----GGGGAGCTAGAGG..... 9847
SYK.KE.x.KE51               ..........................................................................................................................................9068
SYK.KE.x.SYK173_COMGNM      ..............GCCCTTGCCAGT-AAGGCTAGTGAGGTCCGGGTTTGGCCGC-CG-TCTTCGGCAGCTCTCCGTTGCTT.....................................-CTTTGCTTACG---TGAT..... 9523
TAL.CM.00.266               ..........................................................................................................................................9338
TAL.CM.01.8023              ..........................................................................................................................................8845
TAN.UG.x.SIVagmTAN1         ---------.T--G----T.....G-GCCTA...ACCTGGTCTGGCCATCCAGGGG--A.....................................GA-CCT-------.T-----GAA..... 9664
VER.DE.x.AGM3               -------T-.T--G----T.....--GCCTA...ACCTGGT.TGGCCA.CCAGGGG--A.....................................---CT-------.TTG-----GAA..... 9189
VER.KE.x.9063               -------T-.T--G----T.....-GCCTA...ACCTGGT.TGGCCA.CCAGGGG--A.....................................---CT-------.TTG----GGAA..... 9689
VER.KE.x.AGM155             -------T-.T--G----T.....-GCCTA...ACCTGGT.TGGCCA.CCAGGGG--A.....................................---CT-------.TTG-----GAA..... 9670
VER.KE.x.TYO1_patent        -------T-.T--G----T.....-GCCTA...ACCTGGT.TGGCCA.CCAGGGG--A.....................................---CT-------.TTG-----GAA..... 9653
WRC.CI.97.97CI14            ..........................................................................................................................................9484
WRC.CI.98.98CI04            ..........................................................................................................................................9461
WRC.GM.05.Pbt_05GM_X02      CA--T--CCTAG----CTTGGTGCCTGAGTGGTGGTCCCACTTGGCCAAGAGACT-AG-TAGGAG.....................................CTGACCACGCTT-CTTGTTAGGGTGT-TGCAG-T..... 8636
```


Draft of 30 September 2015, not for public release
plvdna_main.tex r8876–10546

```
                                                                                                    Poly-A signal                3' LTR R repeat end 3' LTR U5 start
H1B.FR.83.HXB2                .......................................................................GCCTC.....AATAAAG...CTTGCC.......TTGAG..TG.CTTCAA.GTAGTG...TGTGCCCGTC.TG.TT..GTG.TGACTCT...GGT..AA. 9672
H1O.CM.91.MVP5180             .......................................................................CG---.....-------...------.......------..--AGAAGC-GTGT---CTCATCTGTTCAACCCTGGTGTCTAGAGA--............. 9792
H1N.CM.95.YBF30               .......................................................................................................................................................................... 9182
H1P.CM.06.U14788              .......................................................................................................................................................................... 9238
CPZ.CD.06.BF1167              .................................................................................-------...------.......-GAGAG.--TTA.TC-GTGT---CCCATCT-ATTC-GC-CCGGCCCTGG-GTAGAGATCCCTC-GA 9892
CPZ.TZ.00.TAN1                .......................................................................CG-.TC...---------.-C----.......-----.A-.TG-T.-.................................................. 9326
GOR.CM.07.SIVgorCP2135con     .......................................................................CG---.....-------...------.......A------.--AA--................................................. 9246
                                                      TAR element end                               Poly-A signal            3' LTR R repeat end 3' LTR U5 start
MAC.US.x.239                  ......................................TGACTCCACGCT.TGCTTGC...TTAA.AG...C----TTC.-------.--.---ATT.....--AGA..A-TAAG-T-..GT---TGT-CCCAT-TCT.CCTAGCC-CC.GCCTGG-...CAAC.TCG10202
H2A.SN.86.ST_JSP4_27          ......................................CGGCTCCACGCT.TGCTTGC...TTAAAAG...A----TT..-------.--.---AG......--AGA..A-CA.............................................................. 9672
H2B.CI.x.EHO                  ......................................CGGCTCCACGCT.TGCTTGC...TTAA.AA...C----TT..------A..--.---TA......--AGA..A-CAAGAC-.AGT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CA-..TCG10169
H2G.CI.92.Abt96               ......................................CGGCTCCACGCT.TGCTTGC...TTACAACG..A----TT..-------.--.---AAA.....--AGA..A-TAAG---GTGT---TGT........................................ 9599
H2U.CI.07.07IC_TNP3           .......................................................................................................................................................................... 9540
H2U.FR.96.12034               .......................................................................................................................................................................... 9607
ASC.UG.10.RT03                .......................................................................................................................................................................... 8881
ASC.UG.10.RT08                .......................................................................................................................................................................... 8905
ASC.UG.10.RT11                .......................................................................................................................................................................... 8874
COL.CM.x.CGU1                 .......................................................................................................................................................................... 8728
COL.UG.10.BWC01               .......................................................................................................................................................................... 8688
COL.UG.10.BWC07               .......................................................................................................................................................................... 8733
DEB.CM.04.SIVdeb04CMPF3061    .......................................................................................................................................................................... 9265
DEB.CM.99.CM40                .......................................................................................................................................................................... 9227
DEB.CM.99.CM5                 .......................................................................................................................................................................... 9158
DEN.CD.x.CD1_CMO580407        TCTAAGGTATTGCAACAAGTACC................................................................................................................................................... 9679
DRL.DE.11.D3                  .......................................................................................................................................................................... 9101
DRL.x.x.FAO                   CTTCTAGTGGCCATAG.AAGTAGGGGAGTAGGCACCGCTTGCTTATCCTAAT...................TG---.....-------..--AC--A.......--T--ATG-ATGAGT-.AGT--ACTC-TGT-TTT-G........T-CT-AGTGAGACTCTGGTT-C 9664
GRV.ET.x.GRI_677_gri_1        .......................................................................AG---.....-----C...C---TCGC.....--AGT..C-CTA-ATTG.G---CAAG--CT-ATTG-TGCGCCGA-CCT.CTAGAGG......TG-A 9576
GSN.CM.99.CN166               .......................................................................................................................................................................... 9414
GSN.CM.99.CN71                .......................................................................................................................................................................... 9410
LST.CD.88.SIVlhoest447        .......................................................................................................................................................................... 8588
LST.CD.88.SIVlhoest485        .......................................................................................................................................................................... 8589
LST.CD.88.SIVlhoest524        .......................................................................................................................................................................... 8594
LST.KE.x.lho7                 .................GGCTCCTAGCTTGCTACTT..TAAA...GCCTTC.........................-------.....AA-TTGCTCATTG-T--AAA-TACAAGCTAGT--CTCA-AGTAGTC--T-TC-CCTTC-CCC.........TGGTTCA 9904
MAC.US.x.251_1A11             .......................................................................C----TTC.-------.--.---ATT.....--AGA..A-TAAG-T-..GT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CAAC.TCG10197
MND-1.GA.x.MNDGB1             ......................................GAGGTAGTGGAACTCCTTGC.TTGCTTGCTA..TTG--TTC.-------TAA--.TAGAA......--AGA...-CA.............................................................. 9215
MND-2.CM.98.CM16              .......................................................................................................................................................................... 9518
MND-2.GA.x.M14                .......................................................................................................................................................................... 9529
MND-2.x.x.5440                GGTCTTGATAGACCTCTAGCAATCCCTTAGGCCTGGAGATTGTGGGATTGACTACCGCTTGCTTGCTATTATTGC-C...----------....C-TA-CTGATTAGAATAGCAAGTAAG-G.TACTC-TCCAT-TTGTC-TCAATAGAA...ACTCTGG-TACC-GAG-.. 9587
MNE.US.x.MNE027               ......................................TGGCTCCACGCT.TGCTTGC...TTAA.AG...A----TTC.-------.--.---AT......--AGA..A-TAAG-C-..GT---TGC-CCCAT-TCT.CCTAGTC-CC.................... 9664
MON.CM.99.L1_99CML1           .......................................................................................................................................................................... 9448
MON.NG.x.NG1                  .......................................................................................................................................................................... 8221
MUS-1.CM.01.CM1239            .......................................................................................................................................................................... 9508
MUS-1.CM.01.SIVmus_01CM1085   .......................................................................................................................................................................... 9420
MUS-2.CM.01.CM1246            .......................................................................................................................................................................... 9484
MUS-2.CM.01.CM2500            .......................................................................................................................................................................... 9430
MUS-3.GA.09.09GabOI81         GGCTCCGAGTCGGGAAGGACTCAGGAAACCTTCCAGTTGGCGCCCGAACAGGGACCCGAGTAGCCCAGAAG--TCGGGCCCGGAGGC-GAGGACCG-GTGAGGAG--CCCCA-C-GAGTGA-CCCGATCTCCGAAGAAA-AGCTGCGCAGCGAAGAAGGTAG-AGACGGG 9366
MUS-3.GA.11.11GabPts02        TCTTGAGTCGGGAAGGACTCGGAAA...CCTTCCAGTTGGCGCCCGAACAGGGACTTGAGCTACCCAGAAGA-TCAAGCCC-CAGGC-GAGGAACG-GAACGAGACTGCCCCAG-CGAGTG--CCG-CCTCC--AAAAA-AGCTGCGCAGCGAAGAGAGGTAA-GAGCCG 9349
OLC.CI.97.97CI12              .......................................................................................................................................................................... 8678
RCM.CM.00.SIVagi_00CM312      GGGTACTTGGCCAGTCCCAGTGGGA..GTAGGCTCCACTTGCTTGCAGCTAATTGT..............TA---.....------.CTA-A----......A-T--TTGAA--G---C.A----TGTGCCCGTTT-A....--CCTCAG-.GG-AACGACTCTGGGG 9242
RCM.CM.02.SIVrcm_02CM8081     GGGTACTTGGCCAGTCCCAGTGGGA..GTAGGCTCCGCTTGCTTGCAGTTAAGATTGT............TA---.....------C.TA------......A-T--TTGAA--G---C.A----TGTGCCCATTT-A....--CCTCAG-.GG-AACGACTCTGGGG 9320
RCM.GA.x.SIVRCMGAB1           TGAGGCTCGGCCGGCCTCAACGGGA..GAGATCACCGCTTGCTTATAGCC..TTGA......................................AGCTC.................................AATAAAGC..AT.GCCA-.....T 9346
RCM.NG.x.NG411                GGGTACTTGGCCAGTCCCAGTGGGA..GTAGGCTCCGCTTGCTTTCTGGTTAAT................TG---.....------..--.--ATGCTCTTAGTT-AACCAGAATC-.AGT---TG..CCCATTT-A......TCTCTCA-TGAAACGACTCTGGGG 9497
SAB.SN.x.SAB1                 GGCGACT.GGCCATTGCCAGTAGCA..G.AGACTCCGCTTGCTTGCTTGATT.................-----.....-------.T.------........--AGTTTACAAG-A--GCA----TGTGCCCATTTATTCCTCAG...................... 9990
SMM.CI.79.SIVsmCI2            .......................................................................................................................................................................... 9147
SMM.LR.89.SIVsmLIB1           ......................................TGGCTCCACGCT.TGCTTGC...TTAA.AG...A----TTT.-------.--.---AG.....A-AGA..A-CAAG---..GT---TGT........................................ 9224
SMM.SL.92.SIVsmSL92A          .......................................................................................................................................................................... 9086
SMM.SL.92.SIVsmSL92B          .......................................................................................................................................................................... 7460
SMM.SL.92.SL92B               .......................................GGCTCCACGCT.TGCTTGC...TTGA........---TC...--------.G--.--AAC......--AGA..A-CAAGTC........................................................ 9613
SMM.US.04.G078                .......................................................................................................................................................................... 9157
SMM.US.05.D215                .......................................................................................................................................................................... 9128
SMM.US.06.FTq                 .......................................................................................................................................................................... 9149
SMM.US.86.CFU212              .......................................................................................................................................................................... 9130
SMM.US.x.H9                   ......................................TRGCTCCACGCT.TGCTTGC...TTAA.AG...A----TTC.-------G.--.---AG......--AGA..A-TA.............................................................. 9638
SMM.US.x.PGM53                ......................................TGGCTCCACGCT.TGCTTGC...TTAA.AG...A----TTC.-------.--.---GC......--AGA..A-TAAG---..GT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CA-C.TCG10138
SMM.US.x.pE660.CG7G           ......................................TGGCTCCACGCT.TGCTTGC...TTAA.AG...A----TTC.-------.--.---AT......--AGA..A-TAAG---..GT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CA-C.TCG10205
STM.US.89.STM_37_16           ......................................TGGCTCCACGCT.TGCTTGC...TTAT.AT...AT---TTC.-------.--.---AAT.....--AGA..A-TAAG---..GT-G-TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CAAC.TCG 9856
SUN.GA.98.L14                 ......................................CTCCTTGCTTTGCAGTACAAG...........C-T--.....---------...-C----.........--AG...-TTTACTGCAAGCAA-TGT-CCCGTGT--TCTC--TCTCC...............TTCT 9934
SYK.KE.x.KE51                 .......................................................................................................................................................................... 9068
SYK.KE.x.SYK173_COMGNM        .......................................................................CAGC-CA...-------T.GTGGTAAACGCAAG-A.....ACGTC-GCT.C-CA..TCTGT-CTG-TGGG-.CACCTCCCGA.-GGTG........... 9597
TAL.CM.00.266                 .......................................................................................................................................................................... 9338
TAL.CM.01.8023                .......................................................................................................................................................................... 8845
TAN.UG.x.SIVagmTAN1           ....................................................GGCTATT............G--.....------....G-----TG...C--A--..A-T--ATCTGAG.CAAGTGCC-CATT-CGCGCCC-TCC-CAG-................ 9736
VER.DE.x.AGM3                 .......................................................................AG-.T......------T...---CG-TG...CA--AG-.CA-AGACTT..--GAGAGA....CATT-C-TG-TGCATCCAGCACAAG............ 9257
VER.KE.x.9063                 .......................................................................AG--.......------AC..T-----TG...CA--AGA..G.CT-ATCTGAGTCAAGTGCCCT-ATTGACG.CC-CAC-C..A-GCAGG.....GG--C 9766
VER.KE.x.AGM155               .......................................................................AG--.......------........------TG...CA--AGA..G.CT-ATCTGAGTCAAGTGCCCT-ATTAATG.CC-CAC-C..TTGAACG.....GG-GA 9745
VER.KE.x.TYO1_patent          .......................................................................AG--.......------........------TG...CA--AGA..G.CT-ATCTGAGTCAAGTG-CCT-ATTGACG.CC-CAC-CTCTTGAACG......GG-A 9729
WRC.CI.97.97CI14              .......................................................................................................................................................................... 9484
WRC.CI.98.98CI04              .......................................................................................................................................................................... 9461
WRC.GM.05.Pbt_05GM_X02        .................................................................................---------..AG-CTTTTGCTTAGCA--AAGCGTG-GGTTAGTT-TTGT---AAGAAGTAGAACCCT-G-AG-CT-............. 8709
```

3' LTR U5 end

```
H1B.FR.83.HXB2            CTAGAGA..............TCCCT..CAGA.....................CCCTTTTAGTCAGTGTGGAAAATCTCTAGCA.......................................................... 9719
H1O.CM.91.MVP5180         ..................................................................................................................................... 9792
H1N.CM.95.YBF30           ..................................................................................................................................... 9182
H1P.CM.06.U14788          ..................................................................................................................................... 9238
CPZ.CD.06.BF1167         TCTCT--AGGCTAAGCAAAAATCT-TAC..C-GGCCC.............................................................................................. 9927
CPZ.TZ.00.TAN1            ..................................................................................................................................... 9326
GOR.CM.07.SIVgorCP2135con ..................................................................................................................................... 9246
```

3' LTR U5 end

```
MAC.US.x.239              G-...-CTCAATA....ATAAGAAGA-CCTGGTCTGTTAGGACCCTTTCTGCTT.TGGGAA-CCG-AGCA--------C------------------------------------------10279
H2A.SN.86.ST_JSP4_27      ..................................................................................................................................... 9672
H2B.CI.x.EHO              G-..GCTCCAC......TA..GAAA-CCTGGTCTGTTAGGACCCCTTCTGCTT.TGGGAA-CCA--GCA--------C------------------------------------------10242
H2G.CI.92.Abt96           ..................................................................................................................................... 9599
H2U.CI.07.07IC_TNP3       ..................................................................................................................................... 9540
H2U.FR.96.12034           ..................................................................................................................................... 9607
ASC.UG.10.RT03            ..................................................................................................................................... 8881
ASC.UG.10.RT08            ..................................................................................................................................... 8905
ASC.UG.10.RT11            ..................................................................................................................................... 8874
COL.CM.x.CGU1             ..................................................................................................................................... 8728
COL.UG.10.BWC01           ..................................................................................................................................... 8688
COL.UG.10.BWC07           ..................................................................................................................................... 8733
DEB.CM.04.SIVdeb04CMPF3061 ..................................................................................................................................... 9265
DEB.CM.99.CM40           ..................................................................................................................................... 9227
DEB.CM.99.CM5            ..................................................................................................................................... 9158
DEN.CD.x.CD1_CMO580407   ..................................................................................................................................... 9679
DRL.DE.11.D3             ..................................................................................................................................... 9101
DRL.x.x.FAO              TGGAGA..............TC--TCAGACT-GGTGAGAGCTATCCTGAGCTTTAG--GAGAGTAAA-A--------C------------------------------------------ 9731
GRV.ET.x.GRI_677_gri_1    -CTCTCTTACTGGGTTCTCCTG-A--CAGGT.....................................G-G--G--A---C------------------------------------------ 9623
GSN.CM.99.CN166           ..................................................................................................................................... 9414
GSN.CM.99.CN71            ..................................................................................................................................... 9410
LST.CD.88.SIVlhoest447    ..................................................................................................................................... 8588
LST.CD.88.SIVlhoest485    ..................................................................................................................................... 8589
LST.CD.88.SIVlhoest524    ..................................................................................................................................... 8594
LST.KE.x.lho7             G-GATC..................T--CA-TAGAGAGATTGGAGCCTTGTGAT--GGGAC--GGCT-CCA------------------------------------------------ 9957
MAC.US.x.251_1A11         G-...-CTCGGTA....ATAAGAAGA-CCTGGTCTGTTAGGACCCTTTCTGCTT.TGGGAA-CCG-AGCA--------C------------------------------------------10274
MND-1.GA.x.MNDGB1         ..................................................................................................................................... 9215
MND-2.CM.98.CM16          ..................................................................................................................................... 9518
MND-2.GA.x.M14            ..................................................................................................................................... 9529
MND-2.x.x.5440            ..........................------.T----TCTGTGGCAGAGATTTGATAT-TAG-G--AGTAGAGAA--------C------------------------------------ 9647
MNE.US.x.MNE027           ..................................................................................................................................... 9664
MON.CM.99.L1_99CML1       ..................................................................................................................................... 9448
MON.NG.x.NG1              ..................................................................................................................................... 8221
MUS-1.CM.01.CM1239        ..................................................................................................................................... 9508
MUS-1.CM.01.SIVmus_01CM1085 ..................................................................................................................................... 9420
MUS-2.CM.01.CM1246        ..................................................................................................................................... 9484
MUS-2.CM.01.CM2500        ..................................................................................................................................... 9430
MUS-3.GA.09.09GabOI81     -GG-CCTATAGAGGGTGGCAGACT-A-AGGTCGGTAGGGTCGGGGAGATCCTA-AGAG--TTAG-AGCGC--G--AGCT-CA--GAGGAAAGGGCCAATGGGGGCGAAGCACTCGGCGATGCTGACTGGTACTAAGTTGGATAAGTATGAAAAAGTGAGGCTCAAACCAA 9536
MUS-3.GA.11.11GabPts02    GAC-G-TCTAGGAAGGGTGGCAGA-TCTTAGACCGAGGACAGAGTGGGACTCTGT-AAAA---G--A-GA-T---GAGGACT-CCAGTGGATAGGGCAAG.......................................... 9449
OLC.CI.97.97CI12          ..................................................................................................................................... 8678
RCM.CM.00.SIVagi_00CM312  TAGAGA..............TC--TCAG........ATACTTGTGGCAGAAGTG.C--G-AGGCTAAGAA--------C--GTTGGCGCCCGAACAGGGACTTGA...GAGTGACTGAGGAACCTCCGACGCGGGTCCGGACGCGAGGCGTTGTAACGCGGACCCAG 9384
RCM.CM.02.SIVrcm_02CM8081 TAGAGA..............TC--TCAG........ATATTGGTGGCAGAAGT..AC-G-AGGCTAAGAA--------C--GTTGGCGCCCGAACAGGGACTTGAAGTGAGTGACTGAGGAACCTCTGACGCGGGTCTGGACGCGAGGCGT.GGAACGCAGACCCAG 9463
RCM.GA.x.SIVRCMGAB1       TAGTTT-C.TGTAAGC.AAGTG-GTGCCTGTTTTA....CCTCTCAGCAGTTAA-GAC-CT-GGGTA-G-A..............TCCCTCAG........ATTCTTGTGGCAGAAGAGCCTTG..GGCTAAGAAAA.TTCCCTACCAGT..................... 9465
RCM.NG.x.NG411            T-G-GA..............TC--TCAG........ATATTTGTGGCAGAAGT.AC-.G-AGGCTAAGAA------C-A-C--G.......................................... 9557
SAB.SN.x.SAB1             ....................GACAA-CCTGGTTACTAAGGATCCC.........TGA.........-AGAA------CT------------------------------------------10036
SMM.CI.79.SIVsmCI2        ..................................................................................................................................... 9147
SMM.LR.89.SIVsmLIB1       ..................................................................................................................................... 9224
SMM.SL.92.SIVsmSL92A      ..................................................................................................................................... 9086
SMM.SL.92.SIVsmSL92B      ..................................................................................................................................... 7460
SMM.SL.92.SL92B           ..................................................................................................................................... 9613
SMM.US.04.G078            ..................................................................................................................................... 9157
SMM.US.05.D215            ..................................................................................................................................... 9128
SMM.US.06.FTq             ..................................................................................................................................... 9149
SMM.US.86.CFU212          ..................................................................................................................................... 9130
SMM.US.x.H9               ..................................................................................................................................... 9638
SMM.US.x.PGM53           G-...-CTCGAAT..CAT.AAGAAGA-CCTGGTCTGTTAGGACCCTTTCTGC.............................................................................10184
SMM.US.x.pE660.CG7G       G-...-CTCGACA..CAT.AAGAAGA-CCTGGTCTGTTAGGACCCTTTCTGCTT.TGGGAA-CCG--GCA--------C------------------------------------------10283
STM.US.89.STM_37_16       G-...-CTCAA.............GA-CCTGGTCTGTTAGGACCCTTTCTGC.................................................................... 9892
SUN.GA.98.L14             AA-CCCTGTTCGACTGCTCTCAG-TAGGGAGC-A.............TTACTGG--GCC-AGTGATCC-G--CTG-CGG-------------------------------------10006
SYK.KE.x.KE51             ..................................................................................................................................... 9068
SYK.KE.x.SYK173_COMGNM    ..................................................................................................................................... 9597
TAL.CM.00.266             ..................................................................................................................................... 9338
TAL.CM.01.8023            ..................................................................................................................................... 8845
TAN.UG.x.SIVagmTAN1       ..............GAACCCTCGTTACTGGG-............TTTCTTATC-AGGC-G-----------------C------------------------------------------ 9784
VER.DE.x.AGM3             .......TGAGGGAAGACCCTGATGGAAT--ACCATGGAGAAGTCTTGGTCTGGAAG.........T--ATCCCATGT---------------------------------------- 9325
VER.KE.x.9063             -GTTCCTTACTGGGTTCTCTCTCTGACCCAG.....................................GCG---G--A---C------------------------------------------ 9815
VER.KE.x.AGM155           AGTTCCTTACTGGGTTCTCTCTCAAACCCAG.....................................GCGA--G--A---C------------------------------------------ 9794
VER.KE.x.TYO1_patent      TCTTCCTTACTGGGTTCTCTCTCT.....................GACCCAGGCGAG-------------A---C------------------------------------------ 9778
WRC.CI.97.97CI14          ..................................................................................................................................... 9484
WRC.CI.98.98CI04          ..................................................................................................................................... 9461
WRC.GM.05.Pbt_05GM_X02    ..................................................................................................................................... 8709
```


Draft of 30 September 2015, not for public release
plvdna_main.tex r8876–10546

```
H1B.FR.83.HXB2                ..........................................................................................................  9719
H1O.CM.91.MVP5180             ..........................................................................................................  9792
H1N.CM.95.YBF30               ..........................................................................................................  9182
H1P.CM.06.U14788              ..........................................................................................................  9238
CPZ.CD.06.BF1167              ..........................................................................................................  9927
CPZ.TZ.00.TAN1                ..........................................................................................................  9326
GOR.CM.07.SIVgorCP2135con     ..........................................................................................................  9246


MAC.US.x.239                  .......................................................................................................... 10279
H2A.SN.86.ST_JSP4_27          ..........................................................................................................  9672
H2B.CI.x.EHO                  .......................................................................................................... 10242
H2G.CI.92.Abt96               ..........................................................................................................  9599
H2U.CI.07.07IC_TNP3           ..........................................................................................................  9540
H2U.FR.96.12034               ..........................................................................................................  9607
ASC.UG.10.RT03                ..........................................................................................................  8881
ASC.UG.10.RT08                ..........................................................................................................  8905
ASC.UG.10.RT11                ..........................................................................................................  8874
COL.CM.x.CGU1                 ..........................................................................................................  8728
COL.UG.10.BWC01               ..........................................................................................................  8688
COL.UG.10.BWC07               ..........................................................................................................  8733
DEB.CM.04.SIVdeb04CMPF3061    ..........................................................................................................  9265
DEB.CM.99.CM40                ..........................................................................................................  9227
DEB.CM.99.CM5                 ..........................................................................................................  9158
DEN.CD.x.CD1_CMO580407        ..........................................................................................................  9679
DRL.DE.11.D3                  ..........................................................................................................  9101
DRL.x.x.FAO                   ..........................................................................................................  9731
GRV.ET.x.GRI_677_gri_1        ..........................................................................................................  9623
GSN.CM.99.CN166               ..........................................................................................................  9414
GSN.CM.99.CN71                ..........................................................................................................  9410
LST.CD.88.SIVlhoest447        ..........................................................................................................  8588
LST.CD.88.SIVlhoest485        ..........................................................................................................  8589
LST.CD.88.SIVlhoest524        ..........................................................................................................  8594
LST.KE.x.lho7                 ..........................................................................................................  9957
MAC.US.x.251_1A11             .......................................................................................................... 10274
MND-1.GA.x.MNDGB1            ..........................................................................................................  9215
MND-2.CM.98.CM16              ..........................................................................................................  9518
MND-2.GA.x.M14                ..........................................................................................................  9529
MND-2.x.x.5440                ..........................................................................................................  9647
MNE.US.x.MNE027               ..........................................................................................................  9664
MON.CM.99.L1_99CML1           ..........................................................................................................  9448
MON.NG.x.NG1                  ..........................................................................................................  8221
MUS-1.CM.01.CM1239            ..........................................................................................................  9508
MUS-1.CM.01.SIVmus_01CM1085   ..........................................................................................................  9420
MUS-2.CM.01.CM1246            ..........................................................................................................  9484
MUS-2.CM.01.CM2500            ..........................................................................................................  9430
MUS-3.GA.09.09GabOI81         AAGGAAAGAAAAA..............................................................................................  9549
MUS-3.GA.11.11GabPts02        ..........................................................................................................  9449
OLC.CI.97.97CI12              ..........................................................................................................  8678
RCM.CM.00.SIVagi_00CM312      TCCAGGTA.AGAGTAAG.CCTGGTGGCTTCATAAGTCTAGCAGAAAAAGGTTGCTAGGAAGGAAGCAAGGCAACCCGGTCAGGTACTGGGCCTCTAGGGAAGGGTCGAAGTCCTAGAGAAGGGAGGAAA  9511
RCM.CM.02.SIVrcm_02CM8081     TCCAGGTAAGGAGTAACACCTGGTGGCTTCACAAGTCTAGCAGAAAAAGGTTGCTAGGAAGGAAGCAGGGCAACCCGGTCAGGTACTGGGCCTCTAGGGAAGAAGCGAAAGCCTAGAGAAGGGAGGAAA  9592
RCM.GA.x.SIVRCMGAB1           ..........................................................................................................  9465
RCM.NG.x.NG411                ..........................................................................................................  9557
SAB.SN.x.SAB1                 .......................................................................................................... 10036
SMM.CI.79.SIVsmCI2            ..........................................................................................................  9147
SMM.LR.89.SIVsmLIB1           ..........................................................................................................  9224
SMM.SL.92.SIVsmSL92A          ..........................................................................................................  9086
SMM.SL.92.SIVsmSL92B          ..........................................................................................................  7460
SMM.SL.92.SL92B               ..........................................................................................................  9613
SMM.US.04.G078                ..........................................................................................................  9157
SMM.US.05.D215                ..........................................................................................................  9128
SMM.US.06.FTq                 ..........................................................................................................  9149
SMM.US.86.CFU212              ..........................................................................................................  9130
SMM.US.x.H9                   ..........................................................................................................  9638
SMM.US.x.PGM53                .......................................................................................................... 10184
SMM.US.x.pE660.CG7G           .......................................................................................................... 10283
STM.US.89.STM_37_16           ..........................................................................................................  9892
SUN.GA.98.L14                 .......................................................................................................... 10006
SYK.KE.x.KE51                 ..........................................................................................................  9068
SYK.KE.x.SYK173_COMGNM        ..........................................................................................................  9597
TAL.CM.00.266                 ..........................................................................................................  9338
TAL.CM.01.8023                ..........................................................................................................  8845
TAN.UG.x.SIVagmTAN1           ..........................................................................................................  9784
VER.DE.x.AGM3                 ..........................................................................................................  9325
VER.KE.x.9063                 ..........................................................................................................  9815
VER.KE.x.AGM155               ..........................................................................................................  9794
VER.KE.x.TYO1_patent          ..........................................................................................................  9778
WRC.CI.97.97CI14              ..........................................................................................................  9484
WRC.CI.98.98CI04              ..........................................................................................................  9461
WRC.GM.05.Pbt_05GM_X02        ..........................................................................................................  8709
```

V

HIV-1/SIVcpz Proteins

Contents

V-1	Introduction		**307**
V-2	Annotated features		**308**
V-3	Sequences		**310**
V-4	Alignments		**318**
	V-4.1	Gag	318
	V-4.2	Pol	326
	V-4.3	Vif	340
	V-4.4	Vpr	344
	V-4.5	Tat	346
	V-4.6	Rev	348
	V-4.7	Vpu	350
	V-4.8	Env	352
	V-4.9	Nef	366

V-1 Introduction

The HIV-1/SIVcpz protein alignments are based on the complete genome nucleotide alignment, but in some cases a few sequences were removed because they were too short to be informative (especially in Nef), had many stop codons or a problematic segment of amino acids. As with the other alignments in this compendium, they are intended to display the genetic variation of the world-wide HIV epidemic in a compact form. They are annotated in more detail than the complete genome nucleotide alignment.

V-2 Annotated features

Features of HIV-1 annotated in the alignment that follows.

Feature	Protein	Location	Page
Gag start, p17 start	Gag	1	318
membrane binding	Gag	1-30	318
phosphorylation site	Gag	112	318
p17 end	Gag	132	318
p24 start	Gag	133	318
CyPA binding	Gag	205-241	320
major homology region	Gag	285-304	320
p24 end	Gag	364	322
p2 start	Gag	365	322
p2 end	Gag	377	322
p7 start	Gag	378	322
Zn motif	Gag	392-404	322
Zn motif	Gag	413-425	322
p7 end	Gag	432	322
p1 start	Gag	433	322
p1 end	Gag	448	324
p6 start	Gag	449	324
Vpr binding	Gag	455-460	324
Vpr binding	Gag	489-494	324
p6 end, Gag end	Gag	501	324
Pol start	Pol	1	326
Gag-Pol TF start	Pol	1	326
Gag-Pol TF end	Pol	56	326
protease start	Pol	57	326
protease end	Pol	155	328
p66, p51 RT start	Pol	156	328
M41L	Pol	196	328
D67N	Pol	222	328
K70R	Pol	225	328
D110 catalytic site	Pol	265	328
polymerase motif	Pol	337-342	330
T215Y	Pol	370	330
K219Q	Pol	374	330
p51 RT end	Pol	595	332
p15 RNase H start	Pol	596	332
p66 RT, p15 Rnase H end	Pol	715	334
p31 Integrase start	Pol	716	334
p31 Integrase end	Pol	1004	338
Pol end	Pol	1004	338
Vif start	Vif	1	340
Vif end	Vif	193	342
Vpr start	Vpr	1	344
oligomerization	Vpr	1-41	344
amphipathic α-helix	Vpr	17-33	344
H(S/N)RIG motifs	Vpr	71-83	344
frameshift in HXB2	Vpr	72	344
Vpr end in HXB2	Vpr	79	344
Vpr end	Vpr	98	344

Feature	Protein	Location	Page
Tat start	Tat	1	346
C-rich region	Tat	22-37	346
nuclear localization	Tat	49-57	346
exon 1 end	Tat	72	346
exon 2 start	Tat	73	346
Tat end	Tat	102	346
Rev start	Rev	1	348
exon 1 end	Rev	25	348
exon 2 start	Rev	26	348
NLS	Rev	34-49	348
Leu-rich effector domain	Rev	75-83	348
Rev end	Rev	117	348
Vpu start	Vpu	1	350
transmembrane domain	Vpu	1-27	350
cytoplasmic domain	Vpu	28-82	350
α-helix	Vpu	43-51	350
phos	Vpu	53	350
phos	Vpu	57	350
α-helix	Vpu	58-70	350
Vpu end	Vpu	83	350
Env start	Env	1	352
signal peptide end	Env	30	352
gp120 start	Env	31	352
glycosylation NVT	Env	88-90	352
CD4 binding	Env	124	352
V1	Env	131-156	352
glycosylation NDT	Env	136-138	352
glycosylation NSS	Env	141-143	354
glycosylation NCS	Env	156-158	354
V2	Env	158-196	354
glycosylation NIS	Env	160-162	354
glycosylation NDT	Env	186-188	354
glycosylation NTS	Env	197-199	354
CD4 binding	Env	196	354
glycosylation NKT	Env	230-232	354
glycosylation NGT	Env	234-236	354
glycosylation NVS	Env	241-243	354
glycosylation NGS	Env	262-264	356
glycosylation NFT	Env	276-278	356
CD4 binding	Env	279	356
glycosylation NTS	Env	289-291	356
glycosylation NCT	Env	295-297	356
V3	Env	296-331	356
glycosylation NNT	Env	301-303	356
V3 tip	Env	312-315	356
glycosylation NNT	Env	339-341	356
glycosylation NKT	Env	356-358	356
CD4 binding	Env	365	356
V4	Env	385-418	356
glycosylation NST	Env	386-388	356
glycosylation NST	Env	392-394	358
glycosylation NST	Env	397-399	358
glycosylation NNT	Env	406-408	358

Feature	Protein	Location	Page
CD4 binding	Env	425	358
glycosylation NIT	Env	448-450	358
CD4 binding	Env	455	358
V5	Env	460-471	358
glycosylation NES	Env	463-465	358
CD4 binding	Env	469	358
fusion peptide	Env	512-527	360
gp120 end	Env	511	360
gp41 start	Env	512	360
immunodominant region	Env	588-607	360
glycosylation NAS	Env	611-613	360
glycosylation NKS	Env	616-618	360
glycosylation NHT	Env	624-626	360
glycosylation NYT	Env	637-639	360
transmembrane domain	Env	685-704	362
gp41 cytoplasmic tail start	Env	705	362
glycosylation NGS	Env	750-752	362
glycosylation NAT	Env	816-818	364
cytoplasmic tail end	Env	857	364
gp41 end	Env	857	364
Env end	Env	857	364
Nef start	Nef	1	366
myristoylation	Nef	2-7	366
acidic cluster	Nef	62-65	366
poly-P helix	Nef	69-78	366
phosphorylation	Nef	77-81	366
HXB2 premature Nef end	Nef	124	366
normal Nef end	Nef	207	368

V-3 Sequences

Sequences included in the HIV-1 protein alignments.

Name	Accession	Proteins	Author	Reference
B.FR.83.HXB2	K03455	All	Wong-Staal, F.	*Nature* **313**(6000):277-284 (1985)
A1.AU.03.PS1044_Day0	DQ676872	All	Li, B.	*J Virol* **81**(1):193-201 (2007)
A1.CH.03.HIV_CH_BID_V3538	JQ403028	All	Henn, M.R.	*PLoS Pathog* **8**(3):E1002529 (2012)
A1.CY.08.CY236	JF683783	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
A1.ES.05.X1608_8	FJ670519	All	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
A1.KE.11.DEMA11KE001	KF716475	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A1.RU.11.11RU6950	JX500694	All	Baryshev, P.B.	*ARHR* **30**(6); 592-7 (2014)

Name	Accession	Proteins	Author	Reference
A1.RW.11.DEMA111RW002	KF716472	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A1.SN.01.DDI579	AY521629	All	Meloni, S.T.	*J Virol* **78**(22):12438-12445 (2004)
A1.UG.11.DEMA110UG001	KF859745	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A1.ZA.04.04ZASK162B1	DQ396400	All	Rousseau, C.M.	*J Virol Methods* **136**(1-2):118-125 (2006)
A2.CM.01.01CM_1445MV	GU201516	All	Carr, J.K.	*Retrovirology* **2010** Apr 28;7:39 doi: 101186/1742-4690-7-39
A2.CY.94.94CY017_41	AF286237	All	Gao, F.	*ARHR* **17**(8):675-688 (2001)
B.BR.10.10BR_MG029	KJ849799	All	Pessoa, R.	*Transfusion* **55**(5); 980-90 (2015)
B.CA.07.502_1191_03	JF320424	All	Rolland, M.	*Nat Med* **17**(3); 366-71 (2011)
B.CH.08.M2_0803101_NFLG8	KC797225	All	Castro, E.	*AIDS* **28**(12); 1840-4 (2014)
B.CN.10.DEMB10CN002	JX140658	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.ES.10.DEMB10ES002	KC473842	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.FR.11.DEMB11FR001	KF716496	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.GB.05.MM45d213_GN1	HM586212	All	Turnbull, E.L.	*J Immunol* **182**(11):7131-7145 (2009)
B.HK.06.HK003	FJ460500	All	Tsui, S.K.W.	*ARHR* **26**(1); 117-22 (2010)
B.HT.05.05HT_129389	EU839602	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Nadai, Y.	*PLoS ONE* **4**(3):E4814 (2009)
B.JP.12.DEMB12JP001	KF716498	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.KR.07.HP_18_07JHS10_3909	KJ140263	All	Kim, B.-R.	*Haemophilia* **21**(1); e1-11 (2015)
B.PE.07.502_0525_wg5	JF320191	All	Rolland, M.	*Nat Med* **17**(3); 366-71 (2011)
B.RU.11.11RU21n	JX500708	All	Baryshev	Unpublished
B.TH.08.MERLBDTRC10	JN860769	All	Rutvisuttinunt, W.	*ARHR* **28**(12); 1703-11 (2012)
B.US.11.ES38	JN397362	All	Buckheit, R.W.3.	*Nat Commun* **3**, 716 (2012)
C.AR.01.ARG4006	AY563170	All	Carrion, G.	*ARHR* **20**(9):1022-1025 (2004)
C.BR.07.DEMC07BR003	JX140663	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
C.BW.00.00BW5031_1	AF443115	All	Novitsky, V.	*J Virol* **76**(11):5435-5451 (2002)
C.CN.10.YNFL19	KC870038	All	Wei	Unpublished

Name	Accession	Proteins	Author	Reference
C.CY.09.CY260	JF683803	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
C.ES.08.X2363_2	EU786681	All	Fernandez-Garcia, A.	*ARHR* **25**(1); 93-102 (2009)
C.ET.02.02ET_288	AY713417	All	Brown, B.K.	*J Virol* **79**(10):6089-6101 (2005)
C.IN.09.T125_2139	KC156210	All	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.KE.00.KER2010	AF457054	All	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
C.MW.09.703010256_CH256.w96	KC156214	All	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.TZ.08.707010457_CH457.w8	KC156220	All	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.YE.02.02YE511	AY795906	All	Saad, M.D.	*ARHR* **21**(7):644-648 (2005)
C.ZA.08.705010534_CH534.w12	KC156221	All	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.ZA.10.DEMC10ZA001	JX140669	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
C.ZM.11.DEMC11ZM006	KF716467	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.CM.10.DEMD10CM009	JX140670	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.CY.06.CY163	FJ388945	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **25**(8); 727-40 (2009)
D.KE.11.DEMD11KE003	KF716476	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.KR.04.04KBH8	DQ054367	All	Cho, Y.-K.	*ARHR* **29**(4); 738-43 (2013)
D.SN.90.SE365	AB485648	All	Takekawa	Unpublished
D.TZ.01.A280	AY253311	All	Arroyo, M.A.	*ARHR* **20**(8):895-901 (2004)
D.UG.10.DEMD10UG004	KF716479	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.UG.11.DEMD11UG003	KF716480	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.YE.02.02YE516	AY795907	All	Saad, M.D.	*ARHR* **21**(7):644-648 (2005)
D.ZA.90.R1	EF633445	All	Jacobs, G.B.	*ARHR* **23**(12):1575-8 (2007)

Name	Accession	Proteins	Author	Reference
F1.AO.06.AO_06_ANG32	FJ900266	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Guimaraes, M.L.	*Retrovirology* **6**, 39 (2009)
F1.AR.02.ARE933	DQ189088	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Aulicino, P.C.	*ARHR* **21**(2):158-164 (2005)
F1.BR.10.10BR_RJ015	KJ849791	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Pessoa, R.	*Transfusion* **55**(5); 980-90 (2015)
F1.CY.08.CY222	JF683771	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
F1.ES.02.ES_X845_4	FJ670516	All	Fernandez-Garcia, A.	*ARHR* **25**(11):1187-1191 (2009)
F1.RO.96.BCI_R07	AB485658	All	Takekawa	Unpublished
F1.RU.08.D88_845	GQ290462	All	Fernandez-Garcia, A.	*ARHR* **25**(11):1187-1191 (2009)
F2.CM.02.02CM_0016BBY	AY371158	All	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
F2.CM.10.DEMF210CM001	JX140672	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
F2.CM.10.DEMF210CM007	JX140673	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.BE.96.DRCBL	AF084936	All	Oelrichs, R.B.	*ARHR* **15**(6):585-589 (1999)
G.CM.10.DEMG10CM008	JX140676	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.CN.08.GX_2084_08	JN106043	All	Liu, W.	*Zhonghua Liu Xing Bing Xue Za Zhi* **34**(1); 53-6 (2013)
G.CU.99.Cu74	AY586547	All	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
G.ES.09.X2634_2	GU362882	All	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
G.GH.03.03GH175G	AB287004	All	Takekawa	Unpublished
G.KE.09.DEMG09KE001	KF716477	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.NG.09.09NG_SC62	JN248593	All	Charurat, M.	*J Infect Dis* **205**(8); 1239-47 (2012)
G.ZA.01.TV546	KJ948662	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Wilkinson, E.	*ARHR* **31**(4); 412-20 (2015)
H.BE.93.VI991	AF190127	All	Janssens, W.	*AIDS* **14**(11):1533-1543 (2000)
H.CF.90.056	AF005496	All	Gao, F.	*J Virol* **72**(7):5680-5698 (1998)
H.GB.00.00GBAC4001	FJ711703	All	Holzmayer, V.	*ARHR* **25**(7):721-726 (2009)
J.CM.04.04CMU11421	GU237072	All	Yamaguchi, J.	*ARHR* **26**(6); 693-7 (2010)

Name	Accession	Proteins	Author	Reference
J.SE.93.SE9280_7887	AF082394	All	Laukkanen, T.	*ARHR* **15**(3):293-297 (1999)
J.SE.94.SE9173_7022	AF082395	All	Laukkanen, T.	*ARHR* **15**(3):293-297 (1999)
K.CD.97.97ZR_EQTB11	AJ249235	All	Triques, K.	*ARHR* **16**(2):139-151 (2000)
K.CM.96.96CM_MP535	AJ249239	All	Triques, K.	*ARHR* **16**(2):139-151 (2000)
U.CA.01.TV749	HM215251	All	Quesnel-Vallieres, M.	*Emerging Infect Dis* **17**(2):271-274 (2011)
U.CA.99.TV721	HM215249	All	Quesnel-Vallieres, M.	*Emerging Infect Dis* **17**(2):271-274 (2011)
U.CD.83.83CD003_Z3	AF286236	All	Gao, F.	*ARHR* **17**(12):1217-1222 (2001)
U.CD.90.90CD121E12	AF457101	All	Mokili, J.L.	*ARHR* **18**(11):817-823 (2002)
U.CY.05.CY090	FJ388921	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **25**(8); 727-40 (2009)
U.CY.08.CY223	JF683772	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
U.ES.10.DEURF10DZ001	JX140679	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
U.GR.99.99GR303	AY046058	All	Paraskevis, D.	*J Gen Virol* **82**(Pt 10):2509-2514 (2001)
U.NL.95.U_NL_95_H10986_D1	EF029066	All	van der Hoek, L.	*ARHR* **23**(3):466-470 (2007)
01_AE.AF.07.569M	GQ477441	All	Sanders-Buell, E.	*ARHR* **26**(5):605-608 (2010)
01_AE.CF.90.90CF11697	AF197340	All	Anderson, J.P.	*J Virol* **74**(22):10752-10765 (2000)
01_AE.CN.10.YNFL03	KC870029	All	Wei	Unpublished
01_AE.HK.04.HK001	DQ234790	All	Tsui	Unpublished
01_AE.IR.10.10IR.THR48F	AB703616	All	Jahanbakhsh, F.	*ARHR* **29**(1); 198-203 (2013)
01_AE.JP.x.JRC77AE	AB565504	All	Umeki-Sakamoto	Unpublished
01_AE.TH.09.AA090a_WG11	JX447727	All	Rolland, M.	*Nature* **490**(7420); 417-20 (2012)
01_AE.TH.90.CM240	U54771	All	Carr, J.K.	*J Virol* **70**(9):5935-5943 (1996)
01_AE.US.05.306163_FL	JX863920	All	Heipertz, R.A. Jr.	*ARHR* **29**(10):1310-1320 (2013)
01_AE.VN.98.98VNND15	FJ185235	All	Liao, H.	*Virology* **391**(1):51-56 (2009)
02_AG.CM.10.DE00210CM013	KF859739	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
02_AG.CY.09.CY256	JF683799	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
02_AG.ES.06.P1423	EU884501	All	Fernandez-Garcia, A.	*ARHR* **25**(1); 93-102 (2009)
02_AG.GW.05.CC_0048	FJ694792	All	Vinner, L.	*APMIS* **119**(8); 487-97 (2011)
02_AG.KR.12.12MHR9	KF561435	All	Cho	Unpublished
02_AG.LR.x.POC44951	AB485636	All	Takekawa	Unpublished
02_AG.NG.09.09NG_SC61	JN248592	All	Charurat, M.	*J Infect Dis* **205**(8); 1239-47 (2012)
02_AG.NG.x.IBNG	L39106	All	Howard, T.M.	*ARHR* **10**(12):1755-1757 (1994)
02_AG.SE.94.SE7812	AF107770	All	Carr, J.K.	*AIDS* **13**(14); 1819-26 (1999)
02_AG.SN.98.98SE_MP1211	AJ251056	All	Toure-Kane, C.	*ARHR* **16**(6):603-609 (2000)
03_AB.RU.97.KAL153_2	AF193276	All	Liitsola, K.	*ARHR* **16**(11):1047-1053 (2000)

Name	Accession	Proteins	Author	Reference
04_cpx.CY.94.94CY032_3	AF049337	All	Gao, F.	*J Virol* **72**(12):10234-10241 (1998)
05_DF.BE.x.VI1310	AF193253	All	Laukkanen, T.	*Virology* **269**(1):95-104 (2000)
06_cpx.AU.96.BFP90	AF064699	All	Oelrichs, R.B.	*ARHR* **14**(16):1495-1500 (1998)
07_BC.CN.98.98CN009	AF286230	All	Rodenburg, C.M.	*ARHR* **17**(2):161-168 (2001)
08_BC.CN.97.97CNGX_6F	AY008715	All	Piyasirisilp, S.	*J Virol* **74**(23):11286-11295 (2000)
09_cpx.GH.96.96GH2911	AY093605	All	McCutchan, F.E.	*ARHR* **20**(8):819-826 (2004)
10_CD.TZ.96.96TZ_BF061	AF289548	All	Koulinska, I.N.	*ARHR* **17**(5):423-431 (2001)
11_cpx.CM.95.95CM_1816	AF492624	All	Wilbe, K.	*ARHR* **18**(12):849-56 (2002)
12_BF.AR.99.ARMA159	AF385936	All	Carr, J.K.	*AIDS* **15**(15):F41-F47 (2001)
13_cpx.CM.96.96CM_1849	AF460972	All	Wilbe, K.	*ARHR* **18**(12):849-56 (2002)
14_BG.ES.05.X1870	FJ670522	All	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
15_01B.TH.99.99TH_MU2079	AF516184	All	Viputtijul, K.	*ARHR* **18**(16):1235-1237 (2002)
16_A2D.KR.97.97KR004	AF286239	All	Gao, F.	*ARHR* **17**(8):675-688 (2001)
17_BF.AR.99.ARMA038	AY037281	All	Carr, J.K.	*AIDS* **15**(15):F41-F47 (2001)
18_cpx.CU.99.CU76	AY586540	All	Thomson, M.M.	*AIDS* **19**(11):1155-63 (2005)
19_cpx.CU.99.CU7	AY894994	All	Casado, G.	*JAIDS* **40**(5):532-537 (2005)
20_BG.CU.99.Cu103	AY586545	All	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
21_A2D.KE.99.KER2003	AF457051	All	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
22_01A1.CM.01.01CM_0001BBY	AY371159	All	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
23_BG.CU.03.CB118	AY900571	All	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
24_BG.ES.08.X2456_2	FJ670526	All	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
25_cpx.CM.02.1918LE	AY371169	All	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
26_AU.CD.02.02CD_MBTB047	FM877782	All	Vidal, N.	*ARHR* **25**(8):823-832 (2009)
27_cpx.FR.04.04CD_FR_KZS	AM851091	All	Vidal, N.	*ARHR* **24**(2):315-321 (2008)
28_BF.BR.99.BREPM12609	DQ085873	All	Sa Filho, D.J.	*ARHR* **22**(1):1-13 (2006)
29_BF.BR.01.BREPM16704	DQ085876	All	Sa Filho, D.J.	*ARHR* **22**(1):1-13 (2006)
31_BC.BR.04.04BR142	AY727527	All	Sanabani, S.	*ARHR* **22**(2):171-176 (2006)
32_06A1.EE.01.EE0369	AY535660	All	Adojaan, M.	*JAIDS* **39**(5):598-605 (2005)
33_01B.ID.07.JKT189_C	AB547463	All	SahBandar, I.N.	*ARHR* **27**(1); 97-102 (2011)
34_01B.TH.99.OUR2478P	EF165541	All	Tovanabutra, S.	*ARHR* **23**(6):829-833 (2007)
35_AD.AF.07.169H	GQ477446	All	Sanders-Buell, E.	*ARHR* **26**(5):605-608 (2010)
36_cpx.CM.00.00CMNYU830	EF087994	All	Powell, R.L.	*ARHR* **23**(8):1008-1019 (2007)
37_cpx.CM.00.00CMNYU926	EF116594	All	Powell, R.L.	*ARHR* **23**(7):923-933 (2007)
38_BF1.UY.03.UY03_3389	FJ213783	All	Ruchansky, D.	*ARHR* **25**(3); 351-6 (2009)
39_BF.BR.04.04BRRJ179	EU735535	All	Guimaraes, M.L.	*AIDS* **22**(3):433-435 (2008)
40_BF.BR.05.05BRRJ055	EU735537	All	Guimaraes, M.L.	*AIDS* **22**(3):433-435 (2008)
42_BF.LU.06.luBF_18_06	EU170139	All	Struck, D.	*ARHR* **31**(5); 554-8 (2015)
43_02G.SA.03.J11223	EU697904	All	Badreddine, S.	*ARHR* **23**(5):667-674 (2007)
44_BF.CL.00.CH80	FJ358521	All	Delgado, E.	*ARHR* **26**(7); 821-6 (2010)
45_cpx.FR.04.04FR_AUK	EU448295	All	Frange, P.	*Retrovirology* **2008** Aug 1;5:69 doi: 101186/1742-4690-5-69
46_BF.BR.07.07BR_FPS625	HM026456	All	Sanabani, S.S.	*Virol J* **2010** Apr 16;7:74 doi: 101186/1743-422X-7-74
47_BF.ES.08.P1942	GQ372987	All	Fernandez-Garcia, A.	*ARHR* **26**(7); 827-32 (2010)
48_01B.MY.07.07MYKT021	GQ175883	All	Li, Y.	*JAIDS* **54**(2):129-136 (2010)
49_cpx.GM.03.N26677	HQ385479	All	de Silva, T.I.	*Retrovirology* **7**(1):82 (2010)
50_A1D.GB.10.12792	JN417240	All	Foster, G.M.	*PLoS One* **9**(1); e83337 (2014)
51_01B.SG.11.11SG_HM021	JN029801	All	Ng, O.T.	*ARHR* **28**(5); 527-30 (2012)
52_01B.MY.03.03MYKL018_1	DQ366664	All	Tee, K.K.	*JAIDS* **43**(5):523-529 (2006)
53_01B.MY.11.11FIR164	JX390610	All	Chow, W.Z.	*J Virol* **86**(20):11398-11399 (2012)
54_01B.MY.09.09MYSB023	JX390976	All	Ng, K.T.	*J Virol* **86**(20):11405-11406 (2012)

Name	Accession	Proteins	Author	Reference
55_01B.CN.10.HNCS102056	JX574661	All	Han, X.	*Genome Announc* **1**(1):E00050-12 (2013)
56_cpx.FR.10.URF5_patient_A	JN882655	All	Leoz, M.	*AIDS* **25**(11):1371-1377 (2011)
57_BC.CN.09.09YNLX19sg	KC899008	All	Han, X.	*PLoS ONE* **8**(5):E65337 (2013)
58_01B.MY.09.09MYPR37	KC522031	All	Chow, W.Z.	*PLoS ONE* **9**(1):E85250 (2014)
59_01B.CN.09.09LNA423	JX960635	All	An, M.	*J Virol* **86**(22); 12402-6 (2012)
60_BC.IT.11.BAV499	KC899079	All	Simonetti, F.R.	*Infect Genet Evol* **2014** Apr;23:176-81 doi: 101016/jmeegid201402007 Epub 2014 Mar 3
61_BC.CN.10.JL100010	KC990124	All	Li, X.	*Genome Announc* **2013** Jun 27;1(3) pii: e00326-13 doi: 101128/genomeA00326-13
62_BC.CN.10.YNFL13	KC870034	All	Wei, H.	*ARHR* **30**(4):380-383 (2014)
63_02A1.RU.10.10RU6637	JN230353	All	Baryshev, P.B.	*Arch Virol* **157**(12); 2335-41 (2012)
64_BC.CN.09.YNFL31	KC870042	All	Hsi, J.	*ARHR* **30**(4); 389-93 (2014)
65_cpx.CN.10.YNFL01	KC870027	All	Feng, Y.	*ARHR* **30**(6); 598-602 (2014)
67_01B.CN.11.2011.ANHUI.HF115	KC183779	All	Wu, J.	*PLoS ONE* **8**(1):E54322 (2013)
68_01B.CN.11.2011.ANHUI.WH73	KC183782	All	Wu, J.	*PLoS ONE* **8**(1):E54322 (2013)
70_BF1.BR.10.10BR_PE004	KJ849758	All	Pessoa, R.	*PLoS One* **9**(11); e112674 (2014)
71_BF1.BR.10.10BR_PE008	KJ849759	All	Pessoa, R.	*PLoS One* **9**(11); e112674 (2014)
72_BF1.BR.10.10BR_MG002	KJ671534	All	Pessoa, R.	*Genome Announc* **2**(3):e00386-14 (2014)
O.BE.87.ANT70	L20587	All	Vanden Haesevelde, M.	*J Virol* **68**(3):1586-1596 (1994)
O.CM.98.98CMA104	AY169802	All	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.98.98CMABB141	AY169807	All	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.98.98CMABB212	AY169804	All	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.98.98CMU5337	AY169808	All	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.99.99CMU4122	AY169815	All	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.FR.92.VAU	AF407418	All	Vartanian, J.P.	*J Gen Virol* **83**(Pt 4):801-805 (2002)
O.GA.11.11Gab6352	JX245015	All	Liegeois, F.	*ARHR* **29**(7); 1085-90 (2013)
O.SN.99.99SE_MP1299	AJ302646	All	Toure-Kane, C.	*ARHR* **17**(12):1211-1216 (2001)
O.US.10.LTNP	JN571034	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Buckheit, R.W.3.	*ARHR* **30**(6); 511-513 (2014)
N.CM.02.DJO0131	AY532635	All	Bodelle, P.	*ARHR* **20**(8):902-908 (2004)
N.CM.02.SJGddd	GQ324959	All	Vallari, A.	*ARHR* **26**(1):109-115 (2010)
N.CM.04.04CM_1015_04	DQ017382	All	Yamaguchi, J.	*ARHR* **22**(1):83-92 (2006)
N.CM.06.U14296	GQ324962	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Vallari, A.	*ARHR* **26**(1):109-115 (2010)
N.CM.06.U14842	GQ324958	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env	Vallari, A.	*ARHR* **26**(1):109-115 (2010)
N.FR.11.N1_FR_2011	JN572926	All	Delaugerre, C.	*Lancet* **378**(9806); 1894 (2011)
P.CM.06.U14788	HQ179987	All	Vallari, A.	*J Virol* **85**(3); 1403-7 (2011)
P.FR.09.RBF168	GU111555	All	Plantier, J.-C.	*Nat Med* **15**(8); 871-2 (2009)
CPZ.CD.06.BF1167	JQ866001	All	Li, Y.	*J Virol* **86**(19):10776-10791 (2012)

Name	Accession	Proteins	Author	Reference
CPZ.CM.05.SIVcpzMT145	DQ373066	All	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.GA.88.GAB1	X52154	All	Huet, T.	*Nature* **345**(6273):356-359 (1990)
CPZ.TZ.06.SIVcpzTAN13	JQ768416	All	Takehisa	Unpublished
CPZ.US.85.US_Marilyn	AF103818	All	Gao, F.	*Nature* **397**(6718):436-441 (1999)
GOR.CM.07.SIVgor2139_287	FJ424866	All	Takehisa, J.	*J Virol* **83**(4):1635-1648 (2009)
GOR.CM.07.SIVgorCP2135con	FJ424863	All	Takehisa, J.	*J Virol* **83**(4):1635-1648 (2009)

```
                              membrane binding                                                                        phosphorylation site
                             Gag start, p17 start                                                                                                          p17 end p24 start
B.FR.83.HXB2                 MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKI...EEEQNKS..........KKKAQQ.....AAAD..TG..............HSN...QVSQNYP 133
A1.AU.03.PS1044_Day0         ------I----R--A--------------R---L-----------L--X---SA--Q--ME----A-K--X--IK--F-------------DV-----------.....-IK---...........-QRT--.....----.--...............N-G...K------ 133
A1.CH.03.HIV_CH_BID_V3538    ------I----K--A-------------A----L-----------I-------A--Q--E----A-K-----YK---------------GV-----------.....-I---...........-QRT--......-TTN.--...............S-S...S------ 133
A1.CY.08.CY236               -----------K--A--R-----------RM--L-----------L-------A--Q--ME----A-N-----K--F-------------DV-----------.....-IK---...........-Q-T--.....--------...NSSTG......N-S....---R--- 138
A1.ES.05.X1608_8             -----------K--A--------------RM--L-----------L--D---A--Q--IE-I--A-K--T-----F-------------NV-----------.....-L---...........-Q........A----.--...............N-S...K------ 130
A1.KE.11.DEMA11KE001         -----------K--A--R---------A-RM--L-----------L--S---A--Q--IE----A-N--T------------------DV-----------.....-I---...........---T--.....T---.--...............NRS...N------ 133
A1.RU.11.11RU6950            -----------K--A--------------RI--L-----------L--S-------Q---E----T-K-----FK------------------------------.....-I---...........-Q-T--......--TG.--..................------ 129
A1.RW.11.DEMA111RW002        -----------K--A------------Q-R---L-----------L------SA--Q---E----A---T--VK-----I----------AVR-----------.....-I-K--...........-Q-TT-A........----.--..........N-S...N------ 134
A1.SN.01.DDI579              .#---------K-NK------------Q-RI--L-----------L--S---T--Q--MQ----A-K--T------------------DV-----------.....-I--N...........-Q-T--.......-T-.--..............N-R........------- 131
A1.UG.11.DEMA110UG001        -----------K--A--------------R---L-----------L-S----A--Q--MA---SA-G--T-G-K--F----------Q--V-----------.....-I-K--...........-Q-T--......-E-G.--...............T-.........---Y-- 131
A1.ZA.04.04ZASK162B1         -----------K--S---T----------RI--L-----------L-------A--Q--IE----A-K--T--IK------------------DV-----------.....-I---...........-Q-T--AAAAQ---N.--...............N-S........------ 138
A2.CM.01.01CM_1445MV         ..---------K--A----------R---RM--L-------KYSI-----------K--IR--HSA-PV-T---K-----I-V--Y---K--V---------L........Y...........-Q-T--.........---A.--...............N-.....S----- 129
A2.CY.94.94CY017_41          ------I----K--A--------------R---L-------K-SI-------P------IR----A----T---K------VV--W----VDV-----------.........-...........-Q-T-H........----.--...............N-.....S----- 129
B.BR.10.10BR_MG029           -----------Q--------------Q-------------------------------A------E----A----T--IK--------------K-D-R------E-V......-...........------------..-.-.S-..............N-S...K------ 132
B.CA.07.502_I191_03          -----------------------------------------------------------A-----I------------------F----V--------------------E--......S-...........---V--.......---TAD--.............N-S...T------ 135
B.CH.08.M2_0803101_NFLG8     -----------------------------R-----------------------G--K---I------K------K-------------K-DV--------------.......C...........--RE--.........-----.IK..............DNS....-I----- 133
B.CN.10.DEMB10CN002          -----------Q--K-----------Q---------------------SS-------K--Q------------K---------------DVR--------------------...........-----------.E---..A-..............N-S.KAS------ 135
B.ES.10.DEMB10ES002          -----------K--K--------------------------------------AG----IE----A----T----------------RG--V-------E-V........G...........--E--P.........---A.--...............N-S.....---H-F- 132
B.FR.11.DEMB11FR001          -----------------------------------------------------G-----E-------V----------I-V------K--V-------E--.........-...........------------.....----.--..............N-G....-A----- 133
B.GB.05.MM45d213_GN1         -----------------------------R-----------------------------------A---------------------DV-------E--.........G...........---T--.......-T-N..--NQAAXASG......S-S...------- 141
B.HK.06.HK003                -----------K--K--------------Q----------------------------------SA-------K------V------G-D---------------------..........R--V--..........----.--...............N-S...------- 133
B.HT.05.05HT_129389          ..------------K-------------------------------------------A----------A---------------------K--V--------R-.........C...........------......----.--...............N-S...------- 131
B.JP.12.DEMB12JP001          -------IN---------------S--------------------L----------K--M-----A----------F-V--V---------------------.......S-...........------........----.--...............NNS...------- 133
B.KR.07.HP_18_07JHS10_3909   -----------K--Q--------------R----------------------------------------K--F-----------------------------------...........-------........-T-A.--..............N-S...------- 133
B.PE.07.502_0525_wg5         ------I----K--S--------------R---L------D------------A-------Q------------IK-------------K--VR--------V........E-...........-------........--TN.--..............K-S...------- 133
B.RU.11.11RU21n              ------I-------------------R----------------------------G--K---E--------------K--F----V---------DV-----------------...........--R---.........------.--NSSPG.......N-S...P------ 138
B.TH.08.MERLBDTRC10          -----------K--Q---------------------------------------A-------------------K--F-------------------------R-..........-...........-Q---P.........----.--...............NNS...------- 133
B.US.11.ES38                 -----------D-----#-----------Q-----------------------K------E----A---------K--F-----------G-DV--------------------...........-----P.......-T--.--...............N-S...------- 132
C.AR.01.ARG4006              ..----I-R--K--T--R-K-K-----T-MM--L-----------D-----T-----IR----A-----K--I------------AK---R-----------...K---E--...........QQ-T--.......-E-A..................DKG..K------ 129
C.BR.07.DEMC07BR003          ------I-R-EK--T--R---------T-MM--L-----------LD------R--K--IQ----A----TK--V--H----V-----EK--VR-----------...K-----.............R--.......-E-A..................DKR..K------ 128
C.BW.00.00BW5031_1           ------I-R--K-----------------R---L-----------L--S---AD--K--IK----A-K--T-------------------RN--VQ---------V........NC...........QQ-T--.......-E-A..................GG...KI----- 130
C.CN.10.YNFL19               ------I-R-RK--A------------R-M---------------L----------K--MK----A----T-----F-----------EG--VR---------V........I...........QQ-T-E.......-KEA..................DE...K--R--- 130
C.CY.09.CY260                ------I-R-EK--K------------T-M---L-----------L----------K--MS------K--T-----F-----------AG--VR-----------------...........QQ-T--KTQQ.VE-A..G-...................K------ 134
C.ES.08.X2363_2              ------I-R--K--A------------H-M---L-----------L----------K--IK----A-K--T--K-------------EK--VR-------------.........A...........QQ---------KTQQ.................DDK...KI----- 131
C.ET.02.02ET_288             ..------R-EK--K------------H-M---L-----------L-------T--K--IK----A-K--T-----F-----------EG-D-R-------R-.........V...........QQ-T--.......-G.....E.............ADKG..K------ 129
C.IN.09.T125_2139            ------I-R-EK--K------------H-M---L-------DK--L---------------MK----A----T-----F-----------EG--VR-------------.........C...........QQ-T--.......-K.....E.............ADG...K------ 130
C.KE.00.KER2010              ..----I-R--K--A------------H-R---L-----------L----------K--IK----A----T-------------------NK--VR-----------------...........QQ-T--.......GK.....A.............ADG...K------ 128
C.MW.09.703010256_CH256.w96  ------I-R--K--K--------------Q---L-------K--L--D--------K--IR--H-A----T---K-------------KN-DVR-------R-.........C...........QQ-T--.......--.....A.............ADG.....---R-F- 130
C.TZ.08.707010457_CH457.w8   ---S--I-R-EK--K------------S-M---------------L-------E---K--MK----A----T----------V------K---VR-------R-...........QQ-T--.......-E.....A.............AD....K------ 129
C.YE.02.02YE511              ..----I-R--K--T--------------M---L----------------------K--IR----A----TA--K-------------KD-AV-------------------...........QQ-T--.......-T.....A.............EDK...K------ 128
C.ZA.08.705010534_CH534.w12  ------I-R--K---------------H-MI--L-----------L--------A----MK----A----T-----F----------EE-DVR-------R-...........QQA-...............A..............ADKG..KI----- 127
C.ZA.10.DEMC10ZA001          ------I-R--K--K------------C-Q---L-----------L--------K----IQ----A----T---K--F----------EG-DVR------------------...........QQ-T--.......-K.....A.............TEGK..A------ 131
C.ZM.11.DEMC11ZM006          ------I-R-EK--K--R-K-------H-M---L-----------L--------D----IN----A----T-------------------A--DVR-----------..-R...........-Q-T--.......-E.....A.............ADKG..K------ 131
D.CM.10.DEMD10CM009          -----------K--Q--S------N-R-------------------L-------AD--K---S--E-AIK------I--H----------K--DV-------E--.......K-I...........--RE.A.......----.--...............N-S...------- 132
D.CY.06.CY163                --------R--K----T-------H----Q---------------I--------D--K---------A---S---F------I------QAG-----------E--.......-N-...........------........----.--...............N-S...------- 133
D.KE.11.DEMD11KE003          -----------Q--T--A------S---------------------L----------K--M-----AV-------K--F------W---EG-KV---------------.......T--...........------........----.-R...............N-S...------- 133
D.KR.04.04KBH8               -----------K--T*--------R--M--------------K--L-------K----I------I-----K-K--F-----------E--NV-------------------...........------........-T--.-K...............NNS...------- 132
D.SN.90.SE365                -----------Q-----------------------------------L--------------I----AI--------------I-------EK--V------VE-M.....R--...........---V--........-T--.L-...............NNS...------- 133
D.TZ.01.A280                 ..---------Q--A--------------Q---------------L----------K-#...---AI-------K--F-----------RK--V-------E-L.......T--...........------........-T--.--...............S-S...------- 127
D.UG.10.DEMD10UG004          --------R--K--E-------S--R---R---L-----------I----------K-VM-L----IK------K---------I----AG-KVT---------------.......T--...........-E-V--.......-TT-..-R...............N-S...------- 133
D.UG.11.DEMD11UG003          -----------K--D--R-Q-----N-R------------------I--C-----------I----AI------IK-------------E--KV---------------.......T--...........------........----.-K...............N-S...------- 132
D.YE.02.02YE516              ..---------K--E----------S---R---------------L--------D--K---------A-K---------F----------E--GV-------------------......I...........----P-........----.--...............N-S...------- 131
D.ZA.90.R1                   -----------R--A--R---------Q-R---------------L----------K--MA-----I-----------------------E--DV---Q--------------...........------........----.--...............N-S...------- 133
F1.AO.06.AO_06_ANG32         -----------K--D------------Q-----L-----------L-----------K-I--------------K------VV--Y----EK-S--------L...-K----...........QQ-T--.....VT-V..-A................A...A----- 130
F1.AR.02.ARE933              -----------K--T--------------R---L-----------ID------A----K-I--H--I-------K-------V--S------V-------E-L...-D----...........QQ-R--.....----.--...................------ 129
F1.BR.10.10BR_RJ015          ------I----K--Q--------------RM--LI-----------D-----P--KK-IA------------------F--I-V--F---KV-V-------E-L.........-...........-Q-T--......-T-A..-...............AAA...G------ 132
F1.CY.08.CY222               -----------K--A------------Q--M--L-----------LD-----------KK-I-------------------IVV--Y---K--V-------E-L...-A--T--...........QQ-T--........----.K-..................------ 129
F1.ES.02.ES_X845_4           -----------K--D------------Q-RM--L-----------ID-----------K--I-----A-K------K--F-L--V--Y---KEK-----D----L.........-...........QQ-T--.......---AA.K-..................------ 130
F1.RO.96.BCI_R07             -----------K--A--------------RM--L-------------------P----K-I----------------------I-V-------V-V-------E-L......R--...........QQ-T--........----.KV..................------ 129
F1.RU.08.D88_845             -----------K--D----------R---R----------------D-----------KK-IA-----P-----I--------V--Y----PV-------E-L.........-...........QR-P--.........---A..-...............AAAADQG------ 135
F2.CM.02.02CM_0016BBY        ..---------K--D--------------R---------K-----L-------T---K--I---S--------IK-------V--Y---K-Q-R---------L...Q--D-Y...........QQ-T-P.........----..K-..................------ 127
F2.CM.10.DEMF210CM001        ------I----K--S--------------R---L-----------L-------T--KK-IA---------------K----A-VV--Y------------------L...Q----...........QQ-T--........----..K-..................------ 129
F2.CM.10.DEMF210CM007        -----------K--A------------Q-R---L-----------L-------TD--K--IL---------------K--F-A-VV--Y------V-----------L...Q--D--...........Q--T--........----..K-..................------ 129
G.BE.96.DRCBL                -----------K--A------------R-RM--L-------D---L-------A--QK-MA----A----T--IK--F-------------K--VR------EEV...-KI-K--...........QQ-E.............N..............S-S...------- 126
G.CM.10.DEMG10CM008          -----------K--A--------------QM--------------L-------T--Q--MT----------T---K--------------R---V-------E-...-KI-Q--...........QQ-T--.......--GG..E-..............N--...------- 133
G.CN.08.GX_2084_08           -----------K--A------------------L-----------L--D---AD--WK-MD----A----T--IK--F-A----------K--V---Q--IKEV...-KIXQ--...........QQX--K.......--.....A..............N-S...------- 130
G.CU.99.Cu74                 -----------R--A--------------RM--L-------K--L--------G--Q--MN----A---AT---K---------------K--V------EEV...-KA-K--...........QQ-T--.......--R-..E-..............NNS...------- 133
G.ES.09.X2634_2              -----------K--K----------R-Q-----------------L-------A--Q--MV----A----T--IK----------------G--V-------EV...-KI-K-G...........Q--.--------M-..E-..............DNS...-T----- 132
G.GH.03.03GH175G             -----------K--A--------------RM--L------M----L-S-----A--Q--MM----A----T-----F------------------V-------E-V...-KL-KN-...........QQ-.---------KA.E-..............NN-...------- 132
G.KE.09.DEMG09KE001          ------I-R--K--A--R-----------R---L-----------L--D---E--Q--MR-----I--T--VK--------W---RK--VR-----EEV...-KA-K--...........QQNV--........--R-..G-..............NNS...------- 133
G.NG.09.09NG_SC62            .#---------K--A------------R-RM--L-----------L------SA--Q--MM-----I--T--IK--F-------------K--V-------EEV...-KI-K--...........QQEK-K.......--...................NS...------- 126
G.ZA.01.TV546                ..---------K--A------------Q-R---L------M----L-------X--Q---V-----X--AT-----------------K---V-------EEV...-KI-KN-...........QQ-T--........--MX..-X..............X-S...------- 131
H.BE.93.VI991                -----------K--A----------R---R---L-----------L--D---AD--Q---------A-K--T-D-Q-----I-V---------DV-------G--...-I--N...........-QRT--...AP----..KE..............KDS...KI----- 135
H.CF.90.056                  -----------K--A--------------R---L-----------L-------P---L--IE-I--AIK--T---K--F-L--V-----RK-DV-----------...-I---...........QQ-T--..........----..KE..............KD-...K------ 133
H.GB.00.00GBAC4001           -----------K--A--------------R---L-------D---L--D---AD--LK----I--A----T--IK--F-L--V------K--VQ--S---N-V...K-I--N...........QQTT--........-T-X..KE..............KDS...KI----- 133
J.CM.04.04CMU11421           -----------K--T------------R-R---L-------D---L--X-----K--Q---V-----Y---T--IK--------------EX--V-----X-------...-L-K-N...........-QQ--K.....-EX-..KR..............XXS...------- 133
J.SE.93.SE9280_7887          ------I----K--D--------------RI--L-------D---L------SAK--Q---V----A----TQ-IK--------------------------M---E--...-I--N...........-QQ--K.....-ET-..KK..............DNS...------- 133
J.SE.94.SE9173_7022          ------I----K--D------------Q-RI--L-------D---L------SAK--Q---V----A----T--IK--------------------------------...-I--N...........-QQT-K.....-ET-..KK..............DNS...------- 133
K.CD.97.97ZR_EQTB11          -----------K--K---Q----------R---L-----------L--N---V------IR----------------F-------W---S-QVR---------L.......-RT...........QQ-T--.....GK--..K-....................------ 129
K.CM.96.96CM_MP535           -----------K--A------------------L-----------L-------T------IT-I---I------IK-----I-V--F---K--V---------L.........-...........QR-T--.....E---..K-....................------ 129
U.CA.01.TV749                -----------K-----------------R---L-----------L-S----A----K-I-----A----T---K---------------G--V-------R-.......K-C..............K-.......--DA.--...............SGR...E----F- 129
U.CA.99.TV721                -----------Q---------------Q-R---L-----------L--SF--A----KLI-----A----T---K--F----V-----------V-------R-.........C..............K-.......-TGA.--...............NGR...E------ 129
U.CD.83.83CD003_Z3           -----------K--A----------R---R---L-------K--I-------A---Q--I------A---------K-I--L--------------V-------------...V--G...........-Q--------------..K-..............SNS...------- 133
U.CD.90.90CD121E12           ...K-------K--A----------R---R---L-----------L-------A---Q--IV----AF-------K--------------------------E--...-A--G...........-Q--------------..K-..............DNS...K------ 130
U.CY.05.CY090                -----------K--K--------------R---L-----------L-------A-------I----SAIN------K----------------G--VR---------L.........-...........-R--------ET--..K.......................E----F- 129
U.CY.08.CY223                ------I----K--A--------------R---L-----------L-------A---Q---E----ST-R------K--F--I-----------V-------------...LK---...........-Q-T----------AD.--...............S-S...K------ 134
U.ES.10.DEURF10DZ001         -----------K--D--------------R---L-----------L-------A---K--IE--EST-K------K--F------F------KVG-------------...MKQ--...........-Q-T----------A.--...............GC.........---F- 130
U.GR.99.99GR303              -----------K--A------------Q-R---L-----------L-----------Q--ME----A-K------K------------------DV-------------...L-KR-...........-Q-T--........----.--...............GTS...K------ 133
U.NL.95.U_NL_95_H10986_D1    --------R--Q--T------------------L-------------------A---Q--IM----A----T-------------------CK-----------------...I--N...........---T.......E--.--...............KTS...------- 130
01_AE.AF.07.569M             -----------K--A---------------M--L-----------L-------A---Q--IE---ST-K------K------V--W-------V----------------...-K--...........RQ-T--.........---G.--...............N-S....-A----- 133
01_AE.CF.90.90CF11697        -----------K--A--------------RM--L-----------L-------A---Q-LIE---SN-K-----IK--F--I---W------DV-------------...V--N...........-Q-T--.........---G.--...............SNS...K------ 133
01_AE.CN.10.YNFL03           ------I----K--A------------------LI----------L-------A---Q-LIE---ST-K------K------V--W------QVN---------L...V-KR-...........QQ-T--.........--TG.--...............S-S...K------ 133
01_AE.HK.04.HK001            ------I-T-EK--A--------------MI--L-----------L-------A---Q--IE---ST-K------K--F-----W-------V-----------L...I--N...........Q--T--.........---G.--...............SNS...K------ 133
01_AE.IR.10.10IR.THR48F      -----------K--A--------------R---L-----------L-------A---Q-LIE---ST-K------K--F-----W-------V-----------L...A-T--...........QQ-T--.......T--G.--...............S-S...K------ 133
01_AE.JP.x.JRC77AE           -----------R--A--------------RI--L-----------L------AI-Q--IE---ST-K------K--F-A-VV-W-------V-----------L...V-K-E...........QQ-T--.........---G.--...............N-S...K------ 133
01_AE.TH.09.AA090a_WG11      --------T--K--A--------------RM--L-----------L-------A---Q-L-E---ST-K---------F-AI-V-----K--------------------...A-K--...........QQ-T--.........---S.--...............SG.....------ 131
01_AE.TH.90.CM240            -----------K--A----------R---R---L-----------L--SF--A---Q--IE---ST-K--L---K--F------W-------V-----------------...V----...........QR-T--.........---G.--...............S-S...K------ 133
01_AE.US.05.306163_FL        -----------K--A--------------RM--------------L-------A---Q-LIE---ST-K---------F-AI---W---K---V---------L...-...E-...........QQ-T--.........----.--...............S-S...K------ 129
01_AE.VN.98.98VNND15         -----------K--A--------------RM--L-----------L-------A---Q-LIE---ST-K------K--F--I-------R---V-----------------...V----...........QQ-T--.........---G..P-..............S-S...K------ 133
```

```
                             membrane binding                                                                      phosphorylation site
                             Gag start, p17 start                                                                                                                      p17 end  p24 start
B.FR.83.HXB2                 MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKI...EEEQNKS..........KKKAQQ.....AAAD..TG..............HSN...QVSQNYP  133
02_AG.CM.10.DE00210CM013     ------I----K--Q-------------R---L-----------L-------A---Q--ME---SA-G--T---K--F--I---W------D----------V....-V---RQKTQ.....QQ-T--.........---A..--..............S-.........---F-  135
02_AG.CY.09.CY256            -----------K--A------------Q-R---L-----------L-------A---Q-LIE---ST-K------K--F--I---W---S--T----------.....-V-----..........-Q-T--.........---A..--..............S-..........------  130
02_AG.ES.06.P1423            -----------K--A--------------R---LI----------L------SA---Q--ME---ST-R------K-------V-------F--R-----------.....-IK---..........-Q-T--.........---A..ST..............GC...........------  130
02_AG.GW.05.CC_0048          -----------K--A--R-----------R---L-----------L------SA---Q-LME---ST-R------K------I---W------D-R---------V....-V---KQKAQ.....QQ-T--.........---A..A-..............S-..........------  135
02_AG.KR.12.12MHR9           -----------K--A--------------R---L-----------L-------A---Q-LIE-I-SA-R-----FK-----IV--W------D----------L....-V-----..........-Q-T--.........---A..--..............I-..........------  130
02_AG.LR.x.POC44951          -----------K--A--R-----------R---L-----------I-------A---Q--ME---ST-S------K--F--I---W---R-LD-T----------.....-M-----..........-Q-T--.........---A..--..............S-..........------  130
02_AG.NG.09.09NG_SC61        .#---------K--A------------Q-R---L-----------L-------A---Q-LME---ST-K------K--F--L---W------D-R---------V....-M-K---..........-Q-T--.........---A..A-..............S-..........------  128
02_AG.NG.x.IBNG              -----------K--A--------------R---L-----------L-------A---Q-LME---SA-R------K--F--I---W------D----------M....-I-----..........-Q-V--......T--A..--..............S-..........------  130
02_AG.SE.94.SE7812           -----------K--A--------------R---L-----------L-------A---Q-LME---ST-R------K--F--I---W------D----------V....-V-K---..........-Q-T--.........---A..A-..............S-..........------  130
02_AG.SN.98.98SE_MP1211      #----------K--A--R-----------R---L-----------L-------A---Q--IE---ST-K------K--F---------------V---------L....-I-E---..........-Q-T--.........---A..--..............S-.........---F-  129
03_AB.RU.97.KAL153_2         -----------K--A-----------E--RI--L-----------L--S-------Q---E---T-K------K-----------------------------.....-I-----..........-Q-T--.........--TG..--..............S-S...K------  133
04_cpx.CY.94.94CY032_3       -----------K--A--R-----------R---L-----------L-------A---Q-LME---ST-K-------------IT--W------DVQ----------.....-I-S---..........-Q-T--.........---A..A-..............G-S...N------  133
05_DF.BE.x.VI1310            -----------K--A--------------R---L-----------L-----------K--IS---AI---T---K-------------E---V-------E--.........-K..........R---------.......-E-G..A-..............N-S...-A-----  134
06_cpx.AU.96.BFP90           -----------K--E---------------M--L-----------I-------A---Q--IE---ST-K------K--F-------------KVT-----I---.....-I-K---..........-Q--H-.........---A..--..............N-S...NL-----  132
07_BC.CN.98.98CN009          ------I-R--K--K------------H-M---L-----------L---------------MK----A----T------F-----------TE-DVR-----------.........-I..........QQ-T--.......-K.....E..............ADG...K------  130
08_BC.CN.97.97CNGX_6F        .#----I-R--K--K------------H-M---L-----------L-----------K--IK----A----T------F-----------EE--VR-----------.........-I..........QQ-T--.......-K.....K..............AEE...K------  128
09_cpx.GH.96.96GH2911        -----------K--A------K--------RM--L-------------D-S----TG--Q---E------R------K--------------K--VR-----------.....-I-G---..........-Q-T--.......-T-N..--..............NNS...K------  133
10_CD.TZ.96.96TZ_BF061       -----------K--E------------Q-----L-------D---I-----------K--I----AI-----IK-------------E--KVA----------..........-T---..........-----------.......-T--..--..............N-S...--------  133
11_cpx.CM.95.95CM_1816       -----------K--K----------R-Q-R---L-----------L--S---A---Q--MV-I--A-N--T--------------------Q--VR-----------.....-I-----..........-Q-T--..........----..--..............N-S...K------  133
12_BF.AR.99.ARMA159          -----------------------------R---------------------------K-I------------------F---V------K--V-------E-L..........----------..........QQ-T--...........-----.K-.............................------  129
13_cpx.CM.96.96CM_1849       -----------K--K----------R-RM--L-----------L-------T---Q---R-------------K-----------------V-------EEV...-KV--S-..........QQRT--.......--KN..E-..............N--...........------  133
14_BG.ES.05.X1870            -----------K--S--------------R---L-----------L--D----A---Q--M-----A----T--I---F-------------------------EEV...-KA-K--..........Q--.--.......--K-..E-..............NNS...-A-----  132
15_01B.TH.99.99TH_MU2079     -----------K--A------------------L-----------I--S---SA---Q---E---ST-K------K---------W--QN---V-----------.....-V---N..........QQ-T--.........---G..--..............S-S...T--H---  133
16_A2D.KR.97.97KR004         -----------K--A--------------R---L-----------I-----------K--I---E--IK-----IK--F-A-V-----------VR---------L.........-H..........-Q-T-P........-----..--..............S-.....G-----  131
17_BF.AR.99.ARMA038          ..----------------------R---R-----------------------AD---K-I-------------------I-V--Y---KVDV-------E-L..........----..........RQ-T--.AAADKTQQ..AA..............ADK...G------  135
18_cpx.CU.99.CU76            -----------K--K--------------R---L-----------L-------A---Q--IE---S-IK-----IK--------------G-DV-----------.....-I-----..........-Q-T--.........---A..A-..............S-S...--------  133
19_cpx.CU.99.CU7             ------L----K--A------S-----Q-RM--L-----------I--S---SA---Q--IE---ST-K---------F------W----Q--V-----------L....-V-----..........-Q-T--......ET--..-R..............N-S...H------  133
20_BG.CU.99.Cu103            -----------K--E--------------RM--L-----------L-------V---Q--IN----A-K--T---K--F--------------GV---------EV...-KL-K--..........QQ-M--.........--TG..E-..............N-S...--------  133
21_A2D.KE.99.KER2003         ..---------K--A------------Q-R---L-----------I-----------K--I------I------K--F----------A--DV----------..........-------..........-N---P.........---N..--..............N-S...--------  131
22_01A1.CM.01.01CM_0001BBY   ..------R--K--A------------Q-RM--L-------D---LD------A---Q--IE------K-----IK-----I------------DV-----------.....-I-----..........-Q-T--......G--G..--..............S-S...N------  131
23_BG.CU.03.CB118            -----------K--A--------------RM--L-------D---L-------A---Q--MS----A-K--T--IK--F-A-------------V---------EV...-KL-K--..........QQ-M--.........--TG..E-..............N-G...--------  133
24_BG.ES.08.X2456_2          -----------K--A--------------RM--L-------D---L-------AG--Q---N----A----T--IK-------------K--VR---------EV...-KL-K--..........QQ-V--.........--TG..E-..............S-S...--------  133
25_cpx.CM.02.1918LE          ..----I----R--A----------R-R-R---L-----------L--S---NG---Q--------SA-K------K-------V---------DV-----------.....-I---N..........-Q-K--......ET--..--S..............N-S...--------  132
26_AU.CD.02.02CD_MBTB047     ------I----Q-----------------R---L---------K-SI-------A---Q---E---AVK------K--F---V--------NV---------V....-M--..........----------.........--DN..--..............S-S...K------  130
27_cpx.FR.04.04CD_FR_KZS     -----------K--A----------R---RM--L-----------I------AAD--LD--R----A--------K--F-L------------DV------EEV...-KV-KNR..........QQ-T--......E.....E-..............S-S...-A-----  130
28_BF.BR.99.BREPM12609       -----------------------------R--------------------------I-----------------------V------K--------------.............................----------.....A-..............N-S...--------  133
29_BF.BR.01.BREPM16704       -------------------------N-----------------------------S---------------K--F---V--------D-------E--.................----------T-----------G...NSRQQAAAGTG...N-S...--------  144
31_BC.BR.04.04BR142          ------I-R-EK--T--R---------Q-MM--L-----------LD-------G--K--IQ----A-K--TK--V-------------EK--VR-----------...K----------..........QQ-T--.......-E.....A..............AGKG..N------  132
32_06A1.EE.01.EE0369         -----------K--E----------R------------------L-------A---Q--VE-V--A-R------K--F--------------KVA-----------.....-I---..........-Q-----------TA..--..............N-...........------  130
33_01B.ID.07.JKT189_C        -----------K--A--------------R---L-----------I-------A---Q--ID---ST-K------Q--H--IV--W-------V---------M....-T---R..........QQ-T--.........---G..--..............SNS...T----F-  132
34_01B.TH.99.OUR2478P        ..---------K----------------S-QM--L-----------L-------A---Q--IE---ST-K------K--F--I---W----G-DV----------L....-V-----..........QQ-T--.........---G..--..............SGX...K------  131
35_AD.AF.07.169H             ------I-R--K--T--------------R---L-----------L--S---T---Q--ME----A-K--------F-----------Q-A----------.....-I-----..........-Q---K.....T--.--..............SG....K------  132
36_cpx.CM.00.00CMNYU830      ..---------K--A--------------RI--L-----------D----SA---Q--IE---T-K------K----------------DV-----------...-KL-SN-..........-Q-T--.........----..--..............S-S...T--R---  131
37_cpx.CM.00.00CMNYU926      ..------T--K--A------------Q-RM--L-----------L-------A---Q--IE---ST-K------K--F------W------DVR--------L....-V-----..........-Q-T--.........---G..--..............S-C...KI-----  129
38_BF1.UY.03.UY03_3389       -----------K--A-----------------------------------S-----K-MAH------------K-------V--F----VDV-------E-L.......-K--..........QQ-T--..........----.K-..............................------  129
39_BF.BR.04.04BRRJ179        -----------------------------R---------------------------IA-------------------I-V--Y---KV-V-------L.....---------....................--------------.............NNS...--------  133
40_BF.BR.05.05BRRJ055        -------------K---------------Q-----------------S---A-----------H---------K-----I--------K--V--------...........--R-----.A-----.A-..............DNS...K------  134
42_BF.LU.06.luBF_18_06       #-----------------------------------------------------M-----A--------K---------------K-DVR-----------.....-----..........----------.........----..--..............N--...........------  132
43_02G.SA.03.J11223          -----------K--A--------------RM--L-----------I------SA---Q--ME-X-ST-R------K--F--I---W------D---------V....-M-----..........-Q-T--.........---A..--..............S-...........------  130
44_BF.CL.00.CH80             ------I----K--A--------------RM--LI-----------D------P----K-I-------------------VV--F----V-V-------E-L..........----------..........QQ----.KTQQ---AADK-...............------  135
45_cpx.FR.04.04FR_AUK        -----------K--M------------------L-------D---L-------A---Q--IE---N-K-----IK------V----------DV-----------.....-I-----..........-QNT--.........---A..--..............S-S...S------  133
46_BF.BR.07.07BR_FPS625      -----------K--A--------------RM--LI-----------LDS-----P----K-IA---S-----------F---VV--F---KV-V-------E-L..........----------..........QQ-T--..........----.K-..............................------  129
47_BF.ES.08.P1942            --------------------------T--------------------L-----------------A-----------F-A---------K-DV-----------.........------------V--..........---..--..............NN-.........------  132
48_01B.MY.07.07MYKT021       ------I----K--A--------------R---L-----------L-------A---Q-LIE---ST-K---------F--IV--L-------V-----------.....-M-K---..........QQ-T--.......-E-G..--..............S-S...K------  133
49_cpx.GM.03.N26677          ------I----K--K------------Q-R---L-------------------A---Q--IE---A-K---------------------------V-----------.....-I-K---..........-QGT--.........---A..--..............NIS...SI-----  133
50_A1D.GB.10.12792           ------I----K--T----------R---R---L-----------L--D----A---L--MD-I--A-R--T------F-----------------V-----------.....-V--R..........-Q-TP-........----..AA..............SCS...-A-----  133
51_01B.SG.11.11SG_HM021      .#---------K--K--R-----------R-------------------------------------X-----F---V------M-DV-----------..........-----..........--Q--.........---N..--..............N-S...--------  131
52_01B.MY.03.03MYKL018_1     ------X----K--A--------------QM--L-----------L-------A---Q--IE---ST-K------K------V--W------DV---------M....-A----..........-Y-T--.........---G..--..............NXS...K------  133
53_01B.MY.11.11FIR164        --------T--K--A--R-----------XI--------------L-------A---Q--IE-I--X-K------K--H-L--V-W-------VX---------M....-X-K-G..........QQ-T--.........-----..--..............S-S...K------  133
54_01B.MY.09.09MYSB023       ------I----Q-----R-----------R-------------Y-------S---K---E----A-----GF---F-----V----------V-----------.....-A-K-G..........QQ-K--.........-----..--..............S-S...--------  133
55_01B.CN.10.HNCS102056      -----------K--A--------------R---L-----------L-------T---Q-LIE---ST-K---------F--IV--W------DV---------L....-V-K---..........QQ-T--.........--TG..A-..............S-S...--------  133
56_cpx.FR.10.URF5_patient_A  -----------K--A--------------R---L-------D---LD-----SA---Q-LIE----A-K------K--F--I---L---L--N----------V...........-----..........-----------.........---A..--..............S-S...--------  133
57_BC.CN.09.09YNLX19sg       ------I-R--K--K-----------R-M---L-----------L--E----A---K-LMK----A----T-GFK--F-----------AE-DVR-------R-...-K---I..........QQ-T--.......-Q.....E..............ADG...K------  130
58_01B.MY.09.09MYPR37        --------T--K--A---------------M--------------L-------A---Q--IE---ST-K------K--F-----W-------V-------RL...-A---G..........-Q-T--.........---G..A-..............N-S...K------  132
59_01B.CN.09.09LNA423        ------I-T--K--A--------------RI--L-----------D------K---Q--IE----A-K---------F-----W-------DV---------M....-A-..........----------.--T--......G--G..--..............S-S...KI-----  129
60_BC.IT.11.BAV499           ------I-R-EK--T--R-K-K-----H-M---L-----------LD-------G--K--MN----A-----K--V--H-A---------EK--VR----------...K---R..........QQTTR-.......-E.....A..............ADKG..K------  129
61_BC.CN.10.JL100010         ------I-R--K--K------------H-M---L-----------L-----------K--IK----A----T------F-----------SE-DVR-----------.........-N-..........QQR---.......-K.....E..............ADG...K------  130
62_BC.CN.10.YNFL13           ------I-R--K--K------------H-M---L-----------L--------D--K-LMK----A----T------F--L-------KG-DVQ-----------.......-.................-T--.......-E.....-..............ADG...KG-----  126
63_02A1.RU.10.10RU6637       -----------K--A--------------R---L-----------L-------A---Q-LME---ST-K------K--F--I---W------D-R-----------.....-V-----..........-Q-----------A..--..............S-...........------  130
64_BC.CN.09.YNFL31           ------I-R--K--K------------H-M---L-----------L-----------K--MK----A----T------F-----------EE-DVR---------V.........-I..........QQ-T--......EK.....-..............ADG...K------  130
65_cpx.CN.10.YNFL01          --------T--K--A--R-----------RM--L-----------I-------A---Q--IE----TVK------K--F-A-VV---------D----------L....-------..........QR-T--.........---S..--..............SGG...K------  133
67_01B.CN.11.2011.ANHUI.HF115-----------K--A------------Q--M--L-----------L-------A---Q---E---ST-K------K--F--I-V-W------A-----------.....-V-----..........QQ-T--.........---G..A-..............S-S...K------  133
68_01B.CN.11.2011.ANHUI.WH73-----------K--A--------------RM--L-----------L-----------Q---E---ST-K------K--F------W--------------------.....-V-KR-..........QQ-T--.........---G..A-..............S-S...K------  133
70_BF1.BR.10.10BR_PE004      I-----I------------------Q-R---V--------------------------A--------R------F-------------DV------E--....................--------------.........---G..--..............N-SQSS-SG----  136
71_BF1.BR.10.10BR_PE008      -----------K--K--------------R---------------------------A-------------------------M-DV-----------......................-R-----------.---G..AD..............N-S...--------  133
72_BF1.BR.10.10BR_MG002      -----------------R------N-----------------------------I-------------K--F--I-V---------DV------E--.........................-----------.......-E-...............KGS...P------  133
O.BE.87.ANT70                ---S----T-SK--A--Q---K--S----R---L-----------C--E----A--NEKL-Q--E-A-K---DS-Q--W-AIVV-W---N-YK-G--QQ-IQ-L...K-VMGSRK.................S-D-AKEDT..............SAR...-AG----  131
O.CM.98.98CMA104             --------T-SK--A--Q-K-K--C----R---L-----------C--E----A--NEKL-Q--E-A-K----S-Q--W--I-V-W---N-FKVE--QQ-IQ-L...R--KGGRK.................S-G-AKEDT..............SAR...-TG----  131
O.CM.98.98CMABB141           ---S----R-SK--T--Q---K--C----R---L----K------C--E----A--NE-L-Q--E-A-K----S-Q--W--I-V-W---N-FKVG--QQ-IQ-L...R-VMGNRK.................S-S-AKEDT..............SAK...-TG----  131
O.CM.98.98CMABB212           --------T-SK--A------K--S----R---LI---K-----EC------YA--NE-L-Q--E-A-A----S-K--W-LL-V-W---S--D-Q--QQ-VQ-V...K-KMANRK.................S-D-AKKET..............SPK...-T-----  131
O.CM.98.98CMU5337            ---S----T-SK--A--H-H-K--C----R---L-----------C--E----A--NEKL-Q--E-A-K----S-Q--W--I-V-W---N-YKVG--QQ-IQ-L...K-VMGSRK.................S-G-AKEDA..............SAR...-TG----  131
O.CM.99.99CMU4122            ---S----K-SK--A--Q---K--S----R---L-----------C--E----A--NEKL-Q--E-A-K---DS-Q--W--IVV-W---N-YR-E--QQ-IK-L...R-VMGSXK.................S-SDAKEDT..............SAR...-TG----  131
O.FR.92.VAU                  --------T-TK--A--Q---K--S--------LI----G-D---C--E----A--NE-L-Q--E-A-K----S-*--W-AI-V-W---N-YKVE--QQ-VR-L...K--MK-RK.................PVSSTKENT..............S-K...-TG----  130
O.GA.11.11Gab6352            ---S----R-SK--S--Q---K--S----R---L-----------C--E----A--NEKL-Q--E-A-K---DS-Q--W--I-V-W---NGFR-N--QQ-IQ-L...K-VM.NKK.................S-S-AKEDT..............SER...-TG----  130
O.SN.99.99SE_MP1299          --------T-SK--A--Q---K--C----R---L-------D-S-C--E----A--NEKL-Q--E-A-K----S-Q--W--I-V-W---N--KVE--QQ-IQ-L...K-VMGSRK.................S-GTAKEDT..............SAR...-TG----  131
O.US.10.LTNP                 ---S----T-SR--A--Q---K--S----R---L-----------CS-E----A--NEKL-Q--E-A-K----S-Q--W--I-V-W---N-YK-G--QQ-IQ-L...K-VMGSRK.................P-G-AKEDA..............SAR...-TG----  131

N.CM.02.DJ00131              --------T--K--Q--A-Y---------RM--L-----------C----MD-AS--A-L-N--E-A-A----G------AL-V-----SN-QVHN-Q------KEIK-KKERHKS........EP-NPE.....-G-A..AA..............AD-...NI-R---  138
N.CM.02.SJGddd               --------T--K--Q--S-Y---------R---L-------D---C----MD-AV--A-LIN--E-A-K----G-------L-V-----S--QVHN-Q--------..K-K-EQHK.......PPEP-NPE.....-G-A..AA..............VXS...NI-R---  136
N.CM.04.04CM_1015_04         --------T--K--K--A-Y----------M--L-----------C----MD-AH--N-L-N--E-A-N----G-------I-V-----ST-PVLN-Q--------..K-KKEQHKS........EA-NPE.....-G-A..AA..............AD-...NI-R---  135
N.CM.06.U14296               --------T--K--Q--A-Y---------RM--L-----------C----MD-AN--A-L-N--E-A-K----G-------L-V-----S--QVQN-Q----XX...KXK-XQHKP........EP-NPE....TE-A..AA..............ADS...NI-R---  135
N.CM.06.U14842               --------T--K--Q--G-Y---------RM--L-----------C----MD-AA--A-L-N--E-A-N----G------XL-V-----NKRQ-HN-Q----H-...K-VKEQHKP........-P-EPE.....-G-A..AA..............VD-...XI-R---  135
N.FR.11.N1_FR_2011           --------T--K--Q--A-Y---------RM--L-----------C----MD-A--XV-L-N--E-A-K----G-------L-V-----S--S-Q--QA-------...K-KKEQHK.......QPEP-NPE.....-G-A..AA..............PD-...NI-R---  136

P.CM.06.U14788               --------T--R--A--R---------------V-----------C----M--A---EKL-E--E-A-R---DG-Q--W--LVV-W---K----S--QQ-IT-W...K--MQ-R...........--T.SEG............................-T...GT------  126
P.FR.09.RBF168               --------T--R--A--R---K--S----M---------------C--E-M--A---EKL-E--E-A-R---DG-Q--W--LVV-W---R-VD-S--QQ-IT-W...K--MQRR...........--A.LEGS..........................TGT...GT------  128

CPZ.CD.06.BF1167             --------R-EK-----SV----------LI--L---GN--Q---L----M--AD---K-IM--S--VEI--P-II--F--I-V-H-I-EGEK-Q--EA-VKIV...K-KLKITSSETTGNEETA-GKKET....-VTSSGQTEKIYPTLP......S-G...-TGR---  154
CPZ.CM.05.SIVcpzMT145        ------I-T--R--A--------------MM--L-----------C----M-NA---EKL-K--E-A-K----G----F--L-V-W---K---VE--QQ--Q-L...KAAV-PE..........QGETPSA.......................................-A-F-  123
CPZ.GA.88.GAB1               --------T--K------V------R-R-MM--L-----------CD---M-SK---TKL-Q--E-A-K----G----F--L-V-W-I-SD-TVE--QK--EQL...KRHHGEQQS........-TESNSG....SREGGASQGAS..........A-A...GI-G---  141
CPZ.TZ.06.SIVcpzTAN13        --------R-EK--T--S-----------LI--L---GS------M----M-NK---WK-IL-----IDI--P-IV------C--F---AGEKVR--E--VKIV...KVKLTTV..........-DD-AATSSGQGQNAGGR-TAVTPSGKSENASRVTVPPSEGGRL--  157
CPZ.US.85.US_Marilyn         --------T--R--A--------------MM--L-----------C----M--AD--L-L-K--E-A-K----G----F--L-V-W---S-VTVE--QQ--V-L...K-VVQLQK.......QQEE-E--QQ........EAS..............G--...IG-S---  135

GOR.CM.07.SIVgor2139_287     --------T--R--A--R---K--S----M---V-----------C--E-M--A---EKL----ESA-K---DG----W--LVV-W---R--DVS--QQ--E-F...K-VMS-RK..............SD...........................SRP...ET------  124
GOR.CM.07.SIVgorCP2135con    --------T--R--A--R---K--S----M---V-C---------C--E-M--A---EKL----E-A-K---DG----W--LVV-W---R--DVS--QQ--E-F...K-AMS-RK..............SD...........................SRP...ET------  124
```

```
                                                                                                                         CyPA binding                                        major homology region
B.FR.83.HXB2                 IVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTN.NPPI...PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDR 299
A1.AU.03.PS1044_Day0         ----A----I--NL-----------I--------------------------V---I------------D-----------L--------P---I----------A---P---LQ---G.----...---D-------------------------------S------- 299
A1.CH.03.HIV_CH_BID_V3538    ----A----I--NL-----------I-------------T------------M---I------------D-----------T--------Q---L------------------M----S.----...-----------------------V------------------- 299
A1.CY.08.CY236               ----A----I--NL-----------I--------------------------V---I------------D-----------L--P-----P---I--------------P---LQ---G.----...---D-------------------------K------------- 304
A1.ES.05.X1608_8             ----A---WT--NL-----------I-------------T--------H---M---I------------D-----------L--------P--------------A---P---LQ---S.---V...---D-------------------V------------------- 296
A1.KE.11.DEMA11KE001         ----A----I--NL-----------I--------I--------D----Y---M---I------------D-----------L---Q----P------------------P--------S.----...---D-------------------V-----K------------- 299
A1.RU.11.11RU6950            ----A----I--SM-----------I-------------T------------M---I------------D-----------L--AQ---FP---------------------------S.----...---D-------------------------K------------- 295
A1.RW.11.DEMA111RW002        ----A-------SL-----------I--------------------------M---I------------D-----------L---P---VP---I-D------------S----R---S.----...---D-------------------V-----K------------- 300
A1.SN.01.DDI579              ----A-------SL-----------I-------------------T------M---I------------D-----------M--------P---------------------------S.---V...------R----------------V------------------- 297
A1.UG.11.DEMA110UG001        ----A--------------------I-----------------D--------M---I------------D-----------L--AQ----P---I-D------------I----T---G.----...---D-------------------V-----K------------- 297
A1.ZA.04.04ZASK162B1         ----AH------PL-----------I---G-N--------------------M---I------------D------M-------------P-----------------N---------S.----...---D-------------------V------------------- 304
A2.CM.01.01CM_1445MV         ----A----------------------------------------------------------------D--------------------P---------------------------S.----...-----------------------V------------------- 295
A2.CY.94.94CY017_41          ----A----------------------------------T-----------------------------D--------------------P---------------------------S.D---...-----------------------V------------------- 295
B.BR.10.10BR_MG029           ----A-----Y-S---------------------------------P-----I---A------------------------L---Q---V-----------------------VA----.--A-...------------------------------------------- 298
B.CA.07.502_1191_03          ----L--------L-----------I-------------------------------------------------------L---Q---V------D----------------------H.----...-----------------------V------------------- 301
B.CH.08.M2_0803101_NFLG8     ----L--------L-------------------------------------------------------------------L-------------------------------------H.----...-----------------------V------------------- 299
B.CN.10.DEMB10CN002          ----L-------P-----------------------------A------------------------------D-------L--QQ---------------------------------.----...-----------------------------K------------- 301
B.ES.10.DEMB10ES002          ----L--------M-------------------------------------------------------------------M-------------------------------------H.----...-----------------------S------------------- 298
B.FR.11.DEMB11FR001          ---------------------------------------------------------------------------------L--------------------------------------.----...-----------M-----------------K------------- 299
B.GB.05.MM45d213_GN1         -------------L-------------------------------------------------------------------L---Q---------------------------------H.----...-----------------------------K------------- 307
B.HK.06.HK003                ----M-------------------------------------A--------------------------------------L-------V----I-------------------A---H.----...-----------M-----------------K------------- 299
B.HT.05.05HT_129389          ----M----------------------------------T--------------------------------------------------------------------------A---S.----...------------------------------------------- 297
B.JP.12.DEMB12JP001          ----L----------------------------------------------------------------------------L-------NP-----------------------A----.----...------------------------------------------- 299
B.KR.07.HP_18_07JHS10_3909   ----L----------------------------------------------------------------------------L--------------------------------------.----...------------------------------------------- 299
B.PE.07.502_0525_wg5         ----LX-------------------------------X-------------------------------------------L--------------------------------------.----...------------------------------------------- 299
B.RU.11.11RU21n              ----L--------L-------------------------T-----------------------------------------L-------V-----------------------------.----...-----------M------------------------------- 304
B.TH.08.MERLBDTRC10          ----L-------P---------------------------------------------------------------------------------------------------------A----.----...---D-------------------------K------------- 299
B.US.11.ES38                 ----L--------L-------------------------T-----------------------------------------L--AQ----------------------------A---S.---V...-----------------------------K------------- 298
C.AR.01.ARG4006              ----L-------PL-----------I-------------T-----------------------------D-----------L-------V------------------N-----A---S.----...---D-------------------V-----K------------- 295
C.BR.07.DEMC07BR003          ----L-------P-AA-----------------------T-----------------------------D-----------L-------V----V-----------------------G.----...-----------------------V-----K------------- 294
C.BW.00.00BW5031_1           ----L--------------------IG----N--I--I-T--------R--------------------D-----------L-------V------------------------A---S.---V...---D-------------------V-----K------------- 296
C.CN.10.YNFL19               ----L-------P------------------------------------------------------I--DV----------L-------------------------------A---G.---V...---D-------------------S-----K------------- 296
C.CY.09.CY260                ----L--------------------I----C----VR-VT-V------S---S----------------D-----------L-------V------------------------A----.----...---D-------------------V------------------- 300
C.ES.08.X2363_2              ----L--------------------I-------------T-----------------------------D-----------L--------------------------S-----A---S.---V...---D-------------------V-----K------------- 297
C.ET.02.02ET_288             ----L-------SL-----------I--------I----T-----------------------------D-----------I--------Q----------------------A---S.----...---D-------------------V------------------- 295
C.IN.09.T125_2139            ----L-------S------------I-------------T-----------------------------D-----------L------------S------------------A---G.---V...---D-------------------------K------------- 296
C.KE.00.KER2010              ----L--------L-----------I-------------T-----------------------------D-----------L-------NP-----------------------A---S.----...-----------------------V------------------- 294
C.MW.09.703010256_CH256.w96  ----L-------PL---------------G-N-------T----------M------------------D-----------L---Q---V----I------------------A---S.---V...---D-------------------V-----K------------- 296
C.TZ.08.707010457_CH457.w8   ----L--------L-----------I-------------T-----------------------------D---------------P---V----L-D----------------A---A.--A-...---D-------------------V-----K------------- 295
C.YE.02.02YE511              ------------SL-----------I-------------T-----------------------------D-----------L-------V------------------S-----A---G.---V...---D-------M-----------V-----K------------- 294
C.ZA.08.705010534_CH534.w12  ----L-------P--------------------------T----------------------------------D-------L-------V----------------S------A---S.----...---D-------------------V-----K------------- 293
C.ZA.10.DEMC10ZA001          ----L-------SL-----------I-------------T-----------------------------D-------------------V------------------------T---S.---V...-----------------------V------------------- 297
C.ZM.11.DEMC11ZM006          ----L--------L-----------I-------------T-----------------------------D-----------L---Q---V----V------------------A---S.----...---D-------M-----------V-----K------------- 297
D.CM.10.DEMD10CM009          ----L--------L-----------I-D--------------------A----------------------T---------L-------V------------------------A---S.----...---D-------------------V------------------- 298
D.CY.06.CY163                ----L------------I-S-----I----------------------H--------------------------------L--------------------------------D---S.---V...-----------------------V------------------- 299
D.KE.11.DEMD11KE003          ----L--------------------I-----N--I---------------------A----------------D-------L---Q---HP--------------------------.----...---D-------------------V------------------- 299
D.KR.04.04KBH8               ----L-------PL-----------------N-------T------------S---A-------------------S----L------------L----------------D------S.----...-----------------------VG----K------------- 298
D.SN.90.SE365                ----L--------L-----------I-------------T-----------------------------------------L--------------------------------------.----...-----------------------V------------------- 299
D.TZ.01.A280                 ----L--------------------I-------------------------------------------------------L-------------------------------A----.---V...-----------------------V------------------- 293
D.UG.10.DEMD10UG004          ----L--------------------I-------------------------------------------------------L-------V------------------------------.----...-----------------------V-----K------------- 299
D.UG.11.DEMD11UG003          ----L--------------------I-------------------------------------------D-----------L---Q---V------------------S---------S.----...-----------------------V-----K------------- 298
D.YE.02.02YE516              ----L--------------------I-----N--------------------------------------------G----L--AQ--------I--------V------E----R---S.----...-----------------------V------------------- 297
D.ZA.90.R1                   ----L-------P------------I-------------------------------------------------------L-------------------------------A----.----...---D-------------------V------------------- 299
F1.AO.06.AO_06_ANG32         ----F--------------------I-------------------------------------------D--------------PQ--------I------------------Q---S.---V...--------L--------------V-----K------------- 296
F1.AR.02.ARE933              V---P-----Y-SL-------------------------------------------------------D-----------L--A-----P------------------P----Q---S.---V...-----------------------V-----K------------- 295
F1.BR.10.10BR_RJ015          ----L-------SL-----------I-------------------------------------------D-----------I--PQ----P---I------------------Q---S.---V...-----------------------V-----K------------- 298
F1.CY.08.CY222               ----L-------S------------I-------------------------------------------D-------D---L--------P----------------------Q---S.---V...---D-------------------V-----K------------- 295
F1.ES.02.ES_X845_4           ----L--------M-----------I-------------V------------S----------------D-----------L--QQ---FP---L------------------Q---G.---V...----------V------------V-----K------------- 296
F1.RO.96.BCI_R07             ----L--------------------I-------------------------------------------D-----------L--------P----------------------Q---S.T--V...-------G---------------V------------------- 295
F1.RU.08.D88_845             ----H-------S------------I-------------------------------------------D--------------------P----------------------Q---S.---V...-----------------------V-----K------------- 301
F2.CM.02.02CM_0016BBY        ----L--------------------I-------------------------------------------D-----------L---Q----P---I------------------A---S.---V...-----------------------V------------------- 293
F2.CM.10.DEMF210CM001        ----L--------------------I-------------------------------------------D--------------------P----------------------A---S.---V...---D-------------------V-----K------------- 295
F2.CM.10.DEMF210CM007        ----L--------------------I--------------------------V----------------D---------------Q----P---I------------------A---S.---V...---D-------------------V-----K------------- 295
G.BE.96.DRCBL                ----A----------------------------------T---------------------------------D-------L--QQ--------I-D-T------A--------R---S.---V...-----------------------V------------------- 292
G.CM.10.DEMG10CM008          ----A---P----L-------------------------T-----------------------------D-----------T--QQ--------I-------------N----VA---G.----...-----------------------V-----K------------- 299
G.CN.08.GX_2084_08           ----A-------SL-----------I-------------T-----------------------------D-----------L--PQ---FP---I-D-----------N-----A---S.----...-----------------------V-----K------------- 296
G.CU.99.Cu74                 ----A---P------------------------------------------------------------D-----------L--PQ----P---I--------------------------.----...-----------------------V-----K------------- 299
G.ES.09.X2634_2              ----A----------------------------------------------------------------DA----------T--PQ----P---I----------------D--R---S.----...-----------------------V-----K------------- 298
G.GH.03.03GH175G             ----A-------P--------------------------------------------------------D-----------L--QQ--------I-D-T---------------R---S.----...-----------------------V------------------- 298
G.KE.09.DEMG09KE001          V---A-------------------A--------------V-----------------------------D-----------M--QQ--------I------------------Q---S.---V...-----------------------V------------------- 299
G.NG.09.09NG_SC62            ----A-----------K--------I-------------T-----------------------------D---D-------L--QQ---V----I------------------T----.----...-----------------------V-----K------------- 292
G.ZA.01.TV546                ----A-------P------------I-------------------------------------------D-----------I--PQ----P---I------------------R---S.----...---D--------------------------------------- 297
H.BE.93.VI991                ----A-----------------------------------------------A----------------D-----------L--------P----------------------VA---G.----...---D-------------------V-----K------------- 301
H.CF.90.056                  ----A-----------------------------------------------A----------------D--------------------P-----------------------A---G.--A-...---D-------------------V-----K------------- 299
H.GB.00.00GBAC4001           ----A---------------------------------------------------S---------X--D-----------T--------P-----------------N-----A---X.---X...---X-------X-----------V-----K------------- 299
J.CM.04.04CMU11421           ----M---P----L-----------X-------------------------------------------D-----------X-------P----A-------------------A---G.----...-----------------------V------------------- 299
J.SE.93.SE9280_7887          ----L---P----L-----------I-------------------------------I-----------D------------------------V----------------------G.----...-----------------------V------------------- 299
J.SE.94.SE9173_7022          ----L---P----L-----------I-------------------------------I-----------D-------------------V----V-------------N---------G.----...-----------------------V------------------- 299
K.CD.97.97ZR_EQTB11          ----L--------L-----------I-------------------------------------------D-----------M---Q----P---I------------------T---S.----...-----------------------V------------------- 295
K.CM.96.96CM_MP535           ----L--------L-----------I-------------T-----------------------------D---D-------L--------P----------------------A---S.---V...-----------------------V------------------- 295
U.CA.01.TV749                -M--A----TY-SLT----------I---N------------------------M-----------L--D-----------L--------P----------------------VA---G.---V...---D--R----------------V------------------- 295
U.CA.99.TV721                -M--A----TY-SLT----------I---N------------------S-----M-----------I--D-----------M-------V------------------------V----S.----...---D--R----------------V------------------- 295
U.CD.83.83CD003_Z3           ----M--------------------I-------------------------------------------D-------D---L--------P---------------------------S.----...---D-------------------V-----K------------- 299
U.CD.90.90CD121E12           ----M-------P------------I-------------------------------------------D-------D---L--------P---------------------------S.----...---D-------------------V------------------- 296
U.CY.05.CY090                ----V--------L-----------I---S---------V-----------------------------D--------------------P----------------------A----.---V...-------K--V------------V-----K------------- 295
U.CY.08.CY223                ----A-------SM---------------G----------------------M---I----------------D-------L-------V------------------------V----S.----...---D-------------------VG----K------------- 300
U.ES.10.DEURF10DZ001         ----A---P----------------I-------------T------------M---I------------D---------------Q----P---I------------------K----.----...----------V------------V------------------- 296
U.GR.99.99GR303              ----A----I--PM-----------I-------------T---X--------M---II-----------D-----------L--TQ---FP---------------------------S.----...----..-----------------V------------------- 297
U.NL.95.U_NL_95_H10986_D1    --R-L--------L-------------------------------------------------------D--------------------P----------------------A---S.---V...-----------------------V-----K------------- 296
01_AE.AF.07.569M             ----A-------PL---------------G-N----S---------------M---I------------D-----------T--------P----------------------------.T---...-----------------------V------------------- 299
01_AE.CF.90.90CF11697        ----A-------PL---------------G-N--------------------M---I------------D-----------L--------P------------X----N---------S.----...---D-------------------V------------------- 299
01_AE.CN.10.YNFL03           -I--A-------PL-----------I---G-N--------------------M---I------------------------L--------P--------------------------.----...---D-------------------V------------------- 299
01_AE.HK.04.HK001            ----A----I--SL---------------G-N--------------A-----M---I----------------------------------P---------------------------S.--X-...---D-------------------V------------------- 299
01_AE.IR.10.10IR.THR48F      ----A-------SL---------------G-N-----------D--------M---I----------------------------Q----P---V------------------A----.----...---D-------------------I------------------- 299
01_AE.JP.x.JRC77AE           ----A-------SL-----------I---G-N--------------------M---I----------------D---D---L--AQ---FP---------------------------S.----...-----------------------VG----K------------- 299
01_AE.TH.09.AA090a_WG11      ----A-------PL---------------G-N--------------------M---I------------------------L--------P-----------------N-----A---G.--T-...---D-------------------V------------------- 297
01_AE.TH.90.CM240            ----A----A--PL---------------G-N--------------------M---I------------------P--------------P--------------------------.----...---D-------------------V------------------- 299
01_AE.US.05.306163_FL        V---A--H----PL---------------G-N--------------------M---I------------------------I--------Q---L--------------I--------.----...-----------------------V------------------- 295
01_AE.VN.98.98VNND15         ----A-------PL---------------G-N--------------------M---I----------------------------------P-----------------N--------.----...---D-------------------V------------------- 299
```

```
                                                                                               CyPA binding                                                      major homology region
B.FR.83.HXB2                    IVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTN.NPPI...PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDR 299
02_AG.CM.10.DE00210CM013        ----A-------SM-----------I-------------T------------M---I------------D--------------------P---------------------------S.---V...----------V------------V------------------- 301
02_AG.CY.09.CY256               ----A--------V----------AI--------------------------M---I------------D-------D---T--------P---------------------------S.----...-----------------------V------------------- 296
02_AG.ES.06.P1423               ----A----T--NL-----------I---G----------------------M---I------------D-----------L---Q----P----------------------M----S.----...---D-------------------V-----K------------- 296
02_AG.GW.05.CC_0048             ----A----I--SM----------AI-------------T------------M---I------------D---D-------T--------P---------------------------S.----...----------V------------V------------------- 301
02_AG.KR.12.12MHR9              ----A----T--PM-----------I--------------------------M---I------------D-----------T--------P---------------------------S.--A-...-----------------------V------------------- 296
02_AG.LR.x.POC44951             ----A----T--S-----------AI-------------T------------M---I------------D---------------Q----P---------------------------S.----...----------V------------V------------------- 296
02_AG.NG.09.09NG_SC61           ----A-------PM-----------I---G---------T------------M---I------------D-----------T---Q----P---L-----------------------G.----...----------V------------V------------------- 294
02_AG.NG.x.IBNG                 ----AK---T--SM-----------I---G----------------------M---I------------D--------------------P---------------------------S.----...----------V------------V------------------- 296
02_AG.SE.94.SE7812              ----A----T--S---------------------------------------M---I------------D--------------------P---------------------------S.----...----------V------------V------------------- 296
02_AG.SN.98.98SE_MP1211         ----A-------S------------I--------------------------M---I------------D-------D---T--------P---------------------------S.----...---D-------------------V------------------- 295
03_AB.RU.97.KAL153_2            ----A----T--SM-----------I--------------------------M---I------------D-----------L--AQ---FP--------------SS-----------S.----...---D-------------------V------------------- 299
04_cpx.CY.94.94CY032_3          ----A-------S------------I--------------------------M---I------------D-------D---T--------P---------------------------S.---V...-------------------T---I------------------- 299
05_DF.BE.x.VI1310               -------------L-----------I-------------------------------------------------------L---Q---V------D-----------------A----.----...-----------------------V------------------- 300
06_cpx.AU.96.BFP90              ----A--------M-----------I-D-----------T------------M---I------------D--------------------P---I-----------------------S.----...-----------------------VG----K------------- 298
07_BC.CN.98.98CN009             ----L--------L----------------------------------------------------I--D-----------L--------------------------N-----A---S.---V...---D-------------------------K------------- 296
08_BC.CN.97.97CNGX_6F           ----P-------PL-------------------------T-----------------------------D-----------L-------V-----------------------------.----...-----------------------------K------------- 294
09_cpx.GH.96.96GH2911           ----A--------------------I-------------T------------M---I------------D-------------------NP------------------V----A---G.----...----------V------------V------------------- 299
10_CD.TZ.96.96TZ_BF061          ----L-------PLT------------------------------------------------------------------L---Q---V----I-------------------R---S.----...-----------------------V------------------- 299
11_cpx.CM.95.95CM_1816          ----A----I---L-----------I--------------------------M---I------------D-----------M---Q----P---L--------------P--------G.---V...------R----------------V------------------- 299
12_BF.AR.99.ARMA159             ----L-------PL-------------------------------------------------------D-------D---L--------P-----------------------Q---S.---V...------------------------------------------- 295
13_cpx.CM.96.96CM_1849          ----A-------S--------------------------------------------------------D-----------M--QR---FP---I-------------------T----.----...-----------------------V------------------- 299
14_BG.ES.05.X1870               ----A-------P---------------------------------------L----------------DA----------M--QQ----P---I-------------------T----.----...-----------------------V------------------- 298
15_01B.TH.99.99TH_MU2079        ----A-------PV---------------G-N--------------------M---I--------------------K------------P----------------------V----S.--A-...---D-------------------V------------------- 299
16_A2D.KR.97.97KR004            ----A----TY-NL-------------------------------------------------------D-----------L--------P---------------------------S.---V...-----------------------V------------------- 297
17_BF.AR.99.ARMA038             ----L-------T--------------------------------------------------------D-----------L---Q----P---I----------A--N-----Q---S.---V...-----------------------V-----K------------- 301
18_cpx.CU.99.CU76               ----A-------S-----------AI--------------------------M---I------------D-----------I---Q--------I-D-----------------R----.----...---D-------------------V------------------- 299
19_cpx.CU.99.CU7                --HKL--------L-----------I-----N-------------------------------------------------L--------------------------------A---S.----...-----------------------V------------------- 299
20_BG.CU.99.Cu103               ----A---PI-----------------------------------------------------------D-----------T--QQ----P---I----------A--------T----.----...-----------------------V-----K------------- 299
21_A2D.KE.99.KER2003            ----A---W----------------------N----------------S---S----------------------------L---Q----------------------------A---G.----...-----------------------V------------------- 297
22_01A1.CM.01.01CM_0001BBY      ----A----I---L---------------G-N-------------T--T---M---II-----------D-----------M---Q---YP-----------------------A---S.---V...-----------------------V------------------- 297
23_BG.CU.03.CB118               ----A---P------------------------------------------------------------D-----------L--QQ--------I-------------------R----.----...-----------------------V-----K------------- 299
24_BG.ES.08.X2456_2             ----A---PI-----------------------------------------------------------D-----------M--QQ----P---I-------------------T----.----...-----------------------V-----K------------- 299
25_cpx.CM.02.1918LE             ----A---PI-----------------------------------------------------------D-----------L---Q----P---I-------------N-----R----.----...-----------------------V------------------- 298
26_AU.CD.02.02CD_MBTB047        ----A--------L-------------------------V--------------M------------I-D-----------L---Q--------I-------------------A---S.----...-----------------------V------------------- 296
27_cpx.FR.04.04CD_FR_KZS        --H-A--------L-----------I-------------------------------------------D-----------L--QQ----------------------------A----.----...-----------------------V------------------- 296
28_BF.BR.99.BREPM12609          ----L--------L-----------I-------------------------------------------D-----------L--------P-----------------N-----A---S.---V...-----------------------V------------------- 299
29_BF.BR.01.BREPM16704          ----L-------P------------I-------------------------------------------D-----------L--------P-----------------------Q---S.---V...-----------------------V------------------- 310
31_BC.BR.04.04BR142             ----L-------P-A------------------------T-----------------------------D-----------L-------V---------------S--------A---S.---V...---D-------------------------K------------- 298
32_06A1.EE.01.EE0369            ----A----T--S------------I--------------------------M---I------------D-----------LY-TQ---FQ---------------------------S.----...---D-------------------V------------------- 296
33_01B.ID.07.JKT189_C           ----A-------PL---------------G-N--------------------M---I--------------------D---T--------P---------------------------S.----...---D-------------------V------------------- 298
34_01B.TH.99.OUR2478P           ----A-------PV---------------G-N--------------------M---I--------------------D-------Q----P---------------------------S.----...---D-------------------V------------------- 297
35_AD.AF.07.169H                ----A----I--SL-----------I----------------------G---M---V------------D-----------L---Q-----------------------P----A----.----...---D--R----------------V-----K------------- 298
36_cpx.CM.00.00CMNYU830         ----A--------L---------------G-N--------------------M---I------------D-------------------HP---I---------E---------A----.----..S-----------------------V------------------- 298
37_cpx.CM.00.00CMNYU926         ----A----I--SL-----------I-------------T---D----H---A---V------------D-------D-----------HP---L------------------VA---S.----...---D-------------------V------------------- 295
38_BF1.UY.03.UY03_3389          ----L-------PL-----------I-------------------------------------------D-------D---L--------P-----------------------Q---S.---V...-----------------------V----LK------------- 295
39_BF.BR.04.04BRRJ179           ----L-------SL-----------I-------------------------------------------------------L-------------------------------VA---S.---V...-----------------------V------------------- 299
40_BF.BR.05.05BRRJ055           ----L--------------------I----------------A--------------------------D-----------L--------P-----------------------Q---S.---V...-----------------------V------------------- 300
42_BF.LU.06.luBF_18_06          ----L-------------------------------------A--------------------------D-----------L--AQ----P---I-------------------Q---S.----...-----------------------V------------------- 298
43_02G.SA.03.J11223             ----A-------SL-----------I-------------T-----------------------------D-----------L--PQ----P---I-------------------T---S.----...----L------------------V-----K------------- 296
44_BF.CL.00.CH80                ----L--------L-------------------------T-----------------------------D-------------------AP---------------------------S.---V...------------------------------------------- 301
45_cpx.FR.04.04FR_AUK           ----A----I---------------I--------------------------M---I------------D-------------------AP----------------------V----G.---V...------R----------------V------------------- 299
46_BF.BR.07.07BR_FPS625         ----L--------L-------------------------------------------------P-----GG--------Y-F-T------P-----------------------Q---S.---V...-----------------------V------------------- 295
47_BF.ES.08.P1942               ----A--H-I--P--------------------------------C-----------------------D-----------L--------P-----------------------Q---S.---V...---D-------------------V-----K------------- 298
48_01B.MY.07.07MYKT021          ----A-------PV---------------G-N--------------------M---I--------------------D---T--------P-----------------N---------G.----...---D-------------------V------------------- 299
49_cpx.GM.03.N26677             ----T-------PM-----------I--------------------------M---I------------D--------------------P-------------------------I-G.----...-----------------------V------------------- 299
50_A1D.GB.10.12792              ----L--------------------I-------------------------------------------D-----------L------#------------------------------.----...-----------------------V-----K------------- 298
51_01B.SG.11.11SG_HM021         ----L--------------------I-------------------------------------------------------L-------T------------------N---------S.----...-----------------------V------------------- 297
52_01B.MY.03.03MYKL018_1        ----A-------PV---------------G-N--------------------M---I---------------------------------X---------------------------X.----...---D-------------------V------------------- 299
53_01B.MY.11.11FIR164           -M--A-------PL-----------X---G-N--------------------M---I------------D-------D---T--------P----------------------------.----...---D-------------------V------------------- 299
54_01B.MY.09.09MYSB023          ----L-------PL-----------------N----------A------------------------------D-------L-------V----------------------------S.----...-----------------------V------------------- 299
55_01B.CN.10.HNCS102056         ----A-------PL---------------G-N--------------------M---I------------------------A--------P----------------------------.----...---D-------------------V------------------- 299
56_cpx.FR.10.URF5_patient_A     ----A-------SM-----------I-------------T------------M---I------------D-----------T--------------------------N----------.----...----------V------------V------------------- 299
57_BC.CN.09.09YNLX19sg          ----L-------P----------------------------------------------------------------D---L-------V-----------------------------.----...-----------------------S-----------S------- 296
58_01B.MY.09.09MYPR37           ----A-------SL---------------G-N--------------A-----M---I------------------------I--------P---------------------------S.----...---D-------------------V------------------- 298
59_01B.CN.09.09LNA423           ----A----T--SL---------------G-N--------------------M---I------------------------L--------P-----------------N---------S.----...---D-------------------V-----K------------- 295
60_BC.IT.11.BAV499              ------------P------------I-------------T-----------------------------D-----------L---Q---VQ---L-------------------A---A.---V...---DL------------------V-----K------------- 295
61_BC.CN.10.JL100010            ----L-------P--------------------------T-----------------------------D-------D---L--PQ---L---LI-----------S-------T---G.---V...---D-------------------S-----K------------- 296
62_BC.CN.10.YNFL13              ----L--------L-----------I-------------T---------------------------I-D------------------------I-------------------A---G.---V...---D-------------------VG----K------------- 292
63_02A1.RU.10.10RU6637          ----A----T--SM-----------I-------------T------------M---I------------D-----------T-------NP----------------------------.----...----------VM-----------V-----K------------- 296
64_BC.CN.09.YNFL31              ----L-------P--------------------------------------------------------D-----------L-------V------------------------A---G.---V...---D-------------------------K------------- 296
65_cpx.CN.10.YNFL01             ----L-------P--------------------------------------------------------D-----------L--------------------------------A---G.---V...---D-------M-----------------K------------- 299
67_01B.CN.11.2011.ANHUI.HF115   ----A--------V---------------G-N--------------------M---I--------------------D---X---Q----P---I-------------------A----.----...---D-------------------V------------------- 299
68_01B.CN.11.2011.ANHUI.WH73    ----A-------PV---------------G-N--------------------M---I--------------------D---T---Q----P---I-------------N-----A----.----...---D-------------------V------------------- 299
70_BF1.BR.10.10BR_PE004         ----M---I----------------I-K---------------D----S---A---I------------D-----------T---Q----PA--IK--------R---N-----H---S.----...-----------------------V-------R----------- 302
71_BF1.BR.10.10BR_PE008         ----H-------P--------------------------T--------S--------------------D-----------L--------P-----------------------Q---G.---V...---D-------------------VG----K------------- 299
72_BF1.BR.10.10BR_MG002         ----L-------P-T---------------------------A--------------------------D---------------Q----P---I-------------------Q---G.----...-----------------------V------------------- 299
O.BE.87.ANT70                   --S-A-------------------A------N--I----M------ISY-I-----AI----G-L-V---V-----V----T--PPV--LP---I---T----------Q----H-T-RP-Q--...---D--RK--V-----M-K----V-----K------------- 298
O.CM.98.98CMA104                V---A--------L----------A------N--I----M-------SY-I-----AI----G-L-V---V----------T--IAV--LP---I---T--------------VH-T-RA-N--...---D--RK--V-----M-K----V-----K------------- 298
O.CM.98.98CMABB141              V-P-A----I--------------A------N--I----M------I-Y-I-----AI----G-L-V---V----------T--IPV--LP---I-D-T----------Q---VH-I-RA-N-L...---D--RK--V-----M-K----V------------------- 298
O.CM.98.98CMABB212              V-A-A--------X----------A------N--I----M------I-Y-I-----A-----G-L-V---V-----G----A--PAM--LP---L-D-T----------Q----A-T-RA-N--...---D--RK--V-----M-K----V-----K------------- 298
O.CM.98.98CMU5337               V-P-A--H-----L----------A------N--I----M------I-Y-I-----AI----G-L-V---V--D---D---T--PAV--LP---I---T----------Q---VH-I-RA-N--...---D--RK--V-----M-K----V-----K------------- 298
O.CM.99.99CMU4122               -IP-A-------P-----------A------N--I----M------I-Y-I-----AI----G-L-V---V-----S----T--PAM--LP-------T----------Q---VH-T-RA-N--...---D--RK--V-----M-K----V------------------- 298
O.FR.92.VAU                     V-A-A--------L----------A--K---N--I----M--------Y-V-----AI----G-L-V-E-V--D---D--KT--PVV--LP---I---T---------------H-T-RA-Q--...---D--RK--VI----M-K----V-----K-R-E--------Q 297
O.GA.11.11Gab6352               -IP-A-------------------A------N--I----M------I-Y-I-----AI----G-L-V---V-----SD---T--PA---YPA--I-D-T----------Q----H-T-RA-N--...---D--RK--V-----M-K----V-----K------------- 297
O.SN.99.99SE_MP1299             --T-A-------SL----------A------N--I----M------I-Y-T-----AI----G-L-V---V----------T--PA---LPV--I---T----------Q---VH-I-RP-Q--...---D--RK--V-----V-K----V------------------- 298
O.US.10.LTNP                    V---A----T-VPL----------A------N--I----M------V-Y-I-----AI----G-L-V---V------D---T--PPV--LP---I---T----------Q---VH-T-RP-Q--...---D--RK--V-----M-K----V------------------- 298

N.CM.02.DJO0131                 L---A----I--PLT----------I-----N-------M--------A---------------------V----------T--XPX--LP---L---------------X--VT---S.---V...-------K--V----R-------V---E-K------------- 304
N.CM.02.SJGddd                  L--TA----I--PLT----------I-----N-------M--------S---------------------V----------T---PV--LP---I---------------A--VA---A.---V...---D--R---V----R-------V---E-K------------- 302
N.CM.04.04CM_1015_04            L---A------HPLT----------I-----N-------M--------S---S-----------------V----------T---PV--LP---L---------------P---T---S.----...---D--R---V----R-------V---E-K------------- 301
N.CM.06.U14296                  L---A--H----PLT----------I-----N-------M--------G---------------------V----------T---PV--LP---L---------------A--VA---X.---V...---D--R---V----R--K--X-V---E-K------------- 301
N.CM.06.U14842                  LI--A---L---PLT----------I--------I----M--------S--------------------DV----------T---PV---P---I-------------NVA--VA---A.---V...-X-D--R---V----R-------V-----K------------- 301
N.FR.11.N1_FR_2011              L---A----I--PLT----------I-----N-------M--------A---------------------V------D---T---PV--MP-------------------A---A---A.---V...------R---V----R-------V---E-K------------- 302

P.CM.06.U14788                  ----A----T-MPL----------A------N--I----M------I-D-I-----A-----G-L-V---V----------T---PV--LP---L---T-G--------K----T---R.-N-V...---D--RK--V-----V-K--C-V-----K------------- 292
P.FR.09.RBF168                  ----A----T-LPL----------A------N--I----M-----SI-D-I-----A-----G-L-V---V----------T---PV--LP---L-D-T-G-------NK----T-T-RN-N--...---D--RK--V-----V-K----I-----K------------- 295

CPZ.CD.06.BF1167                --FDAAQ-PR--P-----------T----N-K----------A---I-H-------AI-D--G---V---V------D---L--T----LV--TL---T----------IV---Q-V-T.PQNQGGV---D------------L-K----V------------------- 323
CPZ.CM.05.SIVcpzMT145           V-R-A---LT-NP-----------A------N-------M-----------------I--------I---V--D-------T--T-----------------------S-E--V----A.----...---D--R---V-------K----V-----K------------- 289
CPZ.GA.88.GAB1                  L---A-----------------------------------------L---V-----A-----G---V---V----------L--T---------L---------------------T-A.----...---DV-R--V------V----C-V------------------- 307
CPZ.TZ.06.SIVcpzTAN13           --TDL--TAR--P-----------T----G------------------Y---S---AI-E----------V----------L--A----QQA-ML---T-A--------T----Q---T.P-AQGGV---D------------L----N-V-----K-----A------- 326
CPZ.US.85.US_Marilyn            VI--A--------M----------A------N-------M----------V-----AI----G---V---A------D*--T--L----------------------------V----A.----...---D--R--VV-----V-K--C-VG----K------------- 300

GOR.CM.07.SIVgor2139_287        V-A-A-------PL----------A------N--I----M-----SI-Y-I-----AI----G-L-V---V----------M--AVV--LPV--L---S-G--------Q---VA-I-RN-N-V...---D--RK--V-----V-K--C-V-----K------------- 291
GOR.CM.07.SIVgorCP2135con       V-A-A-------PL----------A------N--I----M-----SI-Y-I-----AI----G-L-V---V----------L--AVV--LPV--L---S-G--------Q---VA-I-RN-N-V...---D--RK--V-----V-K--C-V-----K------------- 291
```

```
                            major homology region                                 p24 end p2 start       p2 end p7 start    Zn motif           Zn motif          p7 end p1 start
B.FR.83.HXB2                FYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVT......N..SAT.I.MM......Q..RGNFR.NQRK..IVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCT....E.RQANFLGKIWPS.YK 442
A1.AU.03.PS1044_Day0        -F-A------T----S----------------S------SG----------------S------------AQ......Q...TS.-.--......-..-----GG-KR...I----------L--------------------------....-.------------.S- 441
A1.CH.03.HIV_CH_BID_V3538   -F--------T-D-------------------S--R---AG------------------------------Q......H...TN.-.--......-..----K.G---...I----------L--------------------------....-.-------R----.N- 440
A1.CY.08.CY236              -F-V------T--------D-------------------SG-S--------------A------------AQ......H...-N.-.--......-..-----.G-K-...I----------L--------------------------....-.------------.S- 445
A1.ES.05.X1608_8            -F-C--------D--T------I---------S--R---AG----------------S------------AQ......Q..HTN.-.--......-..K---K.G---...I----------L--------------------------....-.-------R----.S- 438
A1.KE.11.DEMA11KE001        -F--------T----G------I---------S--R---TG------------------------------Q......H...TN.-.--......-..----K.G-KR...I----------L----------------R---------....-.------------.S- 440
A1.RU.11.11RU6950           -F--------T-D------D---------------R---SG------------------------------Q..........-N.-.--......-..KS---.GPKR...I----------L-------------R------------....-.-------R----.S- 436
A1.RW.11.DEMA111RW002       -F-------------G---D--I---------S--R----G-S--------------S------------AQ......Q...TS.-.--......-..----K.G-KR...I----------L--------------------------....-.-------R----.N- 441
A1.SN.01.DDI579             -F-C------T-D-------------------S--R---TG------------------------------Q......H...TN.V.--......-..-----.G-KR...I----------L--------------------------....-.------------.N- 438
A1.UG.11.DEMA110UG001       -F--------T----G------I---------S--R---AG----------------S------------AQ......Q...TN.-.--......-..----K.G-KR...I----------L--------------------------....-.-------R----.S- 438
A1.ZA.04.04ZASK162B1        -F-A---G--T--------G------------Q------TG-S---V----------S-------------Q......Q...PN.T.--......-..----K.G-KR..NI----------L--------------------------....-.------------.S- 446
A2.CM.01.01CM_1445MV        -F--------T-D------D---------------R----------------------------------IH......S.TNQN.V.--......-..-----.GPKR...I----------L--------------------------....-.-----------P.N- 438
A2.CY.94.94CY017_41         -F--------T--------D-----------RS--R----G-S--------------S-----------H-Q......S.TNTN.-.--......-..-----.G-KR...I----------L--------------------------....-.------------.N- 438
B.BR.10.10BR_MG029          ------------D-------------S-------------G-------------------------------......-..TGN.-.--......-..-----.T--R..T-----------I-K-----------------------V....-.------------.H- 441
B.CA.07.502_1191_03         ----------------------------------------G-------------------------------......S..PGN.-.--......-..-----.----..T-----------I-K------------------------....-.------------.H- 444
B.CH.08.M2_0803101_NFLG8    ------------D---------------------------G----------------S------------AA.........P--.-.--......-..G-K--.---R..T-----------I-K-----------------------A....-.------------.H- 441
B.CN.10.DEMB10CN002         ----------T-------------------------------------S--------S-------------Q.........TT-.-.--......-..-----.----..T--------D--I-K------------------------....-.------------.H- 444
B.ES.10.DEMB10ES002         ------------D----------------------------------------------------------N......T...TS.V.FV......-..-----.--K-..T-----------I-K--------------R---------....-.------------.H- 440
B.FR.11.DEMB11FR001         ------------------------------------------------------------------------......-..---.-.--......-..K----.--KR..SI----------L---------R----------------....-.------------.-- 442
B.GB.05.MM45d213_GN1        --R-------T----------------------------TG-------------------------------......-...-A.-.--......-..K----.--KR.RP--------V--I-K--K----R---------------N....-.------------.H- 451
B.HK.06.HK003               --R------------------------------------------------------A--------------......-..---.V.--......-..-----.----..------------I-K--------------Q---------....-.------------.H- 442
B.HT.05.05HT_129389         ---------------G------------------------G-------------------------------......-..TNA.-.-I......-..GSH--.----..F-----------I-K------#-----------------....-.------------.-- 439
B.JP.12.DEMB12JP001         ------------------------------------------------------------------------.........PT-.-.--......-..-----.--KR..TI------------K------------------------....-.------------.H- 442
B.KR.07.HP_18_07JHS10_3909  ----------------------------------------G------------------------------M.........PT-.-.--......-..-----.----..T-----------I-K------------------------....-.------------.H- 442
B.PE.07.502_0525_wg5        ------------D---------------------------------------------------------A-......S.SA-A.-.--......-..K----.----..-I----------L-K-------R----------------....-.------------.P- 443
B.RU.11.11RU21n             ------------------------------------------------------------------------......S..---.-.M-......-..---YK.----..T-----------I-K------------------------....-.------------.H- 447
B.TH.08.MERLBDTRC10         --R---------D-------------S------------------------------S------------A-........SA-A.-.-QRNNFK.-..----K.G--R..V-----------I-K--------------Q---------....-.------------.H- 449
B.US.11.ES38                ------------D------------------R-----M--G-----------------------------M-......-..-T-.V.--......-..K----.----..TI----------L--------------------------....-.------------.H- 441
C.AR.01.ARG4006             -F--------T------------------------R----G-S----------------------------N......-...TN.-.--......-..----K.GPKR..TI-----R---L--------------------------....-.------------.N- 437
C.BR.07.DEMC07BR003         -F--------T--------D------------S--R----G------------------------------N......-...TN.-.--......-..-S--K.GPKR..TI----------L----------------Q---------....-.------------.H- 436
C.BW.00.00BW5031_1          -F--------T-D------D------G-----N--R----G-----------------------------AN......-...IN.-.--......-..-N--K.GPKR..------------I-K------------------------....-.------------.N- 438
C.CN.10.YNFL19              -F--------T-D------D---------------R----G-SI--------------------------AN......S....N.-.L-......-..-S--K.GPKR..------------I-K------------------------....-.------------.H- 437
C.CY.09.CY260               -F--------T------------------------R----G----------------S------------AN......H....-.-.--......-..KS--K.GSKR..------------I---------R----------------...V-.------------.H- 443
C.ES.08.X2363_2             -F--------T----G-------------------R----G----------------S------------TN......-....A.-.--......-..-S--K.GPKR..------------I-K-----------R-----------S....-.------------.H- 438
C.ET.02.02ET_288            -F--------T-D----------------------R----G-S---------------------------AN......Q...-N.-.--......-..K---K.GP-R..TI----------L--------------------------....-.--------L---.H- 437
C.IN.09.T125_2139           -F--------T-D------D---------------R----G----------------S------------TN......-....A.-.--......-..-S--K.GA-R..------------I-K------------------------....-.------------.H- 437
C.KE.00.KER2010             -F--------T--------D---------------R----G--------S--------------------AN......-...-N.-.-I......-..-S--K.SPKR..------------I----K---------------------....-.------------.H- 436
C.MW.09.703010256_CH256.w96 -F--------T-D------D---------------R----G----------------S------------TN......-...TN.-.L-......-..-S--K.GT-R..------------I-K------------------------....-.------------.H- 437
C.TZ.08.707010457_CH457.w8  -F--------T--------D--I-----------------G----------------S------------AN......-...-N.-.--......-..-S--K.GTKR..-I----------I--------------------------....-.------------.H- 437
C.YE.02.02YE511             -F--------T-D------D---------------R----G-S----------------------------N......-...PN.-.--......-..-N--K.GPKR..------------I-K-----------------------N....-.------------.H- 436
C.ZA.08.705010534_CH534.w12 -F-------ST---------------------G--R----H--I-------------S-------------N......-...TN.-.--......-..KS--K.GPKR..-I----------I-K-------R----------------....-.------------.H- 435
C.ZA.10.DEMC10ZA001         -F--------T------------------------RG---G-----------------------------AN......S....N.-.--......-..KS--K.GP-R..------------I-K------------------------....-.-------R----.H- 438
C.ZM.11.DEMC11ZM006         -F-C-----CT-D------D--------------------G-S--------------S------------AN......H...-N.-.--......-..KS--K.GAKR..---------V--I-K--------------Q---------....-.--------L---.H- 439
D.CM.10.DEMD10CM009         ------------D---------------------------G----------------S---------I--AN......-..---.V.--......-..----K.GSKR..T-----------I---------R----------------....-.-------R----.N- 440
D.CY.06.CY163               ------------D---------------------------G----------------------------KAA......T..NTA.-.--......-..----K.SPKR..------------I-K------------------------....-.----------S-.H- 442
D.KE.11.DEMD11KE003         ------------D--R----------------N----M--G----------------S------------A-......S.VN-A.-.--......-..-N--K.GPKR..------------V-K-------R----------------....-.------------.Q- 443
D.KR.04.04KBH8              ---------------------------------------AG-S--------------S------------A-......S.SN-A.-.--......-..----K.GP--..------------I-K-------R----------------....-.------------.HN 442
D.SN.90.SE365               -------------------D------------N-------G----------------S------------A-......S.SA-A.-.--......-..----K.G---..------------I----K-----------Q---------....G.------------.S- 443
D.TZ.01.A280                ------------D--------------------------------------------S------------A-......-.VN-A.-.--......-..----K.GP--..-I----------I-K------------------------....-.------------.H- 437
D.UG.10.DEMD10UG004         ------------D--------------------------------------------S------------A-......-.VN-A.-.--......-..--KYT.GP--..SI----------L----K---------------------....-.-------RF---.Q- 443
D.UG.11.DEMD11UG003         ---------G--D------------------------M--.----------------S------------A-......-.AN-A.-.--......-..----K.GPKR..-I----------I-K-------R---R------------....-.-------N----.R- 441
D.YE.02.02YE516             ----------------------------------------G-S--------------S------------IA......S.NN-A.-.--......-..----K.-P--..------------I-K------------------------....-.-------R----.S- 441
D.ZA.90.R1                  ------------D---------------------------Q-----------------------------A-......-.LATA.V.--......-..----K.G---..-I----------I-K--------------R---------....-.------------.H- 443
F1.AO.06.AO_06_ANG32        -F-C------T----G---D-------------------SGGSI---------------------------Q......-...-S.-.--......-..KS--K.G---..TI----------L----------------R---------....-.-------R----.H- 438
F1.AR.02.ARE933             -F-A------T-D------D----------------G--SG----------------S------------A-......S...T-.V.--......-..KS--K.G-KR..---------V--I-K--------------R---------....-.------------.N- 437
F1.BR.10.10BR_RJ015         -F--------T----G---D----------------------S-----------------------------......-...P-.V.--......-..KS--K.G--R..P-----------I----------------R---------....-.------------.H- 440
F1.CY.08.CY222              -F--------T----G---D--------------------G---------------------I-------A.......-...TA.-.--......-..KS--K.G---..T-----------I----------------R---------....-.------------.N- 436
F1.ES.02.ES_X845_4          -F--------T--------D--------------------G------------------------------Q......-...TS.-.--......-..K---K.G---..-I----------L----------------R---------....-.-------R----.H- 438
F1.RO.96.BCI_R07            -F--------T----G---D-----H--------------G----------------S----I-------T-......S...-A.-.--......-..KS--K.G---..-I----------I-K------------------------....-.------------.H- 437
F1.RU.08.D88_845            -F-------CT----G---D-----------R--------G----------------S----I-------A-......S...TA.-.--......-..KS--K.G-K-..------------I----------------R---------....-.-------R----.S- 443
F2.CM.02.02CM_0016BBY       -F--------T----G---D--------------------G---------------------I---------......A...TS.V.L-......-..KS--K.G-KR..------------I-K-------R----------------....-.------------.N- 435
F2.CM.10.DEMF210CM001       -F-I------T--------D-------------------------------------S----I-------A-......AVTA-S.M.--......-..KS--K.G--R..TI----------L---------R------Q---------....-.------------.N- 440
F2.CM.10.DEMF210CM007       -F--------T----G---D--------------------G-------------------------------......-...T-.-.--......-..KS--K.G--R..---------V--I-K------------------------....-.--------L--P.S. 436
G.BE.96.DRCBL               -F--------T----S---D--I---------I---G--QG----------------S------------AS......G.AA-A.-.--......-..KS--K.GP-R..TI----------L-----------------------E--....-.------------.N- 436
G.CM.10.DEMG10CM008         -F--------T-D------D---------------R----G-------------------------------......-.TGTA.-.--......-..----K.GPKR..NI----------L--------------------------....-.------------.S- 442
G.CN.08.GX_2084_08          -F--------T-D----------------------RG--TG----------------S-----------NAS......G..A-A.-.--......-..KS--K.GP-R..XI----------L----------------R---------....-.------------.N- 439
G.CU.99.Cu74                -F--------T----G---D------------A--R----G----------------S------------AS......G.AG-A.V.L-......-..-S--K.GP-R..NI----------L--------------------------....-.--------L---.N- 443
G.ES.09.X2634_2             -F--------T--------D---------------R----G--------S--------------------A-......G...-A.-.--......-..KS--K.GPKR..MI-------R--L--------------------------....-.-------R----.N- 440
G.GH.03.03GH175G            -F-A------T----H------I-------------S--TG------------------------------S......G..ATS.-.--......-..NS---.GPKR..NI----------L-----------------------E--....-.------------.N- 441
G.KE.09.DEMG09KE001         -F-V--------D--S---D-------------------QG----------------S------------AS......L...PS.-.--......-..KS--K.GP-R..TI----------L----------------R---------....-.------------.N- 441
G.NG.09.09NG_SC62           -F--------T----G---D---------------R----G----------------S------------A-......-.A--A.-.--......-..KS--K.GP--..NI----------L----------------R---------....-.-----------P.HS 436
G.ZA.01.TV546               -F-C------T--------D----------Y--------QG----------------S------------AS......G..ATA.-.--......-..KS--K.GP-R..MI----------L--------------------------....-.------------.H- 440
H.BE.93.VI991               -FRV------T-D------D-----------R-------RG-SI-------------S--------------......-.A--A.-.--......-..K---K.GP-R..T---S-------I----------------Q---------....G.------------.S- 445
H.CF.90.056                 -F--------T-D----------------------R---QG-SI-------------S--------------......-.TNTA.-.--......-..K---K.G---..F-----------I----------------R---------....-.------------.S- 443
H.GB.00.00GBAC4001          -F--------T-D------D-------------------QG-SI--------------------------M-......-.MNTA.-.-V......-..K---K.G--R..T-----------I----------------R---------....-.------------.S- 443
J.CM.04.04CMU11421          -F--------T--------D--I--------R--------G-----------------------------M-......-..T-N.-.--......-..X---K.G-KR..MI-------P--L----------------Q---------....-.------------.N- 442
J.SE.93.SE9280_7887         -F-A------T-D------D-------------------SG-------------------------------......-...TN.-.--......-..-----.DHKR..---------Q--I-K------------------------....-.------------.S- 441
J.SE.94.SE9173_7022         -F-A------T-D------D-------------------SG-------------------------------......-...TN.-.--......-..-----.DHKR..---------Q--I-K------------------------....-.------------.S- 441
K.CD.97.97ZR_EQTB11         -FRV------T--------------------R-------SG-------------------------------......-...SA.V.--......-..----K.G--R..-I----------L-------------------------S....-.--------F--L.N- 437
K.CM.96.96CM_MP535          -F--------T--------D--------------------G-S--------------S----I---------......-...PV.V.--......-..K---K.GH--..------------I--------------------------....-.------------.H- 437
U.CA.01.TV749               -F----------D------D------------S------TG----------------S----I-------AQ......-...TN.-.--......-..-----.G-KR...I----------L----K---R-----------------....-.-------R----.H- 436
U.CA.99.TV721               -F----------D------D-----------RS------AG----------------S----I-------AQ......-...TN.-.--......-..-----.G-K-...I----------L--------------------------....-.-------R----.H- 436
U.CD.83.83CD003_Z3          -F-V-----S---------D--I----------------TG----------------S------------T-......-...TA.-.--......-..----K.GP-R..------------I-K--------------R---------....-.------------.N- 440
U.CD.90.90CD121E12          -F-------S---------D--T----------------TG-----------------------------M-......-...T-.V.--......-..----K.G---..------------I-K--------------R---------....-.------------.N- 438
U.CY.05.CY090               -F--------T--------D------------H-------G----------------T----I------KAA......T..--N.-.-I......-..----K.GP--..M-------------K-------R----------------....-.------------.QS 438
U.CY.08.CY223               -F--------T--------D---------------R----G----------------S------------AG......G..A-A.-.--......-..K---K.GP-R..NI----------L-K--------------R---------....-.------------.N- 443
U.ES.10.DEURF10DZ001        -F--------T------------------------R----G----------------S------------AQ......Q...HN.-.-V......-..-----.GP-T...I----------L----K----R----------------....-.------------.S- 437
U.GR.99.99GR303             -F----GK--T----------------------------SG------------------------------Q......S...PN.-.--......-..K----.G-KR...I----------L---------X---------------P....-.-------R----.S- 438
U.NL.95.U_NL_95_H10986_D1   -F-I--------D------D----------------------S-------------------I-------A-......K...TV.V.--......-..----K.GTKR..-I----------I--------------------------....-.-------RV---.N- 438
01_AE.AF.07.569M            -F--------T--------------------RS------SG----------------S------------AQ......H...--.-.--......-..----K.G-KR...I----------L--------------------------....-.------------TN- 441
01_AE.CF.90.90CF11697       -F--------T---------------------S------TG------------------------------Q......H...-N.-.--......-..----K.G-KR...I----------L---------R------Q---------....-.-----------L.N- 440
01_AE.CN.10.YNFL03          ----------T---------------------S------TG-----V----------------------H-Q......Q...-S.-.--......-..----K.GPK-...I----------L--------------------------....-.-------R----..- 439
01_AE.HK.04.HK001           ----------T---------------------S------AG----------------S----------N-AQ......Q...T-.V.--......-..---XK.G-KR...I-------R--L--------------------------....-.-----------P.N- 440
01_AE.IR.10.10IR.THR48F     ----------T---------------------S------SG----------------S-----------HTQ......Q...-N.-L--......-..----K.G-KR...I----------L--------------------------....-.------------..- 440
01_AE.JP.x.JRC77AE          --------------------------------S------TG----------------S------------AQ......H...TN.-.L-......-..----K.G-KT...I----------L---------------------L----....-.-------R----.H- 440
01_AE.TH.09.AA090a_WG11     ----------T---------------------S------TG----------------S-------------Q......P...TN.-.-I......-..G-K-N.GP-R...I----------L--------------------------....-.------------.H- 438
01_AE.TH.90.CM240           ----------T---------------------S------TG----------------S-----------HAQ......H...--.-.--......-..----K.G-KR...I-------R--L---------Q----------------....-.------------.N- 440
01_AE.US.05.306163_FL       ----------T---------------------S--RG--SG----------------S------------AQ.....AQ...VN.-.--......-..----K.G-KR...I----------L--------------------------....-.-----------P.N- 437
01_AE.VN.98.98VNND15        ----------T---------------------S------TG----------------S-----------H-Q......H...TN.V.--......-..----K.G-KR...I----------L--------------------------....-.------------..- 439
```

```
                                   major homology region                              p24 end p2 start      p2 end p7 start     Zn motif          Zn motif        p7 end p1 start
B.FR.83.HXB2                  FYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVT......N..SAT.I.MM......Q..RGNFR.NQRK..IVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCT....E.RQANFLGKIWPS.YK 442
02_AG.CM.10.DE00210CM013      -F--------T---------------------S--R----G------------------------------Q......Q...TN.-.--......-..-S---.GR-T...I----------L----K---------------------....-.------------.N- 442
02_AG.CY.09.CY256             -F--------T------------------------R----G----------------S-------------Q......Q...PN.-.--......-..-----.G-KT...I----------L----K----R----------------....-.-------R----.S- 437
02_AG.ES.06.P1423             -F--------T---------------------S------TG-X----------------------------Q......Q...-G.V.--......-..-----.G--T...I----------L----K----R----------------....-.-------RX---.S- 437
02_AG.GW.05.CC_0048           -F--------------------M------------RG---G----------------S------------AQ......Q...SN.-.--......-..-----.G--T...I----------L----K----R----------------....-.------------.N- 442
02_AG.KR.12.12MHR9            -FR-------T---------------S-----A--R----G----------------S-------------Q......Q...PN.-.--......-..-----.G--T...I----------L----K----R---------------P....-.------------.S- 437
02_AG.LR.x.POC44951           -F--------T---------------------S------SG------------------------------Q......Q...-N.-.--......-..-----.GR-T...I----------L----K---------------------....-.------------.S- 437
02_AG.NG.09.09NG_SC61         -F--------T---------------------N--R---QG--------------------K---------Q......Q...TN.-.--......-..-----.G--T...I----------L----K----R----------------....-.------------.S- 435
02_AG.NG.x.IBNG               -F--------T---------------------S--R---TG------------------------------Q......Q...-N.V.--......-..-----.G--T...I----------L----K----R----------------....-.------------.S- 437
02_AG.SE.94.SE7812            -F--------T-D-------------------S--R----G----------------S-------------Q......Q...PN.-.--......-..-----.G--P...I----------L----K---------------------....-.------------.S- 437
02_AG.SN.98.98SE_MP1211       -F--------T---------------S----R--------G-------S---------------------AQ......Q...SN.-.--......-..-----.G--T...I----------L----K----R----------------....-.------------.S- 436
03_AB.RU.97.KAL153_2          -F--------T-D----------------------R---SG--------V---------------------Q......-...-N.-.--......-..KS---.GPKR...I----------L--------------------------....-.-------R----.S- 440
04_cpx.CY.94.94CY032_3        -F-C------T---------------------S------TG----------------S------------AS......-.AA-A.-.--......-..KSK-K.G--R..TI----------L--------------------------....-.-------RM---.S- 443
05_DF.BE.x.VI1310             ----------T----------------------------QQ----------------S------------A-......-.SA--.A.--......-..----K.GP--..-I----------I-K-----G--------R---------....-.--------V---.H- 444
06_cpx.AU.96.BFP90            -F--------T--------D---------------R----G-----------------------------AS......V...GA.-.--......-..KS--K.GPKR..SI----------L--------------------------....-.------------.N- 440
07_BC.CN.98.98CN009           -F--------T-D------D---------------R----G-SI-------------S------------TN......S......-.L-......-..-S--K.GSKR..------------I--------------------------....-.------------.H- 437
08_BC.CN.97.97CNGX_6F         -F--------T-D------D---------------R---SG-S--------------S------------TN......-......-.L-......-..-S--K.GSKR..------------I-K------------------------....-.--------F---.H- 435
09_cpx.GH.96.96GH2911         -F--------T-D---*------------------R----G------------------------------Q......-...TN.-.L-......-..----K.G-KR...I---------XL--------------------------....-.------------.N- 439
10_CD.TZ.96.96TZ_BF061        ------------D---------------------------------------------S------------A-......G..GN-.-.--......-..----K.GPK-..SI------------K------RR------R----L---P....-R------------.N- 443
11_cpx.CM.95.95CM_1816        -F-C------T----Q----S-----------S--R---TG------------------------------Q......Q...TN.-.--......-..----K.G-KR...I----------L--------------------------....-.------------.S- 440
12_BF.AR.99.ARMA159           -F--------T--------D-------------------SG-----------------------------A-......S...TA.V.L-.........KSS-N.G--R..---------T--I-K--K--------------------P....-.------------.N- 437
13_cpx.CM.96.96CM_1849        -F--------T-D------D---------------R----G-----------------------------AS......SAAA-A.-.--.........KS--K.GPKR..-I-------R---L-------------------------M....-.------------.N- 444
14_BG.ES.05.X1870             -F--------T----G---D---------------R----G------------------------------S......G...-S.V.--.........K---K.G--R..NI----------L--------------------------....-.-------R----.S- 440
15_01B.TH.99.99TH_MU2079      ----------T------------A--------S------SG-S--------------S-----------H-Q......Q...TN.V.--......-..----K.G-KR...I----------L----------------E---------....-.-------RL---..- 439
16_A2D.KR.97.97KR004          -F--------T---------------------S--R----G-S--------------S-------------Q......-.TNSNIM.--.........K----.G-K-...#----------L--------------------------....-.------------.HS 440
17_BF.AR.99.ARMA038           -F-------------G---D--------------------G-Q-----------------------------..........P-.V.L-.........KS--K.GP--..N-----------I----------------R---------....-.------------.N- 443
18_cpx.CU.99.CU76             -F--------T---------------------N--R----G----------------S-----------H-Q......-...TN.-.--......-..----K.G-KR...I----------I--------------------------....-.------------.N- 440
19_cpx.CU.99.CU7              ------------D---------------------------G----------------S------------A-......-.SH--.-.--......-..K---K.GL--..V-----------I--------------------------....-.------------.H- 443
20_BG.CU.99.Cu103             -F--------T----G---D---------------R----G------------------------------S......G.AG-S.-.L-.........KS--K.GPKR..NI----------L--------------------------....-.--------L---.H- 443
21_A2D.KE.99.KER2003          -------------------D------------S-------G----------------S------------A-......-.SAV-.V.--......-..----K.GPKR..-I----------I-K--------------------R---....-.------------.S- 441
22_01A1.CM.01.01CM_0001BBY    -F--------T-D-------------------S--R---TG------------------------------Q......H...TG.-.--......-..----K.G-KR...I----------LT-------R-----------------L.TL-.------------.S- 441
23_BG.CU.03.CB118             -F--------T----G---D---------------R----G-S---------------------------IP......G.AG-S.-.--......-..-S--K.GPKR..NI-------D--L-----------------D--------....-G--------L---.H- 444
24_BG.ES.08.X2456_2           -F--------T----G---D---------------R----G-S---------------------------IS......G.AG-S.-.--......-..KS---.GP-R..NI----------L-K-----------------------S....-.--------L---.H- 443
25_cpx.CM.02.1918LE           -F--------T----G---D---------------R----G-S---------------------------AS......G.TA-A.-.--......-..----K.GP-R..NI----------L-----------------D--------....-.-------R----.S- 442
26_AU.CD.02.02CD_MBTB047      -F--------T----A---D---------------RG--QG-S----------------------------Q......-...-N.-.LV.........G---K.G--G.K-I----------L----K-----------R---------....-.-------RF--P.N- 439
27_cpx.FR.04.04CD_FR_KZS      -F--------T--------S---------------R----G----------------S------------AS......G.AVTA.-.--......-..K---K.G-KR..-I----------L----------------R---------....-.----------SP.N- 440
28_BF.BR.99.BREPM12609        -F-------CT----G---D--------------------G----------------S------------A-......-...T-.V.--......-..----K.G---..------------I----------------R---------....-.------------.H- 441
29_BF.BR.01.BREPM16704        -F-------CT----G---D--------------------G----------------S------------A-......-...T-.V.--......-..KS--K.G--R..------------I-----------------------E--....-.------------.H- 452
31_BC.BR.04.04BR142           -F--------T--------D------------S--R----G------------------------------N......-...TN.-.--......-..-S--K.GPKR..TI----------L-------------------------N...N-.-------R----.H- 441
32_06A1.EE.01.EE0369          -F--------THD----------------------R----G-----------------------------AQ......-...-N.-.--......-..KS---.GPKR...I----------L----------------Q---------....-.------------.S- 437
33_01B.ID.07.JKT189_C         ----------T---------------------S------SG----------------S-------------Q......Q...TG.-.--......-..----K.G-KR...I---------.L----------------Q---------....-.--------L---.H- 438
34_01B.TH.99.OUR2478P         ----------T---------------------S------TG----------------S------------TQ......Q...-NMM.--......-..----K.G-KR...I----------L--------------------------....-.K-I-----F---.N- 439
35_AD.AF.07.169H              -F-V------T----G----------------S------TG----------------S------------AQ......H...TN.-.--......-..----K.G-KR...I----------L--------------------------....-.------------.NN 439
36_cpx.CM.00.00CMNYU830       -F--------T-D-------------------S------TG------------------------------Q......H...-G.-.-I.........-S---.G--T...I----------L--------------------------...N-.------------.S- 440
37_cpx.CM.00.00CMNYU926       -F-V-----ST---------------------S------TG---------------------------T-AQ......Q...TS.-.--......-..-S--K.GPKR..NI----------L--------------------------....-.-------R----.S- 437
38_BF1.UY.03.UY03_3389        -F-VV----ST--------D------------S-------G--------S---------------A--------------SA.V.--.........K---K.GP-R..T-----------I-K--------------R---------....-.------------.H- 437
39_BF.BR.04.04BRRJ179         ---------------A---D--------------------G----------------S------------AN......Q...TN.-.-V.........----K..P-R..T-----------I-K--------------------L---....-.------------.N- 440
40_BF.BR.05.05BRRJ055         -F--------T--------D--------------------G----------------S----I-----N---......-...TA.V.--.........KS--K.G--R..------------I-K--------------R---------....-.-----------P.-- 442
42_BF.LU.06.luBF_18_06        -F--------T----G---D--------------------G-S---------------------------A-......G...TA.-.L-.........-S--K.GP--..------------I-K--------------R---------....-.------------.N- 440
43_02G.SA.03.J11223           -F--------T-D------D------------N--R---SG----------------S------------AS......G.AA-A.-.--.........K---K.GP-R..NI----------L--------------------------....-.------------.H- 440
44_BF.CL.00.CH80              -F--------T--------D---------------R----G-------------------------------..........TA.V.--.........KS--K.G--R..------------I--------------------------....-.------------.N- 443
45_cpx.FR.04.04FR_AUK         -F--------T------------------------R----G----------------S---KI-------AQ......S.TNLM.M.--......-..----K.G-K-...I----------L--------------------------....-.------------.SR 442
46_BF.BR.07.07BR_FPS625       -F--------T----G---D--------------------G----------------S----I---------......H...TN.V.L-.........KS--K.G--R..------------I----------------R---------....-.------------.N- 437
47_BF.ES.08.P1942             -F-V------T-D-------------S-------------Q-----------------------------A-......A..-NV.-.--.........K--Y-.G---..--------R---I-K------------------------....-.------------.S- 441
48_01B.MY.07.07MYKT021        ----------T----S----------------S------TG----------------S------------AQ......Q...GG.-.--......-..----K.G-KR...I----------L--------------------------....-.--------L---.H- 440
49_cpx.GM.03.N26677           -F--------T------------------------RG---G-S----------------------------Q......-...SN.-.--.........K---K.G---...I----------L-K-----------R------------....A.G------R----.H- 440
50_A1D.GB.10.12792            --R---------D------------------------------------------------------------Q......S...SN.-.--......-..----K.G-KR...I----------L--------------------------...N-.------------.H- 440
51_01B.SG.11.11SG_HM021       ------------D---------------------------------------------S------------A-......S..ATS.-.--......-..-----.---R..T-----------I-K------------------------....-.------------.H- 440
52_01B.MY.03.03MYKL018_1      ----------T---------------------S------SG----------------S------------AQ......Q..HXN.-.--......-..----K.GPKR...I----------L----------------Q---------....-.------------.N- 441
53_01B.MY.11.11FIR164         ----------T---------------------S------TG----------------S-----------NAQ......H....-.-.--.........----K.G--R...I----------L--------------------------....-.--------L--X.H- 440
54_01B.MY.09.09MYSB023        ------------D---------------------------------------------S----I--------------......-..PP-.-.-I.........--------GR..T-----------V--------------------------...N-.------------.H- 443
55_01B.CN.10.HNCS102056       ----------T---------------------S------TG----------------S-----------HMQ......S...TN.-.--......-..----K.G---...I-------R---L--------------------------....-.-------RL---.N- 440
56_cpx.FR.10.URF5_patient_A   -F--------T------------------------RG---G-----------------------------AQ......Q...-N.-.-I.........-----.G--S...I----------L----K----R----------------....-.------------.N- 440
57_BC.CN.09.09YNLX19sg        ------------D---------------------------------------------S------------TN......K......-.--.........-S--K.GPKR..------------I--------------------------....-.--------L---.H- 437
58_01B.MY.09.09MYPR37         ----------T---------------------S------AG----------------S-----------HAQ......Q...TG.-.--......-..----K.GNXX...I----------L-K------------------------....-.--------F--P.N- 439
59_01B.CN.09.09LNA423         ----------T---------------------S------TG----------------S------------AQ......Q...T-.-.--......-..----K.G-KR...I-------R---L--------------------------....-.------------..- 435
60_BC.IT.11.BAV499            -F--------T---------S-----S--------R----GX----------------------------LN......-...-N.-.--......-..-S--K.GPKR..TI----------L--------------------------....-.-------R----.H- 437
61_BC.CN.10.JL100010          -F--------T-D------D---------------R----G----------------S------------TN......S....A.-.L-.........-S--K.GPKR..------------I--------------------------....-.------------.H- 437
62_BC.CN.10.YNFL13            -F--------T-D------D---------------R----G-S--------------S------------TN......S....A.-.--.........-S--K.GPKR..------------I-----------------------A--....-.------------.H- 433
63_02A1.RU.10.10RU6637        -F--------T---------------------A--R----G------------------------------Q......Q...SN.-.--......-..-----.G-KT...I----------L--------------------------....-.-------RL---.NN 437
64_BC.CN.09.YNFL31            -F--------T-D------D---------------R----G-SI-------------S----------G-AN......-....V.-.L-.........----K.----..------------I-K------------------------....-.------------.H- 437
65_cpx.CN.10.YNFL01           -F--------T-D------D---------------R----G-SI-------------S------------AN.............-.--.........----K.GPKR..------------I-K------------------------....-.------------.H- 440
67_01B.CN.11.2011.ANHUI.HF115 ----------T---------------------S------TG----------------S-----------X-Q......H...TN.-.--.........K---K.GPKR...I----------L--------------------------....-.------------.N- 440
68_01B.CN.11.2011.ANHUI.WH73  ----------T---------------------S------TG----------------S-----------X-Q......Q...-S.V.--.........K---K.GPKR...I----------L--------------------------....-.------------.N- 440
70_BF1.BR.10.10BR_PE004       -F-A-K----T----G---D---------------R----E--I-K------------------------AN.............-.--.........KS--K.G-GR..-----------QR-I-K--K-------------R-L-----....-.------R-----.H- 443
71_BF1.BR.10.10BR_PE008       -F--------T--------D------------A--R----G------------------------------N..........T-.-.--.........KS--K.G---..-I-------R---L----K-----------R---------....-.-------RM---.S- 440
72_BF1.BR.10.10BR_MG002       -F--------T----G---D---------------R----G----------------S------------A-......K...TA.V.--.........KS--K.G---..V-----------I----------------R---------....-.------------.H- 441
O.BE.87.ANT70                 ----------T---------------------Q---S---G-------V--------T----------ATAQ...QDLKGGYTA.V.F-.........--QNP.IRKG..TI----------I----------------Q--------R...NG.K-------Y--P.GG 447
O.CM.98.98CMA104              --------------------------------Q-------G-------V--------T----------ASAQ...QDLKGGYTA.V.F-.........--QNP.IR-G..P-----------I-------------N--QD-------R...NG.K-------Y--P.GG 447
O.CM.98.98CMABB141            --------------------------------Q---S---G-------V--------T---KI-----ASAQ...QDLKGGYTA.V.F-.........--QNP.IRKG..PI----------L-K-------R------Q-----R--K...NG.K-------Y--P.GG 447
O.CM.98.98CMABB212            ---------------S----------S-----Q-------G-------V--------T---KI-----VSAQ...QDLKGGYTS.-.F-.........--QSP.-K-R..TI----------V-K-------R------Q--------N...NG.--------Y--P.GG 447
O.CM.98.98CMU5337             ----------T---------------------Q---S---G-------V--------T----------ATAQ...QDLKGGYTA.V.F-.........--QNP.IRKG..PI----------L----------------Q--------R...NA.K-------Y--P.GG 447
O.CM.99.99CMU4122             --------------------------------Q---S---G-------V--------T----------ATAQ...QDLKGGYTA.V.F-.........--QNP.GRKG..PI----------L----------------Q--------R...NG.K-------Y--P.GG 447
O.FR.92.VAU                   ----------T---------------------Q-------G-S-----V--------T----------ASAQ...QDLKGGYTS.V.F-.........--QKP.IRKG..PI----------VTK--------------Q--------K...NR.--T---K-Y--P.RG 446
O.GA.11.11Gab6352             ----------T---------------------Q-------G-S-----V--------T----------ATAQ...QDLKGGYTS.V.F-.........--QNP.VK-G..P-----------I-K---S--R-------Q--------K...NG.--------Y--P.GG 446
O.SN.99.99SE_MP1299           ----------T---------------------Q---S---G-------I--------T----------AAAQ....DLKGGYTA.V.F-.........--QNP.SRKG..PI----------L----------------Q--------K...NG.--------Y--P.GG 446
O.US.10.LTNP                  --------------------------------Q---S---G-------V--------T----I-----ATAQ...QDLRGGYTA.V.F-.........--QNP.-R-G..AI----------I-K--------------Q---L----R...NG.K-------Y--P.GG 447
N.CM.02.DJ00131               ------------D-------------------Q------------------------A-------------Q......T..ATS.V.FA.........GR--K.GI-R..TIR---------L----K---RG------Q--------K..NEG.--------G-SP.F- 449
N.CM.02.SJGddd                ----------T---------------------Q---S-----------------------------------Q......A..TTS.V.FV.........----K.GI--..TI----------L----K---R-------Q--------N..NEG.-------RG-SP.F- 447
N.CM.04.04CM_1015_04          ----------T-D-------------------Q-------S-S-------------------------A-AQ......A..ATS.V.FV.........----K.GT-R..TI----------L----K---RX------Q--------K..NEG.--------G-SP.F- 446
N.CM.06.U14296                --R---------X------D------------Q-------G----------------A-------------Q......A..TTS.V.FA.........----K.GI-R..PI----------L----K---RG------Q--------K..NEG.-------RG-SP.F- 446
N.CM.06.U14842                ----------X--------D------------QL------G----------------A-------------Q......G..ATS.V.FV.........----K.GM--..TI----------V----K---RG------Q--------K..NEG.--------G-SP.F- 446
N.FR.11.N1_FR_2011            ----------T--------D------------QL------G----------------A-------------Q......P..TTN.-.FA.........----K.GI--..PI----------L----K---R-------Q--------K..SEG.-------RG-SP.F- 447
P.CM.06.U14788                ---------------S---D------------Q-------G-------N--------T----------AAANQASQELKGGYT-.V.F-......-S.......G---..P--------V--I-K--K---RR------Q--------K...SG.-----------P.GG 440
P.FR.09.RBF168                ---------------S---D---I--------Q-------G-------N--------T----------AAANQANQDLKGGYT-.V.F-......-S.......G--X..P--------I--I-K--K---R-------Q--------K...SG.-----------P.GG 443
CPZ.CD.06.BF1167              ----I-------P---------------------------G------L---------A----------Q-TQ......Q..NVN...-V......-G.-AAPQ.GK-GAGNI-------I-----------R-------Q----L------..QSK.SGV-----RT-W.WG 470
CPZ.CM.05.SIVcpzMT145         ------------D------D---I--------N--R----G----------------S-------------N......K...TE.-.F-......-R.G-QN-.PPN-..KI----------L--------R-------Q--------N..QNN.GRI-----N--P.WG 434
CPZ.GA.88.GAB1                -------------------D------------Q-------G----------------S-----------M-Q......-QGR-D.V.FF.........K-QGA.GPKR..KI----------L----K---R----R--Q---------....G.--V-----G---.RS 452
CPZ.TZ.06.SIVcpzTAN13         ----I------P---T-------I--------H------QG------L---------T---KI-----ATA-......-TVGVN...-L.........G-KKP.LRKG..QLQ-------L----K-----------R--Q---------ARTNN.TGV-----RT-L.WG 474
CPZ.US.85.US_Marilyn          -----------P---T-------------------R----G------V---------A----------C-MK.......-.PSS.V.FL.........K--AG.KPGR..KI----------L--------R-------Q---------..AGN.--------H-SP.SW 445
GOR.CM.07.SIVgor2139_287      ------------D----------I--------Q-------G-------N--------T----------AAAH...SDLKGGYTA.V.F-......-KGGN-QG.IKKG..P--------V--I-K------R-------QD-------R...SG.--------Y--P.GG 442
GOR.CM.07.SIVgorCP2135con     ------------D----------I--------Q-------G-------N--------T----------AAAH...SDLKGGYTA.V.F-......-KGGX-QG.IKKG..P--------V--I-K------R-------Q--------R...SG.--------Y--P.GG 442
```

Draft of 30 September 2015, not for public release

hiv1prot_main.tex r8876–10546

```
                                       p1 end p6 start          Vpr binding                                                                                        Vpr binding
                                                                                                                                                               p6 end, Gag end
B.FR.83.HXB2                   ....GRPGNFLQSRP............EPT..........APPEESFRS.....GVETTT..PP......Q.....KQEPI....D.....KE........LY.PLTSLRSLFGNDPSSQ*     500
A1.AU.03.PS1044_Day0           ...-------P----............---..........---A-L-GM.....-E-I-..S-......P........---QK....-.....--........QVP--V--K----------     500
A1.CH.03.HIV_CH_BID_V3538      ...-------P-N-L............---..........---A--GV.....RE--.-..SS......M........---QR....-.....R-........QPP--V--K-------L---     499
A1.CY.08.CY236                 ...-------P----............--S..........---A-I-GM.....-E-I.-..S-......P........---QK....-.....S.........P--V--K----S------     501
A1.ES.05.X1608_8               ...-------P----............--S..........---A-X-GM.....-E-M.-..S-......P........---QKQEQGN....R-........QAP--I--K-------L---     501
A1.KE.11.DEMA11KE001           ...-------P-N-L............---..........---A-ICGM.....-DGI.A..F-......P........---QK....-.....R-........QGP-SI--K------QL---     499
A1.RU.11.11RU6950              ...-------P----............--S..........---A-N-GM.....-E-I.-..-S......L........---QK....-.....R-........QHP-SI--K-------L---     495
A1.RW.11.DEMA111RW002          ...-------P----............--S..........---V-LLGM.....-E-I.A..S-......P.....R--QR....-.....R-........QGP--V--K-------L---     500
A1.SN.01.DDI579                ...-------P----............---..........---ADLLGM.....-E-I.-..SS......P........---QK....-.....R-........P.P--V---------L---     496
A1.UG.11.DEMA110UG001          ...-------P----............---..........---A-L-GM.....-E-I.-..S-......P.....N--QK....-.....R-........QAP-FA--K-------*---     496
A1.ZA.04.04ZASK162B1           ...-------P-N-L............---..........---AG-YGM.....-E-I.-..S-......-......-K-QT....-.....R-........Q.P--A---------L-P-     504
A2.CM.01.01CM_1445MV           ...-------P---T............---..........---A-NVGM.....-E-I.A..SS......S........---LR....N.....R-........QHT-TI---------L---     497
A2.CY.94.94CY017_41            ...-------P---T............---..........---A-NL-M.....-E-I.-..SS......L........---LE....T.....R-........P-N-AI--K-------LL--     497

B.BR.10.10BR_MG029             ...-----------.............---..........--------F.....-E----.-------.-.......-----.....-.....--........--.--A--K-------L---     500
B.CA.07.502_1191_03            ...---------N--............--S..........-------KF.....-E----.-------.-.......----T.....-.....--........M-.-FS--K-------L---     503
B.CH.08.M2_0803101_NFLG8       ...-------F----............---..........--------F.....-E----..-S..QKQE-........---QK....T.....E-........K-.--A--K-------Y---     504
B.CN.10.DEMB10CN002            ...---------N-T............--S..........--------F.....-E----..-A.......-.......----T.....-.....--........--.--S--K-------L---     503
B.ES.10.DEMB10ES002            ...-----------.............---..........--------F.....-E-K-A..--.......-.......---Q-.....-.....--........--.---------------     499
B.FR.11.DEMB11FR001            ...-----------.............---..........--------F.....-E----..-S.......-.......---QT.....-.....E-........K-.-------------*---     499
B.GB.05.MM45d213_GN1           ...-----------.............---..........--------F.....-E-.--..-S......P........----T.....-.....-G........--.--A------------     509
B.HK.06.HK003                  ...-----------.............---..........--------Y.....-E----.---------KQG..L-G--.....-.....-D........--.---------------     504
B.HT.05.05HT_129389            ...---------N-T............--S..........--------F.....-E-.--.-------.-.......---SR.....-.....--........--.--A--K-------*---     496
B.JP.12.DEMB12JP001            ...-----------.............---APP.......--------F.....-EK-E-..-S.......-.......-----.....-.....--........Q-.--A--K-------*---     503
B.KR.07.HP_18_07JHS10_3909     ...-----------.............--S..........--------F.....-E----..-S.......-.......-----.....-.....--........--.--A------------     501
B.PE.07.502_0525_wg5           ...-----------TQSRP........---..........--------F....VE----..-S.......-.......----T.....-.....--........--.--A--K----------     506
B.RU.11.11RU21n                ...-----------.............---..........--------F.....-E----..-S.......-.......---QK....-RDRD.--........--.--A------------     510
B.TH.08.MERLBDTRC10            ...---------N--............---..........--------F.....-E----.-------.-.......-----.....-.....--........--.--A---------L---     508
B.US.11.ES38                   ...-------------ESRP.......--S..........--------F.....-EG-A-..-S.......-.......-----.....-.....-D........--.--A------------     504

C.AR.01.ARG4006                ...-----------.............---..........---A---F.....EE.-.-..-S......P........----K.....-.....R-........------K----S--L---     492
C.BR.07.DEMC07BR003            ...-------P---L............---..........---A---F.....EE.I.-..-S......P........----K.....-.....R-........------K----S--L---     491
C.BW.00.00BW5031_1             ...-----------.............---APP....MPT---A---F.....-E.-.-..-A......P........---MK....E.....R-........----I--K----S--L---     499
C.CN.10.YNFL19                 ...---------N-T............---..........---A---F.....-E----..-S......-......-S---.....-.....--........------K--------L---     496
C.CY.09.CY260                  ...---------N--............---APPTV..VPT---A--KF.....EER-....-A......P........----T.....-.....RD..........-I--K----S--W---     505
C.ES.08.X2363_2                ...-----------.............---APPA...EPT---A---F.....EG.-.A..-A......P........---LK.....-.....R-..........T----K-------W---     500
C.ET.02.02ET_288               ...---------N--............A--APL....EPT---A---F.....EE.A.-.--------P........---LK.....-.....R-..........A----K-------L---     498
C.IN.09.T125_2139              ...-----------.............---..........---A---F.....EE.-.-P.-A......P........----K.....-.....R-..........T---------S--L---     493
C.KE.00.KER2010                ...---------N--............---..........---A---F.....EE.-.A..LA......P........---LK....E.....R-........------K----S--L---     491
C.MW.09.703010256_CH256.w96    ...---------N--............---APPL...EPT---A--ST..............-A......L........---QK....-.....R-..........-SI--K-------L---     494
C.TZ.08.707010457_CH457.w8     ...---------N-T............---..........---A---EF.....EE.-.-..H-......P........---QK....N.....R-..........T----K----S--L---     492
C.YE.02.02YE511                ...-----------.............---..........---L---F.....EE.-.-..-A......P........---AK.....-.....--........------K----S--LLP-     489
C.ZA.08.705010534_CH534.w12    ...-----------.............---APPA...EPT---A---F.....EG.-.-..-A......P.....R--TK.....-.....R-........------K-------L---     497
C.ZA.10.DEMC10ZA001            ...-----------.............---..........---A---F.....EE.-.-..-A......P........---SK.....-.....R-..........---A-K----S-LL---     493
C.ZM.11.DEMC11ZM006            ...-----------.............---APL....QPT---A---F.....EE.-.-..-A......P........----K.....-.....R-..........AI----------L---     499

D.CM.10.DEMD10CM009            ...-------------EPTAPL.....---..........---A--GF.....-E-I.-..LS......-........---QR.....-.....-N........KD.--A--K-------L---     504
D.CY.06.CY163                  ...---------N--............---..........---A--GF.....-E-I.-..-S......-........---QR.....-.....--........--.-----------L---     500
D.KE.11.DEMD11KE003            ...---------G--............---..........---A--LGF.....-E-I.-..-S......-........---QK....G.....--........QC.----K-------W---     501
D.KR.04.04KBH8                 ...-------P---T............---..........---A--GF.....-E-I.-..-S......-......-R-QK.....-.....N-........--.--D--K-------L-K-     500
D.SN.90.SE365                  ...-------P----............---..........---T--GF.....-E-M.-..-SQKQEQA-.......---QK.....-.....--........--.--A--K-------L---     507
D.TZ.01.A280                   ...-----------.............---..........---A-I-GF.....-E-I.K..-S.......-.......---QKDK.....-.....--........--.-SA--K-------L---     497
D.UG.10.DEMD10UG004            ...-----------.............---..........---A-C-GF.....-E--.N..-S.......-.....Q--QK.....-.....--........--.--A------------     501
D.UG.11.DEMD11UG003            ...---------R--............---..........---A--GF.....-E-I.-.--------.-.......---QK.....-.....EK.DE......--.-----K----S--L---     501
D.YE.02.02YE516                ...-----------.............---..........---A--GF.....-E-I.-..-S......-........---QKDK.....-.....-D........--.-----------L---     501
D.ZA.90.R1                     ...-----------.............---..........---A--GF.....-E-I.-..-S......-........---QK.....-.....--........--.-----K----S--L---     501

F1.AO.06.AO_06_ANG32           ...-----------.............---..........---A--GF....RE-M.P..--......P........---QK....E.....EGQGQG.....--P--A--K------Q*---     500
F1.AR.02.ARE933                ...-----------.............---..........---A--GF....RE-I.-..-S......P........---QK....-.....EG........--P--A--K-------*L--     495
F1.BR.10.10BR_RJ015            ...-------I-N--............--S..........---A--GF....RE-I.P..S-...RQE-........---QK....E.....EG........--P--A--K------QWL--     502
F1.CY.08.CY222                 ...-----------.............--S..........---A--GF....KE-I.A..-S......P.....E--QK....-.....-G........--P--A--K-------*L*-     493
F1.ES.02.ES_X845_4             ...-----------.............---..........---A--GF....RE--.-..-S......P........---QK....E.....EGQG......--P--A--K------Q*---     498
F1.RO.96.BCI_R07               ...---------N--............---..........---A--GF....RE--.-..-S......P........---QK....-.....EG........M-P--A--K-------*---     495
F1.RU.08.D88_845               ...-------P---L............--S..........---A--GF.....-E-I.-..-S......P.....R--QK....-.....-D........QCP----K------Q*---     501
F2.CM.02.02CM_0016BBY          ...-------I----............---..........---A-G-GF.....-E-I.-..-S......P........---QK....-.....-G........--P----K-------*---     493
F2.CM.10.DEMF210CM001          ...-----------.............-QSRP.....EPS---A--GF.....-E-I.-..-S......L........---QK....-.....--........QLP-----------Q*---     503
F2.CM.10.DEMF210CM007          ...---------N--............---..........---A--GF.....-E-I.-..-S......P........---QK....-.....--........PAY---A-K----S--*---     494

G.BE.96.DRCBL                  ...---------N--............---..........---A-N-GF.....-E-I.A..-S......P........---QK....E.....--........--.--S--K------Q*-K-     493
G.CM.10.DEMG10CM008            ...---------N--............---..........---A--GF.....-E-I.A..-N......-........---QK....E.....--........--.--A--K----S--*-M-     499
G.CN.08.GX_2084_08             ...---------N-L............---..........---A--GF.....-EGI.-..-S......P........----R....EQEPRE-G.........-P.--A--K-------*---     501
G.CU.99.Cu74                   ...---------N--............--S..........---A--GF.....-Q-I.A..-S......P........---QK....E.....-D............-I--K----S--*---     497
G.ES.09.X2634_2                ...---------N--............---..........---A--GF....RE-I.A..-S......L.....-P--R....EPEQRE--........--.--A--K----S--*---     502
G.GH.03.03GH175G               ...---------N--............--S..........---A--GF.....-R--.A..-S......P........----R....E.....--E.......--.-----K-------*---     499
G.KE.09.DEMG09KE001            ...---------N--............---..........---A--GL.....-E-I.A..-S......P......--DQK....G.....R-..........P----K----S--*---     497
G.NG.09.09NG_SC62              ...-------P-K-L............---..........---A--GF.....-E-I.A..-S......P........----R....E.....--........SP.--A--K----S--*---     493
G.ZA.01.TV546                  ...-----------.............---..........---A--GF.....-E-I.A..-S......P.....R---K....E.....--........TP.--A--K----S--*---     497

H.BE.93.VI991                  ...-------P-K-L............---..........---A--GF.....-E-I.-..-S......P.....R--LK....E.....Q-..........P-----------Q*---     501
H.CF.90.056                    ...-----------.............---..........---A--GF.....-E-M.-..-S......P........---QLK....-.....--..........P--A------S--LL--     501
H.GB.00.00GBAC4001             ...-----------.............---..........---A-C-GF.....-E-M.-.--------P........---MK....E.....--..........P----------S--L---     500

J.CM.04.04CMU11421             ...-----------.............---..........---A--GF.....-E--....XX......P........---AK.....-.....--........--.-----K----X--X---     497
J.SE.93.SE9280_7887            ...-----------.............---..........---A--LGL.....-E-I....-S......P........----K.....-.....--........--.-----K----S--L---     498
J.SE.94.SE9173_7022            ...-----------.............---..........---A--LGF.....-E-I....-S......P........----K.....-.....--........--.-----------S--L---     498

K.CD.97.97ZR_EQTB11            ....E-------N--............---..........---A--GF.....-EKI.-..-S......L.....R--MK.....-.....Q-........QGP----K----S--L---     496
K.CM.96.96CM_MP535             ...-----------.............---..........---A--GF.....-E-I.-..-S......P.....R--TK.....-.....--........QSP----K-------L---     496

U.CA.01.TV749                  ...-------P----EPTAPPFPQSSL---..........---A--LGK....AEGI.-..SS......P.....L-D...........N.........--.--V----------L---     500
U.CA.99.TV721                  ...-------P----............A--..........---A--LQ-....RP-P.-..A-......PAENLW-G-E-....I.....RQ........EQ.--V----------L---     499
U.CD.83.83CD003_Z3             ...---------N--............---..........---A--GF.....-E-.-.-.-S......P........----R.....-.....--....S....--P----K----S---L--     500
U.CD.90.90CD121E12             ...---------N-T............---..........---A--GF.....-E-V.-T.-S......P........---LR.....-.....--........SP.--A--K----S--LL--     497
U.CY.05.CY090                  ...-----------.............--S..........---A--GF.....-E-I.-..SL......P.....R--QK.....-.....--........QGP----------S---L--     497
U.CY.08.CY223                  ...-------P----............---..........---A--LGF.....-E-I.AA.-S......P.....-P-QK....V.....--........ELY----K----S-Q*-E-     502
U.ES.10.DEURF10DZ001           ...-------P----............---..........---M--LGM.....-E--.-..SS......L........---QG.....-.....-G........--P----K----S--*---     495
U.GR.99.99GR303                ...-------P----............---..........---A--GF.....-E-I.V..-S......P.......--G-K.....-.....--........--.--A--K----S--*---     495
U.NL.95.U_NL_95_H10986_D1      ...-----------.............---..........---A--GF.....-E-I.-..-S......P........---MK.....-.....--........QDP----K----S--W---     497

01_AE.AF.07.569M               ...-------P----............---..........---AD.WGM.....-E-I.-..SL......L........---QK.....-.....--........HLP-TV--K-------L---     499
01_AE.CF.90.90CF11697          ...-------P---L............---..........---A--LGM.....-E-I.-..SF......P........---QK.....-.....--........HPS--V--K-------L---     499
01_AE.CN.10.YNFL03             ...-------P----............---..........---A-NWGM.....-E-I.P..SL......L........---QT.....-.....--........SPP----K-------LL--     498
01_AE.HK.04.HK001              ...-------P----............--S..........---A-NW-E........I.-..SL......L........---QK.....-.....--........HPS-SI--K-------L---     496
01_AE.IR.10.10IR.THR48F        ...-------P----............--S..........---A-NWGM.....-E-I.A..SS........................T-........HPP-SI--K-------L---     492
01_AE.JP.x.JRC77AE             ...-------P----............--S..........---A-NWGM.....-E-M..................---QK.....-.....--.........HP--V--K-------L---     494
01_AE.TH.09.AA090a_WG11        ...-------P----............---..........---A-NWGM.....-E-V.-..S-......L.....R--QK.....-.....--........HPP--V--K-------L---     497
01_AE.TH.90.CM240              ...-------P----............---..........---A-NWGMGEEIT-E-I.-..SL......P........---QK.....-.....--........HPP--V--K-------L---     504
01_AE.US.05.306163_FL          ...-------P----............---..........---A-NWGM.....-E--.-..SL......L........---QK.....-.....--.........RP-VV--K-------L---     494
01_AE.VN.98.98VNND15           ...-------P---T............---..........---A-NWGM.....-E-I.-..SL......L........---QK.....-.....--........HPP-SI--K-------L---     498
```

```
                               p1 end p6 start            Vpr binding                                                                  Vpr binding
                                                                                                                                      p6 end, Gag end
B.FR.83.HXB2                   ....GRPGNFLQSRP............EPT..........APPEESFRS.....GVETTT..PP......Q.....KQEPI....D.....KE........LY.PLTSLRSLFGNDPSSQ*    500
02_AG.CM.10.DE00210CM013       ....------P----............---.........---LA--LGM.....-E-R.P..SS......P.....----K....-.....-G........--P--A--K-------*L--    500
02_AG.CY.09.CY256              ....------P----............---..........---A---GM.....-E----..--......L.....---TG....-.....-G........--P-----K-------*---    495
02_AG.ES.06.P1423              ....------P----............---..........---A--LGM.....-E-X.P..-S......P.....R---K....-.....--........--P--A--T------Q*---    495
02_AG.GW.05.CC_0048            ....------P----............---..........---A--..M.....-E-M.-..S-......P.....T---G....-.....-G........--P-----K----S--*---    498
02_AG.KR.12.12MHR9             ....------P----............---..........---A--LGM.....-E-I.-..S-......PKQEQ.RE--R....-.....-G........--P-----K-------*---    499
02_AG.LR.x.POC44951            ....------P----............---..........---A---GM.....-E-I.-..S-......P.....----R....-.....-G........--P-----K-------*---    495
02_AG.NG.09.09NG_SC61          ....------P-N-L............---..........---A--VGM.....RE-A.-..SS......-.....-P--G....-.....-D........--P-----K-------*---    493
02_AG.NG.x.IBNG                ....------P----............---..........---A---GM.....-E-I.P..-S......P.....Q---R....-.....-G........--P-----K-------*L--    495
02_AG.SE.94.SE7812             ....------P----............---..........---A--LGI.....-E-I.-..SS......-.....----G....-.....-G........--P--A--K-------*---    495
02_AG.SN.98.98SE_MP1211        ....------P----............---..........---A---GM.....-E-I.-..SS......P.....----G....-.....-G........--P-----K-------Y---    495
03_AB.RU.97.KAL153_2           ....------P----............--S..........---A-N-GM.....-E-I.-..-S......L.....---QK....-.....R-........QHP-SI--K---D--L---    499
04_cpx.CY.94.94CY032_3         ....--------N--............---..........---A-CLER.....KE----..SS......L.....----R....-.....--........--.-----K----S--L---    501
05_DF.BE.x.VI1310              ....-----------............--S..........---A----F.....-E-I.A..SS......P.....---QK....-.....EG........--P--A--K-------L---    503
06_cpx.AU.96.BFP90             ....--------N--............---..........---I---GF.....-E-I.A..-S......P.....---SK....E.....--EKG.....--.--A--K----S--*---    500
07_BC.CN.98.98CN009            ....-----------............---..........--------F.....-E----..-S......-.....---T-....-.....--........--.-----K---------    496
08_BC.CN.97.97CNGX_6F          ....-----------............---..........---A----F.....EE.---..-A......P.....----K....-.....R-..........----------S--L---    490
09_cpx.GH.96.96GH2911          ....------P----............--S..........---A---GM.....RE----..SS......X.....--XKD....-.....G-........XPP--X--K-------L---    497
10_CD.TZ.96.96TZ_BF061         ....-----------............---..........---A---GF.....-E-I.-..-S......-.....---QK....-.....--........PH.--A--K-------L---    501
11_cpx.CM.95.95CM_1816         ....-----------............---..........---A---GF.....-E-V.A..-S......-.....----K....-.....E-........--.--A--K----S------    497
12_BF.AR.99.ARMA159            ....--------N--............---..........---A---GF.....-E-I.-..SS......P.....-R-QK....-.....EG........Q-P--A--K-------*-*-    494
13_cpx.CM.96.96CM_1849         ....-----------GPTAPPESRP..---..........---A---GF.....-E-I.A..-S......-.....----K....E.....--........T---AA-K----S--*-*-    508
14_BG.ES.05.X1870              ....------P-N-L............---..........---A---GF.....-E-I.A..-S......P.....----K....E.....--........M---A--K----S--*---    497
15_01B.TH.99.99TH_MU2079       ....------P----............--S..........---A-NWGM.....-E---...SL......L.....---QK....-.....R-........HTP--V--K-------LL--    498
16_A2D.KR.97.97KR004           ....------P---T............---..........---A-D-GM.....-E---...-L......-.....---LK....N.....R-........QHT-AI--K-------LL--    499
17_BF.AR.99.ARMA038            ....--------N--............---..........---L---GF.....-ETI.-..-Q......P.....----K....-.....EG........--P--A--K-------*-K-    501
18_cpx.CU.99.CU76              ....--------K--............---..........---A--LGF.....-E-I.A..-S......P.....---SK....E.....EK...........Y-FS--K---S------    497
19_cpx.CU.99.CU7               ....-----------............---..........-------GF.....-E-I.-..-S......-.....--DQK....-.....--........--.--A--K-------L---    501
20_BG.CU.99.Cu103              ....-----------............---..........---A---GF.....-E-I.A..-S......P.....--GQK....E.....-D........ .....I--K----S--*---    497
21_A2D.KE.99.KER2003           ....----------S............---..........---A-M-GF.....-E-I.-..-S......-.....---QK....-.....--........--.-SV--K-------LL--    499
22_01A1.CM.01.01CM_0001BBY     ....------P-G--............---..........---A--VGM.....-E-I.-..SS......-.....---QR....E.....--........QPP--V--K-------L---    500
23_BG.CU.03.CB118              ....------P-N-L............---..........---A---GF.....-E-I.A..-S......P.....--G-K....E.....QD.............I--K----S--*---    498
24_BG.ES.08.X2456_2            ....------P-N-L............---..........---A--LGF.....-E-I.A..-S......P.....--GQP....E.....-D.............I--K----S--*---    497
25_cpx.CM.02.1918LE            ....--------N--............---..........---A--LGF.....-E-I.-..-S......P.....---QK....-.....--........--.-MA--K----S--*---    499
26_AU.CD.02.02CD_MBTB047       ....------P---L.........V--..........---A-NLGM.....EEG-.-..-S......P.....--KEG....E.....--........QSP--I--K----------    496
27_cpx.FR.04.04CD_FR_KZS       ....--------N--............---..........---A--LGI.....ME-M.-..SY......P.....---LK....-.....--..........P--V--K-------L---    497
28_BF.BR.99.BREPM12609         ....--------N--............---..........---A----F.....-E---...-S......P.....---QK....E.....EG........--P--A--K-------*---    499
29_BF.BR.01.BREPM16704         ....--------N--............--S..........---A--LGF.....-E-R.-..-S......P.....---QK....-.....EG........K-P--A---------*---    510
31_BC.BR.04.04BR142            ....-----------............---APPASRLEPT---A----F.....EE.---..-S......P.....---QR....-.....R-........ ........K----S--L---    506
32_06A1.EE.01.EE0369           ....--------N--............---..........---A----F.....EE--AA..-S......S.....----K....K.....--........--.-SA----------*---    495
33_01B.ID.07.JKT189_C          ....-----------............---..........---A----F.....-EQ---..-S......-.....---Q-....-.....--........ .LY-VA--K----------    497
34_01B.TH.99.OUR2478P          ....------P----............---..........--------F.....-EG---..-S......-.....-----....-.....--........ .LY-SA--K-------LL--    498
35_AD.AF.07.169H               ....------P-G--............---..........---A---GF.....-E-I.K..-S......-.....---QK....-.....G-........--.-----K----------    497
36_cpx.CM.00.00CMNYU830        ....------P----............---..........---A-N-GI.....-E-M.S..SS......-.....---SR....-.....-G........--P-----K-------*L--    498
37_cpx.CM.00.00CMNYU926        ....------P----............---..........---A---EV.....KE-V.A..SS......L..........G.....--........--P--A--K---S--*P--     493
38_BF1.UY.03.UY03_3389         ....-----------............--S..........---A---GF.....-E-I.-..--......PKQEQ.---QK....-.....EG........M-P--A--K-------*---    499
39_BF.BR.04.04BRRJ179          ....----------T............---APT.......---A--LMF.....-E--.A..-S......-.....----R....-.....--........--.-------------     502
40_BF.BR.05.05BRRJ055          ....--------N-L............---..........---A---GF.....-E-I.-..--......P.....---QK....N.....EG........--P-----K-------...-    498
42_BF.LU.06.luBF_18_06         ....--------N--............--RT.....EPT---A---GF.....-E-I.-..SS......L.....---QK....-.....EG........--P-----------------    503
43_02G.SA.03.J11223            ....-----------............---..........---A---GF.....-E-I.A..-S......-.....----R....E.....--........SP.--A--K-------W-K-    498
44_BF.CL.00.CH80               ....------P-N-L............---..........---A---GF.....-E-I.-..-S......P.....---QK....-.....EG........K-P-FA--K-------*-K-    501
45_cpx.FR.04.04FR_AUK          ....------P----............--S..........---A---GM.....-E-I.-..SC......P.....---QK....-.....R-........QV.YP--K-------L---    499
46_BF.BR.07.07BR_FPS625        ....--------N--............---..........---A---EI.....RE-I.-..-S......-.....---QK....-.....EG........--P--A--K-------W---    496
47_BF.ES.08.P1942              ....--------N--............---..........---A--LGF.....-E---...-S......P.....---QK....-.....EG........--P--A----------*---    498
48_01B.MY.07.07MYKT021         ....------I----............---..........---A----F.....-E---...-S......-.....-----....-.....--........ .LY--A-------S------    499
49_cpx.GM.03.N26677            ....--------N--............---..........---A---GF.....-E-I.A..-S......P.....---QK....-.....E-........--.-SA--K----S--*LP-    496
50_A1D.GB.10.12792             ....-----------............---..........---A---GF.....-E-V.I..-S......-.....---QK....-.....--........--.-SA-----------    498
51_01B.SG.11.11SG_HM021        ....-----------............---..........---A----F.....-E---...-S......-.....-----....-.....--........Q---A--K------LL---    499
52_01B.MY.03.03MYKL018_1       ....------P---X............---..........---A-NWGM.....-E-.X-..SS......L.....---QX....-.....-D........PPP-SV--K----S--LL--    500
53_01B.MY.11.11FIR164          ....--------N--............---..........--------F.....-EG---..-S......-.....-----....-.....--........ .LY-----K----------    499
54_01B.MY.09.09MYSB023         ....--------N--............---..........--------L.....-E----..--......-.....-----....-.....--........--.-VSA-K-------*---    501
55_01B.CN.10.HNCS102056        ....------P----............---..........---A-DWGM.....-E-I.-..SS......P.....---QK....-.....-G........HRP-VV--K-------L---    499
56_cpx.FR.10.URF5_patient_A    ....------P----............---..........---A--VGM.....-E-A.-..YS......P.....R--SRDK..-.....--........--P-----K----S--*---    500
57_BC.CN.09.09YNLX19sg         ....-----------............---..........--------F.....EE.---..-A......P.....---LK....-.....R-........--------K-------L---    492
58_01B.MY.09.09MYPR37          ....-----------............--X..........--------F.....-E----..-X......-.....--G--....-.....--........ .LY-----K----S------    498
59_01B.CN.09.09LNA423          ....------P----............---..........---A-NWGM.....-E-S.....L......L.....---QK....-.....--........QPH-SI--K-------L---    492
60_BC.IT.11.BAV499             ....-----------............---..........---A----F.....EE.---..--......P.....---QK....-.....R-........--------K----S--L---    491
61_BC.CN.10.JL100010           ....-----------............---..........---A----F.....EE.-.A..SA......P.....---MK....-.....R-........-------------S--L---    492
62_BC.CN.10.YNFL13             ....-----------............---..........---A----F.....EE.---..-A......P.....---SK....-.....R-........-------------S--L---    488
63_02A1.RU.10.10RU6637         ....------P----............---..........---A--VGM.....-E-I.-..SS......P.....Q---R....-.....-G........--P--A--K-------*---    495
64_BC.CN.09.YNFL31             ....-----------............--S..........---A----F.....-E----..-S......-.....-P-GT....-.....--........--.--A--K-------LC--    496
65_cpx.CN.10.YNFL01            ....----------L............---..........---A----F.....EE.---..-A......P.....R--SK....-.....RG........--------K-------L---    495
67_01B.CN.11.2011.ANHUI.HF115  ....------P----............---..........---A-DWGM.....-G-I.-..FS......P.....---QK....-.....--........ .H-SV--K-------L---    497
68_01B.CN.11.2011.ANHUI.WH73   ....------P----............---..........---A-DWGM.....-G-I.-..SS......P.....---QR....-.....R-........ .Q--V--K-------L---    497
70_BF1.BR.10.10BR_PE004        ....---R--I-N--............---..........---A----F.....-E-M.-..--......PK....--EQK....-.....EG........R-P-----K-------W---    504
71_BF1.BR.10.10BR_PE008        ....--------N-T............---..........---A---GF.....-E-M.-..-S......-.....---QK....S.....EG........--P--A--K-------W-K-    499
72_BF1.BR.10.10BR_MG002        ...G--------N-T............---..........---A---GF.....-E-K.-..-S......P.....----K....E.....EG........--P--A--K-------WL--    501
O.BE.87.ANT70                  ....T----YV-RPA............H-S..........---M-EEVK.....-Q-NQ.................E-KGG....P.....N-........--.-FA--K----T-Q...-    499
O.CM.98.98CMA104               ....T-----M-KQV............A-S..........---M-EVVK.....-Q-NQ.................E-KED....S.....N-........--.-FA--K----T-Q...-    499
O.CM.98.98CMABB141             ....A----YM-KQV............P-S..........---M-EMMK.....-Q-NQ.................E-KGD....T.....T-........--.-FA--K----T-Q...-    499
O.CM.98.98CMABB212             ....T----Y--KQV............S--..........---X-EIVK.....-Q-NQ.................S-GXX....Q.....X-........--.-FA--K----T-Q...-    499
O.CM.98.98CMU5337              ....T----YA-RQV............S--..........---M-EEVK.....RQ-NQ.................E-KGD....Q.....N-........--.-FA--K----T-Q...-    499
O.CM.99.99CMU4122              ....T----YA-RQV............S-S..........---M-EE-K.....-P-SQ.................E-KGD....P.....N-........--.-FA------T-Q...-    499
O.FR.92.VAU                    ....T----YV-EQV............S-S..........---M-EMMK.....-Q-EDQ................N-KGE....Q.....N-........--.-FA------T-Q...-    499
O.GA.11.11Gab6352              ....T----YV-KQV............H-S..........---M-EEVQ.....KQ-NQ.................E-KGD....P.....N-........--.-FA--K----T-Q...-    498
O.SN.99.99SE_MP1299            ....T----YA-RQV............S-S..........---MTEEMK.....-Q-NQ.................E-KED....Q.....N-........--.-FA------T-Q...-    498
O.US.10.LTNP                   ....T----YV-KQV............A-S..........---M-EEVN.....RQ-NQ.................E-KGS....P.....T-........--.-FA------T-Q...-    499

N.CM.02.DJ00131                ....------P-TTTTRK.........---..........---L-NYG-.....QE-R.S..TQ......G.....-ERQE....N..............TENS--P-----------L----    513
N.CM.02.SJGddd                 ....------P-TTARK..........---..........---L-..........E-R.S..TQ......G.....-EMQE....N.....Q-...KTENS--P-----------L----     508
N.CM.04.04CM_1015_04           ....------P-TATRK..........---..........---L--YG-.....QE-G.S..TQ......G.....-EMQE....N.....Q-...KTENS--P-----------L----     512
N.CM.06.U14296                 ....------P-TTTKK..........---..........---L-NYGF.....QE-R.S..TQ......G.....-EMQE....N.....Q-...RTENS--P------------L---     512
N.CM.06.U14842                 ....------P-TTTTRK.........---..........---L-DYGC.....QE-R.N..TQ......R.....-EMQE....-.....Q-...KTENS--P---------------     513
N.FR.11.N1_FR_2011             ....------P-TTTRK..........---..........---L-NYGF.....QE-K.S..TQ......G.....-EMQE....-.....Q-...RTENS--P---------------     513

P.CM.06.U14788                 ....K----YV-KQV............Q--..........---M-EEEM...........-.........-NKQ..XE-KE....-.....EK...E.....--.--A--K----T-Q*Y--    495
P.FR.09.RBF168                 ....K----YV-KQV............Q--..........---M-EDEM...........-.........RDKQ..REXEE....X.....E-........E.----------T-Q*Y--    497

CPZ.CD.06.BF1167               ....C-----V-NQSQTQNQEQ.................SQTQ-KLVE.....KA-..P..SAP.....P.....ME-VNHW..NS....-D...........GF--K-I--S-Q*WK-    531
CPZ.CM.05.SIVcpzMT145          ....-----YV-G-K............---..........---M---GF.....QE-K.K..VK......E.....EVREE....S.....-P........PS.-----K----------    492
CPZ.GA.88.GAB1                 ....------V-N-T............---..........---I--YGY.....QE-E.K..SQ......-.....EKKEG....E.....SS........--.-P--K---S------    509
CPZ.TZ.06.SIVcpzTAN13          ....C-----V-NN-............-KE..........K-Q-QEG.......VQ.--.VPTA-.....PM....EMG-M....K.....EG..............F--K-I--S-Q...-    525
CPZ.US.85.US_Marilyn           SGGSK------EN-K............---..........---I-D-GY.....QE--.V..TQ......E.....--GKE....-.....--........PF.Q---K---S----E-    506

GOR.CM.07.SIVgor2139_287       ....E----YV-KQV............Q--..........---M-EGEK.....................E.....MRK-E....K.....E-........--.-----K----T-Q*YL-    492
GOR.CM.07.SIVgorCP2135con      ....E----YV-KQV............Q--..........---M-EGEK.....................E.....MRK-E....K.....E-........--.-----K----T-Q*YL-    492
```

```
                              Pol start
                              Gag-Pol TF start                                                                                              Gag-Pol TF end protease start
B.FR.83.HXB2                  FFREDLAF.LQ....GKAREFSSEQT............RAN..........SPTRRELQV.....WGRDNN..SP......S.....EAGAD....R.....QG........TV.SFNFPQVTLWQRPLVTIKIGGQLKEALLDTGADDTVL.EEMSLPGRWKPKMIGGI  106
A1.AU.03.PS1044_Day0          ----N---.Q-.....-E--K-P----............---..........---SGA-WD....G---.-..L-......-.....----E....G.....-.........-GPTLS---I------I--VR-E---R-------------.-DID---K---R-----  106
A1.CH.03.HIV_CH_BID_V3538     ----N---.Q-.....-E---------............G--..........---S---WG....E--N.-..LL......D.....----E....G.....-.........-AP--S---I------V--V----------------------.--IN---K---------  106
A1.CY.08.CY236                ----N---.Q-.....-E--K--P---............--I..........---S-D-WD....G---.-..L-......-.....----E....-.....-..........P-LS---I---------VE-----R-------------.-DID-------R-----  103
A1.ES.05.X1608_8              ----N---.Q-.....-E--K-PT---............--I..........---S-X-WD....G---.D..LT......-.....----EAGAGK....-.....-.........-GPTL----I------I--VR------------------.-DIN---K---------  110
A1.KE.11.DEMA11KE001          ----N---.Q-....RE-----T---............G--..........---S-D-WD....G---.S..L-......-.....-T--E....G.....-.........-GSTL----I------V--VR------------------.-DIE---K---R-----  106
A1.RU.11.11RU6950             ----N---.Q-....RE--K------............--I..........---SGK-WD....G---.-..PL......P.....-T--E....G.....-.........-AS------I--------AVR------------------.-DID---K---------  106
A1.RW.11.DEMA111RW002         ----N---.Q-.....-E--K--P---............--I..........---SGT-GD....G---.S..L-......-.....-T--E....-.....-.........-GPT-S---I---------V--E---R-------------.-NIN-T-K---------  106
A1.SN.01.DDI579               ----N---.Q-.....-E--K------............---..........---S-SPGD....GR--.-..LL......P.....----E....G.....-..........-A--S---I------I--VR------------------.-DIN---K---------  105
A1.UG.11.DEMA110UG001         ----N---.Q-....RE--K---K--............---..........---G-A-WD....G---.-..L-......-.....-S--E....-.....-.........-GPTLC---I---------V-------------------.-DIN---K---------  106
A1.ZA.04.04ZASK162B1          ----N---.Q-.....-E---------............G--..........---S---WD....G---.-..L-......T.....-E---....-.....-.........A.A-LS---I-----------V----------------.-DIN---K---------  105
A2.CM.01.01CM_1445MV          ----N--S.Q-....RE--KL-P--N............---..........--AS--RWN....G---.S..FL......V.....---TE....K.....-.........-TH-Y----I---------V--E-----------------.-DIN---K-R-------  106
A2.CY.94.94CY017_41           ----N---.Q-....RE--K-----N............---..........---S---EN....G---.-..LL......P.....---TG....D.....-.........-IQ-C----I---------V--E-----------------.--IN---K---------  106

B.BR.10.10BR_MG029            ---------.P-.........L-P---............---..........-------------G---.-..-.......-.....----.....-.....-.........---LS---I---------V----------------------.---N------------  106
B.CA.07.502_1191_03           ----N---.P-.....-E--K------............--I..........-------------G---.-..-.......-.....----.....-.....-.........N-.LL---I------V-------------------------.-DIN---K---------  106
B.CH.08.M2_0803101_NFLG8      ----N---.P-........FP-P-..............---..........----G---------N--.PQ..SEAGA.....---TE....N.....R-........E-.--S---I------VL-V-----VR-----------II.-D-T---K----IV--V  110
B.CN.10.DEMB10CN002           ----N---.P-...........D............--I..........-------------G--D..-S.......-.....----.....-.....-.........--.PLI---I------I--V-----VK-----------.-DIN------------  106
B.ES.10.DEMB10ES002           ----N---.P-.....-E--K------............---..........-------------GE-S..-T.......-.....----.....-.....-.........--.-L----I------I--V-V--------------------.---N------------  106
B.FR.11.DEMB11FR001           ----N---.-.........G--P---............---..........-------------G---.-L.......-.....--R--.....-.....R-........K-.---L--I-----------------------------.---N------------  105
B.GB.05.MM45d213_GN1          ---------.P-.............-T..........----G-------------.-L......P.....----.....-.....-.........--.--S---I-----------V-----------------.---N---K---------  105
B.HK.06.HK003                 ----N--L.P-.............---..........----G-------------.-EAG..T--T-.....-.....-.........P-.---L--I---------V-V-----------------E--IN---K---------  110
B.HT.05.05HT_129389           ---------.-.........LP--D............--I..........----G---------.D-.......-.....---VE....G.....-.........--.--S---I-----------------------------.---N---K---------  105
B.JP.12.DEMB12JP001           ---------.P-.........LP---A............---SPT..........-------------K-RD..-L.......-.....-T---.....-.....-.........--.--S---I-----------V-----------------.-D-N---K---------  109
B.KR.07.HP_18_07JHS10_3909    ----N---.P-.............--I..........--S-G---------G---.-L.......-.....---T-.....-.....-.........--.--S---I------I----------------------.-----------------  106
B.PE.07.502_0525_wg5          ----N---.P-.....-E--KL----DSEQA............---..........----G--I....C-G---.PL.......-.....----.....-.....-.........--.--S---I-----------------------------.---N------------  110
B.RU.11.11RU21n               ---------.P-.....-E----P---............---..........----G-F---------G---.-L.......-.....----E....-QGQG.-.........A-.-LS---I---------R-----------------V.-D-Y------------  110
B.TH.08.MERLBDTRC10           ----N--L.P-.........KL-----............---..........--S-----------------.-.......-.....-T-TN.....-.....-.........-I.-LS---I--------A-------R-------------.---N---K---------  106
B.US.11.ES38                  ----N---.P-.....-E--KLP----REQT............--I..........----G---------S-..-L.......-.....-T---.....-.....-.........P-.--S---I---------V---------------------.---N------------  110

C.AR.01.ARG4006               ----N---.Q-.........K-XP---............---..........---S--------R-.-..PL.......-.....----E....-.....-.........TL----I-----------V--------------------.---K---N---------  102
C.BR.07.DEMC07BR003           ----N---.P-.....-E--K-PP---............---..........---S--------R-.-..PL.......-.....----E....G.....-.........TL-L--I------V--V-V--------------------.--TK---N---------  102
C.BW.00.00BW5031_1            ----N---.Q-.....-E---LPP---............-T-SPT...NAN--S--------R-.-..PR.......-.....---DE....-.....E-.........TL----I--------S--V---I------------------.-DIN---K---------  108
C.CN.10.YNFL19                ----N---.P-.........G------N............---..........-S-S--------------.-I.......-.....-V---.....-.....-.........-I.PL----I---------V-V---I------------------.--I----K---------  106
C.CY.09.CY260                 ----N---.P-.....-E----P---A............---SPTDS..GAN--S--------R-E-...PR.......-.....----.....-.....-.........L----I--------S--V---I------------------.-DIN---K---------  109
C.ES.08.X2363_2               ----N---.P-.....-E-G-LP----............---SPTS..RTN--S--------R-..S..-R.......-.....---VE....-.....-R.........DL----I----------RV---I----X-----------.-DIN---K---------  109
C.ET.02.02ET_288              ----T---.P-.....-E----P----............S--SPT...RANG--S--------R-.G.-..PS.......-.....----E....-.....-.........-L----I----------R---------------------.--IN---K---------  108
C.IN.09.T125_2139             ----N---.P-.....-E----PP---............---..........---S--------R-.--P.PS.......-.....----E....-.....-.........NL----I--------S-RV---I------------------.--IN---K---------  103
C.KE.00.KER2010               ----N---.P-.....-E----P--A............---..........---S--------R-..S.PC.......-.....----E....-.....E-.........TL----I---------V-----IR------------------.--IN---K---------  102
C.MW.09.703010256_CH256.w96   ----N---.P-.....-E----P--A............---SPTT...RTN---S---H.........PR.......-.....----E....-.....-.........TL----I----------RV---I------------------.--L----K---------  104
C.TZ.08.707010457_CH457.w8    ----N---.P-.....-E-GK-P--N............---..........---S---R-.....R-.-..P-.......-.....----E....K.....-.........NL----I--------S--V---I------------------.-DIN---K---------  102
C.YE.02.02YE511               ----N---.P-.........P----............---..........---I---------R-.N.-..PR.......-.....K--GE....-.....-.........TL----I------V----V---I------------------.--IN---K---------  100
C.ZA.08.705010534_CH534.w12   ----N---.P-.....-E----P--A............-T-SPTS..RAN---S--F--------R-.-..PC.......-.....---DE....G.....-.........T-----I--------S-----I------------------.--IN---K---------  109
C.ZA.10.DEMC10ZA001           ----N---.P-.........P----............---..........---S--------R-.-..PC.......-.....---VE....-.....-.........TL-C--I--------S-----I------------------.--I----K---------  102
C.ZM.11.DEMC11ZM006           ----T---.P-.....-E---LPP---............---SSP...SAN--S--------R-.-..PR.......-.....----E....-.....-.........-Y----I--------A--V---T------------------.--IN---K---R-----  107

D.CM.10.DEMD10CM009           ---------.Q-.......GKL-----RANSPT............---..........---S---R-.........PL.......-.....-T-TE....G.....-K.........QG.-LS---I------F--V-----------------------.-DID---K---------  111
D.CY.06.CY163                 ------V-.P-.............---..........---S---R-.....--G-.-..PL.......-.....-T--E....G.....-R.........--.-L----I--------S--V---------------------.-DIN---K---------  105
D.KE.11.DEMD11KE003           ---------.PE.......G-LP-G-A............-TH..........---S--PW-.....--G-.-..TL.......-.....-T--E....-.....-.........A-.-S----I-----------E---------------------.--IN---K---------  105
D.KR.04.04KBH8                ---------.P-....W---KLP---N............---..........R--S---R-.....--G-.-..PL.......-.....-T--E....-.....-R.........--.-LG---I--------KV-----I------------------.-DIE---K---------  105
D.SN.90.SE365                 ----N---.Q-.....-E---L-----............---..........---N---R-.....--G-.D..PLSETGASA.....-T--E....-.....-.........--.--S---I------V-AV-------------------------.-DIN-S-K---------  111
D.TZ.01.A280                  ----N---.P-....R----LP----............---..........---S-D-R-.....--G-.K..TL.......-.....-T--ERQ..G.....-.........--.--S---I-----------------------------.--------K---------  107
D.UG.10.DEMD10UG004           ----I---.PE..........LP----............---..........---S-V-R-.....--G-.K..PL.......-.....AT--E....G.....-.........A-.-LSL--I---------V-----------------------.--IN---K---------  105
D.UG.11.DEMD11UG003           ----Y---.P-..........LP--KA............---..........---S---R-.....------.-..T-.......-.....-T--E....-.....RE.RR....A-.-L-L--I---------V-----------------------.---N---K---------  107
D.YE.02.02YE516               ----N---.Q-.....-E-GKL-----............---..........C--S---R-.....------.-..PL.......-.....-T--EGQ..-.....-.........S-.-S----I----------N----Q--------------.-DID---K---------  107
D.ZA.90.R1                    ----N---.P-.........G-------............---..........---S---R-.....--G-.-..PL.......-.....-T-TE....-.....-.........A-.--DL--I---E---------------------------.---N---K---------  105

F1.AO.06.AO_06_ANG32          ----N---.P-.....-E-GK------............---..........---S---R-.....Q-G-.A..--.......-.....----E....G.....R-TGTG.....--S-LS---I------V--V-VE------------------.-DIN---K---------  110
F1.AR.02.ARE933               ----N---.Q-.....-E--K-P----............---..........--AS---R-.....QRG-.-..PL.......-.....----E....-.....R-.........--P-LS---I--------A-N-----------------------.-DIN---K---------  106
F1.BR.10.10BR_RJ015           ----N---.P-.........WKLHP--A............-TI..........--AS---R-.....Q-G-.P..IS...EAGA.....---TE....-.....R-.........--P-LS---I------V--V-V--E-R-------------.-DIN---K---------  109
F1.CY.08.CY222                ----N---.Q-.....-E--KL-----............--I..........--AS---R-.....QRG-.S..PL.......-.....G---E....G.....-.........A-S-LSL--I------V-VV-V-------------------.-DIN---K-R-------  106
F1.ES.02.ES_X845_4            ----N---.P-.....-E-GK------............---..........---S---RI.....QRGN.D..PL.......-.....----E....G.....G-TG.........--S-LSL--I------V----VE------------------.-DIN-S-K---------  108
F1.RO.96.BCI_R07              ----N---.P-.....-E--KL-----............---..........-SA----W-.....Q-GN.-..-L.......-.....----E....G.....R-.........N-S-LSL--I--------S--V-------------------.-DIN---K---------  106
F1.RU.08.D88_845              ----N---.Q-....RE------X--............--I..........-AGG--R-.....R-G-.-..PL.......-.....----E....-.....-.........S-P-LD---I------V--V-V---R-------------.-DIN---K---------  106
F2.CM.02.02CM_0016BBY         ----N---.Q-....RE-W--H----............---..........G-AS-G-R-.....RR--.-..-L......P.....----E....-.....-.........--S-LD---I------V----VE--R-------------.-DIN-S-K---R-----  106
F2.CM.10.DEMF210CM001         ----N---.Q-.....-E-G-------............--EQT.....RAF--AS---W-.....R-G-.-..PL......P.....----E....-.....-.........-AS-L----I------I--V--E-H-R-------------.-DIN-T-K---R-----  111
F2.CM.10.DEMF210CM007         ----T--S.Q-....RE---------............---..........--AS---R-.....-RG-.-..-L......P.....----E....-.....-.........AGV-L-C--I-----------V----R-------------.-DIN---K---------  105

G.BE.96.DRCBL                 ----N---.Q-.....-E----P--A............---..........------R-.....R-G-.S..PL......P.....----E....G.....K-.........-I.-SI---I------I-KVR-----I------------.--ID---K---------  105
G.CM.10.DEMG10CM008           ----N---.Q-.............---..........------W-.....R-G-.S..PQ......P.....-T--E....G.....K-.........A-.-LSL--I--------NVRV----I------------.-DIN---K---------  105
G.CN.08.GX_2084_08            ----N---.Q-.....-E-G--P--A............---..........------R-.....RRG-.-..PL......P.....-T--E....GTGAEGE-.........VT.--SL--I---------VR-----I------------.-QID---K---------  110
G.CU.99.Cu74                  ----T---.Q-....RE-G----K-A............--I..........---C---R-.....RTG-.S..PL......P.....--R-E....G.....E-..........L-L--I---------VR-----I------------.-DIN---K---------  102
G.ES.09.X2634_2               ----N---.Q-..........P--A............---..........---S---R-.....QRG-.S..PL......P.....----E....GAGAEGK-.........VI.PLSL--I---------V------I------------.-DIN---K---------  110
G.GH.03.03GH175G              ----N---.Q-.....-E-------A............--I..........-S-S---R-.....REG-.S..PL......P.....----E....G.....K-G.........-I.-L----I---------VR-----I------------.--IN---K---------  106
G.KE.09.DEMG09KE001           ----N---.Q-....R------P--A............---..........---C---RA.....RRG-.S..PL......P.....--RPE....G.....E-..........A-L----I---------V------I------------.--I----K---------  104
G.NG.09.09NG_SC62             ----N--S.P-....WE-------EA............---..........---S---R-.....R-G-.S..PL......P.....----E....G.....K-.........VT.-LSL--I---------VR-----I------------.-QID---K---------  105
G.ZA.01.TV546                 ----N---.P-.....-E-------A............---..........---S---R-.....R-G-.S..PL......P.....---TE....G.....K-.........DT.--S---I--------AVRVE---I------------.-DLN---K---------  105

H.BE.93.VI991                 ----N---.Q-..........PP-EA............---..........---S---R-.....RRG-.H..PL.......-.....----E....-.....T-.........T------I------I--V--E-----------------.-DIN---K---------  104
H.CF.90.056                   ----N---.Q-....RE--K--P--A............-T-..........---S---R-.....RRG-.D..PL.......-.....----AE...G.....-.........T-LS---I---------V--E---R-------------.--IN---K---------  105
H.GB.00.00GBAC4001            ----N---.Q-....RE---L-P--A............-T-..........---S-M-R-.....R-G-.D..T-.......-.....---DE....G.....K-.........T-L----I---------V--E---R-------------.--IN---K---------  104

J.CM.04.04CMU11421            ----X---.Q-....XE---L-X--X............---..........---S---R-.....R------.XX......P.....-TRSE....-.....-.........--S----I----------------R-----------I.--ID---X---------  102
J.SE.93.SE9280_7887           ---------.Q-....RE---L-P---............---..........---S--PRA.....RRG-...PL......P.....-T--E....G.....-.........--S----I----------R-----R-------------.-DID--RK---------  104
J.SE.94.SE9173_7022           ---------.Q-....RE-----P---............---..........---S--PR-.....RRG-...PL......P.....-T--E....G.....-.........--S----I----------R-----R-------------.--ID---K---------  104

K.CD.97.97ZR_EQTB11           ----V--S.Q-....RE--K------............---..........---S---W-.....R-E-.-..PL.......-.....-T-NE....-.....S-.........-GS------I------V--V-V---R-------------.--IN---K---------  106
K.CM.96.96CM_MP535            ----N---.P-.....-E---------............---..........---S---R-.....R-G-.-..PL.......-.....---DQ....-.....-.........-EP------I------I----V---R-------------.--IN---K---------  106

U.CA.01.TV749                 ----N---.P-.....-E---LPT---RANSPTIPSEQP--..........---S--PGE.....G---.-..LL......P.....AT-.....-.....Q..........--S---IS--K-----VS-----I--------------.-NIN---K---------  111
U.CA.99.TV721                 ----N---.P-.........L-----............S--..........---S--PSE.....QT-A.-..--......TSREPL-G-R-.....N.....-A.........GT.A-S---IS-----I--VN-----I--------------.-NIN---K-R-------  110
U.CD.83.83CD003_Z3            ----N---.Q-.....-E---------............---..........---S---R-.....RRG-.-..PL......A.....----E....G.....-.....V....PLP------I---------V-------------------V.---T---K---------  107
U.CD.90.90CD121E12            ----N---.Q-...............N............---..........-S-S---R-.....RRGS.-H.PL......A.....----E....G.....-.........VS.--S---I---------V-------------------V.---T---K---------  106
U.CY.05.CY090                 ----N---.PE....WE--KLPP---............--I..........---S---RI.....RRG-.-..L-.......-.....-T--E....-.....-.........-GS------I-----------VE---R-------------.KNIN-S-K---------  106
U.CY.08.CY223                 ----N---.Q-.....-E---L-----............-T-..........---S--SG-.....RRG-.SR.PL......P.....-T-TE....G.....-.........RII-LD---I------I-RVT-E---I--------------.-NIN---K---------  107
U.ES.10.DEURF10DZ001          ----N---.Q-....RE--K------............-T-..........---NG--GD.....G---.-..LL.......-.....---TG....G.....-.........-IP-L----I---------VR-E---I--------------.-DIN---K---------  106
U.GR.99.99GR303               ---------.Q-.............--D..........---S---R-.....RRG-.S..PL......P.....----K....G.....-.........AI.--S---I---------VR-----I--------------.-GI----K------R--  105
U.NL.95.U_NL_95_H10986_D1     ----S---.Q-....RE----P----............---..........---SG--R-.....RRG-.-..PL.......-.....---DE....-.....-.........-GS------I------M--VN----V---------------.-DIN---K-R-R-----  106

01_AE.AF.07.569M              ----N---HQ-.........GK------............---..........--AS-.-GD.....G---.-..LI......T.....-T-TE....-.....-.........-SS-YS---I------I-PV-----I--------------.-DIN---K---------  106
01_AE.CF.90.90CF11697         ----N--S.Q-.....-E---------............---..........---S---GD.....G---.-..LL......P.....----E....-.....-.........-PF--S---I-F-------V--E-----------------.-DIN------------  106
01_AE.CN.10.YNFL03            ----N---.-..........K------............---..........---S-K-GD.....G---.T..IL......T.....-----....-.....-.........KSS------I-----------------------------.-DIN---K---------  105
01_AE.HK.04.HK001             ----N--S.Q-.....-E---------............--F..........---SGK-EG.....-...-..LL......T.....----E....G.....-.........-PF---L--I---------V---------------------.-DIN---K---------  103
01_AE.IR.10.10IR.THR48F       ----N---.-.....R-----------............G-I..........--AS-K-GD.....G---.S..LL.......-.....-........-.....D-.........ASS------I---------V--E---R-------------.-DIN---K---------   98
01_AE.JP.x.JRC77AE            ---Q----.P-.........GKL-----............-TI..........---S-K-GN.....G---.......-.....----E....-.....-.........TS--S---I--------AV-------------------V.-DIN---K---------  101
01_AE.TH.09.AA090a_WG11       ----NM--.P-.........G-------............---..........---S---GD.....G--S.-..L-......T.....----E....-.....-.........-PS-LS---I---------V------R-------------.-DIN---K---------  106
01_AE.TH.90.CM240             ----N---.Q-.........G-------............---..........---S-K-GDGGRDNG---.-..LL......T.....----E....-.....-.........-SS--S---I---------V---------------------.-DIN---K---------  111
01_AE.US.05.306163_FL         -------S.Q-.....-E-------A............---..........---S---GD.....G---.-..LL......T.....----E....-.....-.........TS-CS---I---------VR--------------------.-DIN---K---------  104
01_AE.VN.98.98VNND15          ----N---.-..........G------N............---..........---S-K-GD.....G---.-..LL......T.....----E....-.....-.........-SS------I---------V---------------------.-DIN---K---------  105
```

```
                               Pol start
                               Gag-Pol TF start                                                                                    Gag-Pol TF end protease start
B.FR.83.HXB2                   FFREDLAF.LQ....GKAREFSSEQT............RAN.........SPTRRELQV.....WGRDNN..SP......S.....EAGAD....R.....QG........TV.SFNFPQVTLWQRPLVTIKIGGQLKEALLDTGADDTVL.EEMSLPGRWKPKMIGGI 106
02_AG.CM.10.DE00210CM013       ----N---.Q-....-E--------A............-T-..............---S--PWD....G--E.T..LL......-.........----E....G.....--.........IIS--S---I---------VR-----I-------------.--IN---K--------- 106
02_AG.CY.09.CY256              ----N---.Q-....-E--KL-P---............-T-..............---SG--WD....G---.-..--......-.........---DG....G.....--.........-IS-L----I------II-VR----PI------------.-DID---K--------- 106
02_AG.ES.06.P1423              ----NX--.Q-....-E--Q-----............GT-..............--AS---GD....G--X.T..PL......P.........---TK....G.....--.........-IS-LS--HI------V--V------X-------------.--ID------------ 106
02_AG.GW.05.CC_0048            ----N---.Q-....-E---------............GT-..............---S-K..H....G---.D..L-......-........NT-TG....G.....--.........--S------I---------VR-E---I-------------.--IN------------ 104
02_AG.KR.12.12MHR9             ----N---.Q-....-E---------............GT-..............---S---RD....G---.-..L-......-EAGA.-G--E....G.....--.........-IS-L----I---------VR-----I-------------.--IN---K-R------- 110
02_AG.LR.x.POC44951            ----N---.Q-....RE--KL-----............GT-..............---G---WD....G---.-..L-......-.........---TE....G.....--.........-IS------I---------VR-----I-------------.-DIE------------ 106
02_AG.NG.09.09NG_SC61          ----N---.Q-....-E--K------............GT-..............---S--CWD....E--G.-..LL......-.........---TR....G.....--.........SLP---L--I---------VN-----I-------------.--I------------- 106
02_AG.NG.x.IBNG                ----N---.Q-....-E--K------............GT-..............-S-S---WD....G---.T..-L......-........T--TE....G.....--.........AIS------I---------VR-E---I-------------.-DIN---K--------- 106
02_AG.SE.94.SE7812             ----N---.Q-....-E--K------............GT-..............--AS--PWD....R---.-..LL......-.........---TG....G.....--.........-IS-LS---I---------V------I-------------.--IN---K--------- 106
02_AG.SN.98.98SE_MP1211        ----N---.Q-....-E--KL-----............GT-..............---S---WD....G---.-..LL......P.........---TG....G.....--.........-IP------I-------F-VR-E---IK-------N----.-KIN---K--------- 106
03_AB.RU.97.KAL153_2           ----N---.Q-....RE--K------............--I..............---S-K-WD....G---.-..PL......P.........-T-TE....G.....--.........-AS------I---R-----VR--------------------.-DIN---K--------- 106
04_cpx.CY.94.94CY032_3         ----NV--.Q-....RE--K-----A............---..............--A-GM-RE....ERG-.-..LL......-.........---TE....G.....--.........-I.------I-----------L---IR-------------.--IN---K--------- 105
05_DF.BE.x.VI1310              ----S---.P-....-E---LPP---............G-L..............--AS------....--G-.S..LL......-.........----E....G.....R-.........--P-LS---I------V---R--------------------.-DIN---K--------- 106
06_cpx.AU.96.BFP90             ----N---.Q-....-E-------A............---..............---H---RF....RRG-.S..PL......P.........-T-VE....G.....E-GKG.....AI.-LSL--I---------VR-----I-------------.-DIN---K--------- 108
07_BC.CN.98.98CN009            ----N--L.P-....----------............---..............----G-----....-------.-I......-.........---DN....-.....--.........-I.------I-------------------------------.-D-N---K--------- 106
08_BC.CN.97.97CNGX_6F          ----I---.P-....-E----PP---............---..............---S------....R-.-.-..PS......-.........---TE....-.....--...........TL----I--------S--V---I--------------.--VN---K--------- 102
09_cpx.GH.96.96GH2911          ----N---.Q-....RE---L-X---............G-I............G-XS---WD....E---.X..LL......-........XXXEG..........R-.........-AP--X---I------I--V-------------------------.--IN---K--------- 105
10_CD.TZ.96.96TZ_BF061         ----N---.Q-....R----LP----............---............-S-S---R-....--G-.-..TL......-.........-T--E....-.....--.........AA.-LS---I-------------------------------.---N---K-----M--- 105
11_cpx.CM.95.95CM_1816         ----N---.Q-....-E-----P--A............---..............---S---R-....RRGG.S..PL......P.........-T--E....-.....G-.........AI.--S---I------I--V--A---R---------------.--IN---K--------- 104
12_BF.AR.99.ARMA159            ----N---.Q-....-E--K-P---A............---..............--AS---W-....RRG-.-..LL......-.........----E....-.....R-.........--P-LS---I--------I--V-----------------------.-DIN---K--------- 106
13_cpx.CM.96.96CM_1849         ----N---.Q-....RE---------RTNSPTREQT..---..............---S---R-....RRG-.S..PL......P.........---TE....G.....--.........DI.-LSC--I------I-IV---E--I-------------.K-IN------------ 113
14_BG.ES.05.X1870              ----N---.Q-....-E-----P--A............---..............--------W-....RRG-.S..PL......P.........--R-E....G.....K-.........D-.PLSL--I---------V------I-------------.-DIN---K--------- 105
15_01B.TH.99.99TH_MU2079       ----T---.-.....-----------............--I..............---S---GD....G---.-..LL......T.........----E....-.....--.........-YS--S---I---------------------P----P--.-DINF--K-N------- 105
16_A2D.KR.97.97KR004           ----N---.P-....RE--------N............-T-..............---S-G-WN....G-G-.-..PL......A.........----E....K.....--.........-TH-C----I---------V--E---R--------------.-DIN---K--------- 106
17_BF.AR.99.ARMA038            ----N---.Q-....-E--K-P---A............---..............--AI---R-....R-D-.-..P-......A.........----E....-.....R-.........--P-LS---I--------E--V-----------------------.-DIN---K-R------- 106
18_cpx.CU.99.CU76              ----N---.Q-....-E-----P-KA............---..............---SG--G-....RRG-.S..PL......P.........--RVE....G.....GE..........I--L---I---------V------I-------------.-DIN---K--------- 104
19_cpx.CU.99.CU7               ----N---.P-....-E--KLP----............---..............--------R-....--GN.-..PL......-.........-T-SE....-.....--.........--.-LSL--I-------------------------------.-DIN---K-R------- 105
20_BG.CU.99.Cu103              ----T---.P-....-------P--A............---..............---C---R-....RRG-.S..-L......P.........--R-E....G.....E-...........L-L--I---------VR-----I-------------.-DIN---K--------- 102
21_A2D.KE.99.KER2003           ----N---.--....-E-G-L----V............---..............---S-DVR-....R----.-..PL......-.........-T--E....-.....--.........--.-LS---I-------------------------------.---N---K--------- 105
22_01A1.CM.01.01CM_0001BBY     ----N---.Q-....-E---L--G--............---..............---S--RGD....G---.-..LL......P.........----E....-.....E-.........-AS--S---I-------------------------------.-DID---K--------- 106
23_BG.CU.03.CB118              ----T---.P-....------L---A............---..............---C---RI....RRG-.S..PL......P.........--R-E....G.....A-...........L-L--I---------VR-----I-------------.-DIN---K--------- 102
24_BG.ES.08.X2456_2            ----T---.P-....----G-L---A............---..............---C--PR-....RRG-.S..PL......P.........--R-A....G.....ER...........L-L--I---------V-V----I-------------.-DIN---K--------- 102
25_cpx.CM.02.1918LE            --------.Q-....RE---L-P--A............--D..............---S--PR-....--G-.-..PL......P.........----E....-.....--.........VI.-YG---I---------V-----II-------------.-DIN---K--------- 105
26_AU.CD.02.02CD_MBTB047       ----I--S.Q-....-E--------V..........S--..............-S-S---GN....GRG-.-..PL......P.........-TERG....-.....--.........--S-S----I------------E---R--------------.-DIN-S-K--------- 104
27_cpx.FR.04.04CD_FR_KZS       ----N-VS.Q-....RE-----P--A............---..............---S--PRD....H-G-.D..LL......-.........-T--E....G.....--..........T--S---I---------V--E---R--------------.--IN-Q-K--------- 104
28_BF.BR.99.BREPM12609         ----N---.P-....-E--K-P---A............---..............--AS------....RRG-.-..PL......-.........----E....-.....G-.........--P-LSL--I-----------V---R--------------.-DIN---K--------- 106
29_BF.BR.01.BREPM16704         ----N---.P-....-E---P---A............--I..............--AS---R-....-R-E.-..PL......-.........----E....-.....R-.........E-P-LS---I-----------V-----------------------.-DIN---K--------- 106
31_BC.BR.04.04BR142            ----N---.P-....-E---PP---............---SPASEQTRAD--AS------....R-.-.-..PL......-.........----E....-.....--...........TL----I-----------V-----------------------.--IK---N--------- 112
32_06A1.EE.01.EE0369           ----N---.Q-....-E---L----A............---..............---C------....R-G-SS..PL......L.........----E....E.....K-.........AI.-LS---I------I--VR--E--I-------------.-DIN---K--------- 106
33_01B.ID.07.JKT189_C          ----A---.P-....-E---LP-KP-............---..............---S------....--T---.-I......-.........-----....-.....--..........AI-CS---I-----------------R--------------.--IN---K--------- 106
34_01B.TH.99.OUR2478P          ----I---.Q-....---G--P----............---..............----G-----....------.-I......-.........-T---....-.....--..........TI--S---I---------V-----------------------.-DIN---K--------- 106
35_AD.AF.07.169H               ----N---.Q-....RE--K---G--............---..............---S---R-....-----K.AL......-.........-T--K....-.....R-.........V-.-S----I---------V-L---QR--------------.-DIN---K--------- 105
36_cpx.CM.00.00CMNYU830        ----N---.Q-....-E---------............-T-..............---S---WD....R---.V.LF......-.........---IE....G.....--.........-IS-L----I---------VR-----I-------------.-DIN---K--------- 106
37_cpx.CM.00.00CMNYU926        ----N---.Q-....RE--K------............G--..............---S---RG....ERGS.S.LL......-.........-----....G.....--.........-IP--S---I------IT-V------I-------------.-DIN---K--------- 104
38_BF1.UY.03.UY03_3389         ----N---.P-....-E-GK-P---A............--I..............---S-KFW-....RRG-.-..P-......-EAGA.----E....G.....R-.........N-P-LS---I-------------------------------.-DIN---K--------- 110
39_BF.BR.04.04BRRJ179          ----N---.Q-....-E--------N............---SPN...........---S--PN-....--G-.-S.-L......P.........-T--E....G.....--.........--.---L--I------I---V---R--------------.-DIN---K-T------- 109
40_BF.BR.05.05BRRJ055          -------S.--....-E---P--A............---..............--AS---W-....RRG-.-..P-......-.........----E....K.....R-.........--S-L----I------V--V----------------N---.-DIN---K--------- 106
42_BF.LU.06.luBF_18_06         ----N---.Q-....-E--K-PP--A............--KD.....RAD--AS---W-....R-G-.-..LL......-.........----E....-.....R-.........S-P-L----I-----------V-E-I---------------.---N------------ 110
43_02G.SA.03.J11223            ----N---.P-....-E---LP---A............---..............---C---R-....RRG-.S..-L......P.........----E....G.....E-.........IT.-ISL--I--------KV---E--I-------------.-QID------------ 105
44_BF.CL.00.CH80               ----N---.Q-....-E--K-P----............---..............---S---W-....RRG-.-..PF......-.........----E....G.....R-.........K-P-LC---I--------K--V---I-------------.-DIN-----T------- 106
45_cpx.FR.04.04FR_AUK          ----N---.Q-....RE---L-----............G-I..............--SS---WD....G---.-..LL......-.........----E....G.....--.........AG..LS---I------V--V-V-----------------------.--IN---K--------- 104
46_BF.BR.07.07BR_FPS625        ----N---.Q-....-E--K-P---A............---..............--AS---RD....QRG-.-..PL......-.........----E....-.....R-.........-LP-LS---I-----------V----R--------------.-DIN---K--------- 106
47_BF.ES.08.P1942              ----N---.Q-....-E-----P--A............---..............--AS---G-....-RG-.-..PL......-.........---TE....-.....R-.........--S-LS---I---------V-V--M--------------.-DIN---K--------- 105
48_01B.MY.07.07MYKT021         ---KA---.P-....-E---LY----............---..............---SG------....------.-I......-.........-----....-.....--..........TI--SL--I-------------------------------.---N-----R------- 106
49_cpx.GM.03.N26677            ----N---.P-....-E-----P---............---..............---SG--R-....RR--.S..PL......P.........----E....-.....G-.........--.--S---I----------R-----------------------.--ID---K--------- 104
50_A1D.GB.10.12792             ----N---.P-....R-----PP---............---..............---S---R-.......---S.-..TF......P.........-T--E....-.....--.........A-.--S---I---------VRV---------------------.---N-Q-K--------- 105
51_01B.SG.11.11SG_HM021        --------.P-....---G-------............---..............---S-------....--G--.-.-L......-.........-----....-.....--.........--.-LSL--I-------------------------------.--ID------------ 106
52_01B.MY.03.03MYKL018_1       ----N---.Q-....R--G------A............---..............---S-K-GD....G--.D-..LL......T.........----X....G.....--.........SSS--S---I---------V-V----R--------------.-DID---K--------- 106
53_01B.MY.11.11FIR164          ----A--X.P-....-E-X-LP-XP-............---..............----G--P-....------.-I......-.........-----....-.....--..........AI------I-----------VA----------------------.---N------------ 106
54_01B.MY.09.09MYSB023         ----N--L.P-....------L-----............---..............----G------....G-----.-T......-.........-----....-.....--.........-I.-CVC--I------I-------E--------------.-DIN------------ 106
55_01B.CN.10.HNCS102056        ----AV--.Q-.......--GK------............---..............---S-G-GD....G---.-..LL......T.........K---E....-.....--.........-SP-CS---I-----------E-----------------------.-DIN---K--------- 106
56_cpx.FR.10.URF5_patient_A    ----N---.Q-....-E--K------............GT-..............---S--RWD....G--S.-..LL......-.........---VKGQ..G.....--.........-IP-L----I---------VR-----I-------------.-DIN---K--------- 108
57_BC.CN.09.09YNLX19sg         ----AV--.P-....-E---P-----............---..............------------....R-.-.-..PS......A.........--RTE....-.....--...........TL----I--------S-RV---I--------------.--IN---K--------- 102
58_01B.MY.09.09MYPR37          ----I--S.Q-....RE---LP-XP-............--X..............----G------....--X---.-X......-.........-----....-.....--..........TIP-D---I-------------------------------.-DID---K--------- 106
59_01B.CN.09.09LNA423          ----N---.-.....R---K--PK--............---..............---S-K-GD....G-GE.....L......T.........---TE....-.....--.........-AS-L----I---------V-----------------------.-DIN---K---R----- 103
60_BC.IT.11.BAV499             --------.P-....-E--K-PP---............GT-..............-T-S-------....R-.-.-.........A.........----E....-.....--............T-----I------V----V-------P------------.--TK---N--------- 101
61_BC.CN.10.JL100010           --------.P-....-E----PP---............---..............--AS------....R-.-.S..LS......P.........---NE....-.....--............T-----I--------S--V---I--------------.--IN---K--------- 102
62_BC.CN.10.YNFL13             ----N---.P-....-E--K--PK-A............---..............---S------....RR.-.-..PS......-.........---IE....-.....--............T-----I--------P-IV--M--------------.--IN---K--------- 102
63_02A1.RU.10.10RU6637         ----T---.Q-....RE--K------............---..............---S--RGD....G---.-..LL......-........A---E....G.....--.........AIS--S---I---------VR-E---I-------------.-DID---K--------- 106
64_BC.CN.09.YNFL31             -------L.P-....----------............--I..............---S--F---....------.-I......-.........---G-....-.....--.........-I.PLS---I--------S--V---I--------------.---N---K--------- 106
65_cpx.CN.10.YNFL01            ----N---.P-....-E----PP---............--D..............---S------....RR.-.-..PS......-.........---VE....-.....--...........TL----I--------S-RV---I--------------.--IN---K--------- 102
67_01B.CN.11.2011.ANHUI.HF115  ----N---.Q-.......---X---X---............---..............---S-G-GD....GR--.-..LL......T.........--R-E....-.....--............T-LS---I---------V-----------------------.-DIN---K--------- 104
68_01B.CN.11.2011.ANHUI.WH73   ----N---.Q-....X-----------............---..............---S-G-GD....GR--.-..LL......T.........--R-E....-.....--............TALS---I---------V-V----------------------.-DIN---K--------- 104
70_BF1.BR.10.10BR_PE004        --K-N---.P-....-E-KK-Y---A............---..............---S------....RR--.D..P-.......-E........---GAE....-.....RR.........--P-L----I-----------V---I--------------.XDIN--XKX---X-X-- 108
71_BF1.BR.10.10BR_PE008        ----NV--.Q-....-E--K-P---D............---..............--AS---R-....R---.D..PL......-.........----E....K.....R-.........--P-LS---I--------K-----------------------.-DID---K--------- 106
72_BF1.BR.10.10BR_MG002        ----N---.P-...R-E---LP---N............G--............GSAS---R-....RRGE.-..PL......-.........----E....G.....R-.........--P-LS---I-------M---RV----------------------.-DIN---K--------- 107
O.BE.87.ANT70                  ---QI--S.GG....HE--QLCA-TS............TPI............---DGGGSE....GTGES.....................GTERG....P.....ER.........AL.-VCL--IP--D--I--ARV--H-C-V------------.NNIQ-E-K--------- 102
O.CM.98.98CMA104               ---QI--S.GG....HE--Q-YA-TS............GPI............--SDGGSSE....GTGEP.....................GTERG....L.....KR.........-L.-VCL--IP--D--V--A-V--H-C-V------------.NDIQ-E-K--------- 102
O.CM.98.98CMABB141             ---QI--S.GG....RE--QLYA-TS............API............---NGGNDE....GT-ES.....................GTKGG....H.....NR.........-L.--C---IP--D--V-IARVE-H-C-V------------.NNIQ-E-K-T------- 102
O.CM.98.98CMABB212             ----V--S.GG....HE--QLXA-TS............VP-............---XGGDSK....GAGES.....................-SRXX....X.....GR.........--.PVCL--IP--D--I--A----H-C--------------.NNIQ-E---T------- 102
O.CM.98.98CMU5337              ----I--S.GG....HE--QLCT-TS............VT-............---NGGGSE....ET-ES.....................GTEGR....P.....ER.........AL.-VCL--IP--D--V--ARV--H-C-V------------.NNIQ-E-K-T------- 102
O.CM.99.99CMU4122              ----I--S.GG....HE--QLCT-TS............VPI............-SSDGGGKE....GT-ES.....................GTEGG....P.....ER.........-L.-VCL--IP--D--V--ARV--H-C-V------------.NNIQ-E-K-T------- 102
O.FR.92.VAU                    --K-I--S.ER....HE--QLCAGTS............VPI............---NGRDDE....GA-GGS....................-SEGG....A.....ER.........AL.-VCL--IP--D--V--ARV----C-V------------.NNIQ-E-K-A------- 103
O.GA.11.11Gab6352              ---QI--S.GG....HE--QLCA-TS............APL............---DGGGSA....ET-ES.....................GTEGG....P.....KR.........AL.-VCL--IP--D--V--ARV--H-C-V------------.NNIQ-E-E-T------- 102
O.SN.99.99SE_MP1299            ----I--S.GG....HE--QLCA-TS............VPI............---DDGGNE....GT-ES.....................GTEGG....P.....ER.........-L.-VCL--IP--D--V--ARV--H-C-V------------.TNIQ-E-K-T------- 102
O.US.10.LTNP                   ---QV--S.GG....HE--QLCT-TS............GPI............-S-DGGGGK....QTGES.....................GTERE....P.....NR.........-L.-VCL--IP--D--II-A-V--H-C-V------------.NNIQ-E---T------- 102

N.CM.02.DJ00131                ---KG-VS.--....RET-QLPPDNNNKE..........--H..............--AT---W-....S-GE.E..HT......G.....-GE-G....E...........HRELSIPTL----I------VI-V---KEVR------------I.--IQ-E-K--------- 111
N.CM.02.SJGddd                 ---QG-VS.--....RET-K-PPDNSKE..........G-H..............---T-.........-GE.E..HT......G.....-RD-G....E.....P-...EDRELS-PTL----I------VI-V---EEVR------------I.--IQ-E----------- 108
N.CM.04.04CM_1015_04           ---KG-VS.--....RET--LPPDSNKE..........--H..............--AT---W-....S-GG.E..HT......G....KGD-G....E.....P-...EDRELS-PTL----I------VISV---NEVR------------I.--IQ-E----------- 113
N.CM.06.U14296                 ---KG-VS.--....RET--LPPDNNKE..........--H..............--AT---W-....S-GE.E..HT......G....KGD-R....E.....P-...EDRELS-PTL----I------VI-V---KEVR------------I.--LQ-E-K--------- 113
N.CM.06.U14842                 ---QG-VS.--....RET--LXPDNNNKE.........--H..............--AT-G-W-....S-GE.E..HT......E....KGD-G....G.....P-...EDRELS-PTL-L--I------II-V---KEIR------------I.--LQ-E-K--------- 114
N.FR.11.N1_FR_2011             ---KG-VS.--....RET-KLPPDNNKE..........GTH..............--AT---W-....S-GE.E..HT......G.....-GD-G....G.....P-...EDRELS-PTL----I------VI-V---EEEI------------I.-DLQ-E---R------- 113

P.CM.06.U14788                 -------S.GG....QE--QLCT-TS............TTD..............---YGGGGD.......-..............-EQA..-G-KG....G.....RE..R....V-.P-SL--I---D--VIPV----HIC------------V.DNLP-E----------- 103
P.FR.09.RBF168                 ---K---S.WG....QE--QLCT-TS............TT-..............---YGGGRD.......-............PGQA..-GXGG....G.....X-.........GI.P--L--I---D--VIPVRV--HIC------------V.DKLP-E-----N----- 102

CPZ.CD.06.BF1167               ----THPL.VG....VQT-K-CA-SESDSESGT.....................ESDTG-TSG....ES-..T..LSS......-......NG-GESL..EQ.....--..............WL--EIH-----MMEVTVQ--VCT-----------F.ADLK-K-SCN-RT---- 109
CPZ.CM.05.SIVcpzMT145          ---KK--S.-G....-E---LCAR-E............GT-..............-S-YGI-RI....S-GE.E..-Q......G.....GGEGG....E.....-A.........A-.------I-------IPV--E---I------------I.-DLN---K-S-------- 105
CPZ.GA.88.GAB1                 ----R---.P-....RE--QLCA--N............-T-............G--D---W-....P-GR.E..E-......-.....GEERG....-.....EQ.........SI.-T-L--I-------IPV-VE---C------------I.-RIQ-Q-L--------- 104
CPZ.TZ.06.SIVcpzTAN13          ----THPL.VG....VQT---CAK-P............-EG............EAPGARG......SA.-.SAYC-......-N....GD-PH....E.....RR.............VL--IN-----MMAVRVQ-H-CQ---------S-F.CNLQ-K-Q----T---- 101
CPZ.US.85.US_Marilyn           ----T-VP.IVERGIKET--LPGK-E............G-H............-S-N-GFR-....SR--.S..DT......G.....-T-KG....-.....K-.........AL.-A----I-------LR--VA--IV--------------.DNIQIE-T-R---M--- 108

GOR.CM.07.SIVgor2139_287       ----I--S.RG....RE--QLCA-AS............AT-..............---NGGGRK.....................G.....DEETR....K.....G-.........-L.-S----I---D--VIPV----H-C------------V.-HL--E-----R----- 98
GOR.CM.07.SIVgorCP2135con      ----I--S.RG....RE--QLCA-AS............AT-..............---NGGGRE.....................R.....NEETR....E.....G-.........-L.-S----I---D--IIPV----H-C------------V.-HL--E-----R----- 98
```

protease end_p66, p51 RT start
M41L
K70R
D67N
D110 catalytic site
B.FR.83.HXB2 GGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD 276
A1.AU.03.PS1044_Day0 276
A1.CH.03.HIV_CH_BID_V3538 276
A1.CY.08.CY236 273
A1.ES.05.X1608_8 280
A1.KE.11.DEMA11KE001 276
A1.RU.11.11RU6950 276
A1.RW.11.DEMA11RW002 276
A1.SN.01.DDI579 275
A1.UG.11.DEMA110UG001 276
A1.ZA.04.04ZASK162B1 275
A2.CM.01.01CM_1445MV 276
A2.CY.94.94CY017_41 276
B.BR.10.10BR_MG029 276
B.CA.07.502_1191_03 276
B.CH.08.M2_0803101_NFLG8 280
B.CN.10.DEMB10CN002 276
B.ES.10.DEMB10ES002 276
B.FR.11.DEMB11FR001 275
B.GB.05.MM45d213_GN1 275
B.HK.06.HK003 280
B.HT.05.05HT_129389 275
B.JP.12.DEMB12JP001 279
B.KR.07.HP_18_07JHS10_3909 276
B.PE.07.502_0525_wg5 280
B.RU.11.11RU21ln 280
B.TH.08.MERLBDTRC10 276
B.US.11.ES38 280
C.AR.01.ARG4006 272
C.BR.07.DEMC07BR003 272
C.BW.00.00BW5031_1 278
C.CN.10.YNFL19 276
C.CY.09.CY260 279
C.ES.08.X2363_2 279
C.ET.02.02ET_288 278
C.IN.09.T125_2139 273
C.KE.00.KER2010 272
C.MW.09.703010256_CH256.w96 274
C.TZ.08.707010457_CH457.w8 272
C.YE.02.02YE511 270
C.ZA.08.705010534_CH534.w12 279
C.ZA.10.DEMC10ZA001 272
C.ZM.11.DEMC11ZM006 277
D.CM.10.DEMD10CM009 281
D.CY.06.CY163 275
D.KE.11.DEMD11KE003 275
D.KR.04.04KBH8 275
D.SN.90.SE365 281
D.TZ.01.A280 277
D.UG.10.DEMD10UG004 275
D.UG.11.DEMD11UG003 277
D.YE.02.02YE516 277
D.ZA.90.R1 275
F1.AO.06.AO_06_ANG32 280
F1.AR.02.ARE933 276
F1.BR.10.10BR_RJ015 279
F1.CY.08.CY222 276
F1.ES.02.ES_X845_4 278
F1.RO.96.BCI_R07 276
F1.RU.08.D88_845 276
F2.CM.02.02CM_0016BBY 276
F2.CM.10.DEMF210CM001 281
F2.CM.10.DEMF210CM007 275
G.BE.96.DRCBL 275
G.CM.10.DEMG10CM008 275
G.CN.08.GX_2084_08 280
G.CU.99.Cu74 272
G.ES.09.X2634_2 280
G.GH.03.03GH175G 276
G.KE.09.DEMG09KE001 274
G.NG.09.09NG_SC62 275
G.ZA.01.TV546 275
H.BE.93.VI991 274
H.CF.90.056 275
H.GB.00.00GBAC4001 274
J.CM.04.04CMU11421 272
J.SE.93.SE9280_7887 274
J.SE.94.SE9173_7022 274
K.CD.97.97ZR_EQTB11 276
K.CM.96.96CM_MP535 276
U.CA.01.TV749 281
U.CA.99.TV721 280
U.CD.83.83CD003_Z3 277
U.CD.90.90CD121E12 276
U.CY.05.CY090 276
U.CY.08.CY223 277
U.ES.10.DEURF10DZ001 276
U.GR.99.99GR303 276
U.NL.95.U_NL_95_H10986_D1 276
01_AE.AF.07.569M 276
01_AE.CF.90.90CF11697 274
01_AE.CN.10.YNFL03 275
01_AE.HK.04.HK001 273
01_AE.IR.10.10IR.THR48F 268
01_AE.JP.x.JRC77AE 271
01_AE.TH.09.AA090a_WG11 276
01_AE.TH.90.CM240 281
01_AE.US.05.306163_FL 274
01_AE.VN.98.98VNND15 275

```
                                                                          K70R
                         protease end p66, p51 RT start                M41L                  D67N                                   D110 catalytic site
B.FR.83.HXB2                 GGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD 276
02_AG.CM.10.DE00210CM013     -------------M----K-------------------M-----------------------------------------------------------------------------------------------------------------R----------------- 276
02_AG.CY.09.CY256            ------------X-----K-------------------X-----------------------------R--------------TD--L---------------------------------------------------------------------------------- 276
02_AG.ES.06.P1423            ------------------K---------------X---M----------------X-----X---------R--------X--TD---D--------R--X--------X------------------------------------------------------------ 276
02_AG.GW.05.CC_0048          ------------------K-------------------M-------------S------------------------------TD--K---------------------------------------------------------------------------------- 274
02_AG.KR.12.12MHR9           ------------------K-------------------M--------------------------------------------TD-----------------------------------------------------------------E------------------- 280
02_AG.LR.x.POC44951          ------------P-----K-------------------M---------L-----K----------------------------TD--A--------------------------------------------------------------E------------------- 276
02_AG.NG.09.09NG_SC61        ------------S-----K-------------------M--------------------R-----------------------TD------------------------------------------------------------------------------------- 276
02_AG.NG.x.IBNG              ------------------K-------------------M--------------------------------------------TD------------------------------------------------------------------------------------- 276
02_AG.SE.94.SE7812           ------------------K-------------------M--------------------------------------------TD------------R----------I------N--R-------------------I------------------------------- 276
02_AG.SN.98.98SE_MP1211      -----L------------K--M-S------------K-I-------------------L--------------------L-H-TD--A-------------------------R-------------------------------------------------------- 276
03_AB.RU.97.KAL153_2         ------------------K-------------------M---L----------------T-----------------------TD--K------------------------------------G--------------------------------------------- 276
04_cpx.CY.94.94CY032_3       ------------P-----K-------------------M---L----------------------------------------T----D-------------------I------------------------------------------------------------- 275
05_DF.BE.x.VI1310            -------------V----Q----------------------------------------------------------------T----D--------R----------I---------R-----------------------------------I--------------- 276
06_cpx.AU.96.BFP90           ------K-----------K------------I------M--------------------------------------------T---K------------------------------------------------------------------------------I--- 278
07_BC.CN.98.98CN009          ---------E--P-----------------------------L----------------------------------------TA--D--------T-----------I-----------------------------------------R------------------Y 276
08_BC.CN.97.97CNGX_6F        ---------E--P-----K-------------------M---L----------------------------------------TA--D--------T----D------I----R---S---------------------------------------------------- 272
09_cpx.GH.96.96GH2911        ------------------Q-------------------M------------------------------I-------------T---A---------------------------------------------------------------------------------- 275
10_CD.TZ.96.96TZ_BF061       ------------------Y----------------------------------------------------------------T---K---------R----------I-----------------------------------------------------------Y 275
11_cpx.CM.95.95CM_1816       ---------EE-----E-K-------------------M--------------VD----------------------------T-------------------------------------------------------------------------------------- 274
12_BF.AR.99.ARMA159          ------K---NV--------------I---------------L----------------------------------------T-------------------------------------------------------------------------------------- 276
13_cpx.CM.96.96CM_1849       ----------NMS---E-K-------I-----------M--------------------------------------------T---K---------------------------------------------------------------------------------- 283
14_BG.ES.05.X1870            ------------------K------------I------M------------------------------I-------------I-------R----------------I-------------------------------------S----------------------- 275
15_01B.TH.99.99TH_MU2079     ------------P-----K-------------------M---L----------------------------------------T---K---E----------------I-------------------------------------------------V----------- 275
16_A2D.KR.97.97KR004         ------------T---EKR-------------------M-V-L----------------------------------------T---K--------------------------------------------------------------M---------R--------H 276
17_BF.AR.99.ARMA038          ------K---D-------------------------------L------------------------------------------------------------------------------------------------------------------------------- 276
18_cpx.CU.99.CU76            ------------------KR------------------M--------------------------------------------T------------------------I------------------------------------------------------------- 274
19_cpx.CU.99.CU7             ------------P-----Q----------------------------------------------------------------T----D-------------------I------------------------------------------------------------- 275
20_BG.CU.99.Cu103            ------------------K------------I------M---------------D------------------------------------------------------------------------------------------------------------------- 272
21_A2D.KE.99.KER2003         ------------PL---------------------------------------------Q-----------------------T---MD--------R----------I--------------------------------------X---------------------S 275
22_01A1.CM.01.01CM_0001BBY   ------------------K-------------------M--------------------A---------I-------------T---K--------------------I-----------------------------------------------------------Y 276
23_BG.CU.03.CB118            ------------P-----K------------T------M-------------------------------------------------L--------------------------N------------------------------------------------------ 272
24_BG.ES.08.X2456_2          -------------M----K------------I------M-----------------------------------------------------------------------------------------------------------S----------------------- 272
25_cpx.CM.02.1918LE          ----------E-VL--S-K------------I------I--------------------------------------------T---ND--------R----------I----------------------------------------------S-------------- 275
26_AU.CD.02.02CD_MBTB047     -----------VT-D---K-------------------M--------------V----------------------------IT---N----------V---------I------------------------------------------------------------- 274
27_cpx.FR.04.04CD_FR_KZS     ---------EEVTVD---K------------------------------------------------------------------------------R---------------R----------------------------------------I--------------- 274
28_BF.BR.99.BREPM12609       ------K---D---D------T----------------M------------------------------------------------------------------------------------------------------------------------------------ 276
29_BF.BR.01.BREPM16704       ------K---D---------------------------M--------------------------------------------M-----------------------------------------------K-------------------------------------- 276
31_BC.BR.04.04BR142          ------------------K-------------------M---L----------------------------------------T---N---R-----------------------------------------------------------N------------------ 282
32_06A1.EE.01.EE0369         ------------H-----K-------------------M---------L-------------------R-----------R--T------------------------I------------------------------------------------------------- 276
33_01B.ID.07.JKT189_C        ------------------K-------------------M--------------------------------------------I-------------------------------------------------------------------------------------- 276
34_01B.TH.99.OUR2478P        ------------P-----------I-----------------L----------------------------------------A---A--------------------I------------------------------------------R------------------ 276
35_AD.AF.07.169H             ------K-----A-----K-------------------M---------------D----R-----------------------T------------------------I------------------------------------------------------------- 275
36_cpx.CM.00.00CMNYU830      ------------P-----K------------I--V---M--------------------A-----------------------T---A-------------------------R--------------------------------------R----------------- 276
37_cpx.CM.00.00CMNYU926      ------------P-----K-------------------M---------------N----S---------------S-------T---N---------RV---------I-------------------------------------------R----------------- 274
38_BF1.UY.03.UY03_3389       ------K---N-C----D--VT----I---------------L----------------------------------------M---AD------------------------------------------K------------------R------------------- 280
39_BF.BR.04.04BRRJ179        --IV-----EN-P------------------A------------------------------------R--------------L---A-L-EA------------------------------M-L-------------------------N--------------I--H 279
40_BF.BR.05.05BRRJ055        ----R-K--ED-P------------------------D---------------------------------------------T-------------------------------N--R--------------------------------------------------- 276
42_BF.LU.06.luBF_18_06       ------K---Y-S-D------T----I---------------L----------------------------------------T---M---------------------------------------------------------------------------------- 280
43_02G.SA.03.J11223          ------K-----P---E-K------------I------M-------------------------------------A------T---L--------------------I------------------------------------------------------------- 275
44_BF.CL.00.CH80             ------K--ND-T---------------------------------------------------------------G------I-------------------------------------------------------------------------------------- 276
45_cpx.FR.04.04FR_AUK        ------------V-----K----A--------------M--------------------------------------------T------------------------I------N------------------------------------------------------ 274
46_BF.BR.07.07BR_FPS625      ------K---D---------------------------M---------------D-------------R--------K-----T---A----------------------------------------------------------------R----------------- 276
47_BF.ES.08.P1942            ------------I-----K----------------------------------------------------------------I-------------------------------------------------------------------------------------Y 275
48_01B.MY.07.07MYKT021       ------------P----------------------------------------------------------------------M------------------------------------------------A-------------------R--S-------------- 276
49_cpx.GM.03.N26677          ---------ND-----E-K--V----------------M---L----------------------------------------T-------------------------------------------------------------------------------------- 274
50_A1D.GB.10.12792           ------------P-------------I--------------------------------------------------Q-----T---AD--E----T-----------I-----------------#------------------------------------------H 274
51_01B.SG.11.11SG_HM021      -----------VSVD---------------------------------------------------------------------------------------------I------------------------------------------------------------- 276
52_01B.MY.03.03MYKL018_1     ------------------K-------------------M---------------D----------------------------I-------------------------------------------------H------------------------------------ 276
53_01B.MY.11.11FIR164        ---------E---L----K-------------------M---L-----------N----------------------------T---K---E----------------I-------G----------------------------------------------------- 276
54_01B.MY.09.09MYSB023       ------------P-----------------------------L--------------------N----R--------------I-----L-------------------------N--R------------------------------E----M--------------- 276
55_01B.CN.10.HNCS102056      ------------------KR------------------M---------------D----------------------------T---K---E----------------------------------------------------------R------------------- 276
56_cpx.FR.10.URF5_patient_A  ------------SV----KR------------------M-------------H------S-----------------------T---KDV-------R------------------------------------------------------------------------ 278
57_BC.CN.09.09YNLX19sg       ------------P-----KR----------------------L----------------------------------------TA--D--------T-----------I------------------------------------------N-----------------Y 272
58_01B.MY.09.09MYPR37        ------------P-----------------------------X----------------------------------------I-------------------------------------------------------------------------------------- 276
59_01B.CN.09.09LNA423        ------------C-----K-------------------M------------------------------------------------------------------------------------------------------------------------------------ 273
60_BC.IT.11.BAV499           --S---------A-----K-------------------I---L----------------------------------------TA--D--------T------------------------------------------------------------------------Y 271
61_BC.CN.10.JL100010         ------------P-----K-----------------------L----------------------------------------TA--D--------T-----------I------------------------------------------------------------- 272
62_BC.CN.10.YNFL13           ------------P-----KR------------------M---L----------------------------------------I-------------------------------------------------------------------------------------- 272
63_02A1.RU.10.10RU6637       -------------M--Y-K-------------------M--------------------------------------------TA-----------------------I------------------------------------------------------------- 276
64_BC.CN.09.YNFL31           ------K-----P-------------I---------------L----------------------------------------TA--D--------T-----------I------------------------------------------R-----------------Y 276
65_cpx.CN.10.YNFL01          ------------A-----K-------------------I---L-----------D----------------------------TA--D--------T-----------I------------------------------------------------------------Y 272
67_01B.CN.11.2011.ANHUI.HF115 ------------------K-------------------M--------C------D----------------------------T-------------------------------------------------------------------------------------- 274
68_01B.CN.11.2011.ANHUI.WH73 ------------------K-------------------M---------------D----------------------------T----------------------------------------------------------------------------------I--- 274
70_BF1.BR.10.10BR_PE004      XX---------N-----X---T----I-----------M-------------S------------------------------T---A---------------------------------------------------------------------------------- 278
71_BF1.BR.10.10BR_PE008      ------K--ENM--------------------------M--------------------------------------------T----D-----------------------------------------------------------------I-----------I--- 276
72_BF1.BR.10.10BR_MG002      ------K-----------Q----------------------------------------------------------------T---M---------------------------------------------------------------------------------- 277
O.BE.87.ANT70                ------KE--NVTV--E-REVQ----------------I--GL-----------AP-------------------SK---E--TA--Q---Q-----R----------I-------G-----------------E-----------G---Q-Q-------------C--- 272
O.CM.98.98CMA104             --C-R-KEFNNVTV--E-REVQ----------------I--GL---------S-AP---R---------------SK---E--TA--Q---Q-----RV---------I----------------------K--------------G-----Q-------------C--- 272
O.CM.98.98CMABB141           ------KE-TNVT--VE-R-VQ----------------I--GL---------H-AP-------------I-----SK---E--TA--Q---Q-----R----------I-------------------------------------G---Q-Q-------------C--- 272
O.CM.98.98CMABB212           ------KE-NNVTV-VQ-KEVQ------------L---I--GL-----------AP-------------------SR---E--TA--Q---Q-----R----------I-------------------------------------G---QRQ-------------C--- 272
O.CM.98.98CMU5337            ------KE-NKVAV--E-REVL----------------I--GL-----------AP-------------------SK---E--TA--Q---Q-----R------------------------------------E-----------G---Q---------------C--- 272
O.CM.99.99CMU4122            ------KE-YXVKV--E-REVQ----------------I--GL-----------AP-------------I-----SK---E--TA--Q---Q-----RV---------I--------S----------------------------G---Q-Q-------------C--- 272
O.FR.92.VAU                  ------KE-YNVAV-LE-REVR------------L---I--GL-----------TP-------------------SR---E--T---K-L-Q-----R----------I-------G-----------------------------G---Q-Q-------K-----C--- 273
O.GA.11.11Gab6352            ------KE-NNVTV--E-R-VQ-V--------------I--GL------------P-------------------SK---E--TA--Q-L-Q-----R----------I-------------------------------------G--QQRQ-------------C--- 272
O.SN.99.99SE_MP1299          ------KE-N-VPV--E-REVL----------------I--GL-----------AP-------------I-----SK---E--TA--Q---Q-----R----------I-------------------------------------G---Q-Q-------------C--- 272
O.US.10.LTNP                 ------KE--NVAV--E-REVQ----------------I--GL-----------AP-------------I-----SR---E--TA--R---Q-----R----------I-------------------------------------G---Q-Q-------------C--- 272
N.CM.02.DJ00131              ----------NVT-D-Q-R--V----------------F-----------------------------R-------A---E--RQ--A---Q-----R----------I----------------------------------------Q----------------C--- 281
N.CM.02.SJGddd               ----------NVTV--Q-R--V------------------------------------------------------A---E--R-------------R----------I---R-------------------------------------Q---------------C--- 278
N.CM.04.04CM_1015_04         ----------N-T-D-Q-R--V----------------F---------------K---------------------A---E--R-------------R----------I---R-------------------------------------Q---------------C--- 283
N.CM.06.U14296               ----------NVTVD-Q-R--V----------------------------------------------R-------A---E--R-------------R----------I---R-------------------------------------Q---------------C--- 283
N.CM.06.U14842               ----------NXT-D-Q-R--V----------------F-------------------------------------A---E--R-------------R----------I-----------------------------------------Q---------------C--- 284
N.FR.11.N1_FR_2011           ----------N-SVD-Q-R--V----------------------------------------------R-------A---E--R-------------R----------I---R-----------------------------------------I-----------C--- 283
P.CM.06.U14788               ------KEFEEVKX--E-RQVY----------------I--L-----S-------I------A-----R------SK---E--RA--Q---Q----T-----------I-------GS----------------------------G--Q-R----I---------C--- 273
P.FR.09.RBF168               ----T-KXFEXVKV--E-REVY----------------I--L-----S----S--I---X--A-X------X---SK---E--RX--Q---Q----T-----------I--------S------X---------------X-----G----R--------------C--- 272
CPZ.CD.06.BF1167             ---VP-QE-YNVP-T-A-KTVKTK--I-----------V-R-L---------TV-I-K----E------------SR---E--T---ENLRA----IEV--D------I------------------------------------------R-VI-----------I--- 279
CPZ.CM.05.SIVcpzMT145        ------K-FERVN---E-K-ITS---------------V--KL-----------D--K-F---------------SK---E--T---K--------------------I-------------------------E-------------------------------C--- 275
CPZ.GA.88.GAB1               ------K-F-NVH---E-R-VV----------------I---L----V----S----------------------SA------T---Q--------------------I---------------------------------------------------------C--- 274
CPZ.TZ.06.SIVcpzTAN13        ----P-SE-YN-PVQ-GNKEVRA-----D--L------I-Q-L---------KV-V-K-E--Q------------SK---E--T---STL-------A----------I--------S--------------------L------------R-V------------I--- 271
CPZ.US.85.US_Marilyn         ------K---HVN---E-R--Q-S--------------I-------------SK-----R--------R-------A------T---Q--------TRV---------I-------------------------------------------R-------------C--- 278
GOR.CM.07.SIVgor2139_287     ------KE-EDVE---EHR-VV----------------I--K----------S-DV---S---------------SK---E--TA--Q---Q----------------I-------------------------------------G-------------------C--- 268
GOR.CM.07.SIVgorCP2135con    ------KE-EDVE---EHR-VV----------------I--K----------S-DV---S---------------SK---E--TA--Q---Q----------------I-------------------------------------G-----R-------------C--- 268
```

```
                                                                                        K219Q
                                                    polymerase motif               T215Y
B.FR.83.HXB2                 EDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTE 446
A1.AU.03.PS1044_Day0         -N-----------T-----------------------------------Y-AK----I--------------------A------A---S--F------------------------------N----E---------------------A----K--------A----D 446
A1.CH.03.HIV_CH_BID_V3538    -S--------------------------------------A----------T---E---------------W--E---A------A---S--F------------------------------Q----------------------------------------A----D 446
A1.CY.08.CY236               -------------T--------------------------C------A---SK-------------------------A------A---K--F------------------------------Q----E---------------------A----K--------A----D 443
A1.ES.05.X1608_8             -S-----------T-------------------------------------SK--E-I--------------------A------A---S--F------------------------------E----E---------------------A----K--------A----D 450
A1.KE.11.DEMA11KE001         -S-----------T--A----------------------------------SK--E-I--------------------A------A---N--F---------------------------------------------------------A----K---R----A----- 446
A1.RU.11.11RU6950            -S-----------K-------------------------------------LK---------------------------V----A---S--F------------------------------M--D----------------------------K--------A----D 446
A1.RW.11.DEMA111RW002        -G------------------V------------------------------L---E-I---------------------------A---S--F------------------------------Q--------------------------A--Q-K-------------- 446
A1.SN.01.DDI579              -S-----------T------V-------------------Y----------S---E-I------------N--------------A---K--F------------------------------Q--------------------------A-------------A----D 445
A1.UG.11.DEMA110UG001        -S-----------T-------------------------------------SK--EMI--------------------A------A---S--F------------------------------Q--D----------------------------K---R---------D 446
A1.ZA.04.04ZASK162B1         -------------T------V------------------------------S----------------------------V--------S---------------------------------Q--D-----------------------S-------------A----D 445
A2.CM.01.01CM_1445MV         --------------------V----------------------I-------TK--E-I--------------------A------A------F------------------------------K--------------------------A--Q-K---R----A----D 446
A2.CY.94.94CY017_41          -------------T------V------------------------------SK-TELI----------------S---V------A---K--FY-----------------------------K--------------------------A----K--------A----D 446
B.BR.10.10BR_MG029           -------------K---------------------------------------H---I----------------E---I-------------F------------------------------E--------------------------A------------------- 446
B.CA.07.502_I191_03          -------------A--------------------------Y----------------------------------E-------------K--F--------------------------------------------------------------T-------------- 446
B.CH.08.M2_0803101_NFLG8     -------------T--------------------------A--------------E------V---------------A----------K---Y-----L------------I----------E-------------------------------K--------A----- 450
B.CN.10.DEMB10CN002          K-------------------------------------------------K---------------L--------E-------------K--FC---------------------N----------------------------------S----K-------------- 446
B.ES.10.DEMB10ES002          -------------T--------------------------C--------------E------------------------------------F------------------------------K--------------------------A----K----------S--- 446
B.FR.11.DEMB11FR001          K-------------------------------------------------------M-----------------------------------F---------------------------------------------------------A----K----------S--- 445
B.GB.05.MM45d213_GN1         KE----------------------------------------------------------------------------I---Q-------------------------------------------------------------------A-------------A----- 445
B.HK.06.HK003                -------------T-------------------------------------R-H------------------------I-------------F------------------------------E--D----------------------------K-------------- 450
B.HT.05.05HT_129389          K--------------------------------------------------X--------------------------X------E------F------------------------------X--N-X---------------------S----K-------------- 445
B.JP.12.DEMB12JP001          K------------L--------------------------C------------Y--------------------E----------R---G--F-----------------------------------------------------------------------A-G--- 449
B.KR.07.HP_18_07JHS10_3909   KN----------------------------------------------------------------------------I-V-----------FY---Q-------------------------M--------------------------A----K-------------- 446
B.PE.07.502_0525_wg5         KE-----------------------------------------------------E------------------KE--E----------K---------------------------------I--------------------------A----K-------------- 450
B.RU.11.11RU21n              P------------T---------------------------------D-------E-----------------------------K----------------------------------------------------------------A-L-----------A----- 450
B.TH.08.MERLBDTRC10          KE-----------V------V-----------------------R-----------------------------E-----------------F------------------------------R------------------------------IK-------------- 446
B.US.11.ES38                 -S-----------------------------------------------------E----------------------A----------K--F------------------------------M--D-E--------------------------KH-------A----- 450
C.AR.01.ARG4006              ------------------------------------S--------------A---E----------------------A---K--E---K---------------------------------Q-------------------------------K-------------D 442
C.BR.07.DEMC07BR003          -S-------------------------------------------------A---E----------------------A------E---K--F----------------------------S-Q-----------------------------R----------A----D 442
C.BW.00.00BW5031_1           -N-----------V-------------------------------------AK--E------------------V---A-V----K------F------------------------------Q----------------------------------------A----D 448
C.CN.10.YNFL19               -E-----------R-----------------------------I-------A---E----------------------A------E------F------------------------------Q---------------------------------------------D 446
C.CY.09.CY260                -X--------------A-----------------------C----------A--------------------------X------X---K--F------------------------------Q----E----------------------------------------D 449
C.ES.08.X2363_2              -N----------------------------------S---A------D---AK--E----------------------A------N---K--F----------------------------A-K---------------------------------------------D 449
C.ET.02.02ET_288             -S-----------T-------------------------------------AK--E-----------------------------E---K--F------------------------------Q----E--------------------------K--------A----D 448
C.IN.09.T125_2139            -------------T------------------------------R------AR--E----------------------E------E---K---------------------------------Q--QQ---------------------------K--------A----D 443
C.KE.00.KER2010              KN-----------R-------------------------------------AK--E----------------------A------E------F------------------------------Q--D--------------------------R--H------------D 442
C.MW.09.703010256_CH256.w96  K--------------------------------------------------AK--E----------------------A------K---K--F----------------------------S-Q-------------------------------K-M------A----D 444
C.TZ.08.707010457_CH457.w8   ---------------------------------------------------VR-------------------------A------E---K--F------------------------------K----N---------------------S----K-------------D 442
C.YE.02.02YE511              KE-------------------------------------------------A---E------------------K---A-V----E---K--F------------------------------Q--D-------------------------------------A----D 440
C.ZA.08.705010534_CH534.w12  -N------------------------------------------R------MK--EM---------------------A------E---K--F------------------------------L---------------------------------------------D 449
C.ZA.10.DEMC10ZA001          -G-------------------------------------------------A---E----------------------A------E---K--F------------------------------Q--N-E---------------------A----K-------------D 442
C.ZM.11.DEMC11ZM006          -S-------------------------------------------------A---E----------------------E---K--E---K--F----------------------------A-Q----E--------------------------K-------------- 447
D.CM.10.DEMD10CM009          KE-----------T--------------------------C---------K----EL----------------------------E---K---------------------------------T--------------------------A--Q-KH-------A----- 451
D.CY.06.CY163                -------------T--------------------------C---------K----EL---------------------A---K--E-------------------------------------T----EN----------------------------------A----- 445
D.KE.11.DEMD11KE003          -------------------------------------------------------EM----------------------------E---K--F------------------------------K--D-ED-----------------------------RCI--A-T--- 445
D.KR.04.04KBH8               -------------------------------------------------------E----------------------A---K--E---K---------------------------------T----E---------------------A----K--------A----- 445
D.SN.90.SE365                -------------T---------------------Q---------------NK--E-----------------------------E------F-----------------------------VE----E---------------------A----KE-------A----- 451
D.TZ.01.A280                 K------------T-----------------------------------------E-----------------------------D---K--F------------------------------K----E---------------------S----K--------A----- 447
D.UG.10.DEMD10UG004          -G---------------A-------------------------------------EM---------------------I------G---K--F------------------------------K----E--------------------------K---RC-K-A-T--- 445
D.UG.11.DEMD11UG003          -N-----------------------------------------------------EM---------------------I------E---K--F----------------------------T-T----E-------------------------------CI----T--- 447
D.YE.02.02YE516              -S-----------T--------------------------C--------------EL-------------------------K--E-------------------------------------T--Q-EN-------------------------K---R----A----D 447
D.ZA.90.R1                   ----------------------------------------C-------------------------------------A------E------F------------------------------K----E----------------------------------------- 445
F1.AO.06.AO_06_ANG32         KE-----------V--------------------------C----------TR--E----------------------A------E---K--F------------------------------Q--D--------------------------Q----------A-G--D 450
F1.AR.02.ARE933              -------------V--------------------------C----------TK---M----------------------------K---T--F------------------------------Q-------------------K-----------K--------A----D 446
F1.BR.10.10BR_RJ015          KE-----------T--------------------------C------D---T--------------------------A-V----K---S--F------------------------------Q----------------------------------------A----D 449
F1.CY.08.CY222               KE-----------L-------------------------------------VK-------------------------I------A---K--F------------------------------Q--D---------------------------I---------A----D 446
F1.ES.02.ES_X845_4           KE--------------K-----------------------C----------IK-------------------------E------N---K--F------------------------------Q--D----------------------------K--------A----D 448
F1.RO.96.BCI_R07             KE-----------L--------------------------Y----------TR----I---------------L----A------T---K--F------------------------------Q--N----------------------------K--------V----D 446
F1.RU.08.D88_845             KE---------------------------------------------K---TK-------------------K-------V----D---K---------------------------------Q--D-E--------------------------K--------A----D 446
F2.CM.02.02CM_0016BBY        KE-------------------------------------------------AR--E--V-------------------A------E------F----------------------------A-Q--N-S------------------------R-KH-------A----D 446
F2.CM.10.DEMF210CM001        KE------------------------------------------R------TK-SE-----------------L----A-V----E---K--F----------------------------T-Q--D-S------------------------RIK-----I--A----D 451
F2.CM.10.DEMF210CM007        KE------------------------------------------H------AK--E-----------------------------G------F------------------------------Q----S------------------------R-KHI------A-T--D 445
G.BE.96.DRCBL                -N-----------T-----------#..-----------------------T---E----------------------A------E------F------------------------------Q--N-EN-------------------------K-----I--A----D 442
G.CM.10.DEMG10CM008          -S-----------V--------------------------C---R-------K--EM---------------------A------E---K--F------------------------------Q--D-E--------------------------K-----I-------- 445
G.CN.08.GX_2084_08           -N------------------------------------------X------IK--EM----------------T----A------D------F------------------------------K--D-E-------------------------IK--------A----D 450
G.CU.99.Cu74                 K------------T-----------------------------------------E-E--------------------A----------S--F------------------------------Q--D-E--------------------------K--------A----D 442
G.ES.09.X2634_2              -S--------------A-----------------------C------K---IK----E--------------K-----A------K---S--F------------------------------Q--D----------------------------K---R----A----D 450
G.GH.03.03GH175G             -S-----------------------------------------I-------TK-------------------------A-V----E---K--F------------------------------Q--D-E--------------------------KH--R----A----- 446
G.KE.09.DEMG09KE001          -S-----------V------K------------------------------IK---M---------------------E------K---S--F------------------------------Q----E--------------------------K--------A----- 444
G.NG.09.09NG_SC62            -N-------------------------------------------------TK--EL---------------------A----------K--F--------------------------I---K--D-ED--------------------S----K--------A----D 445
G.ZA.01.TV546                K--------------------------------------------------TK--EM-----------------E---A------G---K--F------------------------------Q--D-EI-------------------------K---R---------D 445
H.BE.93.VI991                -------------T-----------------------------------------EVI--------------------E------A------F----Q------------------------VK-------------------------------K---X----A----- 444
H.CF.90.056                  KE---------------------------------------------A---E---EM---------------------A------A---K--F----------------------------TVK---------------------------N---K--------A----D 445
H.GB.00.00GBAC4001           -------------T--------------------------N--------------EM---------------------A------A---K--F----------------------------TVK----------------------------------------A----D 444
J.CM.04.04CMU11421           -X------------------X-------------------C----------TK--XL-----------------E---A--X---E---K--F------------------------------Q----EX-------------------X---R-K-M------A-S--D 442
J.SE.93.SE9280_7887          ----------------------------------------C------K---ER--E------------------E---R--K---E---K--F------------------------------Q----ED-------------------------K------K-A----D 444
J.SE.94.SE9173_7022          ----------------------------------------C------K---ER--E------------------E---R--K---E---K--FY-----------------------------Q----ED------------------------IKE----I--A----D 444
K.CD.97.97ZR_EQTB11          K---------------------------------------C----------RK---M-L-------------------A------E------F------------------------------Q--D----------------------F-----K--------V----D 446
K.CM.96.96CM_MP535           K-------------------V-------------------H----------IK--EM-------------------P--------E---K--F------------------------------Q--D----------------------------K--------V----D 446
U.CA.01.TV749                -N--------------------------------------C----------R---EL---------------------V------E---K--F----------------------------T-K----------------------------VR-KH--R----A----D 451
U.CA.99.TV721                ----------------------------------------C--------------EL---------------------V------E---K--F-----------------------------VT-------------------------------K---R----A----D 450
U.CD.83.83CD003_Z3           -N--------------------------------------------------E--E----------------------A------A------F------------------------------Q--N--D-----------------------Q---------------D 447
U.CD.90.90CD121E12           -N----------------------------------S--------------TK--E-A--------------------A------A------F------------------------------Q--D--D-----------------------Q----------A----D 446
U.CY.05.CY090                KE-----T-----T--A-----------------------C----------TK--E-------N------NK------A--K---T---K--F------------------------------Q--D-------N------------------RIKH--R----A----D 446
U.CY.08.CY223                -------------T---------------------------------D---AK--EM---------------------A-V----N---K--F------------------------------Q--D-----------------------A----K-----I-------- 447
U.ES.10.DEURF10DZ001         -E-----------V--------------------------A----------TK--EV---------------------A------E---K--F------------------------------Q--D--------------------------R-------I-------D 446
U.GR.99.99GR303              K------------T-------------------------------------IK--E-I--------------------A------E---K--FF-----------------------------Q----VN------TEVS------N---A--RD---------A----- 445
U.NL.95.U_NL_95_H10986_D1    PG--------------------------------------C--------Y-AK--EVI----------------E---A---G--D---K--F------------------------------Q--D-E--------------------F---R-KN-------A----D 446
01_AE.AF.07.569M             -S-----------T------------------------------------KIK--E-----------------------------A---S--F-----------------------R------E--------------------------A----X--------A----D 446
01_AE.CF.90.90CF11697        -G-----------------------------------------------Y-AK--E-----------------------------D---K--F-----------------------R------E--------------------------A----K--------A----- 444
01_AE.CN.10.YNFL03           -S-----------------------------------------I-------IK--EM---------------------V------A---S--F------------------------------T--------------------------A----K-------------- 445
01_AE.HK.04.HK001            KX-------------------------------------------------AN--E-I--------------------I------A---S--F-----------------------R------E--------------------------A---IK---R----A----D 443
01_AE.IR.10.10IR.THR48F      -S-----------T--K---------------------------------KIK--EMI--H------------------------A---N---------------------------------K-----H-------------------------K--------A----D 438
01_AE.JP.x.JRC77AE           -S-------------------------------------------------IK--EV----------------------------A---S--F------------------------------E--------------------------A----K--------A----D 441
01_AE.TH.09.AA090a_WG11      -S-------------------------------------------------LK-S-----------------------A------A---S--F-----------------------R------E-------------------------------K--------A----D 446
01_AE.TH.90.CM240            -S-------------------------------------------------IK--EM-----K----------------------A---S--F-----------------------R------E--------------------------A----K--------A----D 451
01_AE.US.05.306163_FL        -S-------------------------------------------------TK--E-I---------------------------A---S--F-----------------------R------E-------------------------------K-----I-------D 444
01_AE.VN.98.98VNND15         -S-------------------------------------------------IK--E-----------------------------A---S--F-----------------------R------E---R----------------------A----K--------A----D 445
```

```
                                                                                                               K219Q
                                                                      polymerase motif                     T215Y
B.FR.83.HXB2                 EDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTE 446
02_AG.CM.10.DE00210CM013     K------------V--------------------------A----------AK--E----------------------A------E---K--F----------------------------H-Q--------------------------A----K---R----A----D 446
02_AG.CY.09.CY256            KE-----------V--------------------------A--------Y-TK--E----------------------A--X---X---X--F----------------------------H-Q--------------------------A----K--------A----D 446
02_AG.ES.06.P1423            -E------------------------X-------------A------X---TK--E------------------EK--A------G---K--F------------------------------Q--D-----------------------A----K--------A----D 446
02_AG.GW.05.CC_0048          KE------------------V-------------------A----------TK--E----------------------A------E---K---------------------------------Q--------------------------A----K---R----A----- 444
02_AG.KR.12.12MHR9           K------------V--------------------------A----------AE--EV---------------------A--DK--E------F------------------------------Q-------------------------------K--------A----D 450
02_AG.LR.x.POC44951          ----------------------------------------A----------TK--EM---------------------A---Q--D---K--F------------------------------Q--------------------------A--R-K--------A----D 446
02_AG.NG.09.09NG_SC61        P------------V--------------------------A----------AK--EV-V-------------------A------E-------------------------------------Q--D-E---------------------A-------------A----D 446
02_AG.NG.x.IBNG              K------------V--------------------------A----------TK--E----------------------A------G---K--F-----------------------------VE--------------------------A---IK---R----A----D 446
02_AG.SE.94.SE7812           K---------------------------------------A----------TK--E----------------------A------E------F----E--------------------------------------------V--QVR--A----K--------A-T--D 446
02_AG.SN.98.98SE_MP1211      -------------V--------------------------A----------TK--EL---------------------A------E---K--F------------------------------Q-------------------------------K-------------D 446
03_AB.RU.97.KAL153_2         Q------------T-----------------------------------------E-----------------------E-----E------F---------------------------------------------------------A-------------A----- 446
04_cpx.CY.94.94CY032_3       PE-----------T--------------------------C----------FK--E----------------------A------E------F------------------------------QPA-----------------------------K--------A----D 445
05_DF.BE.x.VI1310            KE--------------------------------------C--------------EV---------------------A--K---E---A--F------------------------------Q--D----------------------------K-------------- 446
06_cpx.AU.96.BFP90           -K-----------T-----------------------------I-------IK--E----------------------A------E---K---------------------------------Q--D-E-----------------------------------A----D 448
07_BC.CN.98.98CN009          -------------T------------------------------------------------------------------------------F------------------------------Q----------------------------------------A----D 446
08_BC.CN.97.97CNGX_6F        K------------V--------------------------C----------------------------------K---------E---K--F------------------------------Q----------------------------------------A----D 442
09_cpx.GH.96.96GH2911        K------------V-------------------------------------IK--E----------------------A--K---E---K--F------------------------------Q--D-E---------------------A----K--------AR---D 445
10_CD.TZ.96.96TZ_BF061       -------------------------------------------------------QM---------------Q--H-TI------G---K--F-A------------------------L---Q--K-N-----------------------------------A-V--D 445
11_cpx.CM.95.95CM_1816       -S-------------------------------------------------T---E----------------------EV-----K------F------------------------------E--D-EC------------------------I--------------D 444
12_BF.AR.99.ARMA159          -A----------CV------T-----------------------------------------------------------------------F--------------------------------------------------------------K---R---------- 446
13_cpx.CM.96.96CM_1849       K------------V--------------------------C----------T---E----------------------EV-----R------F----------------------------A-Q--D----------------------------K--------A----D 453
14_BG.ES.05.X1870            -S-------------------------------------------------RK--EM---------------------A------K---N--F------------------------------Q--D-E--------------------------K--------A----D 445
15_01B.TH.99.99TH_MU2079     -S--------------------------------------C-----------K--EM----------------------------A---S--F-----------------------R------E--------------------------A----K--------A----D 445
16_A2D.KR.97.97KR004         -------------T------V----------------------------------E----R------------------------N---K--FY---------------E-------------K--------------------------AE---K--------A----D 446
17_BF.AR.99.ARMA038          KE------------------------------------------------------L--------------------------------K---------------------------------I-------------------------------K-------------- 446
18_cpx.CU.99.CU76            PE--------------------------------------------------E--E---------------------------------K--F------------------------------Q-------------------------------K-----I-------D 444
19_cpx.CU.99.CU7             -S---------------------------------------------D-------E---------------------------------K--F------------------------------S----E-----------------------------------A----- 445
20_BG.CU.99.Cu103            --------------------W----------------------------------G---F----------------------------WK--FY-----------------------------Q--D------------------------------------------- 442
21_A2D.KE.99.KER2003         -------------------------------------------------------EM----------------------------E-------------------------------------K----E----------------------------------------- 445
22_01A1.CM.01.01CM_0001BBY   -G-----------T------------------------------R--D---AK--E------------------E---I--D---A---S--F------------------------------Q--N-----------------------A-------------A----- 446
23_BG.CU.03.CB118            -E--------------------------------------C--------------EL---------------------A----------K--F--------------------------------------------------------------K-------------- 442
24_BG.ES.08.X2456_2          K------------L--------------------------C------X-------EL---------------------A----------K--F-------------Q---------------------E------L----------------------------A----- 442
25_cpx.CM.02.1918LE          KE-----------T---------------------------T-----GT--AK--E-----------------------------E---K--F------------------------------K--------------------------A--Q----------A----D 445
26_AU.CD.02.02CD_MBTB047     KE------------------V------------------------------A---E----------------------A------A---K--F------------------------------K----EN--------------------A----K--------A----D 444
27_cpx.FR.04.04CD_FR_KZS     K---------------------------------------C----------T---E----------------------A------A---K--F------------------------------Q----E--------------------------K---R----A----D 444
28_BF.BR.99.BREPM12609       KE------------------------------------------R-----------------------------------V--------E--F-----------------------------V------I-------------------------K-------------- 446
29_BF.BR.01.BREPM16704       KE-----------V------------------------------------------------------------EK----------------F---------------------------------------------------------A----K-------------- 446
31_BC.BR.04.04BR142          -S---------------------------------------------------------------------------------------K--FD-----------------------------E--D-E----------------------------------------D 452
32_06A1.EE.01.EE0369         K---R--------------------------------------------------EM---------------------V------E---K--F------------------------------E--D-E-----------------------------------AR---D 446
33_01B.ID.07.JKT189_C        -S--------------------------------------R---R------IK-------------------D------------A---N--F------------------------------E-------------------------------K--------A----D 446
34_01B.TH.99.OUR2478P        K---------------------------------------C---------------M----------------------------D------F------------------------------E--------------------------A----K--------A----D 446
35_AD.AF.07.169H             K------------------------------------------I-----------E-I--------------------I------A---K--F------------------------------T----E----------------------------------------D 445
36_cpx.CM.00.00CMNYU830      K---------------------------------------A----------TK--E----------------------A------K---N--F------------------------------Q-----N--G-----------------A----K--------A----- 446
37_cpx.CM.00.00CMNYU926      -S-----------T--------------------------C----------A---E----------------------A------D---K--F------------------------------Q--D--I--------------------A--------R---------D 444
38_BF1.UY.03.UY03_3389       K--------------------------------------------------Q----------------------E-----------------F--------------------------------------------------------------K-------------- 450
39_BF.BR.04.04BRRJ179        -E-----------V----------C------------------------------E------V-----A-----K---A---------W---FY------------H----------------M-------------------------------K-------------- 449
40_BF.BR.05.05BRRJ055        KE--------------------------------------C------D---AK--EL----------------------------E---K--FF--EQ-------------------------M--D-------------------------------------A----D 446
42_BF.LU.06.luBF_18_06       K------------T-----------------------------------------E----------------------I--------------------------------------------Q--N-----------------------A----K-------------- 450
43_02G.SA.03.J11223          -S-----------T---------C---------------------------TK--E------V----------L----A------E---X--F------------------------------Q--D-E--------------------------K---R----A----D 445
44_BF.CL.00.CH80             KE-----------T------V--------------------------------------------------------------------K-----------------------------------------------------------------K---R----P----- 446
45_cpx.FR.04.04FR_AUK        K--------------------------------------------------TR--EM---------------------A------E---K--F----Q-------------------------Q--D-E---------------------S----K--------A----D 444
46_BF.BR.07.07BR_FPS625      Q------------V--------------------------C----------AK--E-----------------------------N---K--F------------------------------Q--N-E------------------------R-K--------A----D 446
47_BF.ES.08.P1942            -E-----------T---------------------------------------------------------------------------K-----------------------------------Q---N-----------------------------R---------D 445
48_01B.MY.07.07MYKT021       K------------V---A----------------------C--------------E-I---------------------------A---N--F------------------------------Q-------------------D------A--Q-K-------------D 446
49_cpx.GM.03.N26677          -N-------------------------------------------------T#--E-----------------------------A---K--F------------------------------Q--D-EA-------------------------K-----I--A----D 443
50_A1D.GB.10.12792           ----------------------------------------A----------SK--E-I--------------------A------A---S--F-----------------------------------E---------------------A----K-----I-------- 444
51_01B.SG.11.11SG_HM021      K----------------------------------------------D---------I----------------E-----V-----------F------------------------------E--------------------------A----K-------------- 446
52_01B.MY.03.03MYKL018_1     -S-----------V---A----------------------C---------------------------------------------------F-----------------------R------E-------------------X------A---XK--------A----D 446
53_01B.MY.11.11FIR164        -S-------------------------------------------------IK--E-----------------------------A---S--F-----------------------R------E---R----------------------A----K--------A----D 446
54_01B.MY.09.09MYSB023       -------------V--------------------------C---------------------V-----A--------------------S--FF---Q-------L----------R------E--------------------------A----K--------V----D 446
55_01B.CN.10.HNCS102056      -S-----------R-------------------------------------------E-------------------------------K---------------------------------A--------------------------A----K--------A----D 446
56_cpx.FR.10.URF5_patient_A  KE-----------T-----------------------------I--------H---------------------R-----V--------K---------------------------------Q--D------------------------------------------- 448
57_BC.CN.09.09YNLX19sg       ----------------------------------------H---R------A----------------------E---A------E---K--F------------------------------M---------------------------------------------D 442
58_01B.MY.09.09MYPR37        -S-----------V-------------------------------------IK--E------------------------V----A---S--F------------------------------M-----N--------------------A-----E-------A----- 446
59_01B.CN.09.09LNA423        -S-----------R-------------------------------------IK--E-I--------------------I-V----A---S--F-----------------------------VE--------------------------A----K---RV---A----D 443
60_BC.IT.11.BAV499           ------X------------------------------V--G----------------I--------------------A------K---Q---A-------------T---------------Q--D-------------------------------------A----D 441
61_BC.CN.10.JL100010         K-----------------------------------------------------------------------------A------E---K--F------------------------------Q----------------------------------------A----D 442
62_BC.CN.10.YNFL13           K------------L--------------------------C---------------------------------A---A----------K--F---------Q--------------------Q----------------------------------------A----D 442
63_02A1.RU.10.10RU6637       K------------L--------------------------A----------RK--E----------------------A------A---N--F------------------------------Q--D-------------------------------------A----D 446
64_BC.CN.09.YNFL31           ---------------------------------------------------A---E----------------------A------E---K--F------------------------------Q--------I-------------------------------A----D 446
65_cpx.CN.10.YNFL01          ----------------------------------------C----------------D--------------------A------E---T--F----------------------N-------Q-------------------------------------I--S----D 442
67_01B.CN.11.2011.ANHUI.HF115 -S-------------------------------------------------NK-------------------------I----------K--------------------------R------E-------------------------------K-------------- 444
68_01B.CN.11.2011.ANHUI.WH73 -N------------------------------------------R---------------------------------I----------K--------------------------R------E-------------------------------K-------------- 444
70_BF1.BR.10.10BR_PE004      K------------V--------------------------C----------AK---M---------------------A------E---K--F------------------------------Q--N-------------------------------------A----D 448
71_BF1.BR.10.10BR_PE008      KE-----------V--------------------------C------D---AK-------------------------A------E---K---------------------------------Q-------------------------------K--------A----D 446
72_BF1.BR.10.10BR_MG002      K---------------------------------------C----------AR--------------------------------E---K--F------------------------------Q--------------------------G----K---R---------D 447
O.BE.87.ANT70                P------------V---------------------------------D---RD--ELE-C------------PLTE--KR--L--E--YQ--F----------------------------S-Q--N--V---------I----------Q--R--E----I----S--- 442
O.CM.98.98CMA104             P------------V------V--------------------------D----D--ELE-C------------PLTE--KRV-L--E--YK--F------------------------------Q--D-EV---------I----------Q--R--E----I----S--- 442
O.CM.98.98CMABB141           P---P--------V--A------------------------T-I---D----D--ELE-C------------PLTE--KRV-L--E--YQ--F------------------------------Q--S--V---------I----------Q--R-KE----I----S--- 442
O.CM.98.98CMABB212           S------------V---------------------------P-----D----N--EVE-C------------PLAE--KRV-L--E--YQ--F------------------------------Q--N-EV--------------------Q--R-KE----I----S--- 442
O.CM.98.98CMU5337            P---X--------V---------------------------------D----D--ELE-C------------PLTE--NRV-K--E--YQ--F------------------------------Q--N-EV-----------------------R--E----I----S--- 442
O.CM.99.99CMU4122            P------------V-----------------------------V---D----D--ELE-C------------PLTE--KRV-L--E--YQ--F------------------------------Q--N-EI-----------------------R--E----I----S--- 442
O.FR.92.VAU                  PE-----------V------V--------------------------D----N--ELE--------------PLAE--KRV-L--K--YK--FY------------Q----------------Q--D-EV--------------------H--R-KE----I----S--- 443
O.GA.11.11Gab6352            PN-----------L------V----------------------V---D----D--ELE-C--V---------PLTE--KRV-L--E--YK--FY-----------------------------Q--D-EV---------I----------Q--R-KE----I----S--- 442
O.SN.99.99SE_MP1299          P------------V---------------------------------D----D--ELE-C------------PLTE--KRV-L--E--YQ--F------------------------------Q--N-EE--------------------Q--R-KE----I--A-S--- 442
O.US.10.LTNP                 P------------V------V--------------------------D----D--ELE-C------------PLTE--KRV-S--E--YQ--F------------------------------Q--N--V---------I----------Q--R-KE----I----S--- 442

N.CM.02.DJO0131              -------------T--------------------------------------KH-E-I-------------S-LA---ETV----S---K--F----------------------------T-K---Q-V-------------------------K---R-I-------- 451
N.CM.02.SJGddd               K---------------------------------------A-----------KH-E-I----------------A---ETV----R---E--F------------------------------K-----V-------------------------K-----I--A----- 448
N.CM.04.04CM_1015_04         -------------V------------------------------T-------KH-E-I--------------DLA---ETV----G---K--F------------------------------K-----T-------------------------------I-------- 453
N.CM.06.U14296               -N---------------------------------------------D----KH-E-I---------------LA---ETV-K--E---Q--F------------------------------K-----V-------------------------K---R-I-------- 453
N.CM.06.U14842               P------------V------V-------------------------------KH-E-I---------------LA---EAVQ---D---K--F-----------------------------VK----EV-------------------------K---R-I-------- 454
N.FR.11.N1_FR_2011           KE------------------V--------------------T----------KHSE-I---------------LT---EAV----G---K--F------------------------------K-----V-------------------------KH--R-I-------- 453

P.CM.06.U14788               P------------V--D---L--V----------------H-----------S--EVE--------------PLSE--QRV-K--E--YV--F------------------------------K----E------------------------R-KE----I----S--- 443
P.FR.09.RBF168               P------------V--------------------------Y----------ES--EVE--------------PLSE--QRV-K--E--YV--F-----------------------------------EA-----------------------R-KE----I----S--- 442

CPZ.CD.06.BF1167             PE-Q---------V---R--K--C-T--------------Y------D----KH-E-DV-------------TAEK-KEVVQ---N--KA---E-----Y-EK------------E-----Q-K---PET-----------------------RT-E----I--V-G--D 449
CPZ.CM.05.SIVcpzMT145        P---F-------------------------------S--------------TKH-E-I--------------P-EE--KRV----A---K--F----------------------------Q-K---QEV------------------------IK-----I--A-G--- 445
CPZ.GA.88.GAB1               K-------------------V---------------S--------------EK----T----------------D---K-V--------K--F------------------------------Q----EV---------I--------------IK-----I----K--D 444
CPZ.TZ.06.SIVcpzTAN13        P------------L--N---K-F------------------------D----KH-EVDV------------YSEEE--KLMK-------A---E-----Y-EK------------------N-T---PEQ----H---------------A---TKE----I--V-G--D 441
CPZ.US.85.US_Marilyn         KE-----------V---------------------------------D-----H--VI--------------NLEK-E-V-L-----Y-T--F--------E---------------------Q--Q-EI------------------------IK-----IK-A----- 448

GOR.CM.07.SIVgor2139_287     P------------V------V-------------------H------D----K--E-E--------------P-TE--Q-V-K--E--YA--F------------------------------K----E--------R---------------R-KE----I----S--- 438
GOR.CM.07.SIVgorCP2135con    P------------V------V-------------------H------D----K--E-E--------------P-TE--Q-V-K--E--YA--F------------------------------K----E------------------------R-KE----I----S--- 438
```

```
                                                                                                                                                         p51 RT end  p15 RNase H start
B.FR.83.HXB2                 VIPLTEEAELELAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPF.KNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTN  615
A1.AU.03.PS1044_Day0         -VT-----------------------------E----------D-----------.----------K-S-X----R--A-V---VVM---------------------D---MD---------------------------D---------------------------D  615
A1.CH.03.HIV_CH_BID_V3538    IVT-----------------D-----------E-V-------PD-----------.---------KKGS------R---AV---VA-------------R----R----A--M----------------------------D--I----------------------I-D  615
A1.CY.08.CY236               IVT-----K----------R------------E-V--------E-----------.----------K-S---------A-V----AM----------------------L--MD---------------------------D--A------------------------D  612
A1.ES.05.X1608_8             -VT-----------------D----------------------D------F----.---------KKGS------R--ANV---VVM------------R------------MD---------------------------D--A------------------------D  619
A1.KE.11.DEMA11KE001         -VT-----------------D----T------E----------D-----V-----.----------KKS-----------V---VVM--------I------------D---MD---------------------------D--E------------------------D  615
A1.RU.11.11RU6950            IVT----------------R-----L--------V--------D-----------.---------KKGS----------AV---VA-------------R---------A--M-----------------------------------------------RI-------D  615
A1.RW.11.DEMA111RW002        -VT-------------------------------V--------D-----------.----------K-S-----------V---V-M-------------------------MD---------------------------D----------------G----------D  615
A1.SN.01.DDI579              IVT-----------------D----A-----------------N-----------.----------K-SS-----R--V-V---VV-------------R------------M------Y---------------------D--I------------------------D  614
A1.UG.11.DEMA110UG001        IVT-----------------D----------------------D-----------.----------K-S------R----V---VV---------------------------D---------------------------D--------------S------------D  615
A1.ZA.04.04ZASK162B1         -V------------------N---------A------------D-----------.---------KK-S-----------V---VV--------------------------M-C--------------------------D--L------------------------D  614
A2.CM.01.01CM_1445MV         IV---R------E-------N-------------T--------D-----------.---------KK-TT---------------VM------------R----------------------------------------Q----------------------------D  615
A2.CY.94.94CY017_41          IVT--K------E-------T----------------------D-----------.---------KR-ST----I-----------M-------------------------A-------------------------------A------------------------D  615

B.BR.10.10BR_MG029           -----K-------------------------------------------------.----------T-----------A------V-------------R-------------------------------------------------------------S-------D  615
B.CA.07.502_1191_03          -V---T-------------R----------T------L-----------------.-----------------------------A-----------------------------------------------------------------------------------D  615
B.CH.08.M2_0803101_NFLG8     -V--SR--------------T-----F----------------------------.-------------T---------------A-------------R----R---D--------------------------------------------------S---------D  619
B.CN.10.DEMB10CN002          -V---R---I--------------------------------E-----------Y.-------H---------------------A-Q-----------R---------A------------------------------------------------D----------D  615
B.ES.10.DEMB10ES002          -----------------------------------------------------Q-.---------K---------R---------AM------E-I---R------------I----------------------------D----------------D----------D  615
B.FR.11.DEMB11FR001          -----Q-----------------------------------EH------------.-----------------------------AQ------------R---------A-----------------------------------------------------------D  614
B.GB.05.MM45d213_GN1         -----------------------------------------EC-----------Y.------------S-----------V----A----------------------QA----------------------------------I------------------------D  614
B.HK.06.HK003                -----K-------------------------------L-----------------.------------------I----------A-------------R--------X-------------------------------------------------------------  619
B.HT.05.05HT_129389          -V----------------------------------------H------V-----.-----------------------------A--G------I---R--------DI-----------------------------------------------------------D  614
B.JP.12.DEMB12JP001          -V---A-------------R--------------V-------L------------.-----------K------I----------A--------R----R---------A--M--------------------------------------------------------D  618
B.KR.07.HP_18_07JHS10_3909   -V------------------V----------------------------------.---------K-------------------AI----------------------A---------*-----L--------------------------------------------  614
B.PE.07.502_0525_wg5         --T---------------------------------------L---S--------.---------K-------------------A-----------------------A----------------------------------L------------------------D  619
B.RU.11.11RU21n              -----A----------------------------V--L-----------------.-----------------------------A---------------------------------------------------------------------------------I-D  619
B.TH.08.MERLBDTRC10          IVS----------------R-----------------------------------.-------------T---------------A--------------------A--A---D------------------------------E----------------I--------  615
B.US.11.ES38                 -V--SK------------I------------------------------------.---------K--S--------------------------I---R--------------------------------------------L----------S-------------D  619

C.AR.01.ARG4006              IV---A-----------------------------------E-------------.---------K--T------R---------AQ---I--------R---------S---D------------------------------T----------------M-------D  611
C.BR.07.DEMC07BR003          IV----------------------------------------E------------.---------KT-T------R---------AQ------------R---------S---D-----------------------------------------------M-------D  611
C.BW.00.00BW5031_1           IV-----------------R----------------------DD-----------.---------KI-T------R---------AL--------I---R-------------D-----------------------------VA-V--------------I-------D  617
C.CN.10.YNFL19               IV-------------------T---A----------------LD-----------.---------K--T---------A-------M------------R------------AD---------------------------D--A-V--------------K-------D  615
C.CY.09.CY260                IV----------------------------------------ED-----------.---------KV-T----------------AK-C----------R-------------D------------------------------A-------------D--I-----I-D  618
C.ES.08.X2363_2              IV----------------------------------------E------------.---------KRKA----------------SM------------R--------D----D------------------------------I------------------R-----D  618
C.ET.02.02ET_288             IV----------------------------------------DD-----------.---------KR-T-------L--------AL--------I---R-------------D------------------------------A----------------I-------D  617
C.IN.09.T125_2139            IV-----------------------A-----------------------------.---------K--T----------------AL--------------------------D---------------------------D--A-V--------------I-------D  612
C.KE.00.KER2010              IV--------------A-------------------------HD-----------.---------K--T----------------A--------R----R-------------D---------------------------D--A----------------V-------D  611
C.MW.09.703010256_CH256.w96  -V----------------------------------------DD-----------.---------KQ-T----------------CM--------I---R-------------D------------I-----------------A----------------K-------D  613
C.TZ.08.707010457_CH457.w8   IV----------E------R-----A----------------HD-----------.---------K--S----------------AM------------R-------------D-----------------------------------------------I-------D  611
C.YE.02.02YE511              IV----------------------------------------YD-----------.---------KI-T----------------AM--------I---R-------------D-----------------------------M--V--------------M-------D  609
C.ZA.08.705010534_CH534.w12  IV----------------------------------------LD-----------.---------K--T---------A------AM------------R----R----I---D------------------------------A------------------------D  618
C.ZA.10.DEMC10ZA001          IV----------------------------------------ND-----------.---------KR-T---------------VAQ-C----------R-------------D------------------------------I-----------S-D--I-------D  611
C.ZM.11.DEMC11ZM006          -VT---------------------------------------HD-----------.---------K--T------R---------AM-C----------R-------------D-------------------------------------------------------D  616

D.CM.10.DEMD10CM009          -V-----------------------------------------S---W-------.---------K---------R---------AM------------R---------A--R----------------------------D--A------------------------D  620
D.CY.06.CY163                IV----------------------------------K------S---W-------.---------K-------------------AI------------R---------A--M-------------------------------I----------S-I-----------D  614
D.KE.11.DEMD11KE003          IV---------------------------Q-------------D---------QY.---------K---T----------------H-C----------R--------D------------------------------------------------------------D  614
D.KR.04.04KBH8               -----K-------------------------------L-----------------.---------K---T---------------AQ-------R----R---------A------------------------------R---I---------------R--R-----D  614
D.SN.90.SE365                IV----A------------------------------------D-----------.---------KR-S----------K------Q------------R---------A--I--------------------------------------------------------D  620
D.TZ.01.A280                 -----K-------------R-T---A---------------E-------------.---------K---T---------------AL-G------I---R--------------------------------------------I---------------R--------D  616
D.UG.10.DEMD10UG004          I------------------R----------A------------D---------QY.----------K--T---------------AQ-C----------R-------------D------------------------------T------------------------D  614
D.UG.11.DEMD11UG003          -V-----------------------------------V----PD---------QY.---------KL--T--------AA-----SQ-C----------R---------I---D------------------------------I------------------------D  616
D.YE.02.02YE516              -----R-------------R-----------------V---------W-------.------------------------V----AQ------------R---------A--M-------------------------------I------------------------D  616
D.ZA.90.R1                   -V----------------------------------------D------------.----------T-----------A----NLS-------------R-------------------------------------------------------------V-------D  614

F1.AO.06.AO_06_ANG32         IV---A------------------------------------HE-----------.---------KR-S---------------VAM-----------------------------------------------------T---A------------K---K-------D  619
F1.AR.02.ARE933              IV---A------------------------------------------------L.------------S----------------AQ----------------L----D---ISI-------------------------T---A----------S-----K-------D  615
F1.BR.10.10BR_RJ015          -V---T------E------------A-----------L-----------------.---------K--S---------------VSL--------I---R---L----D-------------------------------T--------------S--G--E-------D  618
F1.CY.08.CY222               IV---A----K-----------------------V-------------------C.---------K--S---------------VQ---------I---------K---A-*ID--------------------------T---T----------S-----I-------D  614
F1.ES.02.ES_X845_4           IV---A-------------------A---------V-------------------.---------K-KS----------------AL----------------------A------------------------------T---A-------------D--K-------D  617
F1.RO.96.BCI_R07             IVT--A-----------------------------------------------T-.---------K--S----------------A--C--------------------A----------------A-------------T---A----------S-----Q-------D  615
F1.RU.08.D88_845             IVT--A------------------------------------------------Y.---------K--S------R---------AV-------------------------A---------------------------T---PE---------S-----K-------D  615
F2.CM.02.02CM_0016BBY        -V---A------------------------------------HD----------H.R--------KR-S-----------V----A--G------V---R----------------------------------------T---I-----------S----I-------D  615
F2.CM.10.DEMF210CM001        -V---A--------------------------E----------D---------TH.---------KR-S-----------V---VA---------I---R------------A---------------------------T----------------------------D  620
F2.CM.10.DEMF210CM007        -VS--A-----M------------------------------HD----------H.---------KR-S----------DV----A--G-----RV---R----------------------------------------TD----V----------------------D  614

G.BE.96.DRCBL                IVSM-A---M---------------------------V----L------V----Y.---------KGGS-----------V----A--G-I----I-------K-----V------------------------------T---P-V--Y-------------------D  611
G.CM.10.DEMG10CM008          IV---A--------------D----------------V----VE----------Y.---------KRGS-----------V----A----I----I---R---R----DA---D--------D--------------R--T---P----Y-------------------D  614
G.CN.08.GX_2084_08           IV----------T------------------------V----PD----------Y.---------KRGS----------------A----I----V-------R-----V---D--------D-----------------T---X----Y-------------------D  619
G.CU.99.Cu74                 IV------------------------------EI---V----ND----------Y.---------KRGS---------V-V----S-----------------R-----V---D--------D--------------R--T---L----Y-------------------D  611
G.ES.09.X2634_2              IV------------------------------E----V----LD----------Y.---------KRGS-----------V---VS----I--------R---RR-----------------D--------------R--T---A-----F-----S------------D  619
G.GH.03.03GH175G             IVS--A---M----------------------E----V----LD----------Y.---------QRGST----------V----A--G------V-------R-----V------------------------------T---P----Y-------------------D  615
G.KE.09.DEMG09KE001          IV---T-------------------------------V----LD----------Y.---------KRGS-----------V----A-----------------R-----V------------D-----------------A---P----Y-----------V-----I-D  613
G.NG.09.09NG_SC62            IV-----------------R---------S--E----V---EP-----------Y.-----E---KRGS-------E---V----SI--------I---R---K----DV------------D--------------R--T---A----Y--------K-----------  614
G.ZA.01.TV546                IV---A--------------------------E----V----LE----------Y.---------KKGS-----------V----AV--------I-------R-----V------------D--------------R--T---P----Y-------------------D  614

H.BE.93.VI991                IV---K-------------------A------E---------PD-----------.---------K--S-----------V----A---------I---R--------------H---------------H---------T---E----Y-----------M-------D  613
H.CF.90.056                  I----K-------------R--I---------------R----------------.---------K--T-----I----------S---------I---R------------------------------H---------T---A----Y-I-----------------D  614
H.GB.00.00GBAC4001           I----T------E------R-----------------------E-----------.---------K--T-----I----------C---------I---R------------------------------H---------T--------Y-----------I-R-----D  613

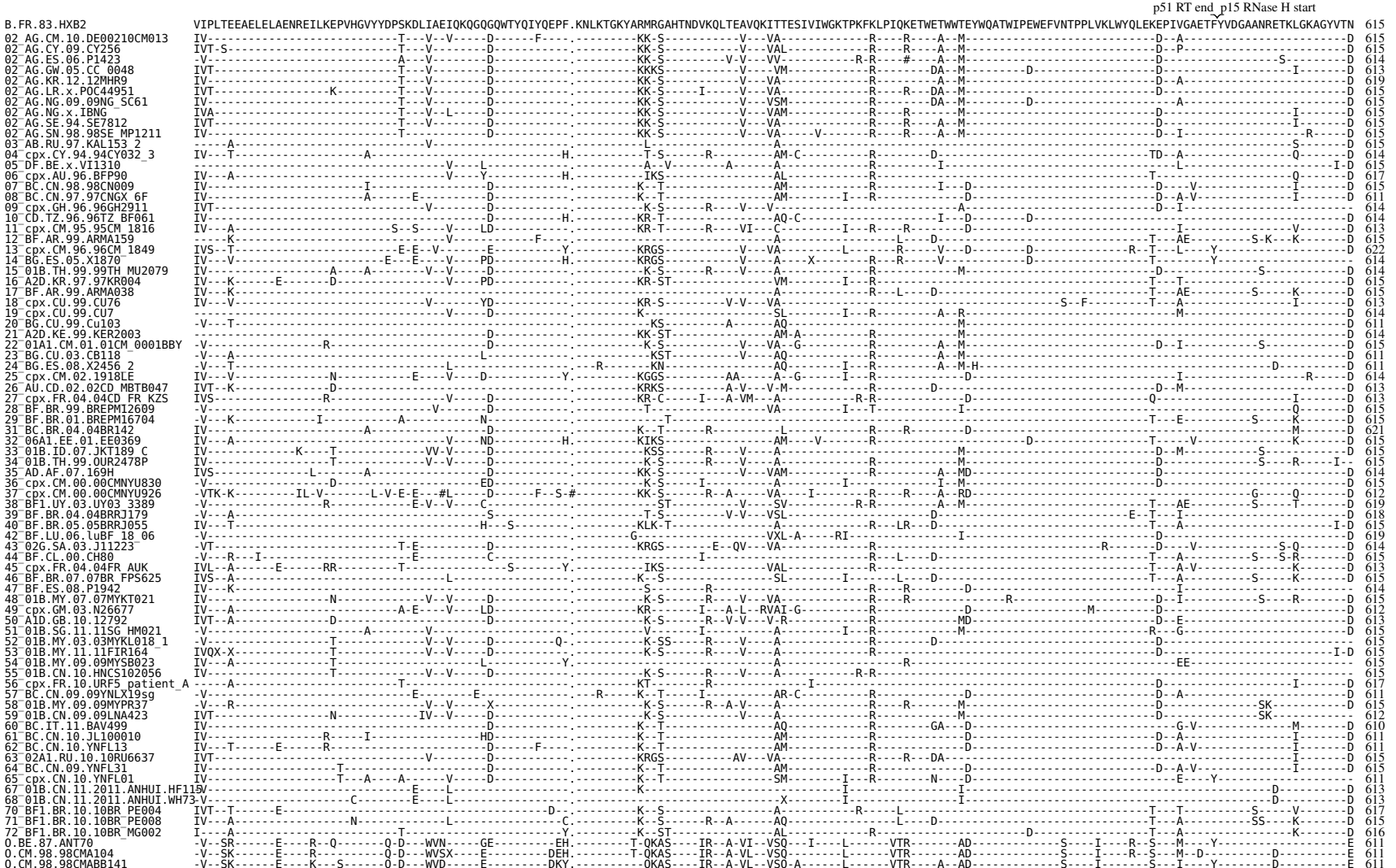
J.CM.04.04CMU11421           IV---K---------K---X-----X----A-E----V----ED-----------.---------KR------------------AL------------R-------------D------------------------------M---------------XK-----I-D  611
J.SE.93.SE9280_7887          IV---R---------K-------------SA-E----V----LD-----------.---------KR-S---------A-V----AL-A----------R----R--------D------------------------------M----------S-----T-------D  613
J.SE.94.SE9173_7022          IV---R---------K--------------ARE----V----LD-----------.---------KR-S---------SQV----AL-A----------R-------------D------------------------------M----------S-----V-------D  613

K.CD.97.97ZR_EQTB11          IV---A------------------------------------H-----------Y.----------I-S-----------V---VAM------------R--------G-------------------------------T--------------------Q-------D  615
K.CM.96.96CM_MP535           IV---A------------------------------------ND----------H.------------S----------------A--G----------R----------------------------------------T---------------H----K-R-----D  615

U.CA.01.TV749                -VT---------------------------------------ND-----------.---------KRKS----------N-----AM-C-----R----R--------DA--M----------------------------D--M------------------------D  620
U.CA.99.TV721                IVT---------------------------------------ND-----------.---------KSKST---------A-----AM-C----------R--------DA--M----------------------------D---------------------------D  619
U.CD.83.83CD003_Z3           IV-M-K----------------------------V--------D--S--------.---------KR-S----------------SM--------I-------------S---D---------------------------D--M----------------K-------D  616
U.CD.90.90CD121E12           IV-M-K-------------E-----------------------D--S--------.---------KR-S----------------SM-------------------------AD------------------------------A-----------S----K-------D  615
U.CY.05.CY090                IVT-----K---------------------------------H------------.---------KI-S--P--I----------S-------------R--------DI-G----------------------------T---A------------I---E-------D  615
U.CY.08.CY223                -VT----------------------A----A---V--------D----------H.---------KR-S-----IR--V------AQ-G------V---R---------A--M---------------------------R---A------------------------D  616
U.ES.10.DEURF10DZ001         IV---A---------------A---A----------------E-----------H.----------IKA----------------AL------------R---------I---D------------------------------I------------------------D  615
U.GR.99.99GR303              IV---A----------D-------------------------H-----------Y.----------V-S-----I-----V------------------R-----G-------D----P-------.------------D-G-------------------R-------D  613
U.NL.95.U_NL_95_H10986_D1    IV---A-------------------------------------D----------Y.----------RKS----------------AM------------R-----------------------------------------D--E----------------M-------D  615

01_AE.AF.07.569M             IV---X-------X------T-------------V--V-----D-----------.----------K-S-----------V----A-------------R------------M----------------------------D--------------S------------D  615
01_AE.CF.90.90CF11697        -V------------------N-----------E-V--------D-----------.----------K-S------R--A-V---VA-------------R----R-------A----------------------------D--A----L------S--N---------D  613
01_AE.CN.10.YNFL03           -V------------------S----A-----------V-----D---------Q-.----------K-S------R--A-V----A-------------R----R------------------------------------D--------------S------------D  614
01_AE.HK.04.HK001            -V------------------T-------------V--V-----D-----------.----------K-S------R--A-V----V---------A--------R-------M----------------------------D--------------S------------D  612
01_AE.IR.10.10IR.THR48F      IV------------------T-------------V--V-----D-----------.----------K-S------R----V----A-------------R------------M----------------------------D--------------S------------D  607
01_AE.JP.x.JRC77AE           I-------------------T-------------V--V-----D-----------.----------K-S------R----V----A-------------R----R------------------------------------D--------------S--N-I-------D  610
01_AE.TH.09.AA090a_WG11      IV------------------T-------------L--------D-----------.----------K-S------R----V---VAM------------R----R-------M----------------------------D--------------S------------D  615
01_AE.TH.90.CM240            IV------------------T-------------V--V-----D-----------.----------RGS------R----V---VA-------------R----R-------M----------------------------D--------------S------------D  620
01_AE.US.05.306163_FL        -V------------------T----------------V-----D-----------.----------K-S-----------V----V-------------R----R-------M----------------------------D--------------S------------D  613
01_AE.VN.98.98VNND15         IV------------------T-------------V--V-----D-----------.----------K-S------R----V----A-------------R------------M----------------------------D--------------S------------D  614
```

p51 RT end p15 RNase H start



```
                                                                                                                 p66 RT, p15 Rnase H end p31 Integrase start
B.FR.83.HXB2                 RGRQKVVTLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQD.EHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLE 784
A1.AU.03.PS1044_Day0         ------IS--E--------H----------S---------------------R-----------K--G-D----S---------------------------S---R-----------E.D--R------T---------II------C------------------------------ 784
A1.CH.03.HIV_CH_BID_V3538    K-----IS--E--------H--H-------S---------------------R-----------E---------S---------------------------K---R-----------E.---R------------H---I-------------------------------------- 784
A1.CY.08.CY236               -------S--E--------H----------S---------------------K-----------K--G-D-I--S---------------------------T---R-----------E.D----------T----------I---------N------------------M--------- 781
A1.ES.05.X1608_8             -------S-AE--------H----------S---------------------R-----------K--E-D----S---------------------------S--------------E.---R----------------I------------------T---------M--------- 788
A1.KE.11.DEMA11KE001         -------S--E--------H--L-------S----------------S----R----I------K--G-D----S---------------------------T--------------E.---R------------------------------------------------------- 784
A1.RU.11.11RU6950            -------P-NE--------H--H-------S---------------------R----I--K---K--E--R---S---------------------------N---R----E------E.---------K----------I-------------------------------------- 784
A1.RW.11.DEMA111RW002        --K----S--E--------H--L-------S---------------------R-DA-V------K--E-D----S--------------------------------------------E.---R----------------I-------------------------------------- 784
A1.SN.01.DDI579              -------S-ME--------H--H-------S-----------------------D--I------K--E------S---------------------------S--------------E.---R------T---------I-------------------------------------- 783
A1.UG.11.DEMA110UG001        -------S--E--------Y----------A-------------V-------K-D---------K--G-D----S---------------------------T--------------E.D--R------T----------I------------------------------------- 784
A1.ZA.04.04ZASK162B1         -------P--E--------H----------P---------------------R--A-I------K--E--R---S---------------------------S---R----------E.D-D-------T---------I-------------------------------------- 783
A2.CM.01.01CM_1445MV         -----I-S--E----R---H----------S---------------------R-----------E-----RI--S---------------------------S--------------E.---R---------H------I------------------I------------------- 784
A2.CY.94.94CY017_41          -----I-S--E--------H--------------------------------ER----I-----K--E--R---S---------------------------S--------------E.---R---------H----------------------------------------------- 784

B.BR.10.10BR_MG029           ------IH----------------------S---M----------------K------S----E------------------------------------X----XX---------E.---X--X----------X--X-------------X--------------X--------- 784
B.CA.07.502_1191_03          K-----IN----------------H---------------------------K------S----------------------------------------N---RI---------E.D-D----------N--------------------------------------------- 784
B.CH.08.M2_0803101_NFLG8     -----AIP-------Q--------------S---------------H-----K------S----------------------------------------T---R----------E.---R-------------I--II---------------------------V--------- 788
B.CN.10.DEMB10CN002          -------SE------------------------------------L------K--------------G------------------------------------------------E.D--R--T------------------------------------------------------ 784
B.ES.10.DEMB10ES002          -------PM---------------H--------------------L------R--------------N-------------------K-------T---R----------E.---R-------------I----------C------------------------------ 784
B.FR.11.DEMB11FR001          K-----IS----------------H---------L-----------------K----------------------------------------T------------------E.----------------------------------------------------------- 783
B.GB.05.MM45d213_GN1         K------SI----------------------K-------------------K------S------------------------------------------------------E.------N-----------I-------------------------------------- 783
B.HK.06.HK003                -------S----------------H----------------------G---K---------------------------------------------------------E.------N-------------------------I------------------------ 788
B.HT.05.05HT_129389          -------S----------------H---------------------------K-----------E------I---------------------NT---R----------E.----------------S------R------------------------------------ 783
B.JP.12.DEMB12JP001          --K----S----------------H---------------------------K------S------------------------------R---T-V---------------.------------------------C------------------------------ 787
B.KR.07.HP_18_07JHS10_3909   -------S---------------------ES-----------I--------K-----------------------------------------N---R-----E--E.D--------K---N------I-------------------------------------- 783
B.PE.07.502_0525_wg5         K-K----SI------------------------------------------K---D------------KR------------------------------------R--E.----------------------------------------------------------- 788
B.RU.11.11RU21n              K------S----------------H---------------------------K------S----------------------------I--------------------.------------------------------------I------------------------ 788
B.TH.08.MERLBDTRC10          K--------P--------------N---------------------------K------S----------------------------I---------------------E.------N-----------I-------------------------------------- 784
B.US.11.ES38                 -------N----------------N-----S---------------------K------SK------R----------------------------------------------E.----------------------------------------------------------- 788

C.AR.01.ARG4006              -----I-A----------------Q---------------------------K--------------Q------S---------------------------S---R----------E.D------------NE-----I------------------T------------------- 780
C.BR.07.DEMC07BR003          -----I-P--E-------------Q-----S---------------------K-----------------R---S---------------------------S--------------E.D------------NE-----I------------------T------------------- 780
C.BW.00.00BW5031_1           -----I-S--E-------------Q-----S---------------------K-----------------RI--S-------------I-------------S--------------E.----------S---------I-----------------------------V--------- 786
C.CN.10.YNFL19               ---K-I-S--E-------------------S---------------------K---------------------S----------------------------------------E.-----------------------------Q----------------------------- 784
C.CY.09.CY260                -----I-S-NE-------------------P---------------------K---G--S-------Q--R---S-------------I-------------S--------------E.D-----------V-E-------------------Q----I------------------- 787
C.ES.08.X2363_2              -----I----E-------------Q-----P---------------H-----K-----------K--N--R---S---------------------------K--------------E.D-----T------E-----I-----I-N----------L------------------- 787
C.ET.02.02ET_288             -----I-S--E-------------Q-----S---------------------K----V------E-----R---S---------------------------S--------------E.------------NE-----I-------------------------------------- 786
C.IN.09.T125_2139            ---K-I-S--E-------------C-----S---------------------K------S----------R---S---------------------------S--------------E.---------------------I----------Q--------------------------- 781
C.KE.00.KER2010              -------SISE-------------Q-----S---------------------K--------------R------S---------------------------S--------------E.--------------E-----I------------------T------------------- 780
C.MW.09.703010256_CH256.w96  -------S--G-------------Q-----S---------------------K-----------------RI--S---------------------------S--------------E.D--R----------E-----I------------------I---------V--------- 782
C.TZ.08.707010457_CH457.w8   -----I-S--E-------------Q-----S-----------------H---K----I---------N--R---S---------------------------N--------------E.--------------E-----I-------------Q----I------------------- 780
C.YE.02.02YE511              -----I-S-NE-----A-------H-----S---------------------K---------------------S---------------------------S--------------E.------N-------------I------------------I------------------- 778
C.ZA.08.705010534_CH534.w12  -----I-P--E-----A-------Q-----S---------------------K----I---------N--RT--S---------------------------K-----------R--E.------N-------------I-------------------------------------- 787
C.ZA.10.DEMC10ZA001          K---------E-------------QI----P---------------------K-----------E-----R---S---------------------------S--------------E.------N-----------------------------Q----T------------------- 780
C.ZM.11.DEMC11ZM006          -----I-S-SE-------------Q-----S---------------------K--------------R--R---S---------------------------S--------------E.---R--N-------E-----I-------------------------------------- 785

D.CM.10.DEMD10CM009          -----A-SF---------------N-----P---------------------K------S----------------------------------------N----I------S----.--DR----------------I------------------------------------- 789
D.CY.06.CY163                -------PF---------------N-----P---------------------K------SL---------------------------------------R-------------S----.--D---N------------I-----------------------W------------- 783
D.KE.11.DEMD11KE003          -------S----------------NM--------------------------K------S--------------S---------------------------N----I-----------E.------N------------------------------------------------------- 783
D.KR.04.04KBH8               ------IP----------------N---------------------------K------S----------------------------------------N------------E--E.D-------------E-------------C-----Q------------------------ 783
D.SN.90.SE365                -----AIP----------------N-----S---------------------K----V-S----------------------------------------K--------------E.------N--------------I-----------------I------------------- 789
D.TZ.01.A280                 K------S----------------N----------I----------------K------S----------------------------------------S--------------E.D--R-------------------S--------C------------------------------ 785
D.UG.10.DEMD10UG004          -------P----------------N-----S-----------------------D-------------------S---------------------------N----I---------E.------N-------------------------------L------------------- 783
D.UG.11.DEMD11UG003          -------S-A--------------N---------------------------K---------------------S---------------------------S--------------E.------N-------------I------------------L------------------- 785
D.YE.02.02YE516              -------PF---------------N-----S---------------------K----V-S----------------------------------------S----------NQ--E.D-D------------I----------------Q------------------------ 785
D.ZA.90.R1                   -------PF---------------N---------------------------K------S----------R--------------------------------Q--------------E.------N------------------------------------------------------- 783

F1.AO.06.AO_06_ANG32         --KE-A-P--E-----A-------H-----S---------------------K--------------Q--R---S------------------------------I----------E.------N------------------------------------------------------- 788
F1.AR.02.ARE933              -----A-P--E-----A--K-VQ-------A---------------------K-------------------------------------------------I----------E.D-----N------------------------------------------------------- 784
F1.BR.10.10BR_RJ015          -----A-F--E-----A--E--Q-------S---------------------K---------------------S------------------------------I----------E.D-----N-------------I----------------------------------------- 787
F1.CY.08.CY222               -------S--E-----A-------H-----S---------------------K------S----R--Q------S--------R----------N----I----------E.------------------------I--K------R-------------------------------- 783
F1.ES.02.ES_X845_4           -----A-S--E-----A-------H-----P---------------------K--------------Q-TRI--S------------------------------I----------E.D--------------------I-----I------------------------------------- 786
F1.RO.96.BCI_R07             -----A-S--E-----A-------H-----S---------------------K--------------Q--RI--S------------------------------I----------E.------N--K----------------------------T-------------------------- 784
F1.RU.08.D88_845             K---A-----E-----A-------------S---------------------K--------------Q--R---S---------------------------S--------------E.------N------------------------------------------------------- 784
F2.CM.02.02CM_0016BBY        -------S--E-------------H-----S-----------------H---K--------------Q------S---------------------------T--------------E.------N-------------I------------------I------------------- 784
F2.CM.10.DEMF210CM001        K------S--E-------------C-----S-----------------H---K--------------Q------S---------------------------S--------------E.-----------------------------------------I------------------- 789
F2.CM.10.DEMF210CM007        -------S--E-------------Q-----P-----------------L---K--------------Q------S---------------------------T--------------E.----F-N-------------------------------------Q------------------ 783

G.BE.96.DRCBL                K---III---E-----A--H--Q------QS---------------------R----I----------------S---------------------------S--------------E.---R-----K----------I-------------------------------------- 780
G.CM.10.DEMG10CM008          --K--IIS--E-------------N---------------------------K----I---------R------S-------------I-------------T--------------E.---R--N-------N------I------------------I------------------- 783
G.CN.08.GX_2084_08           K----II-IQE-----A--H--Q-------S---------------------R-------X------X---I-----------------------------S--------------E.---R----------------I-------------------------------------- 788
G.CU.99.Cu74                 K-K---I-----------------Q-----S---------------------R---------------------S---------------------------T--------------E.---R----------------II------------------------------------- 780
G.ES.09.X2634_2              --K---I--SE-----A-------QI----S---------------------R---------------------S---------------------------S--------------E.---R--N------------------------------------------------------- 788
G.GH.03.03GH175G             K-K--II---E-----A-------Q------KS---------------------R---------------------S---------------------------N--------------E.---R----------------I------------------L------------------- 784
G.KE.09.DEMG09KE001          K-K--II---E-----S--H--Q-------S---------------------R-D-----K---------------S---------------------------N--------------E.---R----------------I-------------------------------------- 782
G.NG.09.09NG_SC62            -EK--I--IQE--------H--Q-------P---------------------R----I----------------T------------K-------S--------------E.D--R----------------I-------------------------------------- 783
G.ZA.01.TV546                --K--I----E-------------Q-----S-----------------I---R--------------N------S---------------------------N--------------E.---R----------------I-------------------------------------- 783

H.BE.93.VI991                --K--I-S--E----------------E--P---------------------K-----------E------F--S---------------------------S--------------V.Q------------------I-------------------------------------- 782
H.CF.90.056                  --K----S--E-----------------------------------------K-----------E---------S---------------------------S-V------------E.---R--N--V--------I-------------------------------------- 783
H.GB.00.00GBAC4001           --K--I-S--E-----------------------------------------K-----------E---------S---------------------------S--------------E.---R--N--K---N------I------------------I------------------- 782

J.CM.04.04CMU11421           K--------------Q------L---X---X-A-------------------K-----------E--X--R---S---------------------------S--------------E.---R----------------I-------------------------------------- 780
J.SE.93.SE9280_7887          K------------------H------R-------------------------K-----------E---------S---------------------------S--------------E.D------------------------------------------------------------ 782
J.SE.94.SE9173_7022          K-----I-----------------------I---------------------K-----------E---------S---------------------------S--------------E.----------------------------------------------------------- 782

K.CD.97.97ZR_EQTB11          K-----ISI-E--------H----------S---------------------K----------------DR---S---------------------------S--------------E#------N------------------------------------I------------------- 784
K.CM.96.96CM_MP535           -------SI-E-----A-------C-----S---------------------K---D-------------RI--S----------------------------------------E.------N-------------I------------------I------------------- 784

U.CA.01.TV749                -------S--E--------H----------S-----------I---------R-----------K--E---I-MS---------------------------S--------------E.------------N--------------------I--------------V--------- 789
U.CA.99.TV721                -----I-S-NG--------H----------S-----------I---------R----I--W---K--E---I--S---------------------------S--------------E.------------N-------------------------I---------V--------- 788
U.CD.83.83CD003_Z3           ------IS--E-----A----L----------A-------------------S---------------------S-------------A---------------------E.---R----------------I-------------------------------------- 785
U.CD.90.90CD121E12           Q----TIS--E-----A----L----------A-------------------S---------------------S------------------------------------E.---------------------I------------------I------------------- 784
U.CY.05.CY090                -------S--E---P---------------S-----------------H---K----------------D----S-------------I-------V----------E.-------T-------------E------------V------------------------ 784
U.CY.08.CY223                K------S-NE--------H----------S---------------------R------S--------------S---------------------------N--------------E.------N-------------I-----------------------------V--------- 785
U.ES.10.DEURF10DZ001         -------S--E--------H--H-------------------I---------K---A--S----K--Q-N----S---------------------------N--------------E.---R---------G--------------------------------G------------ 784
U.GR.99.99GR303              ---X---A--E-------------H-----S---------------------T-----------K--G--R---S-------------I-------------N--------------E.D-XR---------N-------I--------AGIR------------------------- 782
U.NL.95.U_NL_95_H10986_D1    -------S--E-------------Q-----SA--------------------N---D----------Q-DQ---S---------------------------N----I---------E.D-----T--------------------------------------------M--------- 784

01_AE.AF.07.569M             ------I---E--------H--H-------S---------------------R----V------E---------S---------------------------S--------------E.---R-----K------------------N------------------------------ 784
01_AE.CF.90.90CF11697        -------S--E--------H--H-------SD--------H-----------R----I------K--E--R---S---------------------------S--------------E.---R------T---------I-------N------------------------------ 782
01_AE.CN.10.YNFL03           -------S-AE--------H--Q-------S---------------------R----V------E---------S---------------------------S--------------E.---R------T-V-----------------N------------------------------ 783
01_AE.HK.04.HK001            -------S--E--------H----------S---------------------R----I-S----E---------S-----------NK------S--------N------E.---R------T---------I-------N------------------------------ 781
01_AE.IR.10.10IR.THR48F      ---KR--S--E--------H--H-------S--------------L------R----V------E---------S---------------------------S--------------E.---R------T------------------------------------------------ 776
01_AE.JP.x.JRC77AE           -------S--E--------H--Q-------P---------------------R----V------E--R------S---------------------------S--------------E.---R------T-----------------N------------------------------ 779
01_AE.TH.09.AA090a_WG11      -------S--EA-------H--Q-----X-----------------------R----V------E----D----S---------------------------S--------------E.---R------T-----------------N------------------------------ 784
01_AE.TH.90.CM240            -------S--E--------H--H-------S---------------------R----V------E---------S---------------------------S--------------E.---R------T---------I------TN------------------------------ 789
01_AE.US.05.306163_FL        -------S--E--------H--Q-------S---------------------R-D--V------E---------S-----------------------------------------E.---R------T---------I-------N------------------------------ 782
01_AE.VN.98.98VNND15         -------S--E--------H--H-------S---------------------R----V------E---------S---------------------------S--------------E.---R------T---------I-------N------------------------------ 783
```

```
                                  p66 RT, p15 Rnase H end  p31 Integrase start
B.FR.83.HXB2                  RGRQKVVTLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQD.EHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLE 784
02_AG.CM.10.DE00210CM013      ----RIIS--E--------H--H-------S--------------------R-----------K--E------S-------------------S---------------E.---R----------------I-------C------------------------------ 784
02_AG.CY.09.CY256             K---------E--------H--H-------S--------------------R--------------G-D----S-------------------S---------------E.D--R---------G------I-------------------------N------------ 784
02_AG.ES.06.P1423             ------XS--E--------H----------S--------------------R-D-----L---E----D----S-------------------T---------------E.------N-------------I-------------------------------I------ 783
02_AG.GW.05.CC_0048           ----------E--------H--H-------S--------------------R----I------E--R-D-I--S-------------------S---R-----------E.---R----------------I-------------Q------------------------ 782
02_AG.KR.12.12MHR9            -------S--E--------H----------S--------------------R-------L---K--E-D----S-------------I-----S---------------E.---R----------------I-------------------------------------- 788
02_AG.LR.x.POC44951           K------S--E--------H--H-------S--------------------R-----------E--G-D----S-------------------N---------------E.D--R----------------I-------------------------------------- 784
02_AG.NG.09.09NG_SC61         -------P--E-----------H-----A-S--------------------R-----------E----DR---S-------------------S---------------E.---R----------------X-------------------------------------- 784
02_AG.NG.x.IBNG               -------S--E--------H--H-------S--------------------R-----------K----D----S-------------------N---------------E.---R-----K----------I-------------M-----------G------------ 784
02_AG.SE.94.SE7812            -------S-AE--------H--H-------S--------------------R------T----E----DR---S-------------------S---------------E.---R------------------------------------------------------- 784
02_AG.SN.98.98SE_MP1211       ----N--S--E--------H--L-----E-S--------------------R-----------K--E-D----S-------------------K---------------E.---R----------E-----I-------------------------------I------ 784
03_AB.RU.97.KAL153_2          -------S--------------H----------------------------K------S--------------------------------------E-----------E.------G---------------------------------------------------- 784
04_cpx.CY.94.94CY032_3        -------S-SE-------------------S-----------I--------R---D----------R-D----S-------------------N---------------E.------N------------S----------N---------------------------- 783
05_DF.BE.x.VI1310             K----A-S--E-----A-----H----------------------------K--------------E-K----S------------------------------------.----------------------------------------------------------- 784
06_cpx.AU.96.BFP90            ----------E-----------N-------P--------------------K------------------I--S-------------------N---------------E.D--R---------N------I------------------I------------------- 786
07_BC.CN.98.98CN009           ---K---S--E--------H--CI------S--------------------K-----------------R---S-------------------S---------------E.---R---------------------------Q--------------------------- 784
08_BC.CN.97.97CNGX_6F         ---K-I-S---------------I------S--------------------K-----------------R---S-------------------N---------------E.--------------------I----------Q--------------------------- 780
09_cpx.GH.96.96GH2911         -----I-S--E--------H--H-------S--------------------R-----------K--E------S-------------------S---------------E.---R------------------------------------------------------- 783
10_CD.TZ.96.96TZ_BF061        ----RAISI-------------N-------S--------------------K---------------------P-------------------R---------------E.------N--------------------------------L------------------- 783
11_cpx.CM.95.95CM_1816        ----------E--------E--H----------------------------K------S--------------S-------------------H---R-----------E.---R----------------I------------------I------------------- 782
12_BF.AR.99.ARMA159           K------S--E-----A--H----------S--------------------K------S--------------S---------------------V--I----------E.------N---------------------------------------------------- 784
13_cpx.CM.96.96CM_1849        --K--II---E-----------Q-------S--------------------R---------------------S-------------------N---------------E.D--R------T----------------------------T------------------- 791
14_BG.ES.05.X1870             K-K--II---E-----A-----QI------S--------------------R-----------------------------------------S---------------E.---R--N----------I----------------------------------------- 783
15_01B.TH.99.99TH_MU2079      -------S--E-----------Q---R---S--------------------R----V------E---------S-------------------S---------------E.---R------T---------I-------------------------------------- 783
16_A2D.KR.97.97KR004          K----I-S-----------H-------------------------------R----I--L---K--E--R---S---------E---------S----------------.-----R-------H--------------C-----V------------------------ 784
17_BF.AR.99.ARMA038           -----A-S--E-----A--H--K-------S--------------------K----I----------------S------------------------I----------E.------N------------------------------------I--------------- 784
18_cpx.CU.99.CU76             K-----ISI-E--------E--H-------Q--------------------K----I------E---------S-------------------S---R-------E---E.--D---N-----------------------------------------------R---- 782
19_cpx.CU.99.CU7              -------S--------------N----------------------------K------S----------R-----------------------S---------------E.--------------------I-------------------------------------- 783
20_BG.CU.99.Cu103             --K----S--------------H----------------------------K------S----------------------------------S---------------E.---R---------N------I-------------------------------------- 780
21_A2D.KE.99.KER2003          -----I-S--E----R---H-------------------------------R--------------G-KR---S-------------------S---------------E.-------------H------I-------------------------------------- 783
22_01A1.CM.01.01CM_0001BBY    -------S--E-I------H--H-------S--------------------R-------L---K--E------S-------------------S----I----------E.D--R------T-V-------I---------N--------I------------------- 784
23_BG.CU.03.CB118             --K----S--------------H-------S--------------------K------S----------------------------------S---------------E.---------K------------------------------------------------- 780
24_BG.ES.08.X2456_2           -------S--------------H----------------------------K----V-S----------------------------------S---------------E.Q--R-----------------------------------I------------------- 780
25_cpx.CM.02.1918LE           -------S--E--------H--H-------S--------------------K----I----------------S-------------------N---------------E.---R----------------I------------------L------------------- 783
26_AU.CD.02.02CD_MBTB047      K---------E--------H-------------------------------R----I------K--E------S-------------I-----S---------------E.D------------N------I-----I------------I---------V--------- 782
27_cpx.FR.04.04CD_FR_KZS      K---------E-----------Q-------S-----------------R--K---------------------S-------------------S---------------E.---R----------------I-------------------------------------- 782
28_BF.BR.99.BREPM12609        -------P--------------H----------------------------K------S---------------------------------------I-----------.----------S-D-------I-------------------------------------- 784
29_BF.BR.01.BREPM16704        -----A-S--E-----A-----Q-------S--------------------K--------------N------S------------------------I----------E.------N---------------------------V------------------------ 784
31_BC.BR.04.04BR142           -----C-S--E-----------Q-------S--------------------K-----------L--N--R---S-------------------N---------------E.--------------E-----I------------------T------------------- 790
32_06A1.EE.01.EE0369          ---------SE-----------N-------S--------------------K---------------------S-------------------S---------------E.---R---------N------I-------------------------------I------ 784
33_01B.ID.07.JKT189_C         -------S--E--------H--H-------S--------------------R----V------E----D----S-------X-----------S---------------E.---R------T---------I-------N-----------------------V------ 784
34_01B.TH.99.OUR2478P         -------S--E--------H--H-------S--------------------K----V------E---------S-------------------S---------------E.---R------T---------I-------N------------------------------ 784
35_AD.AF.07.169H              -------P--E--------H--H-------S--------------------K-------K---K-VE-D----S-------------------N----I------E---E.---R----------------I-------N--------------------M--------- 783
36_cpx.CM.00.00CMNYU830       -------S--E--------H--H-------SK-------------------R-----------K--E-D----S-------------------K----I----------E.D--R----------------I-------------------------------------- 784
37_cpx.CM.00.00CMNYU926       K----I-SF-E--------H--H------ES--------------------R-D--I------KI-E--R---S-------------I-----S----GV---------E.D-KR---------N------I-------------------------------------- 781
38_BF1.UY.03.UY03_3389        -----A-S--E-----A--H--S-------S--------------------K---------------------S------------------------I----------E.------N---------------------------------------------------- 788
39_BF.BR.04.04BRRJ179         -------S-E--------------------S-----------------K--K------S----E------I--T-----------------------------------E.------N-------------I-------------------------------------- 787
40_BF.BR.05.05BRRJ055         -------P--------------H----------------------------K------S---------------------------------------I-----------.----------------------------------------------------------- 784
42_BF.LU.06.luBF_18_06        -----A-S--E-----A-----Q----------------------------K----I----------------S-------------------T----I----------E.D----------------I--------I-C------------------------------ 788
43_02G.SA.03.J11223           -------S--E--------H----------S--------------------R-----------K--E-D----S-------------I-----S---------------E.---R------------D-----------------Q------------------------ 783
44_BF.CL.00.CH80              -----A-S--E-----A--H--Q-------P--------------------K---G----------E--R---S-----------------------------------E.D-------------------II----I-------------------------------- 784
45_cpx.FR.04.04FR_AUK         -------S--E-----------H-------S--------------------K---D----------M-D----S-------------------S---------------E.----------------------------------------------------------- 782
46_BF.BR.07.07BR_FPS625       -----A-S--E-----S-----Q-------S------------------------D-------L----------------------------------------------.----------------------------A------------------------------ 784
47_BF.ES.08.P1942             K------S--------------L----------------------------K--------------N------S-----------------------------------E.------N----------I----------------Q------------------------ 783
48_01B.MY.07.07MYKT021        -------S--E--------H--H-------S--------------------R----V------E---------S-------------------S---R--------R--E.D-------------------II------------------------------------- 784
49_cpx.GM.03.N26677           K------S--E--------H----------S--------------------K-----------------R---S-------------------K---------------E.--------------E-----I------------------I------------------- 781
50_A1D.GB.10.12792            -------S-NE----M---H--H------R------------------L--K---G-------K--E-D------------------------S---------------E.D--R------T------------------------------------------------ 782
51_01B.SG.11.11SG_HM021       -------S--------------H------------------------------------K------Q--------------------------S---R-----------E.D--------------------------------------L------------------- 784
52_01B.MY.03.03MYKL018_1      -----I-S--E----Q---H--H-------S--------------------R----V------E---------S-------------------S---------------E.---R-----------------I------------------------------------- 784
53_01B.MY.11.11FIR164         -------S--E--------H--H-------S--------------------K----V------E---------S-------------------K---------------E.---R------T---------I-------N------------------------------ 784
54_01B.MY.09.09MYSB023        K--------P------A-----H----------------------------K------S-------Q--R-----------------I-----S---------------E.------N---------------------------------------------------- 784
55_01B.CN.10.HNCS102056       K------S----------------------S--------------------K------------M-S--R-----------------------S---------------E.D--R------T-------------------N-------------G-------------- 784
56_cpx.FR.10.URF5_patient_A   -------S--E--------H----------S--------------------K----V-S--------------S-------------I-----S---R-----------E.D---------T------------------------------------------------ 786
57_BC.CN.09.09YNLX19sg        ---R-I-S--E------------I------S--------------------K-----------------R---S-------------------S---------------E.D--R----------------I----------Q--Q------------------------ 780
58_01B.MY.09.09MYPR37         -------S--E--------H--H-------S--------------------R----V------E---------S-------------------S---------------E.D--R------T---------I-------N------------------------------ 784
59_01B.CN.09.09LNA423         -------S--E--------H--H-------S--------------------R----V------E----D----S-------------------S---------------E.---R----------------II------------------------------------- 781
60_BC.IT.11.BAV499            ------IS--E-----------Q-------S-----------------K--K---------V-------R---S-------------------K---R-----------E.D------------NE---L-I------------------T------------------- 779
61_BC.CN.10.JL100010          ---K-I-S--E-----------CI------A--------------------K---------------------T-------------------S---------------E.---R----------------I---------NQ--Q----R------------------- 780
62_BC.CN.10.YNFL13            ---N-I-S-NE-----A------I------S--------------------K-----------------R---S-------------------H---------------E.---R-------------I--I----------Q--------------------------- 780
63_02A1.RU.10.10RU6637        ------IS--E--------H--H------RS--------------------R----I--K---K--E------S-------------------S---------------E.D--R---------N------I------------------I------------------- 784
64_BC.CN.09.YNFL31            ---K-I-S--E------------I------S--------------------K--------------R--R---S-----------------------------------E.--------------------I----------Q--------------------------- 784
65_cpx.CN.10.YNFL01           K---------------A-----H-------V--------------------K-----------------R---S-------------------S---------------E.------N-------------I-------------------------------------- 780
67_01B.CN.11.2011.ANHUI.HF115 -------S--------------H-----------------------H----K------S----------------------------------T---------------E.------------V---------------------I----L------------------- 782
68_01B.CN.11.2011.ANHUI.WH73  -------P--------------H-----------------------H----R---------------------------------------------------------E.------------V---------------------I------------------------ 782
70_BF1.BR.10.10BR_PE004       -------SI-E-----A-----Q----------------------------K------S----------------------------------N-----------E---E.----------------------------N------------------------------ 786
71_BF1.BR.10.10BR_PE008       -------S--E-----A-----Q-------S--------------------K-----------------RI--T-------------------T---------------E.------N---------------------------------------------------- 784
72_BF1.BR.10.10BR_MG002       K----A-S-IE-----A-----Q------ST--------------------K----I------E---------S------------------------I-------R--E.------N---------------------------------------------------- 785
O.BE.87.ANT70                 Q-K--IIK-DE-----A--M--L------KET------------V-SS--T----PI-Q----E-T---Q---T------------KI-----KD--R----E---Q--E.D--------K-L--E-G---------I---P--HI----I--------EV--I----M- 780
O.CM.98.98CMA104              Q-K--IIK-DE-----A--M-VL------KEK---------V----SS--T----PI-Q----E-T---Q---T-------------I-----KD--R----E---Q--E.D------X---L----G---------I-N-P--HI----I--------EV--M----T- 780
O.CM.98.98CMABB141            Q-K--IIK-EE-----A--M-VL------KEKI-------------SS--T----PI-Q----E-T-------T------------KI-----KD--R----E---Q--E.D----------L----G---------I---P--HIR------------EV--I----V- 780
O.CM.98.98CMABB212            Q-K--IIE-KE-----A--M-VL------KEK---------V----SS--T--D-PI-Q----E-T---Q---T------------KI-----KD--R----E---Q--E.D----------L----G---------I---P--HI----T---------E---M------ 780
O.CM.98.98CMU5337             Q-K--IIK-EE-----A--M-VL------KEX--------------SS--T----PI-Q----E-TR--R----------------KI-----KD--R----E---Q--E.D----------L----G---------I-N-P--HI----T-----Y--EV--I------ 780
O.CM.99.99CMU4122             Q-K---IK-EE-----A--M--L---K--KET--------F-----SS--T--D-PI-Q----E-T---Q----------------KI-----KD--R----E---Q--E.D----------L----G---------I-N-P--HI----I--------EV--I----T- 780
O.FR.92.VAU                   Q-K--IIK-DE-----A--M-VL------KEK---------V----SS--T--D-PI-Q----E-T-R-Q---T------------KI-----KD--R----E---Q--E.D----------L----G---------I---P--HIRR-----------E---M------ 781
O.GA.11.11Gab6352             Q-K--I-K-EE-----A--M-ML------KET---------T----SS--T----PI-Q----E-T--D----T------------KI-----KD--R----E---Q--E.D----------L----GI-A------I---P---I----I-----Y------M------ 780
O.SN.99.99SE_MP1299           K----IIK-EE----RA--M-VL------KET----Q----V----SS--T----S--Q----E-T---Q---T------------KI-----KD--R----E---Q--E.D----------L----G---------I-N-PQ-HI----I--------EV--M------ 780
O.US.10.LTNP                  Q-K--IIN-DE-----A--M-VL------KEK---------V----SS--T----PI-Q----E-TR--Q---T------------KI-----KD--R----E---Q--E.D--------K-L--E-G---------I---P--HIR---I--------EV--M----V- 780
N.CM.02.DJO0131               --K----SIE------------L-------Q---------------HSH--K------S----E-----R---S-------------------S-----------E---E.D--R---------------S-------------------I----N----V--------- 789
N.CM.02.SJGddd                K------SIA------A-----L-------Q-----------M---HS---K------S----E-----R---S-------------I-----S-----------E----.D--R----------------I-----------------------N----V--------- 786
N.CM.04.04CM_1015_04          -------S-EN-----------LM----X-Q-----------M---HS---K------S----E-----R---S-------------------S-----------E---E.---R-----K----------I-----------------------N----V--------- 791
N.CM.06.U14296                -------SIX------A-----LM---E-EK-----------M---H----K----V-S----E-----R---S-------------------S---------------E.D--R----------------I-----------------------N----V--------- 791
N.CM.06.U14842                K------SIA------------LM-----QD-----------M---HS---K------S----E-----R---S-------------------S------------X--E.D--------K----------I-----------------------N----V--------- 792
N.FR.11.N1_FR_2011            ------ISIEN-----A-----LM---E--R-A---------M---HS---K------S----E-----R---S-------------------S---------------E.D--R----------------I-----------------------N----V--------- 791
P.CM.06.U14788                -----IIK-EE-----A--E-VL---KE--EQA--------V----S-T----D-P--QK---EMT-------S------------NI-----KD--R----E---Q--E.D----------L--E-G---I-----INN-P--HV----------------------M- 781
P.FR.09.RBF168                -----IIK-EG-----A--E-VL---KE--KQA--------V----S-T----D-P--QK---EMT-------S------------DI-----KD--R----E---Q--E.D--------K-L----G---I-----INN-P--HV----------------------T- 780
CPZ.CD.06.BF1167              --KSR-IE--E----QA-----L---K-AKEK---------V---L-GT----D-P--EE--QE-MN--A---N------------E------KN-------E---Q--E.D--RF----KML-DEY----I-----I-Q-N--HR----I--------E--------I- 787
CPZ.CM.05.SIVcpzMT145         -----IHS-E------A----VLM------Q---V------V---L-SH--H---DI------E------I--S-----R-------------S-----------E---E.---R-----K---Q------I-------Q-----------------------I------ 783
CPZ.GA.88.GAB1                K-K--IIS-EN----QA--K-LL-----DQQ----------V-----S--H------------E------I--S--------------------------------R--E.---R-----K----------I-------H-----V-----------------V------ 782
CPZ.TZ.06.SIVcpzTAN13         ---Y-TRE-EN----QA--W-VD---K--SQ----------VM-VL-GM----D-PI-EE--QK-MQ-NA----------------E------KN-------E--TE--E.D-D------K---DEY----I-----I-Q-S--HK----I--------EV--------- 779
CPZ.US.85.US_Marilyn          -----II--EN-----A--T-V----K--ENT--V--G---V----HS---------------E------S-IS-------------I-----S---------------E.D-D------T-----------------------P-----I------------------- 786
GOR.CM.07.SIVgor2139_287      K----IRK-EG-----A--E-VL---EE--KN---------V----A-S----D-P--QK---EMTT-------------------EI-----KD--R----E---Q--E.D---------TL----G---------IN--P--HV----R------------------- 776
GOR.CM.07.SIVgorCP2135con     -----IRK-EG-----A--E-VL---EE--KN---------V----A-S----D-P--QK---EMTT-------------------EI-----KD--R----E---Q--E.D----------L----G---------IN--P--HV----R------------------- 776
```

```
B.FR.83.HXB2          GKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTGATVRAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRNPLWKGPA 954
A1.AU.03.PS1044_Day0  ------------------------------------VV---------S-A-K-----N-Q-------------------------E-------------------------I-------------H-------------D-I----- 954
A1.CH.03.HIV_CH_BID_V3538 ----I-----------------------------------RV---------S-A-K-----N-Q-------------------------E-------------------------I-------------------------D-I----- 954
A1.CY.08.CY236        --I-------------------------------A----VV-----P--S-A-K-----D-Q-------------------------E-------------------------I---S----------I-------------D-I----- 951
A1.ES.05.X1608_8      --I-----------------------M------VV---------S-AFK--------Q-------------------------E-------------------------I-------------------------D-I----- 958
A1.KE.11.DEMA11KE001  ------------------------------I---VV---------S-A-K-----Q-------------------------E-------------------------I-------------------------D-I----- 954
A1.RU.11.11RU6950     ----I-----------------------------V-----P--SNA-K-----N-Q-------------------------E-------------------------I-------------I-------------D-I----- 954
A1.RW.11.DEM111RW002  ------------------------------------VV---------S-A-K-----D-------------------------E-------------------------I-------------------------D-I----- 954
A1.SN.01.DDI579       ----I-------------------A----------V---------S-AMK-----N-------------------------E-------------------------I-------------------------D-I----- 953
A1.UG.11.DEMA110UG001 ------------------------------------MV---------S-AFK------VQ-------------------------E-------------------------I-------------I-------------D-I----- 954
A1.ZA.04.04ZASK162B1  --I-------------------------------A---RV---------S-A-K-----NVQ-------------------------------------------------I-------------------------DSV---- 953
A2.CM.01.01CM_1445MV  ------------------------I---------V-----P--S---K------VQ-------------------------------D-------I-----------R-------------E-I----- 954
A2.CY.94.94CY017_41   --------------------T---D---I---------V-----P--IS---K------Q-------------------------------------I-----------R-------------D-I----- 954

B.BR.10.10BR_MG029    --X---------------------X---XX-------X------X--ISXA-K--X----X-------------------X-------------X-------I----N-------------------D------- 954
B.CA.07.502_1191_03   --I-------------------------I-R-------------P--IS---K-------------------------------------------------I-------------------------D------- 954
B.CH.08.M2_0803101_NFLG8 --I-----------I-------------------------P---SN--K-------------------------------------------------I-------T-------N----------D------- 958
B.CN.10.DEMB10CN002   --I-------------------I-------Q----------STA-K------V-------------N----------Q-------------------I---S--L---A----------------D------- 954
B.ES.10.DEMB10ES002   --------A-------------------------------V----P---SN--K-------------------------------------------------I-------------------------D------- 954
B.FR.11.DEMB11FR001   --I-------------------------------------ST--K---------D-------------------------------R-----------I-------------------------D------- 953
B.GB.05.MM45d213_GN1  ------------------------I---------A---------S-A-K-------------------L-------------------------------------EM--R-------------D------- 953
B.HK.06.HK003         --I-------------------------I---------V-------SNA-K-------------------------E-------------------------I-------------------------D-I----- 958
B.HT.05.05HT_129389   ------------------------I---------R---------S-V-K-------------------------E--E-------------------------I-------------------------D------- 953
B.JP.12.DEMB12JP001   --I-------------------------I---------------ST--K-------------------N-------------------------M-------------------------D------- 957
B.KR.07.HP_18_07JHS10_3909 --I-------------------------I---------T-------STA-K------V-------------------T-------------------I----N--Q---R--Q---------N-D-I----- 953
B.PE.07.502_0525_wg5  --I-------------------------I---------------------S---K-------------V-------------------------------------------------------------D------- 958
B.RU.11.11RU21n       --I-------------------------I---------V---------STA-K-------------------------Q--Q-----------------------Q-------------------N-D-I----- 958
B.TH.08.MERLBDTRC10   --I-------------------------------------------SN--K-------------------------E-------------------------------------H--M-------N-D------- 954
B.US.11.ES38          --I-------------------------------------------ST--K-------------------------E-------------------------M-------------------N-D------- 958

C.AR.01.ARG4006       --I-------------------------I---------V---------SNA-K-------Q-------------------------------------------------I-------------I-------D-I----- 950
C.BR.07.DEMC07BR003   --I-------------------------I---------------P--ISN--K-------Q-------------------------------------------------I-------------N-I-------D-I----- 950
C.BW.00.00BW5031_1    --I-----------------------V-I---------A---------S-A-K-------N-------------------------E-------------E-------I-------------------------D-I----- 956
C.CN.10.YNFL19        --I----------M------------I---------V---------S-A-K-------------------L----------E-------------------------I-------R-------I-------D-I----- 954
C.CY.09.CY260         --I-------------------------I---------V---------S---K-------Q-------------------------E-------------------------I-------R-------I-------D-I----- 957
C.ES.08.X2363_2       -R-------------------------YI---------V-----P--S-A-K-------Q-------------------I-------------------------------I-------------I-------D-I----- 957
C.ET.02.02ET_288      --I-------------------------YI---------V---------SSA-K-------T-------------------V-------------------R---D-C----I-----------N--L-------D-I----- 956
C.IN.09.T125_2139     --I----------M------------I---------V---------S-A-K-------Q-------------------------------------------------I-------------V-------D-I----- 951
C.KE.00.KER2010       --I-------------------------YI---------V---------S-A-K-------Q-------------------------------------------------I-------R-----IQ----------D-I----- 950
C.MW.09.703010256_CH256.w96 --I----------------T---YI---------V-----P--S-A-K-------Q-------------------------------------T---I-------Q----I-------------D-I----- 952
C.TZ.08.707010457_CH457.w8 --I----------------E---YI---------V-----R--SN--K-------------------------------------D-T----I-------------I-------------D-I----- 950
C.YE.02.02YE511       --I-------------------------YI---------V---------S-A-------Q-------------------------E-------------------------I-------------I-------------D-I----- 948
C.ZA.08.705010534_CH534.w12 ----------M---------D---I---------V---------S-A-K-------Q-------------------------------------------------I-------------I-------------D-I----- 957
C.ZA.10.DEMC10ZA001   --I-------------------------YI---------V-----T--S-A-K-------Q-------------------------------------------------I-------------I-------------D-I----- 950
C.ZM.11.DEMC11ZM006   ------------------------------YI---------V---------S-A-K-------Q-------------------------------------------------I-------------I-------------D-I----- 955

D.CM.10.DEMD10CM009   --I-------------------------I---------VV--------S---K-------------------------------------------------------------------------D-I----- 959
D.CY.06.CY163         ----------------K-------------------------VV--------S-A-K------V-------------------------R-------L-------------I-------R-------------------D-I----- 953
D.KE.11.DEMD11KE003   ------------------------------------------VV--------S---K-------------------------------Q-------------------------I-------------------------D-I----- 953
D.KR.04.04KBH8        ------------------------------------------VV--------S-V-K-------------------------D-------E-------------------------I-------R-------------------D-I----- 953
D.SN.90.SE365         ------------------------------------------VV--------S---K-------------------------------------------------------I-------R-------I-------------D-I----- 959
D.TZ.01.A280          ------------------------I---------V---------S---K-------------------------R-------------------------------I----S-L-T-------------------------D-I----- 955
D.UG.10.DEMD10UG004   --I-------------------------------VV----R--STA-K-------------------------------------R-------------------I-------------V-------------D------- 953
D.UG.11.DEMD11UG003   ------------------------I---------VV--------IS-A-K-------------------------E-------------------------I-------------------------D-I----- 955
D.YE.02.02YE516       --I-------------------------------VV--------SGA-K-------------------------------------------------------I-------------------------D-I----- 955
D.ZA.90.R1            ----I-------------------------------VV--------S-A-K-------------------------------L-------------------I-------------------------D-I----- 953

F1.AO.06.AO_06_ANG32  ----M-------------------------I---------I---------S-A-K-------Q-------------------------------------------S-----------------------N-D-V----- 958
F1.AR.02.ARE933       ------------L-----------------I-------A-------P--S-A-K-S-----Q-------------------------------------------S-----R-------I-----------D-V----- 954
F1.BR.10.10BR_RJ015   --I----------L-----------------I---------------T---S---K-----Q-------------------I-------------------------I---S-------------------------D-V----- 957
F1.CY.08.CY222        E-I---------------------K----I---------I---------SGA-K-------Q---E---------K-I--K-------K-------------------I-------------------------D-I----- 953
F1.ES.02.ES_X845_4    --I-----------------------------------M---------S-A-K-------Q-------------------------------------------------I-------------I-------------D-V----- 956
F1.RO.96.BCI_R07      ------------------------I---------R---------SGA-K-------Q-------------------------------------D-------I---S-------------I-------------D-V----- 954
F1.RU.08.D88_845      --I-------------------------I---------I---------ST--K-------Q-------------------------R-------------------I---S-------------------------D-V----- 954
F2.CM.02.02CM_0016BBY --I-------------------------I---------I---------STV-K-------Q-------------------------------------------------I-------------S-------F---D-I----- 954
F2.CM.10.DEMF210CM001 --I-------------------------I---------------------STV-K------VN-------------------------------------C-------I---S-L---A----S---K---F---D-I----- 959
F2.CM.10.DEMF210CM007 --I-------------------------------------I---------S-A-K-------Q-------------------------------------------I---S-------------------------D-I----- 953

G.BE.96.DRCBL         --I-------------------------I---------I---------S-A-K-----S-T-------------------------------------------I---S-------------------------D-I----- 950
G.CM.10.DEMG10CM008   -QI-----------------------A--I---------V-----R--SNA-K-----N-T-------------------------E-------------------I---S-------------------------D-I----- 953
G.CN.08.GX_2084_08    --I-------------------------I---------I---------SN--K-----N-T-------------------E-------------------------I---S----------SR----------D-I----- 958
G.CU.99.Cu74          --I-------------------------I---------V---------S-A-K-----N-T-------------------------------------------I---S-------------------F---D-I----- 950
G.ES.09.X2634_2       --I-------------------------I---------I---------S---K-----N-T-------------------M-------R-------------------I---S-L-------------------D-I----- 958
G.GH.03.03GH175G      --I-I-----------------------I---------VL--------S-A-K-----N-T-------------------------------------------I---S-------------------F---D-I----- 954
G.KE.09.DEMG09KE001   ----M-----------------A--I---------V---------IS-A-K-----N-T-------------------------R-------------MI--S-------------------------D-I----- 952
G.NG.09.09NG_SC62     --I-----------T---------I---------V---------SN--K-----N-T-------------------K-------T-------------I---S-------------------------D-I----- 953
G.ZA.01.TV546         --I-I-----------------------------V---------SQA-K-----S-T-------------------------------------------I-------------------------D-I----- 953

H.BE.93.VI991         --I-------------------------I---------M---------S-A-K-----D-H-------------------------R-------------R--I-------P-------SQ--K---------D-I----- 952
H.CF.90.056           -Q-----------------------K---------S---V---------S-A-K-----D-Q-------------------------------------------I-------------SN--K---------D-I----- 953
H.GB.00.00GBAC4001    ------------------------------------I---------I---------SGA-K-------VQ-------------------------------------------I-------------L---K---------D-I----- 952

J.CM.04.04CMU11421    --I-------------------------F-I---------V---------S---K-------X-------------------------E-----R-------------D-T----I---------R-------------D-I----- 950
J.SE.93.SE9280_7887   ------------------------A-F-I---------V---------SGA-K-----D-------------------------E-------------------------I---------R-------------D-I----- 952
J.SE.94.SE9173_7022   ------------------------A-F-I----G----A---------SGA-K-----D-------------------------E-------------------------I-------------------------D-I----- 952

K.CD.97.97ZR_EQTB11   --I-------------------------I--------RV---------S-V-K-----D-------------------------E-------------------------I-------------------------E-I----- 954
K.CM.96.96CM_MP535    --I-------------------------I---------V-----T--STV-K------V-------------------------------------------------------------LN--K---------E-I----- 954

U.CA.01.TV749         --I-------------------------I---------V-----G--IS-A-K-----N-------------------------E-------------------I-------RA------------------D-I----- 959
U.CA.99.TV721         --I-------------------------I---------IL----P--IS-A-K-----N-------------------------E-------------------I---M---L---I-------------D-I----- 958
U.CD.83.83CD003_Z3    ------------L-------------V---------TV---------S-A-K-------Q-------------------------E-------------------I-------------------------D-I----- 955
U.CD.90.90CD121E12    --I---------------------D---V---------T---------S-A-K-------Q-------------------------E-------------------I---------------IQ----------D-I----- 954
U.CY.05.CY090         ------------------S---------I---------V---------S-V-K-------------------------------E-------------------I---------RL----LQ----------D-I----- 954
U.CY.08.CY223         ----------------------A---------------V---------SSALK-----DV-------------------------------------------I---------Q----S-------------D-I----- 955
U.ES.10.DEURF10DZ001  --I-------------------------I---------V-----P--S-A-K-----NVT-------------------------E-------------------I---S-------------R-------D-V----- 954
U.GR.99.99GR303       ----P-----N-----------------I---------V---------S-AMK-----NVT------S-------------------------E-------------------I-------R-----I-------D-I----- 952
U.NL.95.U_NL_95_H10986_D1 --I---A-----------------YI---------V-----P--S-A-K-------------------------------------------------I-M-------T-------------D-I----- 954

01_AE.AF.07.569M      --I-M-------------------------------V---------S-A-K-----NVR-------------------------E-------------------I-M-------------IN----------D-I----- 954
01_AE.CF.90.90CF11697 ------------------------------*---V---------S-AMK-----NVQ-------------------V----E-------------------I----E-------------------------D-I----- 951
01_AE.CN.10.YNFL03    ------------------------------------V---------S-A-K-----NVR-------------------------E-------------------I-------T-------------------D-I----- 953
01_AE.HK.04.HK001     ------------------------------------V---------S-A-K-----NVR-------------------D-------E-------------------I-------------------------D-I----- 951
01_AE.IR.10.10IR.THR48F --I-----------------------------------V---------S-A-K-----NVR-------------------------E-------------------I-------R-------------------D-I----- 946
01_AE.JP.x.JRC77AE    ------------------------------------V---------S-A-K-----NVR-------------------------I-E-------------------I-------R---NR------------D-I----- 949
01_AE.TH.09.AA090a_WG11 ------------------------I---------V---------S-A-K-----NVQ-------------------------E-------------------I-------------------------D-I----- 954
01_AE.TH.90.CM240     ------------------------------------V---------S-A-K-----NVQ-------------------------E-------------------I-------------------------D-I----- 959
01_AE.US.05.306163_FL --I-I-----------------------------V---------STA-K-----NVR-------------------------E-------------------I-------------------------D-I----- 952
01_AE.VN.98.98VNND15  ------------------------------------V---------S-A-K-----NVR-------------------------E-------------------------E-----I-------------------D-I----- 953
```

Draft of 30 September 2015, not for public release
hiv1prot_main.tex r8876–10546


```
B.FR.83.HXB2          GKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTGATVRAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRNPLWKGPA 954
02_AG.CM.10.DE00210CM013  ------------------------------F-I----------V----------S-A-K-----NVT-------------------------------------------------I----S---I-------------------D-I----- 954
02_AG.CY.09.CY256         ---------------------TDP------I----------I----------S-ALK----R-NVT-------------K-------------------------------------I---SS----------H---------D-I----- 954
02_AG.ES.06.P1423         --I-I----------------------I----------I----------S-A-K-----NVT-------------------------------------------------I----S---------I-------------D-I----- 953
02_AG.GW.05.CC_0048       ----------------------------I----------V----------S-A-K-----NVT-------------------------------------------------I----S---------I-------------D-I----- 952
02_AG.KR.12.12MHR9        --I----------------------I----------V----------S-A-K-----NVT-------------N-----E-------------------------------MI--L-S--------H-I-------G---H-I----- 958
02_AG.LR.x.POC44951       --I----------------------I----------V----------S-A-------NVTH-------------------------------------R-----I----S---------Q---------G---H-I----- 954
02_AG.NG.09.09NG_SC61     --I----------------------I----------V----------S-A-K-----DVT-------------------------------------------------I----S---------Q---------D-I----- 954
02_AG.NG.x.IBNG           --I----------------------I----------V----------S-A-K-----NVT-------------------------------------------------I----S---------I-------------D-I----- 954
02_AG.SE.94.SE7812        --I----------------------I----------V----------S-A-K-----NVT-------------------------------L-----R-----------I----S---------------------D-I----- 954
02_AG.SN.98.98SE_MP1211   --I-I----------------------I----KMAS-VV---------S-A-K-----KVTPGI-NS-------------------------------R-----------I----S---------N--K--------D-I----- 954
03_AB.RU.97.KAL153_2      --I----------------------V----------I----------STA-K-------------------------Q--QT-------------------------------I---------------I-------------D-I----- 954
04_cpx.CY.94.94CY032_3    ----M----------------------I----------M--A--P--S-A-K-----D-N-------------------------------------------------I----S---------------------E-I----- 953
05_DF.BE.x.VI1310         ----------------------D---I----------M----P---SGA-K-------Q-------------------------T----------G----I---T---------I---------------D-V----- 954
06_cpx.AU.96.BFP90        --I----------------------I----------V----------S-A-K-----N-T-------------------------E-------------------I----S---------I-------------D-I----- 956
07_BC.CN.98.98CN009       --I----------M-----------I----------V----------STA-K-------Q-------------------------------------------------I---------------------------D-S----- 954
08_BC.CN.97.97CNGX_6F     --I----------------------I----------V----------S-A-K-------Q-------------------L-------------------------------R-----I---------I-------------D-I----- 950
09_cpx.GH.96.96GH2911     --I----------------------I----------V----------S-A-K-----NVT-------------------------------------R-----I----S---------I-------------D-I----- 953
10_CD.TZ.96.96TZ_BF061    ----------------------------I----------VV---------S-A-K-------------------------Q-------------------I---------------I-------------G-I----- 953
11_cpx.CM.95.95CM_1816    ----------------------------I----------V----------S-I-K-----D-Q-------------------------E-------------------L---------K-------------D-------- 952
12_BF.AR.99.ARMA159       --T----------L-------------I----------I----------P--IS-M-K-----Q-------------------S-------------------I----S---R-------I-------------D-V----- 954
13_cpx.CM.96.96CM_1849    --I----------------------I----------V----------S-A-K-----N-T-------------------------E-------------------I----S---L---------I-------------D-I----- 961
14_BG.ES.05.X1870         --I----------------------I----------T----S-A-K-----N-T-------------------------------------------------I----S---------------F----D-I----- 953
15_01B.TH.99.99TH_MU2079  ----------V----------------------V----------S-AMK-----NVR---WD----------I---------I-E-------P-------------------I-L----P----PN-S-----------D-I----- 952
16_A2D.KR.97.97KR004      --I----------------------I----------V----P--IS-P-K-----VQ----------------------I----------S-------------I----------R---------------D-I----- 954
17_BF.AR.99.ARMA038       --I----------L-------------I----------I----------P--SSNA-K-------Q-------------------R-I-------------------I---S----R-----I-------------D-V----- 954
18_cpx.CU.99.CU76         --I-I----------------L-------V----P---S-A-K-----N-T-------------------------I-------------------------SE----F-------------------D-I----- 952
19_cpx.CU.99.CU7          ----------------------------A----V----P---S-A-K-----NVQ-------------------------------------------I---S---------S-------I-------D-I----- 953
20_BG.CU.99.Cu103         --I----------S----------I----------V----P--IS-A-K-----N-T-------------------------------------------I---S---------------------------D-I----- 950
21_A2D.KE.99.KER2003      ----------------------------I----------V----P--IS-A-K-------------------------------------D-----I----------R---------F----D-I----- 953
22_01A1.CM.01.01CM_0001BBY ----------------------I----------V----------S-A-K-----E-Q-------------------I-E-------------------I---------------------------D-I----- 954
23_BG.CU.03.CB118         --I----------------------I----------V----P--IS-A-K-----D-T-------------------------------------------I---------------------------D-I----- 950
24_BG.ES.08.X2456_2       ----------------------------I----------V----P--IS-A-K-----D-T-------------------------------------------I---S---------------F----D-I----- 950
25_cpx.CM.02.1918LE       --I----------------------------I----------VV---P---S-A-K-------------------------------T---E-------------------I---S---------I-------------D-I----- 953
26_AU.CD.02.02CD_MBTB047  ----------M----------I----------RVV----------S-A-K-----N-Q-------------------------E-------------------I---------R-----I-------------D-T----- 952
27_cpx.FR.04.04CD_FR_KZS  --I----------------------------I----------V----------S-AMK-----V-------------------------------------------I---S----R-------------------D-V----- 952
28_BF.BR.99.BREPM12609    --I----------------------------I----------V----------SNA-K-------E-------------I-------------------------I---S----R-------------------D-V----- 954
29_BF.BR.01.BREPM16704    ----------L----------I----------I----------T----S---K-------Q---V-------------------------------------------I---------------I-------------D-V----- 954
31_BC.BR.04.04BR142       --I----------------------I----------V----------SN--K-------Q-------------------------------------------------I---------------------------D-I----- 960
32_06A1.EE.01.EE0369      --I-I----------------------I----------V----------S-A-K-----N-T-------------------------------------------------I----S---------------------D-I----- 954
33_01B.ID.07.JKT189_C     ----------------------------I----------V----------S-A-K-----NVQ----------I---------E-------------------------I---------------------------D-I----- 954
34_01B.TH.99.OUR2478P     --------------T-------------I----------V----------S-A-K-----NVR-------------------------------------------------I---------------------------D-I----- 954
35_AD.AF.07.169H          ----------------------------I----------VV---------S-AFK-----S-Q-------------------------E-------------------I---------------------------D-------- 953
36_cpx.CM.00.00CMNYU830   --I-I----------S-------------I----------V----------S-A-K-----D-Q-------------------------------------------I----S---------------------D-I----- 954
37_cpx.CM.00.00CMNYU926   ----------------------------I----------RV---------S-A-K-----NV-------------------------------------------I----S---------------------D-I----- 951
38_BF1.UY.03.UY03_3389    ----------L----------I----------I----P--IS---K-------Q-------------------------N-----------------F-------------I---SS----R-F---I-------FS----D-V----- 958
39_BF.BR.04.04BRRJ179     ----------------------------I----------I----------ST--K-------------------------------L----V-------------------I---------------------------D-------- 957
40_BF.BR.05.05BRRJ055     ----------------------------I----------I----P--IS-A-K-------R-------------------------I-------------------I---S----R-------------F----D-V----- 954
42_BF.LU.06.luBF_18_06    --I-I----------L-------------I----------I----------S-V-K-------------------------------E-------------------I---------------------------D-I----- 958
43_02G.SA.03.JL1223       ----M----------------D---I----------V----------S-A-------SVT-------------------------------------------I---S-L---------------------D-I----- 953
44_BF.CL.00.CH80          ----I----------L-------------I----------V----P---S-A-K-----D-Q-------------------------R-------------I---------R-----I-------------D-V----- 954
45_cpx.FR.04.04FR_AUK     --I-I----------------------I----------V----------S-A-K-----DV-------------------------------E-------------R-------I---------------------D-I----- 952
46_BF.BR.07.07BR_FPS625   --I----------L----------I----------T---S-A-K-------Q-------------------------K-------------------------I---S----R-----I-------------D-V----- 954
47_BF.ES.08.P1942         ----------------------------------------ST--K-------------------------I-------I-------------R-------------L---------------------------D-V----- 953
48_01B.MY.07.07MYKT021    ----------------------G---P----------V----------A-K-----NVR-------------------------E-------------------I-M---------------------------D-I----- 954
49_cpx.GM.03.N26677       --I----------M----------YI---------V----------S-A-K-------Q-------------------------E-------------------I---------------------------D-I----- 951
50_A1D.GB.10.12792        ----------------------------------VV---------S-A-K-----N-Q-------------------------T-------------------I---SS----------------------D-------- 952
51_01B.SG.11.11SG_HM021   --I----------------------I----------T------P--STAAK-------------------------------------------I-------------T------V-------------------N-D-I----- 954
52_01B.MY.03.03MYKL018_1  ----------------------I----------V----------S-A-K-----NVQ-------------------------E-------------------I---------------------------D-I----- 954
53_01B.MY.11.11FIR164     ----------------------------I----------V----------S-A-K-----NVQ-------------------------E-------------------I-M---------------------------D-I----- 954
54_01B.MY.09.09MYSB023    D-I----------------------I----------V----R---SNA-K-------V-------------------------E-------------------------R---------------------D-------- 954
55_01B.CN.10.HNCS102056   --I----------V----------I----------V----------S-A-K-----NVQ-------------------------E-------------------I---------------------------D-I----- 954
56_cpx.FR.10.URF5_patient_A ----------------------I----------IS---K-------------------------------------------------I---S---S-------L---------------------D-I----- 956
57_BC.CN.09.09YNLX19sg    --I----------------------I----------VV---Y-----S-A-K-------T-------------------------E-------------------I---S---------I-V-------------D-I----- 950
58_01B.MY.09.09MYPR37     ----------------------D----------V----------S-A-K-----NVR-------------------------E-------------------I---------------------------D-I----- 954
59_01B.CN.09.09LNA423     ----------------I----------I----------V----------S-A-K-----NVQ-------------------------E-------------------I---------------------------D-I----- 951
60_BC.IT.11.BAV499        --I----------------------I----------V----------SN--K-------Q-------------------------------------------------I---------------I-------------DAI----- 949
61_BC.CN.10.JL100010      --I----------M----------I----------V----------S-A-K-------V-------------------T-------------------------I---------------------------D-I----- 950
62_BC.CN.10.YNFL13        --I----------------------I----------V----------SSA-K-------Q-------------------R-------------------------I---------------I-V-------------D-I----- 950
63_02A1.RU.10.10RU6637    --I----------------------------I----------RVV----P---SSA-K-----N-Q-------------------------------------------I---------------------------D-I----- 954
64_BC.CN.09.YNFL31        --I----------M----------I----------V----------S-A-K-----D-------------------------E-------------------I---------R---------------------D-S----- 954
65_cpx.CN.10.YNFL01       --I----------M----------I----------V----------STA-K-------Q-------------------------------------------I---------------------------D-I----- 950
67_01B.CN.11.2011.ANHUI.HF115 ----------------------I----------------------SS--K-------------------------N-------E-------------------I---------------------------D-I----- 952
68_01B.CN.11.2011.ANHUI.WH73 ----------------------I----------------------SGA-K-------------------------N-------I-E-------------------I---------------------------D-I----- 952
70_BF1.BR.10.10BR_PE004   --I----------------------I----------------------SNA-K-------------------------I-------------L-----------I---------------------------D-I----- 956
71_BF1.BR.10.10BR_PE008   ----------------------------I----------------------S---K-------------------------I-------------------------------------I---------E--D-V----- 954
72_BF1.BR.10.10BR_MG002   ----------L----------I----------I----------I----------S---K-------------------------I-------------------I---------------------------D-V----- 955
O.BE.87.ANT70             --I-I------F----------------A----V----P---ST-MK-----N-QH----------A-----S-Q-------R-------V-----------T----I--L-SQ--T-----L-XHK--------D-I----- 950
O.CM.98.98CMA104          ----I------F----------------A----V----P---S--MK-----N-XH----------A-----S-E-------R-------V-----------T----LI--L-SQ--T-----L---------------D-I----- 950
O.CM.98.98CMABB141        ----I------F----------------A----V----P---S--MK-----TN-QH----------A-----S-Q-------------------V-----------T----I--L-SH--T-----S---------------D-I----- 950
O.CM.98.98CMABB212        --I-I------F--------S----------A----V----P---S-AMK----T--QH----------------N-S-Q-------------------V-----------T----LI--L-SQ--T-----L---------------D-I----- 950
O.CM.98.98CMU5337         --I-I------F--------S----------A----I----P---S-IMK-----TN--H----------A-----S-Q-------R-------V------L--T-----I--L-SQ--T-----F---K-Q-------D-I----- 950
O.CM.99.99CMU4122         ----I------F----------------A----V----P---S--MK-----TN--H----------A-----S-Q-------R-------V-----------T-----I--L-SQ--T-----F---K-Q-------D-T----- 950
O.FR.92.VAU               --I-I------F----------------A---I-IL---P---S-AMK-----N-QH----------------ES-MQ------------------V-----------T----LI--L-SQ--T-----L---------------D-I----- 951
O.GA.11.11Gab6352         ----I------F----------------S----V----P---ST-MK-----TN--H----------A-----S-Q-------------------V-----------T----I--L-SH--T-----L---------------D-I----- 950
O.SN.99.99SE_MP1299       --I-I------F----------------A----V----P---S-AMK-----N--H----------A-----S-Q-------------------V-Y---------T----I--L-SQ--T-----F---K-Q-------D-I----- 950
O.US.10.LTNP              --I-I------F----------------A----I----P---S--MK-----N-QH----------A-----S-Q-------R-------V-----------T----I--L-SQ--T-----L---------------D-I----- 950
N.CM.02.DJO0131           --I----------L----------I----------V----P---S---K-----N-T-------------------------I-------------E-T-----I-------TK--T--L---------------D-I----- 959
N.CM.02.SJGddd            --I----------L----------I----------V----P---S---K-----N-Q----------A-------------------R---------T----I-------TK--T--LQV---------------D-I----- 956
N.CM.04.04CM_1015_04      --I----------L----------X----I----------V----P---S---K-----S-------------A-------------------------T----I-------TK--T--L-V---------------D-I----- 961
N.CM.06.U14296            --I----------L----------A--I----------V----X---S---K-----N-------------X-------------I-------------T----I-------N--T--L-V---------------D-I----- 961
N.CM.06.U14842            --I----------L----------I---A----V----------V----S---K-------------------A-------------I-------------T----I---------T--L-V-------F----D-I----- 962
N.FR.11.N1_FR_2011        ----I--------L----------M----------V----P---S---K-----N-------------------A-------------I-------------T----I-------TK--T--L-V---------------D-I----- 961
P.CM.06.U14788            --I----------FM-----V----A--V---S----V----P---S-A-K-----LN-TH----------------------------H-------------V-----------T--D--I--L--Q---T-----L-----Q-------D-I----- 951
P.FR.09.RBF168            --I----------FL----------V---A----V----P---S-A-K-----LN-TH----------------------------L-------------V-----------T--D--I--M--Q---T-----L---------------D-I----- 950
CPZ.CD.06.BF1167          ---V---------F-----LSS---K---F-I----------K----A---S-A-K-----N-Q-------------Q--I-VD-I-E----R---A--Y-----------TP-D--L--LT----------F-L---------D-V----- 957
CPZ.CM.05.SIVcpzMT145     --I-I----------------------V----------P---SNA-K-------R-------------------------E-------------------I---------------L---Q-------KD-I----- 953
CPZ.GA.88.GAB1            ----I----------------------I----------P---S-A-K-----D-------------L-------------------------T----I---------S-----L-V-K---------D-I----- 952
CPZ.TZ.06.SIVcpzTAN13     ----I--------F------E---K-----I----------R----P---STA-K-------Q-QH----------Q--Q--E-I----Q----V--Y-----------E-T----LL--LT-N--QQ-----L-V---------A-D-I----- 949
CPZ.US.85.US_Marilyn      --I-I----------L----------I----------V----------SS---------Q-------------------------I-------------T----I---SEL--DL-----L-V---------D-I----- 956
GOR.CM.07.SIVgor2139_287  --I----------F----------------A----I----P---S---K-----LNVTH----------------------Q-------------V-----------T--D--I-ML--Q---T-----L---------------D-I----- 946
GOR.CM.07.SIVgorCP2135con --I----------F----------------A----I----P---S---K-----LNVTH----------------------Q-------------V-----------T--D--I-ML--Q---T-----L---------------D-I----- 946
```

p31 Integrase end
Pol end

		Pol
B.FR.83.HXB2	KLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED.....*	1003
A1.AU.03.PS1044_Day0	----------T----------G----------	1004
A1.CH.03.HIV_CH_BID_V3538	----------G----------	1004
A1.CY.08.CY236	----------G----------	1001
A1.ES.05.X1608_8	----------G----------	1008
A1.KE.11.DEMA11KE001	----------K----------G----------	1004
A1.RU.11.11RU6950	----------N----------	1004
A1.RW.11.DEMA111RW002	----------V----------G----------	1004
A1.SN.01.DDI579	----------I----------G----------	1003
A1.UG.11.DEMA110UG001	----------E----------G--G----------	1004
A1.ZA.04.04ZASK162B1	R---------E----------	1003
A2.CM.01.01CM_1445MV	----------V----------G----------	1004
A2.CY.94.94CY017_41	----------G----------	1004
B.BR.10.10BR_MG029	----------E----------	1004
B.CA.07.502_1191_03	----------E----------	1004
B.CH.08.M2_0803101_NFLG8	----------V----------	1008
B.CN.10.DEMB10CN002	----------E----------	1004
B.ES.10.DEMB10ES002	----------I----------G----------	1004
B.FR.11.DEMB11FR001	----------V----------	1003
B.GB.05.MM45d213_GN1	----------	1003
B.HK.06.HK003	----------	1008
B.HT.05.05HT_129389	----------	1003
B.JP.12.DEMB12JP001	----------	1007
B.KR.07.HP_18_07JHS10_3909	----------V----------M----------	1003
B.PE.07.502_0525_wg5	----------	1008
B.RU.11.11RU21n	----------V----------	1008
B.TH.08.MERLBDTRC10	----------E----------	1004
B.US.11.ES38	----------	1008
C.AR.01.ARG4006	----------L----------V--K--------A----------	1000
C.BR.07.DEMC07BR003	----------L----------V--K--------A----------	1000
C.BW.00.00BW5031_1	----------A--------K--Q---A----G---N----------	1006
C.CN.10.YNFL19	----------XK-------A----G----------	1004
C.CY.09.CY260	----------E-------A----G-----QNME------	1011
C.ES.08.X2363_2	----------A----G----------	1007
C.ET.02.02ET_288	----------A----G---N----------	1006
C.IN.09.T125_2139	----------K-------A----G----------	1001
C.KE.00.KER2010	----------K-------A----G----------	1000
C.MW.09.703010256_CH256.w96	----------K-------A----G----------	1002
C.TZ.08.707010457_CH457.w8	----------K---------V--KH-------A----D-----QNME------	1004
C.YE.02.02YE511	----------V--K-------A----G----------	998
C.ZA.08.705010534_CH534.w12	---------I----------K-------A----D----------	1007
C.ZA.10.DEMC10ZA001	----------K----------A----G----------	1000
C.ZM.11.DEMC11ZM006	----------E----------A----G-----QNME------	1009
D.CM.10.DEMD10CM009	----------G----------	1009
D.CY.06.CY163	----------V----------G---G----------	1003
D.KE.11.DEMD11KE003	----------V----------	1003
D.KR.04.04KBH8	----------S----------	1003
D.SN.90.SE365	----------E----------	1009
D.TZ.01.A280	----------H----------V----------	1005
D.UG.10.DEMD10UG004	----------E----------	1003
D.UG.11.DEMD11UG003	----------V----------	1005
D.YE.02.02YE516	----------V----------G----------	1005
D.ZA.90.R1	----------V----------	1003
F1.AO.06.AO_06_ANG32	----------E----------G----------	1008
F1.AR.02.ARE933	----------E----------M-G----------	1004
F1.BR.10.10BR_RJ015	----------E---L----------G----------	1007
F1.CY.08.CY222	----------E---I----------G----------	1003
F1.ES.02.ES_X845_4	----------RE---------G----------	1006
F1.RO.96.BCI_R07	----------E----------V-------G----------	1004
F1.RU.08.D88_845	----------E--E----------G----------	1004
F2.CM.02.02CM_0016BBY	----------NE-------V----------G----------	1004
F2.CM.10.DEMF210CM001	----------NE----------X--------G----------	1009
F2.CM.10.DEMF210CM007	----------NE----------A----------	1003
G.BE.96.DRCBL	----------NE----------G----------	1000
G.CM.10.DEMG10CM008	----------NE----------G----------	1003
G.CN.08.GX_2084_08	----------NE----------G----------	1008
G.CU.99.Cu74	----------NE----------A--G----------	1000
G.ES.09.X2634_2	----------NE----------G----------	1008
G.GH.03.03GH175G	----------NE----------K----------G----------	1004
G.KE.09.DEMG09KE001	----------N----------G----------	1002
G.NG.09.09NG_SC62	----------E----------G----------	1003
G.ZA.01.TV546	----------NE---------VL---------G----------	1003
H.BE.93.VI991	----------E----------G----------	1002
H.CF.90.056	----------E------E----------	1003
H.GB.00.00GBAC4001	----------E----------G----------	1002
J.CM.04.04CMU11421	----------V---XX----------G--X----------	1000
J.SE.93.SE9280_7887	--P---------E----------G----------	1002
J.SE.94.SE9173_7022	----------E----------G----------	1002
K.CD.97.97ZR_EQTB11	----------#.Y--E----------G----------	1002
K.CM.96.96CM_MP535	----------E----------G----------	1004
U.CA.01.TV749	----------E--E----------D----------	1009
U.CA.99.TV721	----------G----------	1008
U.CD.83.83CD003_Z3	----------E----------N----------	1005
U.CD.90.90CD121E12	----------E----------N----------	1004
U.CY.05.CY090	------R-------E----------G----------	1004
U.CY.08.CY223	----------E----------G----------	1005
U.ES.10.DEURF10DZ001	----------E----------G----------	1004
U.GR.99.99GR303	----------E----------G----------	1002
U.NL.95.U_NL_95_H10986_D1	----------E----------V----------	1004
01_AE.AF.07.569M	----------E----------G----------	1004
01_AE.CF.90.90CF11697	----------G---N----------	1001
01_AE.CN.10.YNFL03	----------K----------G--N----------	1003
01_AE.HK.04.HK001	----------G----------	1001
01_AE.IR.10.10IR.THR48F	----------G----------	996
01_AE.JP.x.JRC77AE	----------G----------	999
01_AE.TH.09.AA090a_WG11	----------G----------	1004
01_AE.TH.90.CM240	----------G----------	1009
01_AE.US.05.306163_FL	----------G----------	1002
01_AE.VN.98.98VNND15	----------G----------	1003

p31 Integrase end
Pol end

```
B.FR.83.HXB2                KLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED.........*  1003
02_AG.CM.10.DE00210CM013    ------------------------------------------GG-T-----------  1004
02_AG.CY.09.CY256           ----------------------------K-------------G--E-----------  1004
02_AG.ES.06.P1423           -------------------------V------------N----G-------------  1003
02_AG.GW.05.CC_0048         ----------------------------K------------G----N----------  1002
02_AG.KR.12.12MHR9          ---------------------------N--------------G-------------  1008
02_AG.LR.x.POC44951         ------------------------V--L-------------G-------------  1004
02_AG.NG.09.09NG_SC61       --------------N---------------------------G-------------  1004
02_AG.NG.x.IBNG             ------------------------------------------G-------------  1004
02_AG.SE.94.SE7812          ------------------------V--V-------------G-------------  1004
02_AG.SN.98.98SE_MP1211     ------------------------T--L-------------GG-N-----------  1004
03_AB.RU.97.KAL153_2        --------------N---------------------------------------  1004
04_cpx.CY.94.94CY032_3      ------------------------------------N----G-------------  1003
05_DF.BE.x.VI1310           --------------E---------------------------G-------------  1004
06_cpx.AU.96.BFP90          --------------E---------------------------G-------------  1006
07_BC.CN.98.98CN009         ----------------------------K-------A----G-------------  1004
08_BC.CN.97.97CNGX_6F       ----------------------------K-------A----G-------------  1000
09_cpx.GH.96.96GH2911       --------------N-------------H-----------G-------------  1003
10_CD.TZ.96.96TZ_BF061      ------------E----------V---K-------A----------QNME------  1007
11_cpx.CM.95.95CM_1816      -------------------I-----------------------G-------------  1002
12_BF.AR.99.ARMA159         -------------D-E-------------------------G-------------  1004
13_cpx.CM.96.96CM_1849      --------------E---------------------------G-------------  1011
14_BG.ES.05.X1870           --------------NE--------------------------G-------------  1003
15_01B.TH.99.99TH_MU2079    ------------------------------------------G-------------  1002
16_A2D.KR.97.97KR004        ----------------------R-------------------G-------------  1004
17_BF.AR.99.ARMA038         --------------E---------------------------G-------------  1004
18_cpx.CU.99.CU76           --------------E---------------------------G-------------  1002
19_cpx.CU.99.CU7            ------------------------------------------G-------------  1003
20_BG.CU.99.Cu103           --------------NE--------------------------G-------------  1000
21_A2D.KE.99.KER2003        ------------------------V-----------------------------  1003
22_01A1.CM.01.01CM_0001BBY  ----------------------------E-----------G----N----------  1004
23_BG.CU.03.CB118           -------------SNE---------------------------G-------------  1000
24_BG.ES.08.X2456_2         --------------NE------------H-----------G-------------  1000
25_cpx.CM.02.1918LE         --------------E---------------------------G-------------  1003
26_AU.CD.02.02CD_MBTB047    ------------------------------------------G-------------  1002
27_cpx.FR.04.04CD_FR_KZS    --------------E----------LKH-------------G-------------  1002
28_BF.BR.99.BREPM12609      ------------------------------------------------------  1004
29_BF.BR.01.BREPM16704      --------------E---------------------------G-------------  1004
31_BC.BR.04.04BR142         ----------L------------V---K-------A--M-------QNME------  1014
32_06A1.EE.01.EE0369        --------------E------------E--------E---G-------------  1004
33_01B.ID.07.JKT189_C       ------------------------------------------G-G-----------  1004
34_01B.TH.99.OUR2478P       ------------------------------------------G-------------  1004
35_AD.AF.07.169H            ------------------------------------------G-------------  1003
36_cpx.CM.00.00CMNYU830     ----------------------------N-------------G-------------  1004
37_cpx.CM.00.00CMNYU926     ------------------------------------------M-G-----------  1001
38_BF1.UY.03.UY03_3389      --------------E---------------------------G-------------  1008
39_BF.BR.04.04BRRJ179       --------------E---------------------------G-------------  1007
40_BF.BR.05.05BRRJ055       --------------E--------V---K-------------------------  1004
42_BF.LU.06.luBF_18_06      ------------------------------------------------------  1008
43_02G.SA.03.J11223         --------------E-------------------N----G-------------  1003
44_BF.CL.00.CH80            --------------E---------------------------G-------------  1004
45_cpx.FR.04.04FR_AUK       ----------------------V---------------------G-------------  1002
46_BF.BR.07.07BR_FPS625     --------------E---------------------------G-------------  1004
47_BF.ES.08.P1942           --------------E---------------------------G-------------  1003
48_01B.MY.07.07MYKT021      ------------------------------------------M-G-----------  1004
49_cpx.GM.03.N26677         --------------K-------------------------A----G-------------  1001
50_A1D.GB.10.12792          ----------------L---------------S--M-G-------------  1002
51_01B.SG.11.11SG_HM021     ------------------------------------------G-------------  1004
52_01B.MY.03.03MYKL018_1    ------------------------------------------G-------------  1004
53_01B.MY.11.11FIR164       ------------------------------------------G-------------  1004
54_01B.MY.09.09MYSB023      ------------------------------------------------------  1004
55_01B.CN.10.HNCS102056     ------------------------------------------G-------------  1004
56_cpx.FR.10.URF5_patient_A ------------------------------------------G-------------  1006
57_BC_CN.09.09YNLX19sg      ----------------------------K-------------G-------------  1000
58_01B.MY.09.09MYPR37       ------------------------------------------G-------------  1004
59_01B.CN.09.09LNA423       ------------------------------------------G-------------  1001
60_BC.IT.11.BAV499          ----------L------------V---K-------A--M-----------------  999
61_BC.CN.10.JL100010        ----------------------------K-------A----G-------------  1000
62_BC.CN.10.YNFL13          --------------K-------------K-------A----G-------------  1000
63_02A1.RU.10.10RU6637      --------------N---------------------------------------  1004
64_BC.CN.09.YNFL31          ----------------------------K-------X----G-------------  1004
65_cpx.CN.10.YNFL01         ----------------------------K-------G----G-------------  1000
67_01B.CN.11.2011.ANHUI.HF115------------------------------------------G-------------  1002
68_01B.CN.11.2011.ANHUI.WH73------------------------------------------G-------------  1002
70_BF1.BR.10.10BR_PE004     ----------------EL------------------------G-------------  1006
71_BF1.BR.10.10BR_PE008     ----------------E--------V---------------G-------------  1004
72_BF1.BR.10.10BR_MG002     ------------------------V-----------------------------  1005
O.BE.87.ANT70               Q------------KG-------------E------T-SM--G-T-SESVEQPSEIP-  1010
O.CM.98.98CMA104            Q------------KG-------------H------A-SM--G-T-SESVEQPGEIS-  1010
O.CM.98.98CMABB141          Q------------QG-------------------T-SM----T-SESVEQPGEIP-  1010
O.CM.98.98CMABB212          Q------------KG---------------------N-SM----T-SENMDQLGETP-  1010
O.CM.98.98CMU5337           Q------------KG-------------H--X----P-SM--G-T-SESVEQPGEIP-  1010
O.CM.99.99CMU4122           Q------------KG-------------N-------T-SM--G-X-SENVEQLGEIP-  1010
O.FR.92.VAU                 Q------------KG---------------------A-S---G-T-GESVEQPSEIP-  1011
O.GA.11.11Gab6352           Q------------KG---------------------S-S---G-T-SENMEQPGEIS-  1010
O.SN.99.99SE_MP1299         Q------------KG-------------H------T-SM--G-T-SENVEQPGEIP-  1010
O.US.10.LTNP                Q------------KG---------------------T-SM--G-T-SESVEQPSEIP-  1010
N.CM.02.DJO0131             --------------G---------------------G----G-A-NQDMA.......  1013
N.CM.02.SJGddd              --------------KG--------------------G----G---NQEME.......  1010
N.CM.04.04CM_1015_04        --------------G---------------------G----G---NQDME.......  1015
N.CM.06.U14296              --------------G---------------------G----G---NQDMD.......  1015
N.CM.06.U14842              --------------G---------------------G----G---NQDME.......  1016
N.FR.11.N1_FR_2011          --------------G---------------------G----G---NQEME.......  1015
P.CM.06.U14788              T------------KG-------------N------------DT-R-SESLEQSG....  1008
P.FR.09.RBF168              T------------KG--------------------------DI-R-SESLE.......  1004
CPZ.CD.06.BF1167            R----------KEGEE-----------K----V--AGG--D--N-------------  1007
CPZ.CM.05.SIVcpzMT145       -------------QDE---------------------A-S-------S.........  1003
CPZ.GA.88.GAB1              T------------QGEL--------------------------N-------------  1002
CPZ.TZ.06.SIVcpzTAN13       R----------KEGE-------------KE-------AGS-DD--N-A.........  999
CPZ.US.85.US_Marilyn        -----------KE-EEV----------K--------A-SM-G----S.........  1006
GOR.CM.07.SIVgor2139_287    T------------KG--------------------------DT-G-SESLEQSS....  1003
GOR.CM.07.SIVgorCP2135con   T------------KG--------------------------DT-G-SESLEQSS....  1003
```

```
                             Vif start
B.FR.83.HXB2                 MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPL.GD.AR....LVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHIVSPRCEYQAGH.NKVGSLQYLALAALIT.....PKK 158
A1.AU.03.PS1044_Day0         ---------------------N-------HI-K-TKK-S----F--RN-KV--------.--.-K....I-VR-----Q---K--Q--H-------L------I--D--------H--N----------I--QV-----------.-----------R--VE.....--- 158
A1.CH.03.HIV_CH_BID_V3538    ----------------K----H--------I-K--KE-C-------R--KV--------.-E.--....--VR---------K-----H-------Q------I--D-------IH------E------I--E--R---------.-----------K--VA.....--Q 158
A1.CY.08.CY236               ------------------K--H----------KR-KN-R----F--R--KV--------.R-.-T....--VK-----Q---K--Q--H-------L------I--DM--H-------N----------I--QV---S-D-----.-----------R--VA.....--- 158
A1.ES.05.X1608_8             ---------------------N----------K--K-------F--R--KV--------.--.-K....--VI---------K-----H-------L------I--D-------MH--N----------I--Q----S-------.S----------K--V-.....QPR 158
A1.KE.11.DEMA11KE001         ---------------------H----Y-----K--KN-C----F-CRN--A--------.--.--....--VR---------KE----H------KLR-----I--DQ----------N---E------I--QV-------HT--.-----------K--V-.....-T- 158
A1.RU.11.11RU6950            ---------------------N----------K---E-R-------G---V--------.--.--....--VK------A--K-----H-------QG-----I--D---------H-----E------IV---GR---N-P---.-----------K--V-.....-TR 158
A1.RW.11.DEMA111RW002        ---------------------H--------I-R--KD-V-K----CRN--E--------.--.-T....IIVK---------KE-Q--H-------L-S----I--DM-------H-------------I---V---S---LS--.-----------K--VA.....-AR 158
A1.SN.01.DDI579              ---------------------N--------I-K--KD-------DCR---A--------.--.-M....--VR------P--Q-----H-------Q------I--D--------H-----------R-I--Q--R------T--.-----------R--VA.....S-- 158
A1.UG.11.DEMA110UG001        --------V------------N--------I-K-VN--C-------R--KV--------.--.-K....--VR---------K-----H-------L---------D--------H------E------I--QV---S-------.-----------K----.....-T- 158
A1.ZA.04.04ZASK162B1         ---------------------N------I-I-K--K----------R--KV--------.--.--....--VR-----QP--K-----H-A-----Q------I--D---------------------VI--QV-R-----S---.-----------K--V-.....--R 158
A2.CM.01.01CM_1445MV         ---------------------N--------I-R--KE------F--R--KV--------.-E.--....-IVR------P--KE----H-------Q-K----I--D---H-----------E------I--EL---K-------.-----------K--VA.....STR 158
A2.CY.94.94CY017_41          ---------------------N--------I-R--K--V-K-----RN-----------.-E.--....IIVR------I--K-----H-------QN--H--I--D---H-----------E------II-E------------.-----------K-VVA.....STR 158

B.BR.10.10BR_MG029           ---------------------N--------I-R-----V-------TN-----------.--.-K....-------------------------------------G-------M--------------I---R-----------.--------------L-.....--- 158
B.CA.07.502_I191_03          ------------------------------I-K--N--------D----K---------.--.-K....---A------------------------R--------NI-------------------N-----R-R---------.-Q---------V--TA.....T-- 158
B.CH.08.M2_0803101_NFLG8     ----------------------------IHI-K--K--I-K-----TN-----------.--.-K....--V------------------------ERK----I--G-------Q-H-----E----N-I--NR-R---------.-E---------T----.....--R 158
B.CN.10.DEMB10CN002          ----------------K-N-------Y---I-----K-S-------S--------Y---.--.-K....--V-----------A----H-A------G--------G--------------------S-IV-R------------.-----------V--V-.....--- 158
B.ES.10.DEMB10ES002          ------------------------------I-----K-I-K-----SN--V--------.--.-K....---K------------------------RS-N-----D---------------E----N-I--R------------.S------H---T----.....--R 158
B.FR.11.DEMB11FR001          --------------------------------K--KK-I-------NN-----------.-E.-N....---------------------------ER---------T------------------KN-I--R----S-------.--------------K-.....--- 158
B.GB.05.MM45d213_GN1         ------A-------------------------K--Q----------T------------.-E.-K....---K----------E------------I---------N---------------E----NTI---R-----------.------------T--A.....-R- 158
B.HK.06.HK003                -A-------------------H----Y--HI-K--K--S-K-----I--K---------.--.ET....---------Q-------M-H---------G-R-----G---K----H---------V---I---R---S-------.-Q---------K----.....--- 158
B.HT.05.05HT_129389          -------------------A---------------KR---------T------------.--.--....----------------------------R--------G---------------E------I--YR-----------.-----------V----.....--- 158
B.JP.12.DEMB12JP001          ------------------NA--------I-I-K--ER-V-------SN-----------.-G.-K....---I------------------------Q--------N---------------E----Q-----R---S---R---.S----------T----....PT-R 159
B.KR.07.HP_18_07JHS10_3909   ------------------------------I-K--KE-V-------T------------.--.-K....--------------E----------------N-----D---K----H-----------H-I--RL-R-E---P---.-T-----K---T----.....--- 158
B.PE.07.502_0525_wg5         ------------------N----I-----H-----KR-S------GT------------.-E.EKLEEK----------------------------DK-------N---------------E----N-IV-R--------H---.-----------T--V-.....--- 162
B.RU.11.11RU21n              ---------------------N------IH--K--S--V-K-----RN-----------.--.-K....---K---------K-------A------E--------NI--------------E----N-I--RR---S-------.--------------R-.....T-- 158
B.TH.08.MERLBDTRC10          ------------------N-------Y--H--K-------------T------------.--.-E....--VI-----N----E------A---------------NV--------------E------I--SR-R---------.-----------T--L-.....--- 158
B.US.11.ES38                 --------------------------------R-------------T------------.--.-K....---K----------------------K-R--------Q---------------E----Q-I---------------.--------------VK.....--- 158

C.AR.01.ARG4006              -------L-----------A-N--------I-K-----Y-------R--KV--------.--.--....---I-----Q---------H-------LXK-------G-------MH-----T-----------R-----------.-----------T--LK.....--- 158
C.BR.07.DEMC07BR003          -------L--------K----N--------I-KR-K--Y-------R--K--------V.-E.-K....---------Q----A----H---V---LR--------D-------IH-----A-----------R-----D-----.-----------T---K.....--R 158
C.BW.00.00BW5031_1           -D----GL-------------N----------RR-N-------H--R--KV--------.--.--....---K---------------H-------LR--N-----G-------MH-----A-------I--R--I---D-----.-Q---------T---K.....--R 158
C.CN.10.YNFL19               -------L--------K----N----------KR-N--------D-R--KV--------.-E.--....---K-----Q---K-----H-------LRK-N--IE-G-------M------A-------I-----I---D-----N-----------T---K.....--R 159
C.CY.09.CY260                -------L--------K----N----------K--N----------R---V--------.-E.-S....---R-----Q----E----H-------W---------G-------MH-----A-----Q-I-----I---------.K----------T---K.....--- 158
C.ES.08.X2363_2              -------L-----------A-N----------KR-N----------R--K----I----.--.--....-I-K-----Q----E----H-------LG--------G-------MH-----T-------I--R--T---D-----.-----------T--VK.....--- 158
C.ET.02.02ET_288             -------L--------K--A-N--------I-RR-N--L-K-----RN--T--------.-E.--....-IVK-----Q---------H-------LRG-T-----C-------MH-----AE------I--QR-----D-----.-----------T--LK.....--- 158
C.IN.09.T125_2139            -------L-------------H----------RR-K--V-------R--KV--------.-E.-K....---K-----QP--------H-------LRS-N-----G-------MH--N--AE------I---L-I---D-P---.T----------T---R.....--Q 158
C.KE.00.KER2010              -------L-------------N-------HI-RR-S-------F--R--KV--------.-E.--....---K-----Q---------H------KL---N-----G-------MH-----A-------I-----I---------.-----------T---K.....-N- 158
C.MW.09.703010256_CH256.w96  -------L--------K----N----------RR-D----------R--K---------.--.--....---K-----Q----A----H-------L---------D-------MH-----A-----Q-I---V-----D-----.-----------T---K.....--- 158
C.TZ.08.707010457_CH457.w8   -------L----I--------N----------RR-N--Y-------R--KV--------.--.--....---K-----Q----E----H-------LR--------G-------MH-----A-------I-----T---D-----.-Q---------T--VK.....--- 158
C.YE.02.02YE511              ------GL-------------N--------I-RR-S----------R--KV--------.-E.--....---K---------------H-------LR--------G-------MH-----T-------I-----I---D-----.T----------T---K.....--- 158
C.ZA.08.705010534_CH534.w12  -------L----I---K----N--------I-K--S----K-----R--KV--------.--.-K....--VK-----Q----A----H-A-----LRE-------G-------MH-----A-----------------------.-----------T---K.....--- 158
C.ZA.10.DEMC10ZA001          ------GL-------------N--------I-KR-K-------F--R--K-----Y---.--.--....-I-K-----Q----E----H-------LRK-------G-------MH-----A-----Q-I-----I---D-----.-----------T---K.....--G 158
C.ZM.11.DEMC11ZM006          ------GL-------------N--------I-R--S----------R--KV--------.-E.--....---K-----Q----E----H-------LR--------G-------MH-----A-------I---------D-P---.-----------T---K.....--- 158

D.CM.10.DEMD10CM009          -------------------I----------I-K-TA--------D----K---------.--.-K....--VI------------------------R--------G-------M-------E------I---R-----------.-----------T---K.....--Q 158
D.CY.06.CY163                --------V-----------------Y---K-K--A--------DNH--K---------.--.-K....--V------N------------------R--------G-------M--------------I---R-----------.-----------T---A.....--Q 158
D.KE.11.DEMD11KE003          --------------------------------KQ-KR-W------C-N-----------.-E.-K....--VK----------E-------------R--------G-------TH-----A-----Q-I---R-IS--N-P---.-----------T--LA.....--- 158
D.KR.04.04KBH8               ------------------N-------Y---I-K-----V-----D-N------------.--.-K....--V------N----K------------L---R--I--G-------MH-------------I---R-----D-----.-----------T----.....--R 158
D.SN.90.SE365                ------------------K-------------K--Q--C-K---D----K---------.-E.--....--VK----------E-------------E--------G-------I-------E----Q-I---------D-----.-----------T--RA.....-R- 158
D.TZ.01.A280                 --------------------------------R--Q--L-----D-H--K---------.--.--....--V-----------E-------------R--------G-------M------T-------I---------D-----.-----------I---K.....--- 158
D.UG.10.DEMD10UG004          ------------------K-------------K--K--W--------------------.-E.--....--VK----------K-------------G--------G-------T------AE----Q-I-----I---N-----.-----------T---K.....--Q 158
D.UG.11.DEMD11UG003          ------------------S---G-----I---KQ-N--L-----DC-N-K---------.-E.--....--VR-----Q----E-------------R--------G-------I------AE------I-----F---N-----.-----------T--V-.....--R 158
D.YE.02.02YE516              ------------------N-----------I-XE-K--------D----K---------.-E.-K....--VI-----N-----------------QR--------S-------T-------E------I--LR-----D-----.-----------T---A.....-RR 158
D.ZA.90.R1                   ------------------N-------Y--H--K--K--------D----K---------.-E.E-....--VK----------K----P--------R--I-----G-------I-------E------I---R-----------N-----------K---A.....S-- 159

F1.AO.06.AO_06_ANG32         ------------------N-------Y-IH--K--K--V----F--R---V--------.RK.DK....---------N----E------------LGK----I--SM--------------E----R-I--YR-----N-----.-----------T---H.....--- 158
F1.AR.02.ARE933              ------------------N-------Y--H--K--KK------F--R---V--------.EE.-K....---I-----------------------XGN-R--I--G-------I-S-----E----G-I---RI----N-----.-RL--------T---A.....-Q- 158
F1.BR.10.10BR_RJ015          ------------X-----N-------Y--HI-R--K--C-------RN-----------.DK.-N....I-V------------------------QG--R--I--G-------I-------E------I--YR-----N-----ST----------T---A.....--G 159
F1.CY.08.CY222               ------------------N-------Y--HI-K--KN------F--R--K--------I.-E.-I....I-V-----------E------A-----QG-----I--G-------I-------E------I---KI----N-----.-----------T---P.....--- 158
F1.ES.02.ES_X845_4           ------------------N-------Y--H--K--KK---------R------------.D-.-K....V--------------------------QG--R--I--GT------IH-----------R-I--YK-----N-----.-----------T---A.....T-- 158
F1.RO.96.BCI_R07             -*----------------N-------Y--HI-K--KK-S----F--R---V--------.EE.-K....--V-----------E------------QG-----I--G-------I--------------I--YR-----N-----.-----------T---A.....--- 157
F1.RU.08.D88_845             ------------------N-------Y-----K--K-------F--R------------.EE.-K....I-V-------P----------------LG--R--I--SM------IH------E----------RI----N-----.-----------T--VA.....--- 158
F2.CM.02.02CM_0016BBY        ---------------------N----------RR-Q----------R---V--------.ENDSK....---V---------K-----H-------Q---R-----G-------K-------E------I---R-----N-----.-----------T--V-.....--- 159
F2.CM.10.DEMF210CM001        -X-------------------N----------R--A--------Q-R---V--------.A-DST....---V-----------------------R---R-----G--------H------E-T--N-I---R-----N-----.-----------T---N.....--R 159
F2.CM.10.DEMF210CM007        ---------------L-----N----------K--K----------R---V--------.K-DSK....---V---------------H-------Q-X-R-----G--------H------E----N----YR-----N-----.-----------T----.....--- 159

G.BE.96.DRCBL                ---------------------N--------I-K--K----------R--KV--------.--.-K....--VR--------DKE----H-------QG--------D---H----H--N---E------I--QT-R-S---P---.-----------KV-VA.....-TR 158
G.CM.10.DEMG10CM008          ---------------------N----------K--K----------R--KV--------.--.--....-IVR-----N---KA----H------SLR-----I--D-------SH-------------I--Q--R---------.-----------K--V-.....-TR 158
G.CN.08.GX_2084_08           ---------------------N----------K--K----------R---V--------.--.-K....-IVR---------K--Q--H-------QR--H-----D--------H-------------I--QL-R-K-------.-----------KV-V-.....-T- 158
G.CU.99.Cu74                 --------L------------N----------K--KD-W-------R--K---------.--.--....-IVR---------K-----H------KQRK----I--D--------H--K--------Q-I--EV-R---------.K----------K--V-.....-T- 158
G.ES.09.X2634_2              ---------------------N----------K--K----------R--KV------L-.--.-K....--VR---------------H-------QG-----I--D--N-----Q--E---N------I--R--------P---.S----------K--V-.....-TR 158
G.GH.03.03GH175G             ---------------------N--------I-K--K-----R----R--KV--------.--.-T....--VK---------QA----N-T-----QR--N--I--D---H----H------E------I--E------------.-----------K--VK.....-AR 158
G.KE.09.DEMG09KE001          ---------------------N----------K--K-------F--R--KV--------.-E.-T....---K-----N---K--Q--H-------QR--I--I--D--------H--N----------II-E--R-K-------.-----------K--VR.....TP- 158
G.NG.09.09NG_SC62            ------------------K--N-------HI-K---K-S-------K---V--------.E-.-K....--VR---------K--Q--H-------QR-----I--D--------H----------------QV-R---------.-----------K--VQ.....-T- 158
G.ZA.01.TV546                ------A--------------N--------X-RX-KD------X--R---V--------.-E.--....I-VR------K--KE-Q--H-------LR-----I--D--------H------E------I--Q------D-----.-----------K--VK.....-VR 158

H.BE.93.VI991                ------------------K--N----------K--KK-V-------TN-KT-------V.--.--....-------------------H-------QE-----I--D--------H-------------I---R--I--D-----.R----------T---S.....--R 158
H.CF.90.056                  ------------------N-------Y--HI-R----------F--T------------.-E.--....---------N----E------------L-------E-G-------MH------E------I--RV-R---N-P---.KQ--T------T--VA.....--- 158
H.GB.00.00GBAC4001           ------------------N-------Y--HI-K--KD------F--T--K---------.-E.--....---I-----N-----------------QR--------D-----------N---E------I---R-Y-S----T--.-----------T--V-.....--- 158

J.CM.04.04CMU11421           ------------------N-------Y-XHX-K---R-L-----D-N--K---V-Y--X.-E.-E....-X-------Q-----------------L-X-R-----G-------MH------E------X---X-----D-P---.-Q---------T---K.....-XX 158
J.SE.93.SE9280_7887          ------------------N-------Y--N--K---K-L-----D-N--K---------.-E.-I....---------Q-----------------QR--R-----G-------MC-------------I--Q------D-----.-----------T---R.....--R 158
J.SE.94.SE9173_7022          ------------------N-------Y--N--K---Q-L-----D-R--K---------.-E.--....--V------Q-----------------R---R-----G-------MH-------------I--Q------D-----.-----------T---K.....--R 158

K.CD.97.97ZR_EQTB11          ------------------N-------Y-----K--NR-R-----D-N--K---------.--.-E....--V-----------E------------L-K-R-----G-------I-------E----------R-----------.TQ---------T---A.....--- 158
K.CM.96.96CM_MP535           ------------------N-------Y--HI-K--NR-Y-------R--K---------.--.-E....--V------L-----------------L---R---E-D-------I-------E--V---I---R------C----.-----------T--VA.....-RR 158

U.CA.01.TV749                ------------I--------N----------K-------------R---V--------.-E.--....--V-----------E------------LG--R--I--G-------TH------E------I--Q------------.-----------K--VX.....--G 158
U.CA.99.TV721                -------------------A-N----------K-----------Q-R--KV---I----.-E.--....-IV-----------E----S------ELR--R--I--G-------TH------E------I--Q------------.-----------K--VA.....--G 158
U.CD.83.83CD003_Z3           ------------------K--N--------I-K--K--V-K-----TN-----------.-N.--....I-V-A----------------------QG-----I--D--------H-------------I------------T--.-----------TT-VA.....-T- 158
U.CD.90.90CD121E12           ---------------------N----------K--NK-V-K-----IN-----------.EN.--....I---A----------------------QG-----I--D-------MH-------------I---------D--T--.-----------T--VP.....-I- 158
U.CY.05.CY090                ----------------K-K--N--------I-K---N-R-K-----RN--V--K-----.KE.-K....---------------------------QG-I-----G-------MH-----T----------LR-----D-PE--.I---------IT---A.....-E- 158
U.CY.08.CY223                ---------------------N--------I-K--K----K-----MN--V--------.-E.-K....IIVR---------K--Q--H-------Q-S----I--D---K----Q------E------I--EV-R---AFS---.-----------K--V-.....-RR 158
U.ES.10.DEURF10DZ001         ------------------N-------Y--HI-K--KK-M-----D-R--K--------I.-E.-V....---------N----A------------L---R--I--S-------VH-------------I--D--Y---------.S-------------VK.....T-- 158
U.GR.99.99GR303              ------------------N-------Y--H--K--KR-V-----D-N------------.RE.-K....--V-----------E------------M-S-R-----S-------MH------E------I--Q------D-----.-----------T---N.....--R 158
U.NL.95.U_NL_95_H10986_D1    ------------------S-------Y--H--K--K----------R---V--------.K-.-K....------R------G--Y---------GQRK-------G-------MH-----------------R-----------.-----------TV-V-.....--R 158

01_AE.AF.07.569M             ---------------------N------K-I-K--KK-L-------QN-KV----Q---.-E.G-....-I-R-----Q---K--Q--H-------LRK-N--I--D--------H-------------I--QV-RH--D-PS--.-----------K---A.....--- 158
01_AE.CF.90.90CF11697        ---------------------Y--------I-K--KN---------Q--KV-------I.--.--....---R---------K-----H-------QRK-------D---RQ---Q-------------M--QV-R-----PT--.-----------K--A-.....--- 158
01_AE.CN.10.YNFL03           ---------------------N----------K--K----------R--KV--------.-E.-K....---I----PQ---K-----H-------Q------I--D---R----Q------E-----RI--QV-RH----PS--.----P------K--T-.....--R 158
01_AE.HK.04.HK001            ---------------------N----------K--KK-Y-------Q--K---------.-E.--....---K-----Q---K--Q--H-------QRN----I---V--R----Q-----A------RI--QV-RYK---PS--.-----------R--TG.....--G 158
01_AE.IR.10.10IR.THR48F      ---------------------N----------K--KK---------W--KV-------I.-E.-K....---R-----Q---K-----H-------QGK----I--D---R----Q-----ET-----GI--KV-R---X-PS--.-----------K--T-.....--- 158
01_AE.JP.x.JRC77AE           ---------------------N----------K--KK-V-------Q--KV--------.-E.--....---R-----Q---K-----H-------QGK-N--I--D---K----Q-------------I--QV-RH----PS--.-----------K--T-.....--- 158
01_AE.TH.09.AA090a_WG11      ---------------------N----------K--KK-C----------KV--------.-E.-K....I--R-----Q---K--Q--H-------QGQ-N-----D---R----H------G-----SI--QV-RH--D-PS--.-----------K--T-.....--R 158
01_AE.TH.90.CM240            ---------------------N--------I-K--KK---------Q--KV--------.-E.--....---R-----Q---K-----H-------QRK----I--D---K----Q--G----------I--QV-RR----PS--.-----------K--T-.....--R 158
01_AE.US.05.306163_FL        ------------------K--N-I----I---K--KK-S-------Q--KV----Q--I.-E.--....---R-----Q---K--Q--H-------QRE-N--I--D---R----Q-----A-------I--QV-RR----PS--.-----------K--T-.....--R 158
01_AE.VN.98.98VNND15         ---------------------N--------I-K--KK-C-------Q--KV--------.-E.-I....---R-----Q---K--Q--H-------QRE-N--I--D---R----Q--E---------RI--QV-RR----PS--.-----------K--T-.....--R 158
```

Vif start

```
B.FR.83.HXB2                MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPL.GD.AR....LVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHIVSPRCEYQAGH.NKVGSLQYLALAALIT.....PKK  158
02_AG.CM.10.DE00210CM013    ---------------L--K--H--------I-R--KD-S-------R---V--------.--.-K....-IVR-----------Q--H------KQES---I--D-------MH------G------I--QV-----------.-----------K--VP.....-V-  158
02_AG.CY.09.CY256           ---------------R-----H-I--Y--HI-R--KD-C----F--R---V--------.--.........--VK----------------H------EQR----I--D-------H---------------QV-R----------.K---------K--V-.....-I-  158
02_AG.ES.06.P1423           --------------------N---------K--GK---------R--K---------.ES.-T...------------------H------M-----I--D-------HH-N---E------I--QV-R----------.-Q---------K--V-.....-TR  158
02_AG.GW.05.CC_0048         ------------------A-N--------R-K-TKD--------TR--KV---------.--.......--VK-----N---------H------KQ-----I--D-------H-----------Q-I--QV-R-S--------.-----------K--V-.....-TR  158
02_AG.KR.12.12MHR9          --------------------N-I------I-K--KK-V-------N--KV---------.--.-K....---R------------------H------KQ-----I--D-------H--------------I--QV-R-I-------.S----------KV-V-.....-TR  158
02_AG.LR.x.POC44951         --------------------N----------R--KD---------R--KV---------.--.-T....---R------------------H-------Q-----I--D-------H--K--T-------I---V---I----E--.-----------K--V-.....-TR  158
02_AG.NG.09.09NG_SC61       --------------------N--------I-K-TQ----------R--KA---------.--.--....-IVR-----------------H------EQ-----I--D-------H----T----------QE-R----------.R----------K--V-.....-TR  158
02_AG.NG.x.IBNG             --------------------N----------K--K-------F--R--KVC--------.--.......--VR-----------------H------KQ-----I--D-------------N----------I--E--R---------.-----------N--VA.....-T-  158
02_AG.SE.94.SE7812          --------------------N----Y---K-R--KD---------S---V---------.--.......--VR-----------------H------KQR----I--D-------H--------------I--Q--R---------.T----------K--V-.....-TR  158
02_AG.SN.98.98SE_MP1211     --------------------N----------K--HR---------R--KV---------.--.-T....-IVR----QP---------H---V-QQ-----I--D-------H------E---------QV-R-K--------.-----------K--V-.....-TR  158
03_AB.RU.97.KAL153_2        --------------------N------I-I-K-----V-K------RN-----------.--.-K......-K-------------------A-----E-------N---------------E----N-I---R---S---R---.--------------R-.....---  158
04_cpx.CY.94.94CY032_3      -A-------------K----N----------K--K-*--------R--KV---------.-E.--....--VR-----QP--Q----H------LR--------D-------MH------E-------I---R------------.---------------S.....---  157
05_DF.BE.x.VI1310           -------------------N-------Y--H--K--NR-C----F--RN----------.--.-K....--V--------------------------R-------S-------V-------E---N-I--R-------F----.-----------T---.....---  158
06_cpx.AU.96.BFP90          -------------------N-------Y--HI-K--KR-N-----D-N--K---------.S.-E...----------N-----K-------------LR--R-----SM-----IH------E----------R-----D-----.-----------T---K.....-E-  158
07_BC.CN.98.98CN009         -------L--------K----N----------RR-N---------D-R--KV-------V.-E.-K....---K-----Q----------H-------LR--T--IE-G---------------A-------I-----I--D-----N-----------T---K.....---  159
08_BC.CN.97.97CNGX_6F       -------L--------K----N----------RR-N---------D-R--KV--------.-E.--....---K-----Q----------H-A-----LRS-N--IE-G-------M------A-------I---V-I--D-----N-----------T---K.......-R  159
09_cpx.GH.96.96GH2911       -------------------N-------Y--HT-K--KR-------D-N--K---------.A.-E....--V----------K-------------L---R-----S-------IH------E-------II-DR-----------.-----------T---N.....-R-  158
10_CD.TZ.96.96TZ_BF061      -------L-------------N----------R--N--S-------R--K---------.-E.-S....-I-K----------E----H-------L-------I--G-------MH-----AE-------I--QV-T---N------.KQ---------T---.......---  158
11_cpx.CM.95.95CM_1816      --------------------N--------I-K-TNR-R-------R---V---------.-P.-T....--VI----------E-------------L---R-----S-------M------TE-------I---R-R-----P---.-----------K--V-.....-T-  158
12_BF.AR.99.ARMA159         -------------------N-------Y--H--K--KR------F--R---V--------.EK.-E..........-QP---------------QG--R--I--G---R---I-------E-------I---KI----N------.-----------T---A.....---  158
13_cpx.CM.96.96CM_1849      -------------------N-------Y--H--K--GR-------D-N--K---------.--.-K....--V------------K-------------L---R-----G-------TH---------------I---R-----D-----.-----------T---H.....---  158
14_BG.ES.05.X1870           --------------------N-I------I-K--K--S-K------R--KV---------.--.-K...X-VR----------K----H------QR-----I--DX--X---QH------E-------I--G-----K---P---.S----------K-VV-......-Q  158
15_01B.TH.99.99TH_MU2079    --------------------N----------K--KK---------Q--KV--Q------.EE.--.....----K-----Q---K----H------QRG----I-----R----Q------E----R-I--QV-RYS-D-PS--.-----------K--T-......-R  158
16_A2D.KR.97.97KR004        --------------------N----------K--KD-C-------IN--V---------.-G.-K...IIVK------P--K------------QE-----I--D--GH----H--------------I--K--R----------.-----------K--VG.....ETR  158
17_BF.AR.99.ARMA038         -------------------N-------Y-IH--KQTKR-V----F--RN--V--------.E-.-K....---A------P------------------QG--R--I--G-------I-------E-------I--NKI----N------.-----------T---A.....---  158
18_cpx.CU.99.CU76           --------------------N-I----------K--KN---------R--KV---------.--.-T....--V----------KE---H------LR--R-----D---------------E----N-I---R---------------...........---  158
19_cpx.CU.99.CU7            --------------------N--------I-K--NK-V--------RN-KV---------.--.-K....--VK-----------K----H------QR-----I--D-------H-------------------QV-R---------.-----------K--V-.....-TR  158
20_BG.CU.99.Cu103           --------------------H----------K--K--W----F--R--KV---------.--.-K....IVR-----------K----H-----KQRE---I--D--HM------N------------I--E--R----FH---.-----------K--V-.....-T-  158
21_A2D.KE.99.KER2003        --------------------N--------I-K---R---------R---V---------.-E.-T....-IVR------S--K----H-------Q-----I--D--H----H-------T-----I--EL------------.-----------KV-VA.....A-R  158
22_01A1.CM.01.01CM_0001BBY  --------------------N----------K--------------R--K---------.--.-.....--VR----------KE---H-------Q---------D-------M--------------I--Q----S----T--.-----------R--VA.....---  158
23_BG.CU.03.CB118           --------------------H------K-I-K--K--C-K-----TN-----------.-E.-K....-I-K------------------------------G------Q-------E---S-I---R---S------------.-----------K--V-.....-T-  158
24_BG.ES.08.X2456_2         ---K----------------N--------I-K--KD-W----F--R--KV---------.--.-N....-IVR----------K----H-----KQRK---I--D---R---QH--K---------------D-F-----P---.-----------K--V-.....-T-  158
25_cpx.CM.02.1918LE         -------------------N-------Y--H--K--KK---------H------------.--.-E....--V----------E-----------L---R-----G---R---MH------E-------I-------K-D-----.K----------R--VA.....---  158
26_AU.CD.02.02CD_MBTB047    --------------------H----Y---R-K--K--L-------R--KV---------.--.........---R-----Q----E----H------LR----------G-------Q--E----------IV-Q-IIHK-------.K----------K--VA.....T-R  158
27_cpx.FR.04.04CD_FR_KZS    ------G-------------NA------Y--HK-K--K--H----D-H--K---------.--.-E....--V-----Q------------------L---R-----G-------IH------E-------I--YR-----D-P---.-Q---------T---.......---  158
28_BF.BR.99.BREPM12609      --------------------N--------I-R--KK-V-K------T---V---------.--.-K....--VI-------------------------E-------D---------------E----N-I---R--R---D--T--.-----------T--TK.....---  158
29_BF.BR.01.BREPM16704      -------------------N-------Y--HI-K--KR------F--R---V--------.EE.DEL...V---------------------------GK-R--I--G-------IN------E-------I--QKI-----------.-----------------.......-R  159
31_BC.BR.04.04BR142         -------L--------K----N--------I-KR-S--S-K-----RN-K---------.-E.--.....---K-----Q----------H---L---LRG------G---R---MH-----AE-------I---R-----D-----.-Q---------T---K.....---  158
32_06A1.EE.01.EE0369        --------N-L-------N-------Y--HI-K--KN-------D-H--K---------.S.-E...----------N----E-------------L---K--I--S-------MH------E-------I-----RH---------.-----------K---K.....---  158
33_01B.ID.07.JKT189_C       ----------G-----K-K--H------V-I-K--K---------Q--KV---------.-E.-K...I--R-----Q---K-----H------QRD----I--D-------N--------------I--QV-RH----PS--.-----------K--T-.....Q-R  158
34_01B.TH.99.OUR2478P       --------------------N-I----------K--KK---------Q--KV---------.-E.--.....---R-----Q---K--Q--H------QRN----I--D---R----Q------E----R-I--QV-RR----PS--.-----------K--T-......-R  158
35_AD.AF.07.169H            -------------------K--N--------I-K--K--V-------RN--V----Q---.--.--...I-VR-----------K-----H------LQ-----I--DV-------H--N---------I--E-----S---------.-----------K---.......-T-  158
36_cpx.CM.00.00CMNYU830     --------------------N--------HI-KT-KD---------R--KV---------.--.......---R-----------------H-----KQRK---I--D-------H--------------I--Q------D-P---.-----------K--V-.....-T-  158
37_cpx.CM.00.00CMNYU926     --------------------N--------HI-K--K-----------RD--V--------.-E.-W....--VR-----QP---------H------LRK-R--I--D-------H--------------T-I--R-R-----P---.S----------K--VP.....-TR  158
38_BF1.UY.03.UY03_3389      -------------------N-------Y--H--K--KR-C-K--F--RN--V--------.QE.DK...--------------------------QGS-R--I--D-------I-------E----N-I---RI----N------.-----------T---A.....-Q-  158
39_BF.BR.04.04BRRJ179       -------------------N-------Y--H--K--KE---------T------------.-E.--....--V------------------------------R--R-----D--------------AE---H-I--R--R---N------.-----------T---A.....---  158
40_BF.BR.05.05BRRJ055       ---------------------------I-R--NK-------K-T--K---------.-E.-A....--V----------E--------K--RK----N-D-----------N---E---N----R----S-N-------.---K------T---K.....---  158
42_BF.LU.06.luBF_18_06      --------------------N------I---K--K--H-------I--K---------.--.S-.....-K--------------------------R---------D---------------E----NTI---R-----------.-Q---------T---.......---  158
43_02G.SA.03.J11223         --------------------N----------K--K---------RN-KV-------X.X-.-T....-XXX-----Q------Q--H------XQ-K-------D-------H------E-------I--Q--R-----------.-----------K--VK.....-TR  158
44_BF.CL.00.CH80            -------------------N-------Y--H--K--KR-L----F--R---V--------.E-.-X......................-A-----QG--R--I--G-------IH-----T-----------YK-----N-P---.-----------T--VA.....---  158
45_cpx.FR.04.04FR_AUK       --------------------H----------K--KR--------TR--KVAA--------.--.-T....--VR-----Q---K-----H-------L---------D-------H-------------I--QV-R-----T--.-----------K--V-.....-TR  158
46_BF.BR.07.07BR_FPS625     -------------------N-------Y----K--KR--------TR---V--------.DK.DT...---------------------------QG--R----D-------I-------------I--QK-----N-----.-----------T--VA.....---  158
47_BF.ES.08.P1942           -------------------SK------Y-IHT-K--KK------F--R--K---K-----.EEENK....--V------N-------------------QGK-R--I--G-------I-------E-------I---R-----N-----.-----------T---A.....T--  159
48_01B.MY.07.07MYKT021      --------------------N----------K--KK-C-------QN-KV---------.EG.-K.....---K-----Q---K--Q--H------QRD----I--D-------HH-------------I--QV-RR----PT--.-----------K--T-.....Q-R  158
49_cpx.GM.03.N26677         -------L-------------N-------HI-R--S----------R--KV---------.-E.--....--VR-----Q----E----H-------LR--------G-------MH-----T-------I-----R---V-P---.-----------T---K.....T-R  158
50_A1D.GB.10.12792          -------V-------------H-I----I-I-K--KN-V----F--RN-KT---------.--.-K....--VK----------K-----H-------LR-----I--Q-------H------E-------I--QV---S-------.-Q---------K--V-.....-TR  158
51_01B.SG.11.11SG_HM021     --------------------N----------K--KK---------Q--KV---------.-K.--.....---R-----Q---Q--Q--H------QRD-N--I--D-------H--E---------I--Q--IH----PS--.-----------KT-T-......-R  158
52_01B.MY.03.03MYKL018_1    --------------------H-V--------K--KK-V-------Q--KV---------.-E.-K....--VR-----Q---K-----H-------QGL----I--D-------Q----A-------I--QV-RH----PS--.-----------K--T-......-R  158
53_01B.MY.11.11FIR164       ---K----------------N-----I---K--K.-V-------QN-KV---------.--.-K...I--R-----Q---Q-----H------QGN-N--I--D-------HH----------NTI--QRIRS----PS--.-R---------T--T-.....Q-R  157
54_01B.MY.09.09MYSB023      -------------------K--------I-I-K--K--V-K---D-M--K---------.-G.-K..............................D-RES--------D---------------E----N-I--RV---S-A-----.-----------K--T-.......---  158
55_01B.CN.10.HNCS102056     --------------------N--------I-K--K----------Q--KV--------I.-E.--.....---R-----Q---K--Q--H------QRT-R--I--D---R----H--N-------R-I--QV-RH----PS--.-----------K--T-.......---  158
56_cpx.FR.10.URF5_patient_A ---------------K----N--------I-K-TK----------R--KV---------.--.--.....-IVR-----Q-----------H-------Q-----I--D-------H------E-------I--QV-R-S-------.S----------K--V-.....-T-  158
57_BC.CN.09.09YNLX19sg      ---------------K----N----------RR-N----K---D-R--KV-------V.-E.-T....-I-K-----Q----------H---V---LR--T--IE-D---H-----------A-------I-----I--D-----K------H----T---K.....---  159
58_01B.MY.09.09MYPR37       --------------------H-------HI-K--KK-S-------Q--KV---------.--.-K....I--R-----Q---K-----H------QGN-N--I--D-------N--------------I--QV-RR----PS--.-----------K--T-.......---  158
59_01B.CN.09.09LNA423       --------------------N----------K--KN---------Q--K--------I.-E.-K....I--K-----Q---K-----H---V---LR--N--I--D---R---Q--------------I--QV-RR----PS--.-----------K--T-.......-G  158
60_BC.IT.11.BAV499          -------L--------K----N----------KR-S--Y-------R-------------.EE.T-...........S-Q----------H-------LR--R-----D-------VH-----A-------I---R-----D-----.-----------T---K.....---  158
61_BC.CN.10.JL100010        -------L--------K----N----------KR-N---------D-R--KV--------.-E.--.....---K-----Q----------H---V---LRK----IE-D---------------A-------I-----I--D-----N-----------T---K.....---  159
62_BC.CN.10.YNFL13          -------L--------K-K--N----------RR-N-----------RN-KV-------V.-E.--.....---K-----Q----------H-------M------IE-GM-------H-----A-------I-----I--D-----..-----------T---R.....---  157
63_02A1.RU.10.10RU6637      --------------------N------I---K---E---------R--KV--------.--.--.....--VR-----QK---------H-------RE--------D-------H--E---ET-----IV-E---T--N-S---.-----------K--V-.....-TR  158
64_BC.CN.09.YNFL31          -------X--------K----N----------KR-K---------R--K---------.-E.--.....---K-----QP---------H-------WR--N--IE-G-------M------A-------I-----I--D-----T-----------T---K.....---  159
65_cpx.CN.10.YNFL01         ---------------K----N--------I-RR-K---------R--K---------.--.--.....--MK-----Q----E----H-------MRG-N--IE-D---------------A-----R-I--Q--L---N-----.-----------T---K.....---  158
67_01B.CN.11.2011.ANHUI.HF115-------------------H-I------I-K---K---K-----Q--KV---------.-E.-I....---R-----Q---K-----H-------Q-N-R--I--G-------H--------------I--QV-RR----PS--.-----------K--T-......--Q  158
68_01B.CN.11.2011.ANHUI.WH73-------------------H----------K---K---K-----Q--KV---------.-E.-M....--VR-----Q---K--Q--H--M---Q-N----I--D-------H--------------I--QV-RR----PS--.-----------K--T-......--Q  158
70_BF1.BR.10.10BR_PE004     -------------------N-------Y--HI-K--KN------F--R--KV---------.-K.-E......-I-----N--------------------QG--R--I--G-------I-------E----N-I--Q-------------.S----------TXXRX.......-X  158
71_BF1.BR.10.10BR_PE008     -------------------N-------Y--H--K--KR------F--R---VG--------.EE.-T....--V-----------------------------RG--R--I--GT------I-------E-------------R------------.---------------TP.........---  158
72_BF1.BR.10.10BR_MG002     -------------------N-------Y--HI-K--KR---------T-------------.--.-E.....--V---K-------------------------R--H-----N-------H------E----S-I--R--R----------.-----------------A.......-R  158
O.BE.87.ANT70               -------L-------QKVKA-N----Y-K-R-R-TEN-W-------RN--V--S-Y--V.-V.-H...V-V------MP----E---H------Y-K-K--I---T--RM---H--T--TA--V---I--QR-LTK---PT--.SQ--T--L---R-VVK.....ARS  158
O.CM.98.98CMA104            -------L-------QKVKV-N----Y-K-R-R--KD-Y-----X-RN--V--G-Y--V.-Q.-C...I-V------MP---EDQ--H------QY-K-R--I---T--RM------T--TE------T--QR-LTK-D-PE--.S---T--L---RVVVK.....T-R  158
O.CM.98.98CMABB141          ---K---L-------QKVKA-N----Y-K-R-R--KD-Y-------RN-----S-Y---.-P.-C...I-VS-----MP---EE-M-H------QY-K-T--I---T--RM------T--TE----ETI--QR-LTK---S---.SQ--T--L---R-VVK.....E-R  158
O.CM.98.98CMABB212          ---------------QKV---T------K-I-K--KN-H-K---FRS---G-G-Y--V.AE.-K...IIV------MP---EE---H------QY-N-T--I---T--RM---H-----T-------I--QR-LTK---P---.SQ--T--F---K-IEK.....V-R  158
O.CM.98.98CMU5337           -X-----L-------QKVKA-N----Y-K-RTR---D-S-K--F-XRN--V--S-Y--V.-M.--....VKV------MP----E-M-H------QY-X-T--I---T--RI------T--TE------I--QR-LTK---P---.SQ--T--L---RVVVK.....A-R  158
O.CM.99.99CMU4122           -------L-------QKVK--N----Y-KHR-R-TKE-H-K--F-FRN--A--G-Y--V.-M.-C...V-V------MP----E---H------QY-N-T--I---T--RI--QH--T--TE------II-QRILTK---S---.SQ--T--L---RVVVK.....E-R  158
O.FR.92.VAU                 -------L-------QKVKA-N----Y-K-R-R--TD-Y-K-----IN--V--R-Y--V.-P.-C...IMV------MP---EEQ--H------QCRK-I--I---T--K----H--T--TE----R-T--QR-L-Q---PE--.SQ--T--L---R-VVK.....A-R  158
O.GA.11.11Gab6352           ------AV-------QKVK--N----Y-K-R-R--KE-Y----F-HRN-----S-Y--V.-E.-W...V-V------MP----E---H------QY-K-K--I---T--RM--I---T--TE------I--QR-LVK--FP---.SQ--T--L---RV-EK.....A-R  158
O.SN.99.99SE_MP1299         -------L-------Q-VK--N----Y-K-R---T-D-Y----F--RN--V--G-Y--V.-G.PW...I-V------MP----E---H--------Y-K-K--I---T--RM--I---A--TE------I--QR-LT----P---.SQ--T--L---RVVVK.....E-R  158
O.US.10.LTNP                -------L-------QKVKA-N----Y-K-R-R-TKD-H------TRN-----S-Y--V.-M.-C...V-V------MP---EDQ--H------QY-K-K--I---T--RI--Q---T---E----R-I--QR-LTK---P---.SQ--T--L---R-VVK.....A-R  158

N.CM.02.DJO0131             --------V------R--K--H----------R--E--Y------TN--K---------V.-L.--...............T----S----H------KLGK-K-------M--K--------N--TA----Q-V--RP-F-----P----.-Q--T--H---T-WVG.....V--  158
N.CM.02.SJGddd              -A------V-------K-KK-N----Y-----K--Q--Y------TR--K---------V.-Q.--.....----V------T---QS----H-------LRK-K-------M------H---N--AA----Q-I--RP-V-K---P----.QQ--T-------R-WVG.....V--  158
N.CM.04.04CM_1015_04        --------V----------I-N-------H--K--K--Y-------R--KT--------V.-Q.-T.....----V------T---QP----H-------LRK-K-------M--K-----------T-----Q-V--RQ-F-----P----.-Q-DT------T-WVG.....I--  158
N.CM.06.U14296              --------V--------K----T-----------KR-K--Y------IN---K--------V.-Q.-K.....--VI-----T---QP----H-------LGK-K-------M--R--------N--TA----Q-A--RP-L-K---P----.KQ--T-------T-WVG.....VQ-  158
N.CM.06.U14842              --------V--------K----N-----------K--T--H------TX--KK--------V.-Q.-K.....---I------T---QP----H-------LGK-K-------M--K--------N--TA----Q-V--KP-L------P----.QQ--T-------T-WVG.....V--  158
N.FR.11.N1_FR_2011          -A------V-----------K-N-----------K--K--Y------TN--K-G----------.P.--.........---I------T---QP----H-------LRK-K-------M--K--------N--TA----Q-V--RQ-I-K---P----.KQ--T---I--T-WVG.....A--  158

P.CM.06.U14788              ------------IP-EKVKA-N----Y-K-R-K-T-K---------TT------A-Y--V.-T.-T...IIV------MP---EEQ--H-A-V---QGK-T--I---T--R----H--Q------V-R-I--DR-LNX---S---.SQ----------KVVVG.....KV-  158
P.FR.09.RBF168              ------------IS-EKVKA-N----Y-K-R-K-TKN--------TV------G-Y--V.-A.-T...I-V------MP---EEQ--H-A-V---QGK-T--I---T--R----H--Q------V-R-I--DR-LNK---S---.SQ-----F---KVVVG.....KV-  158

CPZ.CD.06.BF1167            -------RV---I------A-N------I-K-KHCKE-K-----DID---KGG-----ITEK.-K.........I------C--KA----H-A------QG--W-E----T---I--SR--P--TER---Q-I--EKILQH-HFPT--.S-------F--FKKVLE.QNK.Q-P  162
CPZ.CM.05.SIVcpzMT145       -------Q----------VK--N----Y---R-K--KN--------DGTN--MA-QIY--I.NQ.-K....-IV------MP----------H--A----ERG----------T--R----H-----TAT---N-I-EKA-L-K--FPE--.S--------W--KQ-VK.....T--  158
CPZ.GA.88.GAB1              -----------------K--N----Y-I-R-K------------DH-N-KVA--I---FRDY.SK....-IV----A-SP---A----H---Q--LGS-V-----FT--R---SQ-----AET---R-I--QL-A------KE--.RQ------F---K---S.....ERR  159
CPZ.TZ.06.SIVcpzTAN13       -------QV--MI----L---N------IFTTKYCKE-K------TDV-KRAG-I----E-K.--.....I-VLH----AC---P----H--G----QG------I---T------TK--A--TAG-V-Q-I--ERILTF-H--T--.RQ--T------FLKVVEAQYKTA--  164
CPZ.US.85.US_Marilyn        -------L----------VKA-N----Y---R-K--K-----------TN--VN------I.DK.-K...I--I-----K---QA----H-------QGT-K--I---M--K--------------TA--------V--RP-I-K---PT--.---------F---K--VG.....QS-  158

GOR.CM.07.SIVgor2139_287    ------------IP-E-VKA-N----Y-K-RTKQ-KQ--------AVN--V--G-Y--V.-P.-E...IIV------MP--KEEQ--H--G----QGK-I--L---T--RI-------Q--TE--V---IV-EKLLYK-N-P---.SQ--T-------R--VG.....KVR  158
GOR.CM.07.SIVgorCP2135con   ------------IP-E-VKA-N----Y-K-RTKQ-KQ--------AVN--V--G-Y--V.-P.-E...IIV------MP--KEEQ--H--G----QGK-I--L---T--RI-------Q--TE--V---IV-EKLLCK-N-P---.SQ--T-------R--VG.....KVR  158
```

Vif end

B.FR.83.HXB2	IKPPLPSVTKLTEDRWNKPQKTKGHRGSHTMNGH...*	192
A1.AU.03.PS1044_Day0	--------R--A-----------P------------	193
A1.CH.03.HIV_CH_BID_V3538	T------IR----------------------C...	193
A1.CY.08.CY236	--------K-------------D------------	193
A1.ES.05.X1608_8	--------R--------------L---R--------	193
A1.KE.11.DEMA11KE001	--------RI----------R---P-E--I---C...	193
A1.RU.11.11RU6950	ER------R-------------R-R--N-----C...	193
A1.RW.11.DEMA111RW002	TN------R--A-------R----P--N-I---C...	193
A1.SN.01.DDI579	TR------R-------------R----------C...	193
A1.UG.11.DEMA110UG001	T-------R-------------R-PKE----------	193
A1.ZA.04.04ZASK162B1	P-------K----------------------C...	193
A2.CM.01.01CM_1445MV	T-------R--V----------R----N-----C...	193
A2.CY.94.94CY017_41	T-------R--V----------------------C...	193
B.BR.10.10BR_MG029	R-------A----------R--T-------------	193
B.CA.07.502_1191_03	T-------R--------------------I------	193
B.CH.08.M2_0803101_NFLG8	--------K----------------....V------	189
B.CN.10.DEMB10CN002	R-------R-----------------N--TS-----	193
B.ES.10.DEMB10ES002	--------R---------H---R-------------	193
B.FR.11.DEMB11FR001	--------K--------V--------------Q...	193
B.GB.05.MM45d213_GN1	-------------------R----------------	193
B.HK.06.HK003	----------------------T--------H-----	193
B.HT.05.05HT_129389	--------Q------------I----E-P--------	193
B.JP.12.DEMB12JP001	R---------------------------------...	194
B.KR.07.HP_18_07JHS10_3909	--------R----------------E----------	193
B.PE.07.502_0525_wg5	--------A----------R----------------	197
B.RU.11.11RU21n	----------------R--D-----IS-----	193
B.TH.08.MERLBDTRC10	--------R----------R--------Q-T------	193
B.US.11.ES38	V----------------------E----------	193
C.AR.01.ARG4006	R-------R--V----------RDR--N-I-------	193
C.BR.07.DEMC07BR003	R-------K--V------*--RD---N---------	192
C.BW.00.00BW5031_1	P-------R--A-------R--R-R--N---------	193
C.CN.10.YNFL19	-------IQR-V-----N----M-R--N---------	194
C.CY.09.CY260	T-------R--V----------N---------	193
C.ES.08.X2363_2	R-------R--V------S--ARDC--N-S-------	193
C.ET.02.02ET_288	R-------S--V----------R-R--N--V------	193
C.IN.09.T125_2139	-------IK--V---------I--R--N-IL------	193
C.KE.00.KER2010	--------Q--V-----N----R-R--N---------	193
C.MW.09.703010256_CH256.w96	--------S--V--K-------R-R--N---------	193
C.TZ.08.707010457_CH457.w8	--------R--V-----N----R-R--N-I-------	193
C.YE.02.02YE511	--------R--V----------M-R--N---------	193
C.ZA.08.705010534_CH534.w12	T-------S--V----------R-R--N---------	193
C.ZA.10.DEMC10ZA001	R-------R----------EI-----N---------	193
C.ZM.11.DEMC11ZM006	R-------R--V----------N-----L...	193
D.CM.10.DEMD10CM009	R-------R----------Q---------T------	193
D.CY.06.CY163	R-------R----------Q----K-----S------	193
D.KE.11.DEMD11KE003	--------K----------R-----E----------	193
D.KR.04.04KBH8	--------R----------E---R--P-----WSF-	196
D.SN.90.SE365	--------R----------------E----------	193
D.TZ.01.A280	--------K----------R------P----------	193
D.UG.10.DEMD10UG004	T-------R----------R----------------	193
D.UG.11.DEMD11UG003	--------K----------R-R--------S------	193
D.YE.02.02YE516	R-------K----------Q----------------	193
D.ZA.90.R1	T-------R----------R----------------	194
F1.AO.06.AO_06_ANG32	V-------R--V--------E--------------...	193
F1.AR.02.ARE933	T-------Q--V----------R---------------	193
F1.BR.10.10BR_RJ015	T-------Q--V--------------#---D--WNFW	196
F1.CY.08.CY222	T------IK--V--------E-R----N---------	193
F1.ES.02.ES_X845_4	T-------R--V--------E-R---E----------	193
F1.RO.96.BCI_R07	T-------K--V-------REIR-------------	192
F1.RU.08.D88_845	T-------K--V--------E----------------	193
F2.CM.02.02CM_0016BBY	V-------K--V--K--N-------------------	194
F2.CM.10.DEMF210CM001	TR------S--V-----N--R----------------	194
F2.CM.10.DEMF210CM007	--------R--V-----N----R-------------	194
G.BE.96.DRCBL	RR------R--------E----R---ENP--------	193
G.CM.10.DEMG10CM008	R-------Q---------S--IRD-KE-P--S-----	193
G.CN.08.GX_2084_08	--------K-----------------NPS-S-----	193
G.CU.99.Cu74	--------K----------R-R---ENP--------	193
G.ES.09.X2634_2	A-------R----------R-R----NP--T-----	193
G.GH.03.03GH175G	K-------R----N----------ENP-----WNC-	196
G.KE.09.DEMG09KE001	--------R--------S--IR-R-ENL--T-----	193
G.NG.09.09NG_SC62	-R------R------*---------ENP--S-Y----	192
G.ZA.01.TV546	RR------K-------E----R---XNP----WSC-	196
H.BE.93.VI991	T-------R--V----------R---------------	193
H.CF.90.056	--------R--V----------R---------------	193
H.GB.00.00GBAC4001	T-------R----------R---------------	193
J.CM.04.04CMU11421	R-------Q--V-----N----R---------------	193
J.SE.93.SE9280_7887	R-------Q--V----------T---E----------	193
J.SE.94.SE9173_7022	R-------Q--V----------RD--E----------	193
K.CD.97.97ZR_EQTB11	T---V---Q--V----------R--------S-Q...	193
K.CM.96.96CM_MP535	P---V---K--V----------R-----Q-------...	193
U.CA.01.TV749	T-------Q--A-G---------A---N---------	193
U.CA.99.TV721	T-------Q--A----------RD-------------	193
U.CD.83.83CD003_Z3	R-------R--V------------K-----H-----	193
U.CD.90.90CD121E12	R-------R--V---------I---K----------	193
U.CY.05.CY090	R-------Q--V-----R-R--R----N---------	193
U.CY.08.CY223	T-------R-----------I----ENQ--------	193
U.ES.10.DEURF10DZ001	R-------Q--V----------RD-----S-------	193
U.GR.99.99GR303	T-------K--V------------E-Q---------	193
U.NL.95.U_NL_95_H10986_D1	--------R--V----------R---E---I------	193
01_AE.AF.07.569M	T-------K------------.G--GENP--------	192
01_AE.CF.90.90CF11697	TR------R------------R---ENP--------	193
01_AE.CN.10.YNFL03	-R------K----------IR---ENP--------	193
01_AE.HK.04.HK001	SR------K-------SE---R---ENP--------	193
01_AE.IR.10.10IR.THR48F	--------K----------I----ENL--------	193
01_AE.JP.x.JRC77AE	TR------KI----------IR---ENP--------	193
01_AE.TH.09.AA090a_WG11	-R------K----------IR---ENP----C...	193
01_AE.TH.90.CM240	-R-------.EI----------.R---ENP--------	191
01_AE.US.05.306163_FL	-R------K----------IR---ENS--------	193
01_AE.VN.98.98VNND15	-R------K----------IR---ENP--------	193

Vif end

```
B.FR.83.HXB2                IKPPLPSVTKLTEDRWNKPQKTKGHRGSHTMNGH...*  192
02_AG.CM.10.DE00210CM013    T-------R-------KE----------S-------  193
02_AG.CY.09.CY256           TR------K--A---------R-----RS-------  193
02_AG.ES.06.P1423           TR------K-----------R------S-------  193
02_AG.GW.05.CC_0048         K-------K-------E----------RS-------  193
02_AG.KR.12.12MHR9          R-------K--A-----E----RD-----S-------  193
02_AG.LR.x.POC44951         T-------K--A-N------IR-----RS-------  193
02_AG.NG.09.09NG_SC61       R-------K---------S----R-----S-------  193
02_AG.NG.x.IBNG             T-------R--A----KE----R-----RP-------  193
02_AG.SE.94.SE7812          T-------K--A-----E----R-----RS-------  193
02_AG.SN.98.98SE_MP1211     R-------K--A-----E----R----NRS-------  193
03_AB.RU.97.KAL153_2        --------------------R--D-------S-----  193
04_cpx.CY.94.94CY032_3      T-------K--V---------R-R-ENQI--------  192
05_DF.BE.x.VI1310           T-------R--------------R--N-----Y...-  193
06_cpx.AU.96.BFP90          R-------Q--V---------R---E----------  193
07_BC.CN.98.98CN009         -------IK--V-----N---IR-R--N--------  194
08_BC.CN.97.97CNGX_6F       -------IK--V-----N---IR-R--N--------  194
09_cpx.GH.96.96GH2911       T-------K--V--------R-R---ENQ-------  193
10_CD.TZ.96.96TZ_BF061      T-------R----------R----------------  193
11_cpx.CM.95.95CM_1816      V-------K-----------R---EN-----CWNC-  196
12_BF.AR.99.ARMA159         T-------K--V------N----R-----------  193
13_cpx.CM.96.96CM_1849      --------K--V----------RD--ENQ-------  193
14_BG.ES.05.X1870           T-------K----------R-R----NP-------  193
15_01B.TH.99.99TH_MU2079    -R----R-K-------D----IR---ENP--D-----  193
16_A2D.KR.97.97KR004        T-------R----------------------------  193
17_BF.AR.99.ARMA038         T-------K--V--------E-R-------------  193
18_cpx.CU.99.CU76           T-------R--A----------R----NQ-------  193
19_cpx.CU.99.CU7            SR------R------------R---ENP-------  193
20_BG.CU.99.Cu103           --------R----------R-R---ENP-------  193
21_A2D.KE.99.KER2003        T-------R-----------I----E--M-------  196
22_01A1.CM.01.01CM_0001BBY  --------R------------R---EYP-----WKC-  193
23_BG.CU.03.CB118           --------R----------R-R---DN--------  193
24_BG.ES.08.X2456_2         --------R----------R-R---EN--------  193
25_cpx.CM.02.1918LE         -------R--------E--------ENQ-------  193
26_AU.CD.02.02CD_MBTB047    T------K---------------E----------  193
27_cpx.FR.04.04CD_FR_KZS    --S----R-------------R---ENQ----Q...  193
28_BF.BR.99.BREPM12609      T------R-----------I----N----------  193
29_BF.BR.01.BREPM16704      --------R---------R----------------  194
31_BC.BR.04.04BR142         --------R--V-----R----RDR--NQ-------  193
32_06A1.EE.01.EE0369        R-------Q--V-----E----R---E------R...  193
33_01B.ID.07.JKT189_C       -R------K-----------IR---ENP--S-----  193
34_01B.TH.99.OUR2478P       -R------K----------QIR---ENP-T------  193
35_AD.AF.07.169H            TR-----IR--A---------R-P-E------C...  193
36_cpx.CM.00.00CMNYU830     R-------K------------R-----RST------  193
37_cpx.CM.00.00CMNYU926     KR------K------------R---ENP-------  193
38_BF1.UY.03.UY03_3389      T-------K--V---------R-------------  193
39_BF.BR.04.04BRRJ179       -------LK------------T-------------  193
40_BF.BR.05.05BRRJ055       -------LK----NK-----R--------------  193
42_BF.LU.06.luBF_18_06      -------K-----------------------------  193
43_02G.SA.03.J11223         K-------A-------S---R----RX-------  193
44_BF.CL.00.CH80            T-------Q--V--------E-----E--------  193
45_cpx.FR.04.04FR_AUK       T-------K--A---------RD--------C...  193
46_BF.BR.07.07BR_FPS625     T-------K--V--------E-R-------------  193
47_BF.ES.08.P1942           A-------Q--VX------------E----------  194
48_01B.MY.07.07MYKT021      --------K------------IR---ENP--S-----  193
49_cpx.GM.03.N26677         -------TR------------R-R--N--------  193
50_A1D.GB.10.12792          --------R--A-----E----R-P-----I------  193
51_01B.SG.11.11SG_HM021     -R------K------------I----ENPI-H-----  193
52_01B.MY.03.03MYKL018_1    -R------K-----------I-D--EH---------  193
53_01B.MY.11.11FIR164       -R------K-----------I--R-ENPI-------  192
54_01B.MY.09.09MYSB023      --------K-----------IR---ENL--S-----  193
55_01B.CN.10.HNCS102056     -R------K-----------M---EN---------  193
56_cpx.FR.10.URF5_patient_A K-------A-----I------R-----RP-------  193
57_BC_CN.09.09YNLX19sg      ------IKQ-V-----N----R-R--N--------  194
58_01B.MY.09.09MYPR37       -R------K-----------IR---ENP--S-----  193
59_01B.CN.09.09LNA423       -R------K-----------IM---E-P-------  193
60_BC.IT.11.BAV499          R-------R--V---------RDR--N---H-----  193
61_BC.CN.10.JL100010        -R------K--V---------IR-R--N---------  194
62_BC.CN.10.YNFL13          --------K--V-----N----R-R--N--X------  192
63_02A1.RU.10.10RU6637      ER------K-------------------RS-------  193
64_BC.CN.09.YNFL31          R------IK--V-----D---IM-R--N---------  194
65_cpx.CN.10.YNFL01         -------IK--V-----N---I--R-..D--D-----  191
67_01B.CN.11.2011.ANHUI.HF115R------K------------I----ENP-------  193
68_01B.CN.11.2011.ANHUI.WH73-R------K------------I----ENP-------  193
70_BF1.BR.10.10BR_PE004     -------XK-----------I---K----------  193
71_BF1.BR.10.10BR_PE008     ---------Q--------------KE----------  193
72_BF1.BR.10.10BR_MG002     R---------------E-------K-N---------  193
O.BE.87.ANT70               R-------Q--------HLRIRDQLK-PS-------  193
O.CM.98.98CMA104            N-------Q--------RHL-IRDQLE--S-------  193
O.CM.98.98CMABB141          N-------Q--------YLRIRDQPE--S-------  193
O.CM.98.98CMABB212          N-------Q--------H---RDQL---S-S-----  193
O.CM.98.98CMU5337           N-------Q-----Q-KGHLRITDQLE--S---Y...  193
O.CM.99.99CMU4122           N-------Q--------HLRIMDQL-N-S-------  193
O.FR.92.VAU                 NR------Q--------RHLRI-DQP--QL-H-----  193
O.GA.11.11Gab6352           N-------Q-----K-S-HLMIRD-LE--S-------  193
O.SN.99.99SE_MP1299         N-------Q-------SRHLRIRDQLE--S-------  193
O.US.10.LTNP                S-------Q--------RHLR-RDQPE--S-------  193
N.CM.02.DJO0131             R----------------H---Q----NPI-------  193
N.CM.02.SJGddd              R----------------RH---Q----NPI-------  193
N.CM.04.04CM_1015_04        R-------A--------RH---Q----NPI-------  193
N.CM.06.U14296              R-------A-------SER---Q-R--NPI-------  193
N.CM.06.U14842              R----------------EH---Q----N-I-------  193
N.FR.11.N1_FR_2011          R----------------EH---Q----NP--------  193
P.CM.06.U14788              R-------QI--Q-I-SN--R---XQE--S-S-CYKP-  196
P.FR.09.RBF168              R-------QI--Q-I-SX--R-X-RQE--S-S-CYKP-  196
CPZ.CD.06.BF1167            HR-------V-A--------RI-A-QENRGI--Y...  197
CPZ.CM.05.SIVcpzMT145       NR------A--A--------R-RD-QE-PMSS-C...  193
CPZ.GA.88.GAB1              HR------A---------H-R--V-QENL-R-------  194
CPZ.TZ.06.SIVcpzTAN13       PGR---A-V---------H--IR----N-II--Y...  199
CPZ.US.85.US_Marilyn        RR----------------C-----QPEN---S-----  193
GOR.CM.07.SIVgor2139_287    VR--F---Q--AA-T--RS-QMR--KEN-S-------  193
GOR.CM.07.SIVgorCP2135con   VR--F---Q--AA-T-SRS-QMR--KEN-ST------  193
```


Draft of 30 September 2015, not for public release
hiv1prot_main.tex r8876–10546

amphipathic α-helix
oligomerization
H(S/N)RIG motifs
Vpr end in HXB2
Vpr start
frameshift in HXB2
Vpr end

```
B.FR.83.HXB2          MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIW.LHGLGQHIYETYGDTWAGVEAIIRILQQLLFIH#FRIGCRHSRIGVTR....QRRA..RNGASRS..*   96
A1.AU.03.PS1044_Day0  --------------Y------------Q--------P-.-------Y--N------E-------T------V-.-----Q-----IX-....G--G..---SG-X..-   97
A1.CH.03.HIV_CH_BID_V3538 --------------Y---M--------H--------Q-.-Q--------D------E-----V--------V-.-----H-----LM-G...G--G..-D--G----   98
A1.CY.08.CY236        --------------Y------------H--------P-.-------Y--N------E--Q---M------V-.-----Q-----I-....G--G..-D-SG-P..-   97
A1.ES.05.X1608_8      ----------------------------H--------P-.-------Y--N------E--V---K------V-.-----Q-----IIP....G--G..---S-----   97
A1.KE.11.DEMA11KE001  --------------Y---M-D---D-----------P-.---------N------E-----V-------V-.-----Q-----II-G...G--G..---S-----   98
A1.RU.11.11RU6950     --------------Y---M-----DI-H--------P-.-------Q--N------E-------T------V-.-----Q-----I-Q....R--V..---SR----   97
A1.RW.11.DEMA111RW002 --------------Y---M-----D--H--------P-.---------N------D-H---E-----------V-.-----Q-----II-....G--V..-D-PG----   97
A1.SN.01.DDI579       --------------Y---M--------Q--------P-.---------N------E------M------V-.-----Q-----II-R...G--V..-D----P..-   98
A1.UG.11.DEMA110UG001 --------------Y---------D----------P-.---------N------E----------I-V-.-----Q-----II-....G--G..---S-----   97
A1.ZA.04.04ZASK162B1  --------------Y---M--------H-----Y--P-.--A------N------T---V--------V-.-----Q-----IIP....G--G..-D--G----   97
A2.CM.01.01CM_1445MV  --------------Y---M--------Q--------P-.--D------DI-----E-------Y------V-.-----Q-----II-....R--V..-D-T--P..-   97
A2.CY.94.94CY017_41   --------------Y---M--------Q--------H-.-------Y--N------E---V--Y------V-.-----Q-----II-....R--V..-D----P..-   97

B.BR.10.10BR_MG029    -------H------Y-----------S--------A-.--S------T----------LL-----------.-----Q-----I-----------.----T-----   97
B.CA.07.502_1191_03   --------------Y-Q---------A--------P-.-------Y----------------T---------.-----Q-----ILQ.........R----N----   98
B.CH.08.M2_0803101_NFLG8 ---------....-S----------S--------P-.-D---Y--------R------T--RM---------.-----Q-----I-----------T----------   93
B.CN.10.DEMB10CN002   ------------------------------S--------P-.-S---Y--------T----------M----.Y---Q-----II---------..#D-------   96
B.ES.10.DEMB10ES002   --------------Y---------------------P-.-------Y--------T--L----M------V-.----------I---------T----------   96
B.FR.11.DEMB11FR001   --R----------Y-----I------R--------T--.-S---Y--Q----------------------------H-----IN-----.-T-T----------   97
B.GB.05.MM45d213_GN1  --------------Y---------------------P-.-------Y--------T----L-K----------.-----Q-----IQ....R---.----------   97
B.HK.06.HK003         -------H--P---Y-A-----------------T-.-------Y--------E----------V-.Y----Q-----I-----------T----------   97
B.HT.05.05HT_129389   --------------Y---A--------S--------V-.-----------------------------.----Q-----I-PTR.-----.---TG----   99
B.JP.12.DEMB12JP001   --------------Y---------------------P-.------L---N----------------.Y-L--Q-----I-----------.-----G----   97
B.KR.07.HP_18_07JHS10_3909 --------------Y---------D--S--------P-.--S-----------------A--------.-----Q-----II----------.----------   97
B.PE.07.502_0525_wg5  --------------Y---------------------V-.-------Y----------------H------.-----Q-----I-----------.----------   97
B.RU.11.11RU21n       --------------Y-Q---------A--------P-.-------Y----------L-T--------.-----Q-----II----.RS---.-H----------   97
B.TH.08.MERLBDTRC10   --------------N----------S--------P-.--S---Y--R------T----L--M--------.-----Q-----I-P....T---.----------   97
B.US.11.ES38          --------------Y-----------------------S---Y--------E----------------.-----Q-----II-----------.----S-----   97

C.AR.01.ARG4006       --------------Y----------Q--------P-.--S---Y--------M----------V-.----------II----------.----------.----   97
C.BR.07.DEMC07BR003   ----L---------Y----------Q--------P-.--S---Y--------T----------V-.-----Q-----IL----------.---PN----.----   97
C.BW.00.00BW5031_1    --------------Y----------Q--------P-.-S---Y--------M----L----------.-----Q-----IIL---------T----------.----   92
C.CN.10.YNFL19        ---S---H-----Y----------Q--------P-.-------Y----------LL-T----------.-----Q-----IV--------G..-T--N----.----   97
C.CY.09.CY260         --------------Y---A--------Q--------V-.--S-----------T----L----------.-----Q-----IL----------.-----N----.----   97
C.ES.08.X2363_2       ---V--G--L---F----------Q--------P-.-------Y--------T---V----------N--Q-----II--------T------G----.----   97
C.ET.02.02ET_288      --------------YS--------D--Q--------P-.--S---Y--N------R----L-TF---S----.----H-----I-----------.----------   97
C.IN.09.T125_2139     --------------YT--A--------Q--------P-.-------Y--------T----L---------T-.-----Q-----IR----------.----------   97
C.KE.00.KER2010       ---P---------Y------I-----Q--------P-.--SI--Y--------T----------------.----H-----IL----------.----------   97
C.MW.09.703010256_CH256.w96 --------------Y----------Q--------P-.-S-----------T----------V-.-----Q-----II----------.-----N----.----   97
C.TZ.08.707010457_CH457.w8 ---P--------------A--I-----Q--------P-.-A---Q--N------T----------.-----Q-----IV----------.----------   97
C.YE.02.02YE511       -------H-----Y----------Q--------P-.-S---Y--------T-----L--------V-.-----Q-----II----.H---.-----N----.----   97
C.ZA.08.705010534_CH534.w12 ----------------------Q--------P-.--S---Y---N----T---VL----------.-----Q-----IL----------.-----N----.----   97
C.ZA.10.DEMC10ZA001   -----G--------Y----------Q--------P-.-S---Y----------------------------.-----Q-----II----------T----------   97
C.ZM.11.DEMC11ZM006   --------------Y---------I-Q--------P-.--S---Y--------T----------------.----H-----IL----------T---S-----   97

D.CM.10.DEMD10CM009   -----A--------Y---A--------SK-------.----------D------T---L----------.-----Q-----IIP.----------.----S-----   97
D.CY.06.CY163         -----A--------Y----------S--------T-.-----------------------L--T---I----.-----Q-----IIP....P---.----S-----   97
D.KE.11.DEMD11KE003   --------------Y----------------------V--------Y----------------------.-----Q-----IL----------I---SN----.----   97
D.KR.04.04KBH8        -----G--------Y---A--------S--------P-.-D------A------V----LL-V------V-.-----Q-----II----------.----S-----   97
D.SN.00.SE365         --------------Y----------S--------T-.-S---Y--------------------------.-----Q-----I-P---------I---S-----   97
D.TZ.01.A280          --------------Y----------S--------L-.--N---F--------------------------.-----Q-----I-L---------T---S--P..-   97
D.UG.10.DEMD10UG004   --------------Y----------S--------P-.-----Y-------------V----------.-----Q-----IM----------L----S-----   97
D.UG.11.DEMD11UG003   --------------Y---A--------R----------.---S-----------T----LT-S-----V-.-----Q-----INP---------T-----R----.   97
D.YE.02.02YE516       -----A--------Y----------S--------A-.-------------------------T-----L--.-----Q-----ISP----------.----S--P..-   97
D.ZA.90.R1            --------------Y----------S------IT---S--------------------------------.-----Q-----II----------.----S-----   97

F1.AO.06.AO_06_ANG32  -----G--------Y----------------------P-.--S---Y--Q------L----L--M---M-----.--V--P-----I-P.------V.----N-----.-   97
F1.AR.02.ARE933       --------------Y---A-D--------------P-.--S-----N----------LL----------.-----H-----I-----------.----------   97
F1.BR.10.10BR_RJ015   -----------#---G----------H--------A-.-------Y--N.....------------------H---M-I---------V..G---G----.-   73
F1.CY.08.CY222        -----G--------Y---A----------------P-.----------N------E----------------.-----Q-----M---------V----------   97
F1.ES.02.ES_X845_4    -----G--------Y----------------------P-.-------Y--N------V----LV----------.--M--Q-----I---------V----S-----   97
F1.RO.96.BCI_R07      -----G--------Y----------------K----P-.----------N------E----------------.-----Q-----I-------A-V----------   97
F1.RU.08.D88_845      -----G--------Y----------------------P-.--S-----D------E----------------.-----Q-----I---------V----------   97
F2.CM.02.02CM_0016BBY ---P----------Y----------H--------E-.----------N------E------T----------.-----H-----II--------I----NG----.-   97
F2.CM.10.DEMF210CM001 ---P----------Y----------Q--------E-.-------S------E------T----------.-----IIQ.------I----SG-P..-   97
F2.CM.10.DEMF210CM007 ---P----------Y----------Q--------E----M--Y--N------E------M----------.-----H-----II--------I----N-----.-   97

G.BE.96.DRCBL         --R----------A--------H--------P-.-------L--N------E----------------.-----Q-----IIPQ...R--V..-D-SG----.-   98
G.CM.10.DEMG10CM008   ---V--N------Y---A--------H--------P-.-------Y--N------E----------V-.-----Q-----IIP....R--V..-D-PG----.-   97
G.CN.08.GX_2084_08    --------------F----------T--------P-.----------N------E----------------.-----Q-----INLR...R--V..-D-SG----.-   98
G.CU.99.Cu74          --------------------------------------P-.-------YV-N------E----------------.-----ISL....R--V..-D-PG----.-   97
G.ES.09.X2634_2       --------------Y-D-A--V--D--R--------P-.-------Y--N------E----L----------.-----Q-----IL----R--V..M--P-----.-   97
G.GH.03.03GH175G      -------N------Y---A----------------P-.-------Y--N------E------M-------.Y----Q-----IIP.------V..-D-P--P..-   97
G.KE.09.DEMG09KE001   ---V---------Y-D----------------P-.----------N----------------------.-----Q-----IIP.------V..-D-PG----.-   97
G.NG.09.09NG_SC62     -K----------Y---I-G------V--------P-.----------N------E-------TK-----.-----Q-----I-P.------V..-D-SG----.-   97
G.ZA.01.TV546         --R----------Y---A-----D----------P-.------X--X------E---X----------.-----Q-----I-P.------V.----P-----.-   97

H.BE.93.VI991         ------------------------S----Y--P-.--D------N------E------T------A-.-----Q-----I---------V.----P-----.-   97
H.CF.90.056           ----------------------I---------V-.--Q------N------V----L--T---------.-----Q-----I---------V.----P-----.-   97
H.GB.00.00GBAC4001    ------------Y----------S--------A-.--Q---Y--N------R----LT--------T-.----Q-----I---------V.----P-----.-   97

J.CM.04.04CMU11421    ---P----------X----------------------X-.----------X------E--RX----------.-----Q-----IMP.------G.----------.-   97
J.SE.93.SE9280_7887   -------H------YH----------------------P-.--S---Y--S------E----------------.-----H-----IIP.------G.----------.-   97
J.SE.94.SE9173_7022   --------------YH----------------------P-.-------YV-S------E----V----------.-----H-----IIP.------G.----------.-   97

K.CD.97.97ZR_EQTB11   --------------Y----I-I-----R--------P-.--D------NN-----E----------------.-----H-----IL----------G..G---------   97
K.CM.96.96CM_MP535    --------------N------I-----R--------P-.--N------T------E-L----------.-----H-----IIP.------G.----S-----.-   97

U.CA.01.TV749         --------------Y----------S--------P-.--D------N------E----------------.-----H-----II--------G.----SN----.-   97
U.CA.99.TV721         --------------Y----------S----------.--D------N------E----------------.-----H-----II--------G.----S-----.-   97
U.CD.83.83CD003_Z3    --------------Y-A--------H--------P-.--N---Y--N------Q--------S----------.-----Q-----I-----------..-H-P-----.-   97
U.CD.90.90CD121E12    --------------Y----------Q--------P-.--N------N------Q----L----------.-----Q-----II----------...-H-P--P..-   97
U.CY.05.CY090         ---T---------Y------I----R--------P-.--N------N------E----L--M--------.-----Q-----II--------R...D-------.-   97
U.CY.08.CY223         --------------N----------H--------P-.----------N------E--R----------.-----Q-----IVP.------G.----------.-   97
U.ES.10.DEURF10DZ001  --------------F----------Q--------P-.----------N------E--R----------.-----Q-----IIP....G--G.----G----.-   97
U.GR.99.99GR303       --------------N----------H--------PC.----------N------E-----L--------.-----Q-----IV--------G.----S-----.-   97
U.NL.95.U_NL_95_H10986_D1 ----------P---Y-Q-A--I-----R--------P-.---------N------E----------T-.-----Q-----ILQ.....--G.----N----.-   97

01_AE.AF.07.569M      -------.R--R---N----------I--------P-.-------Y--N------E-------M------V-.-----Q-----IVP....G--G.-----G----.-   96
01_AE.CF.90.90CF11697 --------------Y----------I--------P-.-------Y--N------E------T------V-.-----Q-----II----.G--G.----G----.-   97
01_AE.CN.10.YNFL03    ------N-------Y----------------------P-.-------YV-N------E----------.-----Q-----IIP....G--G.----G----.-   97
01_AE.HK.04.HK001     --R----------Y----------X----XP---S---Y--D------E------T-T------V-.-----Q-----IIP....G--G.----------.-   97
01_AE.IR.10.10IR.THR48F ----------------------D----------H-.-------Y--NQ-----E----------V-.-----Q-----IIP....G--G.----------.-   97
01_AE.JP.x.JRC77AE    --------------------------------------P-.-------Y--N------E----------V-.-----Q-----ILP....G--G.-----G----.-   97
01_AE.TH.09.AA090a_WG11 ----------------M--I-----I--------P-.--S---Y--D-----LE--Q----------V-.Y----Q-----IIP....G--G.-----G----.-   97
01_AE.TH.90.CM240     -------.E------Y----------------------P-.-Q----Y--NN-----E-------M------V-.-----Q-----IMP....G--G.----TG----.-   96
01_AE.US.05.306163_FL ----------LN----------------------P-.-------Y--T------E----------V-.-----Q-----ILP....G--G.----G----.-   97
01_AE.VN.98.98VNND15  ---------S----Y---------D----------P-.-------Y--NN-----E-------T------V-.-----Q-----IIP....G--G..G-E-G----.-   97
```

```
                                        amphipathic α-helix
                                     oligomerization
                               Vpr start                                            H(S/N)RIG motifs
                                                                              Vpr end in HXB2
                                                                   frameshift in HXB2                  Vpr end
B.FR.83.HXB2                   MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIW.LHGLGQHIYETYGDTWAGVEAIIRILQQLLFIH#FRIGCRHSRIGVTR....QRRA..RNGASRS..*   96
02_AG.CM.10.DE00210CM013       --R-----------F------------H--------P-.----E-----------Q--Q-----------V-.-----Q-----IIP....G--G..---------..-   97
02_AG.CY.09.CY256              --------------F---A-----------------P-.--E------N------E--R-L-----S---V-.-----Q-----II-....G--G..----G----..-   97
02_AG.ES.06.P1423              --------------F----------I-Q--------P-.---------N------E--V--------------.-----Q-----IIP....RXGT..GK.-G----..-   96
02_AG.GW.05.CC_0048            --R-----------F------------H--------P-.----------------E--G----L------V-.-----Q-----II-....G--G..---------..-   97
02_AG.KR.12.12MHR9             --R-----------F-------V-------------P-.---------N------E--G---T--H----V-.-----Q-----II-....G--G..----G----..-   97
02_AG.LR.x.POC44951            --------------F------------Q--------P-.-----EY--N------E-------T------V-.-----Q-----II-....G--G..---S-----..-   97
02_AG.NG.09.09NG_SC61          ---V----------F---A--------H--------P-.---------N------E--G-----------V-.-----Q-----II-....G--G..----G----..-   97
02_AG.NG.x.IBNG                --R-----------S------------H--------P-.----------------E--K-----------V-.-----Q-----IIQ....G--G..----G----..-   97
02_AG.SE.94.SE7812             --R-----------F------------H---T----P-.----------------E--G-----------V-.-----Q-----II-....G--G..----G----..-   97
02_AG.SN.98.98SE_MP1211        --R-----------F------------H--------P-.----------------E--G-----------V-.-----Q-----IIP....G--G..----G----..-   97
03_AB.RU.97.KAL153_2           --------------Y------------S--------V-.--S---Y--------------------------.-----Q-----IIQ....R---..---------..-   97
04_cpx.CY.94.94CY032_3         --------------N---------------------P-.---------N------E----------------.-----Q-----I-P....---R#AGD--------..-   98
05_DF.BE.x.VI1310              --------------Y---I--------H--------P-.--S---Y--N------E-----V----------.-----------IL-....---I..---S-----..-   97
06_cpx.AU.96.BFP90             --------------Y------------S--------P-.--NI--Y--S------E----LL-T--L---V-.-----H-----IIP....---G..---------..-   97
07_BC.CN.98.98CN009            ---S----------Y------------Q--------P-.--S---------------------T--------.-----Q-----IL-....---T..---------..-   97
08_BC.CN.97.97CNGX_6F          ---P-----------------------Q--------P-.------YV--------T---T-L----------.-----Q-----IV-....------.--------..-   97
09_cpx.GH.96.96GH2911          --------------N---------------------P-.---------D------E--------------V-.-----H-----II-G...R--G..-D--G----..-   98
10_CD.TZ.96.96TZ_BF061         --------------Y-----------------------.--N-------------T----------------.-----P-----IV-....----..---T-T---..-   97
11_cpx.CM.95.95CM_1816         --------------Y---M--------H--------P-.---------N------V--------------V-.-----Q-----IIP....G--G..---------..-   97
12_BF.AR.99.ARMA159            ---P----------Y---------------------P-.---------N------E----------------.-----Q-----IM-....---V..---------..-   97
13_cpx.CM.96.96CM_1849         --------------N------------H---K----P-.---------N------E-----T-M--------.-----Q------V-....---G..---S-----..-   97
14_BG.ES.05.X1870              --------------Y---------------------P-.------Y--N------E----XL-X--------.-----Q-----I-P....---V..-D-S-----..-   97
15_01B.TH.99.99TH_MU2079       -G------------Y-G-P------------K.-S-P-.------S-TN------E-----V---H----V-.-----H----AIIP....G---.R---PG----..-   97
16_A2D.KR.97.97KR004           --------------Y---------D--R--------P*.--S---Y--A------T-----------M----.-----Q-----IM.....P-G-..---S-----..-   95
17_BF.AR.99.ARMA038            -----G--------Y---------------------P-.--S------N------E----------------.-----H-----II-....---V..-D-------..-   97
18_cpx.CU.99.CU76              --------------N---------------------P-.------Y--N------E----------------.-----Q-----IV-....---G..---------..-   97
19_cpx.CU.99.CU7               ------N-------Y---A-----D-----------P-.------YV-T------E----------------.-----Q-----I-P....R--V..-D-PG----..-   97
20_BG.CU.99.Cu103              ------------------------------------P-.-----V---N------E----L-----------.-----Q-----ISL....R--V..---P-----..-   97
21_A2D.KE.99.KER2003           ---------------D-----------S--------P-.--S------A-----------M-----------.-----H------I-....-------S-------..-   97
22_01A1.CM.01.01CM_0001BBY     ------------V-Y------V-------------GP-.-----E---N------E--------------V-.-----Q-----II-....G--G..----G----..-   97
23_BG.CU.03.CB118              ------------Q-----------------------P-.-----V---N------E--------------T-.-----Q-----ISP....R--G..---PG----..-   97
24_BG.ES.08.X2456_2            ------------------------------------P-.-----A---N------E----------------.-----Q-----ISP....R--R..-D-SG----..-   97
25_cpx.CM.02.1918LE            --RT----------N---------------------P-.------Y--N------E--V-------------.-----Q-----IIP....---G..---P-----..-   97
26_AU.CD.02.02CD_MBTB047       --------------Y------------Q--------P-.----------------G--------------V-.-----S-----IF-....S--G..----G-P..-   97
27_cpx.FR.04.04CD_FR_KZS       --------------N----I----------------P-.------YV-N------E--V-------------.-----Q-----IIP....---G..---P--P..-   97
28_BF.BR.99.BREPM12609         --------------Y------------S----------.---------------------L-------M---.-----------IN-....----..---------..-   97
29_BF.BR.01.BREPM16704         --------------Y------------S----------.---------------------------------.-----------INP....----..---------..-   97
31_BC.BR.04.04BR142            --------------N------------Q--------P-.------Y---------T--------------V-.-----Q-----II-....----..---S-----..-   97
32_06A1.EE.01.EE0369           --R------------------------S---K----P-.--S---Y--N------E----------------.-----N-----IIP....---G..---T-----..-   97
33_01B.ID.07.JKT189_C          -------------------------------K----P-.------X--N------X--------------X-.-----Q-----I-P....G--G..----G----..-   97
34_01B.TH.99.OUR2478P          -----A------------------------------P-.------Y-DN------Q--------------V-.-----Q-----ILP....G--G..---------..-   97
35_AD.AF.07.169H               --------------Y---M-----D--H--------P-.-Q----Y--N------------------------------L----I--....G--V..-D-SG-P..-   97
36_cpx.CM.00.00CMNYU830        --------------F------------H--------P-#------Y---------E-----M--------V-.-----Q-----II-....G--G..----G----..-   97
37_cpx.CM.00.00CMNYU926        --------------Y---A-----------------P-#FMA*-NISMKPM-ILGKELKP**KYYSNYCLSI.-----Q-----I-P....R--V..---P-----..-   94
38_BF1.UY.03.UY03_3389         ---------------------------A--------P-.--S------N------E-----T-T--------.-----H-----IH-....---V..---N-----..-   97
39_BF.BR.04.04BRRJ179          -------H------Y------------T--------V-.------Y---------------T-T--------.-----Q-----I--....----..---S-----..-   97
40_BF.BR.05.05BRRJ055          --------------Y------------S--------P-.--N---Y---I-----T----------------.-----H-----II-....L--S..---S-----..-   97
42_BF.LU.06.luBF_18_06         --------------Y---------------------P-.------Y---------E----------L---T-.-----------L-L....R---..---------..-   97
43_02G.SA.03.J11223            ---V----------X------------H--------P-.---------N------E-------X------V-.-----Q-----II-....G--G..-H--G----..-   97
44_BF.CL.00.CH80               -----G--------Y------------S--------P-.--S---Y--N------E----------------.-----H-----II-....---V..----N----..-   97
45_cpx.FR.04.04FR_AUK          --------------Y---M-----------------P-.----------------------V----------.-----Q-----IIP....G--V..-D-PG----..-   97
46_BF.BR.07.07BR_FPS625        -----G--------Y---------------------P-.--S------N------E----------------.-----Q-----II-....---V..----G----..-   97
47_BF.ES.08.P1942              --------------Y---------------------P-.---------N------E-------M--H-----.-----N--K--II-....---T..---------..-   97
48_01B.MY.07.07MYKT021         ------------------------------------T-.------Y--N------E--------------V-.-----Q-----IIP....G--G..----R----..-   97
49_cpx.GM.03.N26677            --------------Y------I-----Q--------P-.--S---Y--N------E--V---K---------.-----H-----II-....---G..------P..-   97
50_A1D.GB.10.12792             --R-----------Y-Q-A-DM--D--H----------.--E------N------E-------T------V-.-----Q-----II-....G--G..---S-----..-   97
51_01B.SG.11.11SG_HM021        --------------Y-A-------------------P-.------Y--D------E-------T--------.Y----Q-----IIP....G--G..------P..-   97
52_01B.MY.03.03MYKL018_1       ------------A-Y---A-----------------P-.-Y----Y-TN------E--------------V-.-----Q-----I-P....G--G..---------..-   97
53_01B.MY.11.11FIR164          --------------N------------H--------Q-.------X--N------E--Q-----------X-.-----Q-----IIP....G--G..---------..-   97
54_01B.MY.09.09MYSB023         -----------------------------------SP-.-L-------N-----RE--------------V-.-----Q-----IIP....G--G..----G----..-   97
55_01B.CN.10.HNCS102056        -------H------Y---------------------P-.---------N------E-----------I----.-----Q-----ILP....G--G..----G----..-   97
56_cpx.FR.10.URF5_patient_A    --------------S------------H--------P-.---------------------------------.-----Q-----ILP....-G-V..-D-SD----..-   97
57_BC.CN.09.09YNLX19sg         ---P-----------------------Q--------P-.------YV--------T------K---------.-----Q-----IL-....----..---------..-   97
58_01B.MY.09.09MYPR37          ---------------------------T--------P-.------Y--N------E--------------V-.-----Q-----I-P....G--G..----G----..-   97
59_01B.CN.09.09LNA423          -------H------Y---------------------P-.------Y--H------E-------T------V-.-----Q-----IIP....G--G..----G----..-   97
60_BC.IT.11.BAV499             --------------Y-A----------Q--------P-.------Y---------T--------------V-.--V-GQ-----II-....-N--..------N..-   97
61_BC.CN.10.JL100010           ---------------------------Q--------P-.--S---YV-DN-----T----------------.Y----H-----IV-....----..---------..-   97
62_BC.CN.10.YNFL13             ---P----------YX-----------Q--------T-.------Y-----------------T------T-.-----Q-----IL-....----..---------..-   97
63_02A1.RU.10.10RU6637         --------------F------------H--------P-.------Y---------E--R-----------V-.-----Q-----I--....G--G..---S-----..-   97
64_BC.CN.09.YNFL31             --RP---H------Y------------Q--------P-.------YV--N-----T---T------------.-----Q-----IM-....----..---------..-   97
65_cpx.CN.10.YNFL01            ---S-------..R--G----------Q--------P-.---------N------E--------------V-.-----Q-----ILP....G--G..---------..-   95
67_01B.CN.11.2011.ANHUI.HF115  ------N-------Y---------------------P-.------Y--N------E-------T------V-.-----Q-----ILP....G--G..----G----..-   97
68_01B.CN.11.2011.ANHUI.WH73   --------------Y---------------------P-.------Y--S------E-------T------V-.-----Q-----ILP....G--G..----G----..-   97
70_BF1.BR.10.10BR_PE004        --------------Y---A--------S----------.---------D-----------------------.-----Q-----IIQ....R---..---TG----..-   97
71_BF1.BR.10.10BR_PE008        ------------------------------------P-.------------R--------LVK---------.---R-Q-----ILL....R--T..----G----..-   97
72_BF1.BR.10.10BR_MG002        --R-----------Y---------------------V-.------Y---N-------------T--------.-----Q-----I-P....S---..---------..-   97
O.BE.87.ANT70                  ------N---AK--F---A--------A--------P-.--A---Y---------V--M-----------T-.Y----Q-----IN...PRG-GR..---S-----..-   98
O.CM.98.98CMA104               ---T------A---F---A-----D--T--------P-.-XA---Y---------L--M--T-----I--A-.-----Q-----INPSNTRG-GR..---S-----..-   101
O.CM.98.98CMABB141             ---V------A---F---A--------A--------P-.-MS---Y--D------V--M-------L-I-A-.-----Q-----INPSNTRG-GR..---S--P..-   101
O.CM.98.98CMABB212             ------N---A---F------------E--------P-.-QAC--Y--D------E--M-----------T-.Y----Q-----IIPSNTRG-GR..---S-----..-   101
O.CM.98.98CMU5337              --R----H--A---F---V--T-----A--------H-.-QA--NY---------V--M--T--------T-.-----Q-----INPSNTRG-GR..---S-----..-   101
O.CM.99.99CMU4122              -------H--A---F---A--T--D--A--------P-.--A---Y---------V--I--M--------T-.-----Q-----INPSNTRG-GR..---S-----..-   101
O.FR.92.VAU                    ----------A---V-A-A--------A--------N-.-QA---Y---------E--I--T-----II-T-.Y----Q-----IIPSNARR-GR..-D-SH----..-   101
O.GA.11.11Gab6352              -----N----P---F---A--------A--------P-.--A---Y---------V--M-----------T-.-----Q-----INPTNPRG-GR..---S-----..-   101
O.SN.99.99SE_MP1299            ----------A---F---S--------A--------P-.-QA---Y--D------V--M-----------A-.-----Q-----I-PSNTRG-GR..---S-----..-   101
O.US.10.LTNP                   ----------A---F---A--------T--------H-.------Y---N-----R--M--------M--T-.Y----Q-----ID...PRG-GR..---S-----..-   98

N.CM.02.DJO0131                -------A------Y---------------------T-.---------N------E----------------.Y----Q-----I-P....RG-...---------..-   96
N.CM.02.SJGddd                 -------A------Y---A-------------------.---------D------E-------M--------.Y----Q-----IIP....---...---------..-   96
N.CM.04.04CM_1015_04           -------A------Y---A-----D-------------.---------D------E-------M--------.Y----Q-----I-P....---...---------..-   96
N.CM.06.U14296                 --R----A------Y---A-----------------M-.---------N------E-------M--------.Y----Q-----I-P....---...---------..-   96
N.CM.06.U14842                 --R----A------Y---X-----------------X-.---------N------E----------------.Y----Q-----I-P....---...---------..-   96
N.FR.11.N1_FR_2011             --R----T------Y-----------------------.---------N------E----------------.-----Q-----I-P....---...-------I-..-   96

P.CM.06.U14788                 ---P------P---F---M-QT-----E--------P-.--S---Y--N------E--T---------I---.Y----Q-----ILTPS..R-GR..-H-P-----..-   99
P.FR.09.RBF168                 ----------P---F---M-QT-----A----P--LQY.--S---Y--N------E--T------X--I---.Y----H-----ILPPS..R-GR..-H-P-----..-   99

CPZ.CD.06.BF1167               ------N---P---W-Q-L-DT---I----M-----NT.-QNI-NWVFQN---S-E--QVL-TL--KA--T-.Y-Y--A-----I.......GPR..-------RS-   96
CPZ.CM.05.SIVcpzMT145          ----------S---YEQ-L-DT---I-Q--------P-.--A--E--------S----Q-----------V-.-----H-----INP....---R..-D-TN-A..-   97
CPZ.GA.88.GAB1                 ----------P---YQ--A--T--------------P-.--Q---F--D------V----------H-----.--L--Q-----ILP....---R..S--SN----..-   97
CPZ.TZ.06.SIVcpzTAN13          --------------Y-Q-I-DV-----Q--------PV.-Q-V-NWVFTV---S-E-AQEL-K---RA--S-.Y-H--V-----...................-..-   84
CPZ.US.85.US_Marilyn           ---V------A---Y---A-DV-----E--L-----P-.-Q----Y--S------E----------------.Y----N-----ISLARRTPQGR..---S-----..-   101

GOR.CM.07.SIVgor2139_287       ---I-T-E------F------M-----Q--------P-.--Q----V-D------E--T--------VI-V-.Y----H-----ILPPS..R-GR..---PN----..-   99
GOR.CM.07.SIVgorCP2135con      ---I-T-E------F------V-----Q--------P-.--Q----V-D------E--T--------VI-V-.Y----H-----ILPPS..R-GR..---PN----..-   99
```

```
                              Tat start            C-rich region            nuclear localization     exon 1 end  exon 2 start                   Tat end
B.FR.83.HXB2                  MEPVDPRLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRQ.RRRAHQNSQTHQASLSKQPTSQPR..GD.PTGPKE*KKKVERETETDPFD...............*  100
A1.AU.03.PS1044_Day0          -D----N----N------T---NK-------Y--PS--LN-G------------.--GPP-GNKNY-XPVP---IP-TQ..R-.S---E-P-----SKA---R--...............-  102
A1.CH.03.HIV_CH_BID_V3538     ------N-----------T-P-S--------W------LK-G----------KH.--GT-PS-KD--NP-----IP-TQ..R-.S--QE-S-----SKA-A-R--...............-  102
A1.CY.08.CY236                -D----N----N------T---SK-------Y---D--LK-G------------.---PP-S-ED--NPIPE--IP-TQ..-V.S---E-S-E---SK--A-R--...............-  102
A1.ES.05.X1608_8              -D----N----N------T-P-SK-------Y------LN-G------------.--GTP-S-KD--NPIP---L--AQ..-I.----E-S--E--SK--A-R-A...............-  102
A1.KE.11.DEMA1IKE001          -D----N----N------T-P-NK----R--Y--PA--LN-G------------.---TPHS-KD--NPVP------TQ..--.S---E-S-----SKA-A-R--...............-  102
A1.RU.11.11RU6950             -D----N----N----------S--------W---L--LK-G-----------H.--GTPHS-KD--NPI----LPHTQ..R-.Q---E-SA----SKA---R-A...............-  102
A1.RW.11.DEMA111RW002         -D----N----N------A---S--------Y--P---LS-G--V-------KR.--GTP-SNKD--NPVP--SIP-TK..-N.----E-S--E--SKA---R--...............-  102
A1.SN.01.DDI579               ------N----H----------S----A--W------LN-G----------K-.--GTP-SNKD--NPVP---LP-TP..--.-------S-----SK----R-A...............-  102
A1.UG.11.DEMA110UG001         -D----N----N------A---SP-------Y------LN-G------------.---TSHS-KD--DLVP--SIP-TQ..-V.---T--S-----SK----RL-Q.....RADS......-  107
A1.ZA.04.04ZASK162B1          ------N-D--N------S-P-S-----Y--W------LN-G-----------R.--GSPPS-KD--NPIQ---I--T-..R-.S---E-S-----SKA---R--...............-  102
A2.CM.01.01CM_1445MV          ------K----N------R---N-----V--W------LN-G-----------P.---PS-S-KD--NPV-E-SLPHAQ..RV.S--Q--S--E--SK---HQ--...............-  102
A2.CY.94.94CY017_41           ------K----N-----------K----R--Y---L---N-G-----------P.--KPSPSNKD--NPIP--SLP-AQ..RV.----E-P--E--SKA---R--...............-  102
B.BR.10.10BR_MG029            -----------Q------R----T-------Y---A--L--G------------.---TP-G-K---D--P---A---Q..-E.----E-S-----K------GY...............-  102
B.CA.07.502_I191_03           ---I--S----------------N----------------K-G------------.---SP-S-KID--P-----A--H-..--.-------P--E------HQVG...............-  102
B.CH.08.M2_0803101_NFLG8      ------------------R---N-----G---------TK-G------------.---S-PD---------------S-..--.-------P---------IH-S-...............-  102
B.CN.10.DEMB10CN002           ----------------------------------L------TK--------------.---PP-D-K---NP-P---A---..--.------S-----------VH...............-  102
B.ES.10.DEMB10ES002           ---------------------------------------TQ-G------------.----P-D-----V------------..--.S-----S-----T------DP...............-  102
B.FR.11.DEMB11FR001           --------------------NK----R--Y---L-----G------------.---TP-D-P----------------..--.------S-E----KA--H-DA...............-  102
B.GB.05.MM45d213_GN1          --------------------------R----------K-G---N--------.---PSEG-----D-------------..--.Q-----S-E----------G-W.....STDS......-  107
B.HK.06.HK003                 ----------------------------L------MQ-G------------.----PPE--N--V--P------C-..--.----E-S-----T--A---VP...............-  102
B.HT.05.05HT_129389           --L---------------R---X-----X--Y------L--G------------.---XS-X------A-----A-------..--.-K----Q-----------Q-W...............-  103
B.JP.12.DEMBI2JP001           ------S-----------------N----------P---TK-G------------.---PP-G-----VP-P---A--L-..--.-------S-E--------QV-...............-  102
B.KR.07.HP_18_07JHS10_3909    ----------------------P-----------L-------K-G------------.----P-DNKN--V---------A-..--.--Q--S-E---K---M--VH...............-  102
B.PE.07.502_0525_wg5          ------N-----------R---N--------W------S--G------------.----PKD-------PI-------------..--.-A----S-----T----H-H-...............-  102
B.RU.11.11RU21n               ----------------------------L------MK-G------------.---SP--N--D----P---A----..--.------S----------H--...............-  102
B.TH.08.MERLBDTRC10           ------------------R-T------------------K-G------------.--GPP-D-------------A-----..--.-E----S-----------GH...............-  102
B.US.11.ES38                  -D----N-----------R-------------------------G------------.----P-G-PNN-VP--E--A---S..--.------P---------K---H...............-  102
C.AR.01.ARG4006               ------K----N--------P--K------SY------L--G------------.--S-PPSGED--NLI-E--L--T-..--.QA-SE-S-----SK-K-----...............-  102
C.BR.07.DEMC07BR003           -D-I-HN----N--------P-NK----R-SY---I--L--G------------.--S-PPS-ED--DLI----L--A-..--.Q---EK------SK--A----...............-  102
C.BW.00.00BW5031_1            ------N----N--------P-NK----Q-SY--P---L--G---Y---------.--S-PSS-EN--NPV----LP-A-..-N.S--SE-S-----SK-------W..........RTDS-  107
C.CN.10.YNFL19                L--I--N----N------D---N-----R-SY--L---QK-G------------.--S-P-S-ED--NPI----L-RTK..--.S--SE-S-----SK-K----A...............-  102
C.CY.09.CY260                 ---I--N----N----------NK----Y--Y--LA--Q--G------------.--S-PR--ED--DLI----L--T-..--.-------E-S*--K-SNK-ETQLV...............-  101
C.ES.08.X2363_2               --L---K----N--------P-NT----S-SY--IA--QK-G------------.--S-PSS-KS--DLIP---LP-T-..--.Q--SE-S-----SK--A--S-...............-  102
C.ET.02.02ET_288              ------N----N--------P---------SY--P---L--G------------.--S-PPS-ES--DPI----LP-T-..-N.---SE-S-----SK------A...............-  102
C.IN.09.T125_2139             ------N-D--N------D---N--F--S-SY--L---QK-G------------.--S-P-S-ED--NPI----L-RAQ..--.---S--S-----SK------A...............-  102
C.KE.00.KER2010               --L---K----N----------NT------SY--L---QK-G---------W--.----PPS-ED--DLI----LP-T-..--.---SE-S-----SK-K-----...............-  102
C.MW.09.703010256_CH256.w96   ---I--K----N----------NK------SY--L---Q--G------------.--S-PPS-KD--NP-P---L--T-..-N.---SE-S-----SQ-------W...............-  103
C.TZ.08.707010457_CH457.w8    --Q---K----N----------NT-F--R-SY--L---Q--G-----------K.--S-PPS-ED--NPI----LP-T-..--.---SE-S--E--SK-T---L-...............-  102
C.YE.02.02YE511               ---I--K----N------R---NK----A--Y--L---QK-G------------.--S-PPD-ED--NLV----LP-T-..--.Q--SE-S-----SK------NW.....RTDS......-  107
C.ZA.08.705010534_CH534.w12   ---I--N----N----------NK------SY--L---Q--G------------.--STPSS-KD--NP-----LP-T-..-N.S--S--S--E--SK---------...............-  102
C.ZA.10.DEMC10ZA001           --Q---K-------------P-NS------SY--L---QK-G------------.--S-PPS-ED--NPI----ST-T-..--.----S--S-----SK--A-QSG...............-  102
C.ZM.11.DEMC11ZM006           -D----K----N----------NK----R--Y--L---Q--G------------.--S-PPG-KD--NPI----SP-A-..-N.---SE-S-----SK-KA--YA...............-  102
D.CM.10.DEMD10CM009           -D----N----N------S---NS----R--Y---R---S-G----------G-.--KPPKDNT---VPVPD--S--H-..--.-----------E--SK--A----W...............-  103
D.CY.06.CY163                 -D----N----N------S---NS----R--Y---C-----G----------G-.---PP-GNK---GPVPE--S--HP..-P.-P--NPS--E-ASK--A----W...............-  103
D.KE.11.DEMD11KE003           -D-I--S----N------R-P-NK-------Y---Q--VS-G------------.----P-GG-V--DPIP---S----..--.--------------SKA-----AW...............-  103
D.KR.04.04KBH8                -D----N-----------R---N--------L------LK-G------------.---PP-GD-A--VP-----P----..--.--------Q--E--SK--ANQ--W...............-  103
D.SN.90.SE365                 -D----N----N--------P-N--------Y---N-----G------------.---TP-GN----V-VP---S--H-..--.--------------SKA---Q--W...............-  103
D.TZ.01.A280                  -D----N----N------R-P-N--------Y---N-----G------------.---TP-GD-A--DPIP---S----..--.Q-----------SKA-A-R--W...............-  103
D.UG.10.DEMD10UG004           -D----NI---N--------P-NK----N--Y---------G------------.---PP-SG-A--DPIP---S----..--.--------------SKA---Q--C............WE.  104
D.UG.11.DEMD11UG003           --L---N-D--N------R-P-NK-------Y---N-----G------------.--GTP-GG-A--VPVP---S--L-..--.--------------S-A----Y-W...............-  103
D.YE.02.02YE516               -D----N---------R-S-P-SP-------------------G----------G-.---SP-SN----VPVP---S----..--.--------------SQAKAP---W...............-  103
D.ZA.90.R1                    -D----N----N--------P-NK-H-----Y---L---S-G------------.--KPPHSD-D--VPIPE--S----..--.--------------S-A--NQ--W...............-  103
F1.AO.06.AO_06_ANG32          --I---N-D--N------T-P--K----A-----YY--G--G------------.---TP-S-KN--NPVPE--L--A-..--.------S--EM-SK---S-*-...............-  101
F1.AR.02.ARE933               --L---N-D--N------T-P-------R-----YR-LAR-G------------.-H-TPHGT-I--DLVP---L--A-..-N.------S--E--SKAK---C-...............-  102
F1.BR.10.10BR_RJ015           --L---N-D--N------A-P-NK----R-----YW--T--G------------.-H-HR----L--DPVP---S--A-..--.------S--E--SK-----CA...............-  102
F1.CY.08.CY222                --------D--N------T-P--R----------YY--A--G------------.-H-TP-S--N--DPV----L--A-..--.------S--E--SK-----C-...............-  102
F1.ES.02.ES_X845_4            --V---N-D--N------T-P--T----V-----YR--T--G------------.---PP-SNKN--NPVPE--L--A-..--.------S--E--SK---G-*-...............-  101
F1.RO.96.BCI_R07              ------N-D--N------T-P-NR----------YH--T--G-D----------.-H-TP-S-PI--NPVP---L--A-..-N.------S--E--NK--A--*-...............-  101
F1.RU.08.D88_845              --L---N-D--N------A-P--K-----------R--TM-G------------.--GTX-S--I--NPVPE--S----..--.--Q--S-----SKA----*-...............-  101
F2.CM.02.02CM_0016BBY         --M---K-D--N-------NP-NK----R------L--TS-G------------.--S-SHS-KD--DPV----L--A-..--.----E-S-----SK-K---*-...............-  101
F2.CM.10.DEMF210CM001         --V---KID--N------E-P-NK-----------L--T--G------------.---TP-S-EI--NPV----L--A-..--.-------S-----SKAK---*-...............-  101
F2.CM.10.DEMF210CM007         --I---N-D--Q------E-P-N------------R--TS-G------------.---SP-S-KV--DPVP---L--A-..--.----E-S-----SK-K-G-*-...............-  101
G.BE.96.DRCBL                 -D----K-------------P-N-----SVAAL.----LN-G----------KH.--GTPHS-KD--TPVP---F-TT-..-N.----Q-S--E--SK--------...............-  101
G.CM.10.DEMG10CM008           -D----K-------------P-NK----A--W------LN-G----------KH.--GSSHG-KD--NPVPE--SPIS..-N.---Q--S--E-ASKA---Q--...............-  102
G.CN.08.GX_2084_08            -D----N----N----K-----NK----V--W---L--LN-G----------KH.--ATPSSNKD--NPVP---IPTT..-N.--D---S-----SK-K-----...............-  102
G.CU.99.Cu74                  -D----N----N----------N-----R--W---S--LH-G----------KH.--GTP-G-KD--NPVP---LPTT..-N.------S--E-ASK-------...............-  102
G.ES.09.X2634_2               -D----N----N-----------K----V--W------LN-G----------KH.---TP--NTD--NPVPE--LPTI..-N.----E-S-----SK---N---...............-  102
G.GH.03.03GH175G              -D----N----N------T-P-NK----V--W------LN-G----------K-.--GTP-G-KD--NPVP---LPIS..-N.Q--*--P--E-ASK---G-CA...............-  101
G.KE.09.DEMG09KE001           -D----N-----------R---N-----V--W---L--LN-G----------KH.----SPG-KD--NPVPE--LPTT..-N.-A---S--E--S--K------...............-  102
G.NG.09.09NG_SC62             -D----KI---N----------NK----V--Y------LN-G----------NH.---TPSG--N--NPVP---FPTT..-N.------S--E--SK-------...............-  102
G.ZA.01.TV546                 -D----N----N--------P-NK----V--W------LN-G----------K-.--GTP-D-KD--NPVP--SLPTT..-I.----E-S-----SK--AN-C-...............-  102
H.BE.93.VI991                 -D----NQ---N------R---N--------Y---L--LK-G---Y---------.--GTPKSL-D--TLIP---L-RTS..--.-----EK-----ASK-------W...............-  103
H.CF.90.056                   -D----K----N------Q---N--------Y---M--LK-G----------S-.-H-TPASL-D--N-I----L-RTH..--.-------Q--E-ASK-----*-...............-  101
H.GB.00.00GBAC4001            -D----K----N----------NG----N--Y---L--LK-G------------.--GTPAGL-DN-DPIP---LPRT-..--.--------------ASKA----LH...............-  102
J.CM.04.04CMU11421            ------NI---N------R----Q-----------------LK-G---N---------.--A-PPG-KX--DXVXE--LPLTS..RK.Q-----Q--E--SKA-PXR--...............-  102
J.SE.93.SE9280_7887           ------NR---N-------------------Y------LQ-G------------.--S-PPG-K---DLIP---L--TQ..RK.----E-S--E--SKA-P-R--...............-  102
J.SE.94.SE9173_7022           ------NR---N-------------------Y------LQ-G------------.--S-PPG-KN--DLIPE--LF-TQ..RK.----E-S--E--SKA-P-R--...............-  102
K.CD.97.97ZR_EQTB11           ------NI---NQ---------NQ-------Y---L--LQ-G---C---E----.-TTTPYA-KN-KDPIP---LP-A-..--.------S--E--SK-K---*-...............-  101
K.CM.96.96CM_MP535            -D----NI---NQ---------NQ----R--Y---I--LK-G---N------P.--TTPY--EN--DP-R---L-----..-E.Q-D---S-----SK-K--Q--...............-  102
U.CA.01.TV749                 -D-I--NI---NQ---------N--------Y---L--LK-G---H---------.--A-PLS-KD--NPVP--S-P-T-..RV.S--S--SQE---SQ--P-R--...............-  102
U.CA.99.TV721                 -D----NI---NQ---------N-----------------LK-G-----------P.--A-PLS-KD--NPVP--SIP-AQ..R-.S--S--SQE---SQA-P-R--...............-  102
U.CD.83.83CD003_Z3            -D---------N----------NS----R--L------M--G----------GK.---TP-SGPN--NIV----S----..--.---QE-P-----KK-T-----...............-  102
U.CD.90.90CD121E12            -D----K----N----------N--------L------VK-G----------GK.---TP-SDPD--NIVP---L----..--.----E-P-----KK-IP----...............-  102
U.CY.05.CY090                 ------N----N----------NK-F--------VA--TR-G------------.----S-SN-N--NPIP---L-HA-..-N.------S-EE--SK-----C-...............-  102
U.CY.08.CY223                 ------NI---NQ---------NQ-------W---L--LK-G------------.--A-SRS-KN--DHIPE-SLP-T-..RV.S--SE-S-E---SK--P-R--...............-  102
U.ES.10.DEURF10DZ001          --L---S-----------T-P-N-----F--W---L--LN-G-----------R.-PGPP-GR-D--D-VP---L-IF-..-N.S--S--Q--E-ASK---GQ--...............-  102
U.GR.99.99GR303               -D----NI---NQ----------------------------R-G------------.--A-PHDNKN--DLIP--SL--A-..--.S-----S-----SK-K-----...............-  102
U.NL.95.U_NL_95_H10986_D1     ---I--KI---NQ---------NQ-------Y---L--LK-G---N------P.--TTPYD-KN--DPIPE--F--A-..--.-------P--E--SK--------...............-  102
01_AE.AF.07.569M              --L---N----N------T-N-S--------W---L--LK-G----------KH.--GTP-S-KD--NPIP---LPII..-N.---Q--S--E-ASKA----C-...............-  102
01_AE.CF.90.90CF11697         ------S----N------T---SK-------W---L--LK-G----------KH.--GPPPG-KD--NPIP---LPTT..-N.------S--E-AKKAK----A...............-  102
01_AE.CN.10.YNFL03            ------N-----------T-P-SP-------W---L--LK-G---H------KH.--GTP-S-KD--NPIP---LPII..-NH--D---S--E-ASKA----*-...............-  102
01_AE.HK.04.HK001             ------N----N----K-T----K----I-SW---L--LK-G----------KH.--GTPRS-EG--NPVP--SL-TS..-N.Q-D-E-S----AGK-----*-...............-  101
01_AE.IR.10.10IR.THR48F       ------N----N------P-V-SK----I--W---L--LK-G----------KH.---TP-SNKN--DPIPE-SLPTS..-N.--D---S--E-ASK---G-CA...............-  102
01_AE.JP.x.JRC77AE            --L---N----N----K-T---SS----R--W---L--LK-G----------KH.--GTP-SRKD--NPVPE--LPII..-N.--D---S--E-ASKA----C-...............-  102
01_AE.TH.09.AA090a_WG11       ------N-----------T-D-SK-------W---L--LK-G----------KH.--GT--SNKD--NPIP---FPIS..-N.S-----S--E-ASKA----YA...............-  102
01_AE.TH.90.CM240             --L---N----N------T---SK-------W---L--LK-G----------KH.--GTP-S-KD--NPIP---LPII..RN.--D---S--E-ASKA---QC-...............-  102
01_AE.US.05.306163_FL         ------N-----------T-P-S--------W---L--LK-G----------KH.--GTP-S-KD--NPVP---L-II..-N.---Q--S--E-ASK---G-C-...............-  102
01_AE.VN.98.98VNND15          -K----N----N------T-N-NT-------W---L--LK-G----------KH.--GTP-S-KD--NPIRE--LPII..-N.--D---S--EGASKA----*-...............-  101
```

Draft of 30 September 2015, not for public release
hiv1prot_main.tex r8876–10546

Tat start | C-rich region | nuclear localization | exon 1 end/exon 2 start | Tat end

```
B.FR.83.HXB2          MEPVDPRLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRQ.RRRAHQNSQTHQASLSKQPTSQPR..GD.PTGPKE*KKKVERETETDPFD.............*  100
02_AG.CM.10.DE00210CM013  ------S----N------A---S-----W--W---L--LN-G----------H.--GTP--RRD--NPVP---LPTT---N.----E-P--E--SK-----Y.............-  102
02_AG.CY.09.CY256         --L---K----------T---NT-------W---L--LK-G--------Q-GR.---TP-SR-D--NPVP---LPTTS..-N.---Q--S--E--SK--A--CA............-  102
02_AG.ES.06.P1423         K.----S----N------S-------W---Y--LN-G----------P.---TP-SR-D--NPVP---LPTT---N.------S--E-ASK----Q*-.............-  100
02_AG.GW.05.CC_0048       ------S----N------T---SK----I--W---L--LK-G----------Q--GTPPS--D--NPVP---FTTI---I.-A----S--E-ASK----RC-.............-  103
02_AG.KR.12.12MHR9        --L---S----N------T---SK------W---L--LN-G----------R.G-GTP-SH-D--NPVP---LPTT---N.------S--E-ASKA--NQC-.............-  102
02_AG.LR.x_POC44951       -D---SI---N------T---SK------A--W---L--LN-G----------KR.--GSPRSH-DR-NPVP---LPTA---N.------S--E-ASK---NLC-Y...........-  102
02_AG.NG.09.09NG_SC61     --L---S----N------T---N-------W---L--LN-G----------R.--GTPKSR-D--NPVP---LPAT---N.------S--E--SK-----C-.............-  102
02_AG.NG.x.IBNG           --L---S----N------T---SK----M--#---L--LN-G----------R.--GTP-SR-D--NPVP---LPTT---N.--D---S--E--SK-K---C-.............-  102
02_AG.SE.94.SE7812        --L---S----N------T---SK----I--W---L--LN-G----------KR.--GTP-SR-DN-DPVP---LPTT---N.-A----S--E-AGK-----C-.............-  102
02_AG.SN.98.98SE_MP1211   --L---K----N------T---SK----L--W---L--LN-G----------R.--GTP-SR-D--NPVP---LPTT---E.Q-----S--E-ASK---G-C-.............-  102
03_AB.RU.97.KAL153_2      ----------------------------------L------MK-G----------.----P-DN--D-V--P---A-----------------M------H---.............-  102
04_cpx.CY.94.94CY032_3    ------D----N------T-D-NK-F-----W------LK-G----------KH.--GSL-G-KG--NLIP---L--QPN.---S---E-Q----ASK--A---A.............-  103
05_DF.BE.x.VI1310         -D----N----N-----R---NQ----M--Y---N-----G----------P.---PP-G--A--DPVPE--P-------------KQ--E--SK--A-Q--W.............-  102
06_cpx.AU.96.BFP90        ------KI---NQ---R------K-------Y--P---LN-G----------.-Q-PPG-KN--DPV----L--TQ..RE.Q---EKS--E--SKA-P-R-.............-  102
07_BC.CN.98.98CN009       ------K--S-------------------------MK-G---F----------.-S-P-S-ED--NLI----LPRTQ.---.---SE-S-----SK-K-----.............-  102
08_BC.CN.97.97CNGX_6F     ------N----N------E---N-----R-SY--L---QK-G----------.-S-P-S-ED--NLI---SLPRTQ.----.QA-SE-S-----SK-KPA-.............-  102
09_cpx.GH.96.96GH2911     --L---N----N------A---NK----R--Y---H--LN-G----------.-A-SCG-KSN-DPIP---LP-T---N.------S--E--SK-KS-Q--.............-  102
10_CD.TZ.96.96TZ_BF061    ----H-N----N---------NK-H--A-SY--L---Q--S----------.-S-PPS-ED--NLI----S-------------------A-A----W....WTDS.....-  107
11_cpx.CM.95.95CM_1816    ------NI---NQ---------NQ------Y---H--LK-G----------.-P-SHS-KN--DHIPE--L--AT..RK.--V-E-S--E--SK--P-R-.............-  102
12_BF.AR.99.ARMA159       --L---N----N------K-S-P--K-------R-----YW--V--G----------.----P-G---N------I--A----N.------S--E-SKAKS--C-.............-  102
13_cpx.CM.96.96CM_1849    -D---KI---NQ---------P--K-------R-----YW--V--G----------.----P-G---N------I--A----N.------S--E-SK---------W.....SADS.....-  107
14_BG.ES.05.X1870         -D----N----N----------K----R--W------LN-G----------KH.--GPP-SXKD--TPVP---LPTT---N.----E-S-----SK-K-----.............-  102
15_01B.TH.99.99TH_MU2079  -D---LN----N----R-T---N-----I--W---I--LK-G----------KH.--GTP-S-KD--NPAP---L-IS---K---D--KS--E-ASKA-----*-.............-  101
16_A2D.KR.97.97KR004      -D----N----N------R---NK----R--Y--P---LN-G----------P.---PP-GGTG--NPIPE-SLPR-Q..RI.---TE-S-TE--SKA----C-.............-  102
17_BF.AR.99.ARMA038       --I---N-D--N------T-P--R-F--E-----YW--T--G---N----------.---PP-GNPP--V--P---AP-A---N.------S--E--SKAK--Q*-W.............-  102
18_cpx.CU.99.CU76         ------NI---NQ---------N-----R--Y---L--LK-G----------.--TTPYD-KH--DLIPE--L-RTS..W-------Q-E--ASKA-A-QC-.............-  102
19_cpx.CU.99.CU7          -D----N----N----------NT----------L--TK-G----------.-GPP-SH-D--NPVP---LPDT..EG.---T--S--E--SK-K-----.............-  101
20_BG.CU.99.Cu103         -D----N----N----------NK----V--W------LN-G----------KH.--GPPPG-KD--NPVP---IPTT---.N.------S-EE--SK-T---CA..........-  102
21_A2D.KE.99.KER2003      -D----N----N------R---NK-H--R--Y--IK--L--G---R----------.---PP-GD-A--VPVP---S-------------------S-A------C.............-  103
22_01A1.CM.01.01CM_0001BBY --L---S----N------T---N--F--V--W------LK-G----------KH.--GTP-G-KD--NPIP---LP-TK..R-.S---E-S--E--SK---NR-A.............-  102
23_BG.CU.03.CB118         -D----N-D--N----------NK----M--W------LK-G----------KH.--GTP-D-KD--IPVP---LPTT---N.------S--E--SK--A----.............-  102
24_BG.ES.08.X2456_2       -D----N----N----------NK-F--V--W------LN-G----------KH.--GSP-S-KD--NPVP---IPTT---N.---EKS-EE-ASK-KA----.............-  102
25_cpx.CM.02.1918LE       -D---NI--N------G-P-N-----R----------K-G----------S-.--GTPPS-KD--NPVP---IPTT---.N.Q---E-S----ASK-----YA............-  102
26_AU.CD.02.02CD_MBTB047  -----RNID--N-------P-NK----Q--Y---L----G----------P.--GTSRSN-D---NP-P---S--AQ..RE.SP--E-S-----SK--Q-R-.............-  102
27_cpx.CM.04.04CD_FR_KZS  -D----N----NQ---------NQ----R--Y---H--LK-G---F----------.-T-PYGNKN---DPIQE-SIP-TQ..R..S--SE-S-----SKARA-R-.............-  102
28_BF.BR.99.BREPM12609    ----------Q------R---NK-------L-------MK-G----------.---P--DN-NN-V------A--S----N.------E-S-----T----A.............-  102
29_BF.BR.01.BREPM16704    ----------------------R---NK----------L----G----------.----T--D---D-V------A--F-----.------S--T----A.............-  102
31_BC.BR.04.04BR142       -D----K----N---------N----R-SY------L--LG----------.-S-PPS-ED--NLI----VP-S------.Q--SEKS-----SK-----C-.............-  102
32_06A1.EE.01.EE0369      -----HKI---NQ---R---P---P-------Y--YM--LN-G---Y----------.--T-PPG-KN--DPE----V--TQ..RE.----EK---E--SKATPNR-.............-  102
33_01B.ID.07.JKT189_C     ------N----N------AT-SK----I--W---L--LK-G----------KH.---PP-S-ED--NPX-E--L-IS---N..-D-E-S----ASK----RC-.............-  102
34_01B.TH.99.OUR2478P     ------N----N------T---SK----N--W---L--LK-G----------KH.--GTP-S-KD--NPIPE--LPII---N.Q-----S--E-ASKA----*-.............-  101
35_AD.AF.07.169H          -D----N-D--N------P-P-S-----I-SY------LN-G----------.---TP-RD-D--NPIP---LPRAK..-I.SA-QE-S-----SK--P-R-.............-  102
36_cpx.CM.00.00CMNYU830   --L---NI---N------R---N--F-----W------LN-G----------R.--G-P-SR-D--DPV----LPAI---KN------S--E-ASK---N-C-.............-  102
37_cpx.CM.00.00CMNYU926   -D----N----------------N----------------TK-G----------.--T-A-SR-DN-H-VP---LPTT---N.------S--E-ASKA----CA............-  102
38_BF1.UY.03.UY03_3389    --I---K----N------T-P-NK----R-----YW--T--G----H---------.---TP-SN-I--DPVP---V--A----N.------S--E--SK-K---CA............-  102
39_BF.BR.04.04BRRJ179     -D----------------------K----------T-G----------.---TP----I--DPVPE--I--AH..RN.------S--E--SKAK---*-.............-  101
40_BF.BR.05.05BRRJ055     -D----------------------K----R----------G----------.---SP-GR-S--G--#--SAP-Q.---------E-S-----------E-.............-  101
42_BF.LU.06.luBF_18_06    -----HK----E----------------R----------R-G----------.-NPPPD-E-----V----AT-H-----------Q----------DL.............-  102
43_02G.SA.03.JL1223       -L---S----N------T---SK------W---L--LN-G----------R.--GTP-SR-D--NPV----LPITS..EN.------S--E--SK-----.............-  102
44_BF.CL.00.CH80          --LI--N-D--N------T-P--R----R---YW--V--G----------.---PP-SD-I--DPVP---L--T-----.------S--E--SKA----CA............-  102
45_cpx.FR.04.04FR_AUK     -D----N----N------T---NK-H-----Y--YS--LN-G----------KH.--GTPYS-KN--NPIP---L--T-----.------S--E--SK-K-----.............-  102
46_BF.BR.07.07BR_FPS625   --L---N-D--N------T-P--K----R------W--V--G----------.--GPP-S--I--DLVP---I--AQ..RN.----E-S--E--SKA--G-W-.............-  102
47_BF.ES.08.P1942         ----------------------R----K----R----------G----------.--XPP-SG----D-------T-L-----.----E-S-----------V-R.....RADS.....-  107
48_01B.MY.07.07MYKT021    ------N---------K-T---SK------W--RL--LK-G----------KH.--GTSHRRED--NPIPE--L-IS---N.--D---S----ASKA--G-C-.............-  102
49_cpx.GM.03.N26677       ------KI---N----------Q-------Y---I--LK-G---C----------.---T-PPG-KN--DLVQ---L--TQ..RK.----E-S--E--SKAGP-R-.............-  102
50_A1D.GB.10.12792        -D----N----N------T-P--K-------Y--PA--LN-G----------K---GPPKGD-A--VPIP---S---------.-K----------SKA-A-Q--W.............-  103
51_01B.SG.11.11SG_HM021   ------N----N------G-P--A----R-S-----A--K----------.---T--D-E----------A-------.E-------S-E-------HSSRQ.......RLDS...-  107
52_01B.MY.03.03MYKL018_1  ------N----N------T---SS----I--W---L--LK-G----H-----KH.--GTP-S-KD--NPVPE--FPII----N.------S--E-ASKAA---C-.............-  102
53_01B.MY.11.11FIR164     ------S----N----X-T---SK------W---L--LK-G----------H.--GTP-S-KD--NPIP---LPIX---NNS-D---S--E-ASKA--X-C-.............-  103
54_01B.MY.09.09MYSB023    ------N----------T---SK----R--W---L--LK-G----------KH.--GPP-S-KD--NPIPE--LPTT---N.------S--E-ASKA--G-*-.............-  101
55_01B.CN.10.HNCS102056   -----N----N----K-T---SK------W---L--LK-G----------KH.--GNP-SR-D--NPVPE--LPII---.N.--D---S--E-ASKA----C-.............-  102
56_cpx.FR.10.URF5_patient_A -D-I--N----N---------NK----A--W---L--LN-G-X--------SH.---TPPG-KD--NPVP---LPTIN..-N.---QE-S----ASK-----*-.............-  101
57_BC.CN.09.09YNLX19sg    ------N----N------T---SK----I--W---L--LK-G----------KH.---STPP-ED--NLI----LPRTQ.----.Q--SE-S-----SK-KP-R-.............-  102
58_01B.MY.09.09MYPR37     ------N----N------T-D-SK----I--W---L--LK-G----------KH.--GTP-S-KD--NPVPE--LPTT---N.Q-N---S--E-ASQ-----CA............-  102
59_01B.CN.09.09LNA423     -----N----N----K-T---SK------W---L--LK-G----H-----KH.--GTP-S-KD--NPVPE-SSPII---.N.------S--E-ASKA----YA............-  102
60_BC.IT.11.BAV499        ---LETN----NQ-E-----P-N-----R-SY------L--G---Y----------.-S-PSSNKD--NPV----L--A------.Q-DLEKS----KSK-K-----.............-  102
61_BC.CN.10.JL100010      ------N----N------E---N--F--R-SY--L---Q--G----------.----P-S-ED--SLI----LPRTQ.---.T--SE-S-----SK--P-----.............-  102
62_BC.CN.10.YNFL13        ------N----N------E---N-----R-S---L--QK-G----------.----P-S-ED--NPIP---LP-IH.-------SE-S-----SK-K----A.............-  102
63_02A1.RU.10.10RU6637    -D----N----N---------S-------W---L--LN-G----------GR.--GTS-SR-D--NPVP---LPTT---.N.----E-S--E-ASK----QC-.............-  102
64_BC.CN.09.YNFL31        ------N----N------E---N--F----SY--L---HR-G--V----------.-----PPSGED--NPI---SLPKAQ.-------SE-S-----SK-RA-Q-.............-  102
65_cpx.CN.10.YNFL01       -----N---------Q---N--F--------X---TK----------.---TP-S-EDY-NP-----L-II---NN--DQ--S--E-ASKA----S-.............-  103
67_01B.CN.11.2011.ANHUI.HF115 -----N----------T-D-SP------VW---L--LK-G----H-----KH.---SPRS-KD--NPIP---L-IS---.N.----E-S----ASKA----C-.............-  102
68_01B.CN.11.2011.ANHUI.WH73 -----N----------T---SK------AW---L--LQ-G----------KH.---SPRS-KGY-NP-PX--L-II---.N.----EKS----ASK----QG............-  102
70_BF1.BR.10.10BR_PE004   --L---K----------R---K-------Y--------G---------W-.-----SHS--N--V--P----------.-------PE---------QDN.............-  102
71_BF1.BR.10.10BR_PE008   ---------Q-------R----T------SY---S--L--G----------.--G-P-SG----------A--L-----------SEE---G----HLGNQ.........WMDS-  107
72_BF1.BR.10.10BR_MG002   -----K----------N--F-----Y---A----G----------.-KPP-D-------P----T-L-----.----E-S-------A--QKH.............-  102
O.BE.87.ANT70             -D---EVP--H------QIP-N-----R--Y--Y---VR-G----------GR...P-AASHPD-KDPVP---SPTITK..RK.QERQE-QEEE--KKAGPGGYPRRKGSCHCCTRTSEQ-  115
O.CM.98.98CMA104          -D----EMP--H------Q-P-N-----S--Y--Y--LVR-G----------R...P-TASHPDNKDPVPE-SP-ITN..RK.*ECQE-QE-E--K--GPG*YHC.........SK.  100
O.CM.98.98CMABB141        -D----EMP--HQ-----PNP-N-----R--Y--Y---TS-G----------R.-SA-AASRPDNKDLVP--SP-ITX..RK.QERQE-QE-E--KKAGP-**HR..........RK.  102
O.CM.98.98CMABB212        -D----EMP--H---K-Q-P-NS----R--Y--L---TK-G----H------R..PA-AASYPDNKDPVPE-SP-NTG..RK.QKRQE-QE----T--DPGGRHCLQDSCNSCTRISGQ-  116
O.CM.98.98CMU5337         -D----EIP--H------PNP-N-----N--Y--Y---TR-G---------Y-R.---TTRHXD-KHLVP--SP-ITN..RK.QER-E-QE-E--KKAGPS*YPR..........LK.  101
O.CM.99.99CMU4122         -D----EMP--H------Q-P-NK------Y--Y---TS-G----------R...P-AAHHPDNKDLVPE-SPTITN..RK.QERQE-QE----K--GP-*YPC.........RQ.  102
O.FR.92.VAU               -D-I--EMP--H------Q-P-NA----Q--Y--YL--TK-G----H------R..PA-AASYPDNKDPVPE-SP-NNK..RK.*ERQE-QE-E--TK-GAGGLHC.........RK.  102
O.GA.11.11Gab6352         -D----EMP--H------Q-P-NT------Y--Y---TS-G----------R..PAVGASRPDNK-FVPE---AITD..RK.QKRQE-QE-E--KKAGP-*QPR..........HK.  102
O.SN.99.99SE_MP1299       -D----EMP--H------Q-P-NK----A--Y--Y---AS-G----------R..PA-AARHPDN-DIVPE-L-YITN..RK.QKRQE-QE-E--N-ACPR*YPG.........RK.  102
O.US.10.LTNP              -D----EVP--H------Q-P-NT------Y--Y---TK-G----------R...P-TASHPD-KDPVP--SL-ITK..WK.QERQE-QE-E--KKAGPGGDHCSKGSWHCCTRTSEQ-  115
N.CM.02.DJO0131           ----------N---------NG----Y--Y--MC--TK-G---------S-.----TPKS-KN--DPIPE--L--QQQP---.Q--Q-KQ--AL-DKA----*-.............-  103
N.CM.02.SJGddd            ----------N---------N-----R--Y--MC--TK-G---L------S-.----TPKS-KN--DLIPE--L--QQ...-X.Q--Q--Q--AL-SK--A--C-.............-  102
N.CM.04.04CM_1015_04      ----------N---------NK----N--Y--V---TK-G----------S-.--KTPNS-KDY-DPIPE--L--QP.----.Q--Q--Q--AL-SK--A-LC.............-  102
N.CM.06.U14296            ----------------R-----N--F--R--Y--MY--TK-G----------S-.----TLRS-KS--DLIPE--L--QQ.----.Q--QEKQE-TL-SK--A--C-.............-  102
N.CM.06.U14842            ----------------------------NK----R---X-I---TK-G----------SR.----TP-S-KN--DPIPE--L--QH.----.Q--Q-KQE-A--SK-----C-.............-  102
N.FR.11.N1_FR_2011        ----Y------N---------N-----R--Y--VC--MK-G-----X----S-.----PKS-KN--DIIPE--L--QH..-I.Q--Q--Q--AL-SK--A--C-.............-  102
P.CM.06.U14788            -D----D-P--QQ-----SSP-N-----A--Y--Y---TK-G----------R.PA--DN-N-N--DPV--*-FACTG..AQ.QE*HQ-QT----KQATA-RS-C.....------WK.
P.FR.09.RBF168            -DQ---D-P--QQ-----SSP-N-----A--Y--Y---TK-G----------R.PA--DN-N-N--DPV--*-LA*TG..IQ.QE*HQ-QT----KQATA-RS-CWTDSCLSSGTT...-  111
CPZ.CD.06.BF1167          ------D----L--S-T-S-P-NS----V--Y---L--TK----------R-G....KPSASNKNN-NPVRE-SL-KQ-----.QVSQ--Q--E-ANQA--GGG-I.............Q-  102
CPZ.CM.05.SIVcpzMT145     ----IE-D----N----K-R-E-NK-F-----Y--IL--TK-G----------G...-KPASSN-DNKDP-R---L---I..R-.Q--SEKQ--E---TS-A-QS-W.............-  102
CPZ.GA.88.GAB1            -D-I--D----------R-V-N-----A--Y--IY--TK-G----------TTR--T-PAG-KNN-D-IP---L--S---.N.KE-SEKST-E-ASK--A-Q*-.............-  102
CPZ.TZ.06.SIVcpzTAN13     -N-I--QVA--M---AA-E-P-------------PL--TK----------R-G...-KSAGDNK---DPVRQ-SLPKRS..RI.QSSQE-SQ-E--T-AGSGGRPRPEDSSASSGRTSGTS  115
CPZ.US.85.US_Marilyn      -D----NI----Q-----------N----------V---TK-G----------GRR--TTAKS-ADN-DP-R---F--QL...N.Q--Q--Q--AL-KQ-TS-QR-.............-  103
GOR.CM.07.SIVgor2139_287  -D-T--E-P--QQ-----P-P-NT------Y--Y---TK-G----------R.PA-TADKD-NN--PV---SLAGT..SQ.QE*YQK-E----GQ-T--RS-C.....------WQ.  103
GOR.CM.07.SIVgorCP2135con -D-T--E-P--QQ-----P-P-NT------Y--Y---TK-G----------R.PA-TADKD-NN--PV---SLPGT..SQ.QE*YQK-E----GQ-A--RS-C.....------WQ.  103
```

```
                            Rev start         exon 1 end exon 2 start   NLS            Leu-rich effector domain                                    Rev end
B.FR.83.HXB2                MAGRSGDS.DEELIRTVRLIKLLYQSNPP.PNP..EG.TRQARRNRRRRWRERQRQIHSISERILGTYLGRSAEPVPLQLPPLERLTLDCNEDCGTS.GTQ..............GVGSPQILVESPTV.LESGTKE*   116
A1.AU.03.PS1044_Day0        ---------.--D-LKAI-I--X------Y.-K-..R-.------K-------A-----D-------S-C---P----------I-------H------.---QSQ.......GTET---R--VS----GI.------N-   124
A1.CH.03.HIV_CH_BID_V3538   ------ST.-----QA--I--I------Y.-K-..R-.------K-------A-----D-------S-C--------A---------R---S--S---.---QSQ.......GAET---R--VSG--SVI.-G---ES-   124
A1.CY.08.CY236              ---------.--D-L-A--I--I------Y.-K-..R-.S----K-------A--K--D-----V-SAC----E---------I---H---C------.---QSQ.......RVET---R--VSG-P-GI.------N-   124
A1.ES.05.X1608_8            ---------.----L-A--I--I------Y.-KR..R-.S---Q--------A-----D-L-----S-C---P-----------G-H---H------.---QSQ.......GVET---R--VX----G-.-G---ET-   124
A1.KE.11.DEMA1IKE001        ---------.-A--LTA--I--I-------.-K-..R-.------K-------A-----D-------S-C----T-----------H---S------.---QSQ.......GVET---R--VP----A-.-G----N-   124
A1.RU.11.11RU6950           -------T.-A--LTA--I--I------Y.-T-..R-.------K-Q-----A-----D-L-----SSC------S-SF----I---H---S------.---QSQ.......GTET---G---SG-PSGI.-G------   124
A1.RW.11.DEMA111RW002       ------SG.----L-AI-I--I------Y.-K-..R-.-----Q--------A-----D---Q---S-H--G-E---HF----------N-S------.---QSQ.......GAAT---RS-VSG--D-I.-D---EN-   124
A1.SN.01.DDI579             ------N-.----L-AI-I--I------Y.-K-..Q-.------K-------A-----D-L-----S------E-----------H---S------.---.......................TGS*-S--CSAI.-GE---KQ   112
A1.UG.11.DEMA110UG001       ---------.-V--LTA--I--I------Y.-K-..K-.S----K-------A--K--D-------S-C------------------H---S--G---.---QSQ.......GTET---G--VS---SII.-G----K-   124
A1.ZA.04.04ZASK162B1        -------D.V-D-LQA--I--I--K---Y.-K-..G-.------K-------A-----D-------S-C---P--------------H---S---.................GTET--EGHS------A-.--P-----   117
A2.CM.01.01CM_1445MV        -------P.--D-LKA--T--I------Y.-T-..R-.S----K-Q-----A---H-S-------S-C---------------HI--S---D--.---QSQ.......GTET---GH--S---SVI.-G---E--   124
A2.CY.94.94CY017_41         T----D-P.--S-LQAI-T--I------Y.-K-..R-.S---Q---------A-----D-----V-R-C---PT------------HI--S------.--LQSQ.......GTET---RS-ES---SVI.-G---E--   124
B.BR.10.10BR_MG029          ---------.----LKA-----T------L.-S-..K-.N------------R-----QD--GWV-CNL---------H---I-------------.---................-----------X-AI.----A---   117
B.CA.07.502_I191_03         ----R---.--D-LKA-----R--------.-ST..--.-----------------H-RLA-DW--R---D-PT---D-P----------SS------.---..............-E-------------.-GT--E--   117
B.CH.08.M2_0803101_NFLG8    -------R.--D--Q--------------.-S-..--.-----------------Y--Q--G-L-S-----P--------------------------.-PE...............-------F----A-.--------   117
B.CN.10.DEMB10CN002         ---------.--D-LK------T-------.---..--.--------------------R----Q--H------V---------EI---------K---N-.---.................--------G--D--.-D--AA-E   117
B.ES.10.DEMB10ES002         ---------.----LK------F-------.-S-..--.--------------Q--K--QTL-AW--SNH-----------------------SK-----.---................---G--V-----A-.--------   117
B.FR.11.DEMB11FR001         ---------.E---LK------R-------.-S-..--.----------------K--H-RML-GW--DN----PE-----------------S-----A.---................---G--V-----A-.----A---   117
B.GB.05.MM45d213_GN1        ---------.--D-LKA-----I-------.---..--.--------------------RG-GQ---D-----XX--------------------------.---................----H-V-----A-.-D-----E   117
B.HK.06.HK003               ---------.----LQK--I--F-------.-SA..--.-------------KQ-----Q-L-GW-ISNH---P------------------DQE-----.---................---N----L-P---.----AE-E   117
B.HT.05.05HT_129389         ----R---.----LK------Q------L.-S-..--.--K-----------------K-G----A------E-----------S---S-G--D-.---................-------V-----A-.--------   117
B.JP.12.DEMBI2JP001         -------N.--D-LKA-----F------L.-SS..--.-----------------------K------A-----PE------------------D----N-.---................---N--------AI.-------E   117
B.KR.07.HP_18_07JHS10_3909  -------N.----L--I-I--F------L.-K-..--.-------------KK--W-R----W--N-H---P------H-----------S-E-----.---................------E------A-.-------K   117
B.PE.07.502_0525_wg5        ---------.----L------R---------.---..--.-------------QK--H-RT--DWL-T------P--------------------K-S----.---................---N-----------.-----E--   117
B.RU.11.11RU21n             ---------.--D-LK-I------------.---..--.-----------------H-R-------IC-----PE-----P-----------S--------.---................-------------A-.-D------   117
B.TH.08.MERLBDTRC10         ---------.--D-LK------------L.-S-..--.--R--------------RD------D-----PP---------I--------S-------.---................---G--V---P-A-.-G------   117
B.US.11.ES38                ---------.----LKA--I--F-------.-S-..A-.-----------------K-R----W--SNS---P-----F----I------SS----I-.---................---N--V-----A-.--------   117
C.AR.01.ARG4006             ---------.--A-LLA--I--T------Y.-K-..--.--R----------A--------------I-C---PT----------I---NIN-S-SG---.---................R--N-*-FRKPCA-.-G-RA-KE   116
C.BR.07.DEMC07BR003         ---------.--A-LQA-KI--I------Y.-K-..--.---------------A--------------T-C---P-----------I---SIN-S-SG---.---QSQ.......GTTE---N-*V-GKPCA-.-G-RA-KE   123
C.BW.00.00BW5031_1          -------N.--A-LQA--I--I------Y.-K-..--.------K-------A-------G----S-C-------------------HI--S-SS--P.---QPQ.......GDTE---R-*VSGK-CA-.-G----KE   123
C.CN.10.YNFL19              ---------.--A-LKA--I--I------Y.-E-..R-.------K-------A--K--R-L-----SSC---PT----F---------HIGDS-SG-A-.---QSQ.......GTTD--E--*-SGR-CA-.-G--A-KE   123
C.CY.09.CY260               ---------.-AA-LE--KI--I------Y.-K-..--.--------------R-ATKKRLN*F-H-L-SSGCVGPP-QL-FF-----E-L-IDS--SRS-C--HPPG.......GTAE-G-HHVFCE*CCV-.----RA--E   123
C.ES.08.X2363_2             -------N.--A-LLA--A--I------Y.-K-..--.---------------A--K---Q------S-C---------------I---YI--S--S-I-.---................----S*VPGK-CA-.-G----KE   116
C.ET.02.02ET_288            ---------.-AA-LQA--A--I------Y.-K-..--.-----Q--------A-------L-----SDC---HP-------------N---S-SG---.---QPQ.......GTTE---N-*VSGQYCA-.-G-RA-KE   123
C.IN.09.T125_2139           ---------.--A-LKA-KI--I------Y.-E-..R-.------K-------A-------L---V-SAC---P-----------I---HISDS-SG---.--PQSQ.......GTTE-----*-SGKPCA-.-G--A--G   123
C.KE.00.KER2010             -------N.-D--L-V--I--I------Y.-K-..--.------K-------A--R-----------SAC---P-----F----I---HI--S-GG---.---QSQ.......GTTE-----*VSGKLCA-.-G--A-KE   123
C.MW.09.703010256_CH256.w96 ---------.--A-LQA--I--I------Y.-K-..--.------K-------A--K-----G----SS----P-----------I---HIGDS-SG---.---QLQ.......GTTE-----*VSGKPCA-.-G-RA-KE   123
C.TZ.08.707010457_CH457.w8  ------E-.--A-LQA--I--I------Y.-K-..--.------K-------A---------------S-C----------------I---HI--S-GG---.---QPQ.......GTTE---N-*VSGK-CA-.-G-R--KE   123
C.YE.02.02YE511             -------N.-AA-LQ--KI--I------Y.-KL..--.---------------A--K-----G----SCC-E-P-----------I---HI--S-SS--T.---QSQ.......GNTE----S*VSRK-CA-.-GARA-KE   123
C.ZA.08.705010534_CH534.w12 -------N.--A-LQA--I--I------Y.-K-..--.------K-------A-------------V-C---P-----F----I---HI--S-SS---.---QHQ.......GTAE---N-*VSGK-CAI.-D-RA-KE   123
C.ZA.10.DEMC10ZA001         ---------.--V-LQA--I--I------Q.-K-..--.--------------A-----NQA-----ASC------------------HI--S-GG---.-E-QSQ.......GTTE-----*-SGK-CA-.-G--A--E   123
C.ZM.11.DEMC11ZM006         ---------.--A-LQA--I--I------H.-KL..--.------K-------A-----TL------S-C---P-----F----------HI--S-SS---.---QSQ.......GTTE-------SGK-Y-I.-G----KE   124
D.CM.10.DEMD10CM009         ----RE--.--N-LK-IQ---F---I---.-ST..--.---------------A-------G---S----------------I---N-N-S---EP-.---................----T--S----A-.-----E--   117
D.CY.06.CY163               ----RE--.--D-LKAI----V--------.-SI..QA.P----TH------A-------G---S-----P---------------N-N-S------.---................----T--S----A-.--T--E--   117
D.KE.11.DEMD11KE003         ---------.----LKA--F--I--------.-S-..A-.-------------A------LG--F-S-----PE----------------------------.---................----H-------S-.-D------   117
D.KR.04.04KBH8              ---------.--D-LKAI----F------L.-S-..--.---------------A-----N--G---IS-----------------I----IN--------.---................----S-VSG--SAI.-DT-----   117
D.SN.90.SE365               ---------.----LKAI----F---N---.-ST..--.---------------A-----N--G---S-----------------------N-N-S------.---...............RM----VS----A-.-DT--EK-   117
D.TZ.01.A280                ---------.----L-AI----I--------.-SL..--.---------------A-----D--G---S-----------------------N-N-HT-----.---................----H-VS--P-AI.-D---E--   117
D.UG.10.DEMD10UG004         ---------.-DD-LKA-----I--------.-S-..--.---------------A-----N--AG--ID-----PE--D----------NIN-I--L---.---................----Q--P--P-A-.-D--S---   117
D.UG.11.DEMD11UG003         ---------.----LKA-----F--------.-SS..--.---------------A-----RT-G---S-----P-----F-----------IN-S-N-----.-A-................---GH--S----P-.-D------   117
D.YE.02.02YE516             ----RE--.--D-LKAI----F---N---.-S-..--.---------------A---HL---G---S-L--G------------I---N-N-S----P-.---................-------S--PS--.--P--ED-   117
D.ZA.90.R1                  ---------.--N-LTAI-I--F--------.-S-..--.---------------A-----N--G--V-F--------------------------S--------.---................-------S----A-.--------   117
F1.AO.06.AO_06_ANG32        -------N.----L-A--I--I------Y.-K-..--.---------------A----VRE------ASC---P---------------HIN-S---EQG.AEE...............R------S---RAI.--------   117
F1.AR.02.ARE933             ---------.-T--LTAL-Y--I------Y.-K-..--.---------------A-----RA--D---SSC---PE--------------HIN-S----PG.TEE...............-L-G---AG--HA-.-----I-N-   117
F1.BR.10.10BR_RJ015         ---------.-TDIV-I-NY--I--------.-K-..--.---------------A-----RAL-----RFG---PE--------------HIN-S----QG.TKE...............EL-----SG-HH--.--Q-----   117
F1.CY.08.CY222              ---------.-T--L-A--T--I------Y.-K-..--.---------------A-----RA----L-LSC---PE--------S----HIN-SK--EQG.TEE...............---N---S---H--.-D------   117
F1.ES.02.ES_X845_4          ---------.--D-L-AI-T--I------Y.-K-..--.---------------A----VRE------SSC----P--------------HIN-S--GEQG.TEE...............R------S---RA-.--------   117
F1.RO.96.BCI_R07            ---------.-T--LKA-QY--I------Y.-K-..--.---------------T-----RE------SSC---P---------------HIN-S---EQG.AER.................-------S---HA-.--------   117
F1.RU.08.D88_845            ---------.-A--LKA-KY--I--------.---..--.---------------A-----RE--Q---DDC---P---------------HIN-S-G--QG.TEK................----S--P---S-I.-G------   117
F2.CM.02.02CM_0016BBY       ---------.--A-LTA--T--I------Y.-K-..--.---------------A-----RE------S-C------------------HI--S--S-QG.TER.................-------S---RS-.-G---EN-   117
F2.CM.10.DEMF210CM001       ---------.----LKA--Y--I------Y.-K-..--.---------------A-----RE------IACV--P-D------------HIN-S---DQG.AAT................---N---P--PHAA.-G-----E   117
F2.CM.10.DEMF210CM007       ---------.--D-LKA--Y--I------Y.-K-..--.---------------A----VRE------RAC---PE--------------HIN-S---EQG.TGK...............--E----S--PRAA.---------   117
G.BE.96.DRCBL               ------ST.----LTA--I--------S.-P-..--.---------------A-------------S-C---PE--------------H---S--G---.---QSQ.......GTEI-------F---SV-.-G------   124
G.CM.10.DEMG10CM008         ------ST.--A-LTA--I--I------H.-SA..--.------K-------A-----N-------SSC---P------V-------HI--S--S-NF.---QSP.......GTEP.......G---SV-.-G------   117
G.CN.08.GX_2084_08          ------ST.--Q-LQAI-I--I------Y.-P-..--.---T-K-------A-----R-------S-CV--LE-----P---I---H---S--SR-P.E--QSP.......GTET---G---S----V-.-G------   124
G.CU.99.Cu74                ------ST.----L-A--I--I------Y.-P-..--.------K-------A--K--------V-S-C----X--------------S---SK-S---.---QSQ.......GTET---R---S----G-.-G------   124
G.ES.09.X2634_2             ------NT.----L-IIQT--I------Y.-PS..--.------K-------A-------------L-S-CV---T-------------HI-----S--L.---QSQ.......GTET---G---P---SV-.-G------   124
G.GH.03.03GH175G            ------SN.----L-A--I--I------Y.-SA..--.------K-------A----VRAL---F-SAC---PT-------------C---S--S---.---QSQ.......GTET---R---S----V-.-G------   124
G.KE.09.DEMG09KE001         ------ST.-G--LQA--I--I------Y.-P-..--.------K-------A--------------SAC---P------P-------HIA-S--S--.---.............GTETEL-G---S----G-.-GA-A---   117
G.NG.09.09NG_SC62           ------TT.----LQA-KI--I------S.-P-..--.------K-------A------------IS-C---CT--------------H---S--SR-P.---QSP.......GTET---G---S----V-.-G-R----   124
G.ZA.01.TV546               ------S-.----L----I-XI------Y.-P-..--.S---Q--------A-----RA----V-SSCV--LE----F----------IN-S--I---.---QSQ.......GTET---G---SE--SGI.-GT-----   124
H.BE.93.VI991               -------N.--G-L-AC-I-R-------Y.-E-..A-.-----Q--------A--------G--V-A-C--GP-----------------------S------.-EK................------TSG---A-.-GT-A---   117
H.CF.90.056                 ------A-.-T--LQVCKI--I------C.-E-..T-.---------------A-----RE------TSC---PP---T-----------N-S------.-EK...............-E-----SL--S-I.-GT-----   117
H.GB.00.00GBAC4001          ---------.--G-LQVCKI--I------Y.-ER..AE.---------------A-------------H-C-AGPV-----------------------S------.-EK................-------S--P-A-.-GT-A---   117
J.CM.04.04CMU11421          ---------.--Q-LLA-XX--I--R---Y.-S-..R-.S-------------A--N-XD--------C-C-----------------I---HI--S----D-.---................---D---SG-PCM-.-GA-----   117
J.SE.93.SE9280_7887         ---------.-DQ-LLA-----I------Y.-K-..N-.S-------------A--N--D------PSSC---P-----------I---R---S----N-.---................---D---SG-PCM-.-GA-----   117
J.SE.94.SE9173_7022         -------N.-DQ-LLA--I--I------Y.SK-..N-.S-------------A--N--D-------SSC---P-----------I---R---S----N-.---................---D---SG-PCM-.-GA-----   117
K.CD.97.97ZR_EQTB11         ----R---.EQQ-LTP--I--I------Y.-K-..--.---------------A-----RE--Q-V-SSC----T-------------S-N-D--S-QG.TEG...............EL-----P--PD--.-G--D---   117
K.CM.96.96CM_MP535          ----R--P.--Q-LT---T--I--E---Y.--L.---.S--T-----------A--K--S-----L-SAC--------------I-K-N-N-D--P-KG.TEG...............-L-----S--PC--.---------   117
U.CA.01.TV749               -------N.--R-LSA--I--I--------.-K-..G-.S----K-------A--N--D-------SAC-E-----I-----------S---R---------.---QPQ.......GTEAR--R---SG--SV-.-G---ED-   124
U.CA.99.TV721               -------R.--Q-LLA--T--I------Y.-K-..R-.------K-------A--N--D---R---S-C----E--------------S---S---------.---QPQ.......GTEAR--R---S---SGI.-G---ED-   124
U.CD.83.83CD003_Z3          ------E-.----L-A--I--I--------.---..--.------K-------R--Q--R-------S-C---P----H---------N---SK.......................GTAT----T--PG--CA-.-G------   113
U.CD.90.90CD121E12          ------E-.----L-AI-I--I------Y.---..--.---------------K--Y--R-------SAC---PT-------------H---S........................GTET----T--SG--CA-.-------K-   113
U.CY.05.CY090               ---------.-A--L-VI-I--I------Y.-T-..--.---------------A-----RA------SSC----T-----P-----N-S-N-SK--E........Q.......GTTGEL-NS--SG--RV-.-G------   117
U.CY.08.CY223               ---------.--Q-LAA--I--II-----Y.-K-..G-.S----K-------A--N--D---R---SDC----E----F----I---HIN-S--S-N-.---................----T--PG--GA-.--PR-E--   117
U.ES.10.DEURF10DZ001        -------A.-QD-L-A--I--I------Y.-SS..--.------K-------A----VN--------SAC---P----------I-K-S---S-S-----.---QSQ.......GTET-----*-SL-YSVI.-G------   123
U.GR.99.99GR303             ---------.--Q-LT-I-I--I------Y.-K-..--.---------------A-------------L-C----E--------SI---H---S----N-GD-G..............#-AD---SGGTFLL#MG---A-E   117
U.NL.95.U_NL_95_H10986_D1   ----R--Q.--Q-LT---I--I------F.-KL..--.---------K---A-------------S-C-------------------------S--SREG.TEG...............EL-----P--PCA-.---------   117
01_AE.AF.07.569M            ------ST.----L-A--I--I------F.-SS..--.---------------A-----RA------S-C----T-------------H-------------.---QSQ.......GTET---RS--SG-PSVI.-GP---N-   124
01_AE.CF.90.90CF11697       ------NT.--D-LQA--I--I------Y.-P-..--.------K-------R-------L-----VACV---T-------------HI--S----------.---QSQ.......GTET---G---SG--SVI.-G----N-   124
01_AE.CN.10.YNFL03          ------ST.----L-A--I--I------Y.-SS..--T---T-K-------A-----RE------R-------------------H---S---------.---QSQ.......GTET---R---SG--SVI.-GP-AEN-   125
01_AE.HK.04.HK001           ------ST.----LGA--A--I------S.-PA..--.---T-K-------A-----RK------SSVV---T-------S----H-------------.---QSQ.......GTET---R---SG--SVI.-GA---N-   124
01_AE.IR.10.10IR.THR48F     ------ST.-D--L-AI-I--I------Y.-PA..--.---T-K-------A----VRAL-D---SDC----T--------I---S---S---------.---QSQ.......GTET-L-R---SG-PSGI.--P--EN-   124
01_AE.JP.x.JRC77AE          ------ST.----L-A--I--I------Y.-SS..--.---T------------A-----RA------SSCV---T------------------S---------.---QPQ.......GTET---R---PGK--DL.-GP---N-   124
01_AE.TH.09.AA090a_WG11     ------NT.------AI-I--I------F.-SA..--.---------------A------TL-----S-C---P-------------H--SS--S---.---................EA---R---SG--SV-.-GP--EN-   117
01_AE.TH.90.CM240           ------ST.----L-A--I--I------F.-SS..--.---T-----------A-----SA------S-C----T--------V---N---S--G---.---QSQ.......GTET---R---SG--SVI.-GP---N-   124
01_AE.US.05.306163_FL       ------ST.----L-A--I--I------Y.-SS..--.------K-------A----VRA------SACV---T--------I---H---S---------.---QSQ.......GTET---R---SG-PSSI.-GP--ED-   124
01_AE.VN.98.98VNND15        ------ST.----L-A--I-RI--E---Y.-SS..--.---T-K----KV-A-----RE------S-C---PT-------------H---S---------.---QSQ.......GTET---R---SG--SVI.--P---N-   124
```

Draft of 30 September 2015, not for public release
hiv1prot_main.tex r8876–10546

Rev start | exon 1 end | exon 2 start | NLS | Leu-rich effector domain | Rev end

```
B.FR.83.HXB2          MAGRSGDS.DEELIRTVRLIKLLYQSNPP.PNP..EG.TRQARRNRRRRWRERQRQIHSISERILGTYLGRSAEPVPLQLPPLERLTLDCNEDCGTS.GTQ.............GVGSPQILVESPTV.LESGTKE*  116
02_AG.CM.10.DE00210CM013  -------T.-----L---EI--I------Y.-P-..G-------K-------A--K--RT------SAF---P----HF----I---C----------.---QSQ.......GTET-------S---SVI.-G------  124
02_AG.CY.09.CY256         ---N---A.----L-A--I--I------Y.-P-..A-.------K-------A--K--RAL-----N-C----E----F----I---C--SS------.---QSQ.......GTETR------SG--SII.-G--A---  124
02_AG.ES.06.P1423         -------P.-A--L-V--I--I------Y.-P-..-.-------K-------A-----SE------SSC---PT-----L---I---S---C-GS-D-.---QSQ.......GTGT-------SG--SII.-G------  124
02_AG.GW.05.CC_0048       -------NS----L-A--T--I------L.-PS..--.S---K-------A-----GA------RAC----------------S-N----S--.---QSQ.......GTKT-------S---SVI.-G------  125
02_AG.KR.12.12MHR9        -------A.----L-VI-I--I------Y.-P-..-.-------K-------A-----SA------S----P---------I--S---SK-------.-RQSQ.......GPETR--G--TP---SSI.-G------  124
02_AG.LR.x.POC44951       ------NA.--G-LGVI-IVRI------Y.-P-..-.-------K-------A-----CA-I--V-S-C----E---------I---C---S-------.---QSQ.......GTET-------S---SII.-G------  124
02_AG.NG.09.09NG_SC61     -------A.----L-A--I--I------Y.-P-..-.-------K-------A--K--RA------S-C----------I---S---SK------.---LSQ.......GAGT-------S---SII.-------  124
02_AG.NG.x.IBNG           -------A.----L-A--I--I------Y.-P-..-----T-K-------A-----RA------S-C----------I---N---S-------.---LSQ.......GTET-------S---YII.-G------  124
02_AG.SE.94.SE7812        ------NA.----L-A--T--I------Y.-P-..-.-------K-------A-----RA------S-C----------I--R---S-------.---------------------------S---SI--G------  117
02_AG.SN.98.98SE_MP1211   -------A.--G-L-A--I-RI------Y.-P-..-.S-----------------A---VRA------S-C--P------P---I--C---S--S---.---QSQ.......GTET----S-TS---SVI.-G------  124
03_AB.RU.97.KAL153_2      ---------.----LK-I---F--------S-..-.-----------------H------Q--S-----PE---L-H----------S-----------.---------------------------D------  117
04_cpx.CY.94.94CY032_3    -------NI--D-FKAA-A--I------YPN---.T-----------------A--K----L-----A-----P----------K---N-S-------.DK.....................VS--L-A--.GT-A---  118
05_DF.BE.x.VI1310         -------R--D-LKA-----I------L.-S-..-.-----------------A-----N--G--L-S------E----------N-N-S-----------.---------------------------S--P-AI.--------  117
06_cpx.AU.96.BFP90        ---------.DR-LLA--I--I------Y.-K-..N-.S-----------------A--N--D------V-S-C----------I---R---S----N-.---------------------N---SG-PDML.-GT--T--  117
07_BC.CN.98.98CN009       -----E-..-A-LKA--T--I------Y.-E-..R-.-----K-------A-----------------SAC----------I---HISGS-SG-----.---QSQ.......GTTE-----*-SGK-CA--.G--A-KE  123
08_BC.CN.97.97CNGX_6F     ---------.--A-L-A--I--I------Y.-E-..R-.-----K-------A--N-L---------SAC---PT----------HITDS-RG-----.---QPQ.......GTTE-----*-SGK-CA--.G--A-KE  123
09_cpx.GH.96.96GH2911     -------N.--Q-LAA--V--I------Y.-K-..-.-----------------A--S--N-------RAC---PE----------H-N-S----N-.-----------------AD---SG--CAI.-G------  117
10_CD.TZ.96.96TZ_BF061    ---------.--A-LQA--I--I------.-S-..-.-----------------------GG---DA----PE----------N-N---------.H.---------------GH--S-----I.-D------  117
11_cpx.CM.95.95CM_1816    -------N.-DQ-LTA--I--II-----Y.-K-..Q-.S--S---------A--N--D---Q---SSC---P----------N------S-N-.-----------------DX--PG--Y---.-PR-E--  117
12_BF.AR.99.ARMA159       ---------.----LKAA-----------Y.-K-..-.-----------------A--S--RA------ID----PE----F-------G-----------.---------------------------------  117
13_cpx.CM.96.96CM_1849    ---------.--V-LT---I--I---N--Y.-SQ..-.-----K-------A-----------GQ---S-C----------H-----------.---QSQ.......GTET---G---P---GI.-G-----N-  124
14_BG.ES.05.X1870         ------ST.--D-L--AX-I---------Y.-P-..-.-----------------A-----------SACV---E----------H-G-S--S--V.---QSQ.......GTET---GS--S---SVI.-G------  124
15_01B.TH.99.99TH_MU2079  ------ST.----L-A--I--I-H----Y.-SA...-.S--T-------A--KR-RE------SPC----------------S-----------.---QSQ.......GTET---R---SG--SV--.GP--N-  124
16_A2D.KR.97.97KR004      -------P--D-L-A--A-RI------S.-D-..R-.S----K-------A-----RA------SAC--PT----------H---S-------.---QPQ.......GTET---R---S--PSV--.G---E--  124
17_BF.AR.99.ARMA038       ---------.--D-LKAI----F--------K-..-.-----------------A-----NK-G----A-----P------L----------SK----------.E----V---PS-I.--------  117
18_cpx.CU.99.CU76         ---------.--Q-LT--SI--I------Y.-E-..A-.-----------------A-----SA------N-C--GP----------------S-----F.---QSQ.......GTET---GS-VP---SG--.GT-D---  117
19_cpx.CU.99.CU7          ---------.--D-L-VI-I--I------Y.-T-..QR.A----K-------A-----------------S-C----E---F-------H-N-S-----------.R--G---S---SV--.G--A-N-  116
20_BG.CU.99.Cu103         ------ST.--D-L-VA-I--I------Y.-P-..-.-----------------A-----RAL-----T-C---PP----------------S--S-----.---QSQ.......GTET---R--VS----G--.G------  124
21_A2D.KE.99.KER2003      ----R---.-DD-LKAI----F--------S.-P-..-.-----------------A---------A---VSYH----E---------I---N-N-R---A-.-----------G-R---S----AI.-G---E--  124
22_01A1.CM.01.01CM_0001BBY  ------ST.----L-A--I--I------Y.-K-..R-.-----K-------A--K--D-L-----SAC----------------H---S-------.---QSQ.......GTET---R---S----DI.-GA---K-  124
23_BG.CU.03.CB118         ------ST.----L-----I--F------Y.-P-..-.-----------------A--------G---ADC---PP----------------S--I-----.---QSQ.......GTETR--G---S----A--.G------  124
24_BG.ES.08.X2456_2       ------ST.--D-L-A--I--I------Y.-P-..-.-----------------A--------Q---NSC---P----------------S-.-----------.---------------GTET---GS-VS---SV--.G------  113
25_cpx.CM.02.1918LE       ------VN.----LQAA-I--I------S.-P-..-.-----------------A-----TL------ASC---LE----------H---S-------.---LSQ.......GTET---R--VS----A--.GT------  124
26_AU.CD.02.02CD_MBTB047  -------P--A-LGAI-I--I------L.-K-..R-.S-----------------A--NK-D-------S-C--GPP---------S---S--F-N-.---QPQ.......GTET---R---S----GL.-GA--.E--  124
27_cpx.CM.04.04CD_FR_KZS  ----NE----.-D-LTAI-I--I--K---Y.-K-..R-.-----K-------A--E--D---K-V-S-C---P----P-----C---S-------.---QPQ.......GAET---R---S----SA--.GT-A---  124
28_BF.BR.99.BREPM12609    ---------.--D-LK-I-I--F------.-S-..-.-----------------------L-RW--SNH---PE----F---I---------.---------------------A--.G------  117
29_BF.BR.01.BREPM16704    ---------.-----K-----F--------.SS..-.-----------------------RW--SNH---PT-------I-------S-----------.---------------------AI.--------  117
31_BC.BR.04.04BR142       ---------.-A-LQA--I--I------Y.-----.-----------------A-----RA------S-C--P-----FL---I---NIN-S-SG-----.HPQ.....GTAE---N-*V-KKPCA--.G-RA-KE  123
32_06A1.EE.01.EE0369      -------N.--Q-LLA--I--I-N----Y.-K-..N-.S-----------------A--H--D---H---S-C--P----------I---R---S----N-.-----------------D---PEKPGVL.-GT-----  123
33_01B.ID.07.JKT189_C     ------ST.-DD-L-A--I--I-X----Y.-SA...-.-T-K-------A-----GA------SAC--P--S----------S---S-------.---QSQ.......GTET---R---SG--SVI.-D----N-  124
34_01B.TH.99.OUR2478P     ------ST.----L-A--T--I------Y.-SS..-.-----K-------A--K--RE--D---ISC----T----------------SK----------.---------------------V-----A--.D------  124
35_AD.AF.07.169H          ---------.-LKGI-I--I------Y.-E-..R-.S--K-------A--N--D---Q---S-C----P----------------Q---Q-------.---QSQ.......GVET---R--VF----V--.G----N-  124
36_cpx.CM.00.00CMNYU830   -------A.----L-V--T--I------Y.-PS...R----K-Q-----A-----RA------S-C--FE----------H---S-------.---QPQ.......GTEA---G--V-----GI.-GT---N-  125
37_cpx.CM.00.00CMNYU926   ---------.--Q-L-A--I-NI-----Y.-P-..-.-----K-------A-----RAL-----S-C----------------H-N-S---RD----.---QPQ.......GTET-L-G---S--P-G--.G----N-  124
38_BF1.UY.03.UY03_3389    ------G-.-----LKAI-Y--I------F.-K-..-.-----------------A-----RAL-----SSC---PE----------HIN-S----QG.TE.----------------N-PV---P-AI.-D------  117
39_BF.BR.04.04BRRJ179     ---------.----LK---Y-I------Y.-K-..T-.-----------------A-----RE----F-SAC---RE----------HIS-S---EQG.AEE.----------------TSG--LA--.--------  117
40_BF.BR.05.05BRRJ055     ---------.--D-L-A--A--V--#-----.SS..-.-----------------------RK---W-ISNH---PT----F----------S-----------.---------------------V-----E--.--------  116
42_BF.LU.06.luBF_18_06    ---------.----T-LQ-----R--------ST..-.-----------------------QTF-GW--N---PE----------------S-----------.---------------------A---.--T-----  117
43_02G.SA.03.JL1223       -------A.----L-X--I--I------Y.-S-..AR.-----------------A-----------SAC---R----------H---S--SR-P.---QSP.....GTET--G---P---V-.-G------  124
44_BF.CL.00.CH80          ---------.-VD-LKVI-Y--I------Y.-K-..-.-----------------A-----RAL---SSC---LE----------HIN-S----QG.TEG.----------------SG---A--.G------  117
45_cpx.FR.04.04FR_AUK     ------NT.----LTA--I--I------Y.-K-..-.-----------------A-----------S-C---P----------H---S-----D-.---------------AD--TSG--CAI.-G------  117
46_BF.BR.07.07BR_FPS625   ---------.--D-LKA--Y--I------Y.-K-..R-.----Q-------A----VRG------RSC---PE----------HIN-S----QG.TEE.----------------TSG--RV--.G------  117
47_BF.ES.08.P1942         ---------.--X-LKA----I--------SS..-.-----------------------R--G----N---D-PT----------S-----------.---------------T-----P--.--------  117
48_01B.MY.07.07MYKT021    ------ST.----LTA--I--I------S.-SA...-.-T-K-------A----VRA------S-C--P----------------S-----P.---QSH.....GTET---R---SG--SDI.-GA---N-  124
49_cpx.GM.03.N26677       -------N.--Q-LLA--I--I--K---Y.-K-..N-.S-----------------A--D--D---Q---IDC---P----------S---C--SS----N-.-----------------D--VSG--CPI.-GE-----  117
50_A1D.GB.10.12792        ------S-.--D-LKAI----F--------.-----.-----K-------A-----N--G----D-----PE--G----------N-N-S----N-.-----------------N-V-----AI.-D------  117
51_01B.SG.11.11SG_HM021   ---------.-----K-----------------.-S-----------H-RVV-DW--S------T----------------S-G-----------.---------------V-----P--.--------  117
52_01B.MY.03.03MYKL018_1  ------ST.----L-A--I--I------F.-SS...-.----K-------A-----RA------C-CV--PT-----P----------SK----------.---------------G---P------A--.D-----  117
53_01B.MY.11.11FIR164     -------T.----L-A--I--I------Y.-SX...T--T-K-------A-----XRA------SXCV---T---------I---N---S-G-----.---QSQ.......GTET---R---SG-PSVI.-G---XN-  125
54_01B.MY.09.09MYSB023    ------NT.--D-L-A--I--I------Y.-P-..-.-----------------A-----VRE------RD---G-T---------I---H---S-------.---QSQ.......GTET-L-G---SG--SG--.G--X-N-  124
55_01B.CN.10.HNCS102056   ------ST.--IL-A--I--I------F.-SS...-.----K-------A-----RA------AACV---T----P-------I--HI--SK--E-----.---T.....GTET---R---SG--SVI.-G---EN-  120
56_cpx.FR.10.URF5_patient_A  -----VT.----LQA--X--I--Y--Y.-PS..T-.----K-------A-----RE------S-C----T---H-P---I--C-----GG-.-----.---LSQ.......GTETR-----TS---IL-.G----E  124
57_BC.CN.09.09YNLX19s9    -------T.----A-L-A--I--I------C.-P-..R-.-----K-------A--N--D---A-S-C--P----F------G-HISDS-SG-A-.--HQPQ.....GTTE-----*-SGKYCA--.---E--KE  123
58_01B.MY.09.09MYPR37     ------ST.----L-A--I--I------Y.-P-..-.-----T-------A-----RAL-----S-C----P----------H-N-S------.---HQSQ.....GTET---R---SG--SVI.-GQ--.N-  124
59_01B.CN.09.09LNA423     ------ST.----LKA--I--I------L.-SS...-.-----------------A-----TL---V-S-CV--PT----------H-GDS-----P.---QSQ.....GTET---R---SG--SGI.--P--N-  124
60_BC.IT.11.BAV499        ---------.--A-LQAI-I--I------Y.-R-..-.---T*---K--*-A-----------------S-C---P----------I---NINHS-SS---.R--QPQ.....RTAEK--N-*VFRKPCA--.--------  124
61_BC.CN.10.JL100010      ---------.----L-A--I--V------Y.-E-..R-.----QK------A--N--R-------SSC----------I-G-HISDS-SG-----.---QSQ.....GTTE----S*-SGKPCA--.G--A-KE  123
62_BC.CN.10.YNFL13        ---------.----L-A--I--I------Y.-KS..T-.-----------------A-------L-----SAC---P----------I---HISGS-SGE----.---QSQ.....GTTE-----*-SGKHCA--.G----KE  124
63_02A1.RU.10.10RU6637    -------V.----L-V--T--I------Y.-P-..-.-----------------A-----SA------SSF----------I---S---S---AA-.---QSQ.....GTAT-------S---SSL.-G------  124
64_BC.CN.09.YNFL31        ---------.----L-V--I--I------Y.-K-..R-.----K-------A--E--N-------R-C---PE--S--------HISDS-SG--P.---QSQ.....GTTD--E--*-SGK-CA--.G--A--E  123
65_cpx.CN.10.YNFL01       ---------.----L-A--T--I------F.-SS...T--T-T-------A------Q-D---RAC---PE----------I---YISD--SG----.---QSQ.....GAET---R---XG-PSV--.GP---N-  125
67_01B.CN.11.2011.ANHUI.HF115  -----NT.--D-LAA--I--I------X.-SA...-.-----------------A-----RA------XPF----E----------H-----------.---QSQ.....GTET---R---SG--SVI.-G---EN-  124
68_01B.CN.11.2011.ANHUI.WH73  -----NT.--D-LAA--A--I------F.-SS...-.-----------------A-----KE------A-X----T---------D--S-----------.---QSQ.....GTET---G---SGK-SVI.-G---EN-  124
70_BF1.BR.10.10BR_PE004   ---------.----LTA--I--F------L.-S-..-.-----------------------RT--DW--SNH---PT----F----I---S-----------.---------------------------P-----  117
71_BF1.BR.10.10BR_PE008   ---------.----L-A--------------SS..-.--R----------KH-WA--GW--SNH------E--H--Q---------S-----------.---------------------------S-I.--------  117
72_BF1.BR.10.10BR_MG002   --------.E-N-LK-----------------SS..-.-----------------RN--GW--N---D-PE----------------S-----------.---------------D---R--HT.--------  117
O.BE.87.ANT70             -----E-D...Q-LQAIQI--I------Q.-S-..R-.S-N--K-------R--A-VDTLAA-V-A-VVHGPQNNNIVD-----Q-SIRDP-GDQL-.EAWTVD.....PRAEDN*CL-N-CSCN-I.-ATRIA--  121
O.CM.98.98CMA104          -----D-D...Q-LQAI-I--I------H.-S-..I-.S-S----------R--A-VDT-AA---A-VVHGPQNNNLVD-----Q-SIRDP-SDQP-.-NWTVD.....PRAEDN*WL-T-*SCN-I.-ATRIA--  120
O.CM.98.98CMABB141        -----D-D.VQP-LQA--I--I------H.-S-..I-.S-N--K-------R--A-D---AA-V-A-VVHGPEDNNLVE-----Q-SIRDP-GDQP-.E-WTVD.....SRAEDN*CL-N-CSCN-I.-ATRIA--  123
O.CM.98.98CMABB212        -----EED..QQ-LQAIQI--I------H.-TQ..A-.S-S--K-------R--T-VDA-AS---A-VVHG-QDNNLVD-----Q-SIRDP-ANQLP.E-WTVD.....PGTKDN*-L-T-WSCN-I.-ATRIA--  122
O.CM.98.98CMU5337         -----T-E.---L-VI-I-NI------H.-S-..I-.S-N---------R--A-VNTLAS---A-VVHGPQDNNLVE-----Q-SIRDPDGDQP-.--WTED.....PRAEDN*CL-N-YSCNSI.-ATRIA--  121
O.CM.99.99CMU4122         -----D-V...P-LH-I-I--I------Q.-S-..I-.S-S----------R--A--DTLAA-F-ASVVHGPQDNNLVE-----Q-SIRDPDGDQP-.--GTVD.....PRTEDN*CL-N-CSCS-I.-ATRIA--  121
O.FR.92.VAU               -----EED..QQ-LQAIQI--I------H.-TT..S-.S-N----------R--A-VD--AA-V-ASVVHGPQDNNLVD-----Q-S-RDP-VDQ-P.--WTVD.....HRAEDN*-L-T-*SYN-I.-ATRIA--  121
O.GA.11.11Gab6352         -----DAD..QQ-VQA--I--------Q.-S-..I-.S-N--K-------R--V--D-LAT---A-VVHGPQDNSCVE--S--H-SIGDS-GDQP-.E-WTVD.....PRREDN*CL-N-YSCS-I.-VTRIAK-  123
O.SN.99.99SE_MP1299       -----DGD..QP-L-AIQI--I-----SH.TS...T-.S-S----------T-HAHVDTLAA---A-VVHGPQDNNLVE-----Q-SIRDPDGDQP-.---WTVD.....SGTEDN*CL-T-HSCN-I.-ATRVA--  122
O.US.10.LTNP              -----D-D...R-LQAIQI--I------Y.-S-..S-.S-N----------R--A-VET-AA-V---VVHGPQNNNIVE-----Q-SIRDS-GDQP-.E-WTVD.....LGTETN*CL-N-YSCN-I.-ATRIA--  121
N.CM.02.DJO0131           ------VN.----L-A--I--I------Y.--NNNQ-.-----------------T-----RE--Q---D-C--GLQ---D-P----D-----PA--S-AP.--EPHQ.....GTATTE*TQNT--GNNSI.-GKRV-G-  125
N.CM.02.SJGddd            ------VN.----L-A--I-RI-----Y.--N..K-.X----------------A-----RA------SSC--GPP---D-P----D-----PK--S--P.E-ESQ.....GTATTE*THNT--GNTSI.-GRRX-R-  123
N.CM.04.04CM_1015_04      ------VN.--K-LTA--I--I------Y.--N..Q-.-----------------A-----CA------S-C--GLQ--GDPP---------NT--S-DP.--EPQQ.....GTATIE*TQKT--GNTSI.-GKRV-R-  123
N.CM.06.U14296            ------VN.----LGA--A--I------Y.--N..K-.-----------------A-----RA------L----GPQ--GD-P----D----PE--S----.--ESQQ.....GTATTE*TQNT--GNNSI.-GKRV-R-  123
N.CM.06.U14842            ------VD.--G-L-A--I-RI-----Y.--N..T-.-----------------A-----RA------S-C--GPQ---D-P----D----ST--S----.--ESQQ.....GTATTE*TQDT--GNNSI.-GXRV-R-  123
N.FR.11.N1_FR_2011        X-----VN.----L-A--I-RI-----Y.--S..T-.S----------------A-----RA------S-C--GPQ---D-P----D----TK--S--P.--ESQQ.....GTATTE*IQN---GNTSI.-GK-V-R-  123
P.CM.06.U14788            -----DED.LR---T-I-I--I------S.-AR..GP.S-NST---------S-----DQ-AG--IASH---PE-LGGAD--DISQ-HIEDQGPPDN-.VDTPGP.....TTEQ..*L-AR-LCCVWL.-D-RI-TX  121
P.FR.09.RBF168            -----DED.LR---T-I-I--I------W.-ER..GS.S-NST--K-----S-----DQ-AG---ASR---PE--GGAD--D-S--HIGDQGPPNNP.VDTPGP.....TADQRA*L-AR-LCCVWL.-D-RI-TK  123
CPZ.CD.06.BF1167          ----EEE....N-LQAIKI--I--E---Y.--S.....-S---------I----VEG-S---A-----LPTAP-VD--D-SK---APLDIDQVD.PSVEPIAVGDCETIAAQAHTP-DTSGPTN*TA  127
CPZ.CM.05.SIVcpzMT145     -----EG....N-L-AI-I-RI--E---Y.-SQ..S-..-R-----K---------NQ-G--V-ASC----S--PQ-P---I---CISDL-TT-DT.--IQPQ.....GSTTNRY--TPTERD-A--.DNRNQ--  123
CPZ.GA.88.GAB1            -----EPQD-AR-LQA-KI--I------Y.-S-..-.-----K----------A--K--SE--G-V-A-----PPK-GD-E--E-DK-S-Q-V-TTQDV.---SNTS.....QPQTAT-ETVPAGGNYSI.-GK-A-N-  125
CPZ.TZ.06.SIVcpzTAN13     ----EE-...AN-LE-I----I--D---Y.-SG..A-.S-A---------Q--D-VDALAQ---QHR--GPP--P--D--D-SQ-R-AAGDHTPSP.E-GDH*LHTDHNNST*KLSA*VETI*RLACQ.CLCDC-DF  124
CPZ.US.85.US_Marilyn      ------GDA--Q-L-A--I--I--A---F.--N..S-----------KN-----NE--Q---S-C----S--PD-P---IGG--INSE--STNP.--EHQQ.....GTATVE*TQST-LGGNRL.-GK-A-S-  124
GOR.CM.07.SIVgor2139_287  -----DED.LQ--LTKI-I-R-------W.-E-..GP.S-NST------KG-----DQ-AG---TSH---YQN-GCVD--DISH--IGDQG-NTQP.DRVPDQ.....TTEKSE*FDKR-LCFYCL.-GTR-*AK  122
GOR.CM.07.SIVgorCP2135con -----DED.LQ--LTKI-I-R-------C.-E-..GP.S-NST------KG-----DQ-AG---AAH----QN-DCVD--DISH--IGDQGNNTQP.DRAPDQ.....TTEKSE*FDKR-LCFYCL.-GTR-*AK  122
```

Draft of 30 September 2015, not for public release
hiv1prot_main.tex r8876–10546

```
                                                                                    phos phos
                                         transmembrane domain                    cytoplasmic domain
                             Vpu start                                       α-helix          α-helix                    Vpu end
B.FR.83.HXB2                 TQPIP.....IVAIVALVVAIIIAIVVWSI.........VIIEYR.......KILRQRKIDRLIDRLIERAEDSGNESEGEISAL.VEM....GV.EMGHH...APWD.VDDL...*   82
A1.AU.03.PS1044_Day0         MT-LE.....-WS--G----L-------T-.........-G---K.......-L-K--------E-IR------------DTDE-.AL.....LI.---NY...DLGN.DIN-...-   82
A1.CH.03.HIV_CH_BID_V3538    MS-LE.....-Y--AG-I--L-------T-.........-G---K.......RL-K--R-----E-IR----------D-DAEE-.ST.....L-.-V-DR...-LLV.-NN-...-   82
A1.CY.08.CY236               MT-LE.....-S---G---VS-------T-.........-G--IK.......-V---------K-IR------------DTDE-.AL.....L-.---NY...DLG-.DNN-...-   82
A1.ES.05.X1608_8             MS-LE.....-S---G---TL-------TL.........-A-----.......-V-K----E---E-IR------------DTDE-.AA.....LI.---DY...DLGN.DNN-...-   82
A1.KE.11.DEMA11KE001         MNSLT.PWN.-C---G-I--L-L-----T-.........-A-----.......---K--------E-IK--E-------D-DTDE-.AA.....LC.---NY...GLG-..NN-...-   84
A1.RU.11.11RU6950            MT-LE.....-Y--A----VF-------T-.........-G-----.......RL-K--------E-IR------------DEEE-.ST.....LM.---NL...-LL-.DNNV...-   82
A1.RW.11.DEMA111RW002        M-LLV.....VW--IG-I--L-L--I--T-.........-GL--K.......-L-K-------LE-IR----------D-DTEE-.WA.....T-.---NY...DLGN.DNN-...-   82
A1.SN.01.DDI579              MTSLE.....-W--IG-I--L-L-----T-.........-G---K.......-V-K-KR-----E-IR----------D-DTEE-.ST.....L-.-V-DD...I-LY.DNN-...-   82
A1.UG.11.DEMA110UG001        M-LLE.....-W--AG---L-L--I--T-.........-G---K.......RLQ--K------E-IR----------D-DTDD-.AK.....L-.---NY...DFG-.-NN-...-   82
A1.ZA.04.04ZASK162B1         MT-LE.....-C--AG----L-------T-.........-G-----.......-L---------K-IT----------D-DTEE-.ST.....L-.D--NY...DLG-.DNNW...-   82
A2.CM.01.01CM_1445MV         ML-LA.....-LS--G-I--F-L-----T-.........-F-Q-K.......ELKK-----W--K-IS----------D-DTEE-.ST.....L-.---NL...DF--.ANN-...-   82
A2.CY.94.94CY017_41          ML-LV.....-L---G-I--L-L-----T-.........-F---K.......--KK-----W--K-IS----------D-DTEE-.SA.....L-.-R--L...DFG-.-NNV...-   82
B.BR.10.10BR_MG029           M-TLI.....-AS------VA----A---V.........-L-----.......---------------IR------------DQEE-.S-.....XL.----L.......-N--...-   82
B.CA.07.502_1191_03          M--LA.....-L--A-----A-L-----T-.........-G-----.......--------N--L--IA------------DQEE-.SALVERGHL.-----.......-N--...-   87
B.CH.08.M2_0803101_NFLG8     MT-LY.....-WS-------A-------T-.........AL-----.......-----------IL--IR------------DQEE-.SA.....L-.---D..............-   82
B.CN.10.DEMB10CN002          MFSLQ.....----------A-L-----T-.........-Y-Q---.......---K----------I---T---------DQEE-.SA.....L-.D---D......N.-N--...-   82
B.ES.10.DEMB10ES002          MTTLV.....-CS-A----VA----I--T-.........-L-----.......------R---I---IR------------DQEE-.SKL....V-.D---L.......I---...-   83
B.FR.11.DEMB11FR001          M-SLT.....-L----------L-----T-.........-F-----.......---------S----IR-----------EDQ#...#A.....LM.----L.......I---...-   78
B.GB.05.MM45d213_GN1         M--LI.....VI-------VG----------.........-L-----.......---------------IR------------DHEE-.AA.....L-.-R--L.......IN--...-   82
B.HK.06.HK003                M-SLE.....-L-V------A-L--I-----.........-F-----.......R-R------------IR------------DQED-.SA.....LL.---N-........I---...-   82
B.HT.05.05HT_129389          M-SLX.....XL--------X-X--I--X-.........-F-----.......X-K--KR-------IRD------------DQEE-.SA.....L-.D---L.......-N--...-   82
B.JP.12.DEMB12JP001          M--LQ.....-L--------A-L-----T-.........-G-----.......---K--------N-IS------------DQEE-.SA.....LM.--------.......-N-M...-   82
B.KR.07.HP_18_07JHS10_3909   M-SLT.....-FT-------A----------.........-F---K.......---K----------IR------------DQEE-.SA.....L-......................-   82
B.PE.07.502_0525_wg5         M-SLV.....-L--------A-L-----T-.........-FL----.......--------------IA------------DHEE-.SA.....LM.---YN.......N.---M...-   82
B.RU.11.11RU21n              M--LV.....-AT---F--VG--------L.........--V----.......------------A-I----------D-DQ#...#A.....LM.----L..........---M...-   78
B.TH.08.MERLBDTRC10          M-ALY.....-L------I-A-------T-.........-------.......-----K--------IR------------DQEE-.SAL....V-.----D..................-   83
B.US.11.ES38                 MPSLY.....--S-------A-L--I-----.........-F-----.......---K-Q-------IR----------D-DQEE-.SA.....L-.......................-   82
C.AR.01.ARG4006              #....LXXDYRLG-G--I--L----I--T-.........AY-----.......-LV--KR--W--E-IR------------DTEE-.ET.....M-.D---L...RLLH.-X--...-   82
C.BR.07.DEMC07BR003          MLTL-DRVDYRLG-G--I--L----I--T-.........AY-----.......-LV---R-----E-IR----------D-DTEE-.ET.....L-.D--NL...RLL-.-HA-...-   87
C.BW.00.00BW5031_1           MFALF.EVDYRLT-G-FI--LFL-----T-.........AYL----.......-LV---R--Q-VK-IR----------D-DTEE-.ST.....M-.D---L...RLL-.-N--...-   86
C.CN.10.YNFL19               MVD....LDYK-TVG-FI--L-L--I--T-.........AY-----.......-LTK-----W--K-IG------------DTEE-.ST.....M-.D--RL...RLL-.-N--...-   83
C.CY.09.CY260                MTGLIEKVDYKI-V--FII-LV------I-.........-YL----.......-LV--------E-IR----------D-DHEE-.SK.....M-.D--NL...RLL-.AH--...-   87
C.ES.08.X2363_2              MLDLLAKVDYRI-V--FLI-L-L-----I-.........AY-----.......-LVK-TR--Q--K-IR------------DTEE-.AT.....L-.D---L...RLL-.-N--XCVG   90
C.ET.02.02ET_288             MFSFTERVDAR-G-G--I--L-L--I--T-.........AY-----.......-L---------E-IR------------DTEE-.ST.....M-.D--TL...RLL-.-H--...-   87
C.IN.09.T125_2139            MVE....LNYKL-VG--I--L-L-----I-.........-Y-----.......-LVK-----W-VK-IR----------D-DTEE-.ST.....M-.D---L...RLL-.-N--...-   83
C.KE.00.KER2010              MLDLLARVDYRLGVG--I--L----I--T-.........-Y-----.......-L-------W--E-IR------------DTEE-.AT.....M-.D---L...RLL-.-N--...-   87
C.MW.09.703010256_CH256.w96  MLDLVARVDYRIGVG--I--L-------IL.........AY---K.......RW-K-K---W--K-IR------------DTEE-.ST.....M-.D-E-L...GLL-.-NG....-   86
C.TZ.08.707010457_CH457.w8   MLDLLARVDYRIGVA--LI-L-------T-.........-Y---K.......-L---KR--W--K-IR----------D-D-EEV.ET.....L-.D--NL...RLL-.-IN*...-   86
C.YE.02.02YE511              MLDLLAKVDYRIG-A-F---L----I--T-.........AY-----.......-LVI-----W-VK-IR----------D-DTEE-.ST.....M-.D---L...RLL-.IN--...-   87
C.ZA.08.705010534_CH534.w12  MLN....LDYR-G-G--I--LN------I-.........-Y-----.......-LV---R--W-VK-IR----------D-DTEE-.ST.....M-.D---L...RLL-.-I--...-   83
C.ZA.10.DEMC10ZA001          MLSLLEKVDYRI-VA-FTI-L-L-----T-.........AY-----.......-VV--T------K-IR----------D-DQED-.ST.....M-.D---L...RLL-.......-   85
C.ZM.11.DEMC11ZM006          MISLIERVDYRLGVG--II-L-------T-.........AYL----.......-L---K---W--E-VR------------D-EE-.ST.....M-.D---L...RLL-.GNE-...-   87
D.CM.10.DEMD10CM009          M--LV.....-IS-----TL-------T-.........-Y-----.......--Q--K---W---IR--E---------DKEE-.ST.....L-.D-----.....IVA--...-   83
D.CY.06.CY163                M-SLE.....-L--A-----L-L-----T-.........-Y-----.......--K-----W---IG------------DKEE-.SQ.....LM.-----.....NVA--I...-   83
D.KE.11.DEMD11KE003          M-SLQ.....-A--------L-L-----T-.........-F--C-.......RLR--#---W---IR------------DEEE-.SA.....F-.G--Y.........IN-I...-   81
D.KR.04.04KBH8               M--VV......I--X----TL----I--T-.........-F--CC.......RLRK--------N-IR------------DEEE-.ST.....LM.-----.....N------...-   81
D.SN.90.SE365                M-SLE.....-L--A-I---L-L-----T-.........-F--V-.......--R-----W---IK------------DEKE-.SA.....L-.D-----.....N------...-   82
D.TZ.01.A280                 MT-LE.....-T--A---I-S-L-----T-.........-Y-----.......R-KK-K---W---IR----------D-DTEE-.ST.....L-.--------.....N.---I...-   82
D.UG.10.DEMD10UG004          M--LV.....-L--------LLL-----T-.........-F--C-.......-LR--KR--W---IR------------DEEE-.SK.....FM.-----.....N.I--M...-   82
D.UG.11.DEMD11UG003          MNLLV......IG---------L-----A-.........-F--W-.......RLK------KI-N-IA------------DREE-.SA.....L-.G---L........IG-I...-   81
D.YE.02.02YE516              M-TLE.....-LS-----I-A----I--T-.........-Y-----.......--R-----Q---IR------------DEEE-.ST.....LM.----A.....NVA----...-   83
D.ZA.90.R1                   M--LV.....-L--------L-L-----T-.........-F-----.......R-RK-----Y---IR------------DGEE-.SK.....L-.---YD.....N------...-   82
F1.AO.06.AO_06_ANG32         MSDLL.....AI--------L-------T-.........AYL----.......-VVK-KR-N--YE-IR----------D-DAEE-.A-.....LG.-V-PF...I-GG.INN-...-   82
F1.AR.02.ARE933              MSYLL.....AIG-T--I--L-------T-.........AY---K.......-LV---R-N--YK-IR------------DAEE-.AA.....LG.---PF...I-R-.I-N-...-   82
F1.BR.10.10BR_RJ015          MTDLL.....VIS-A--I--L--V----T-.........-YL----.......-LV-----N--YQ--R------------DAEE-.AA.....LG.---PF...L-G-.INN-...-   82
F1.CY.08.CY222               MTNLL.....AI--A--IL-L-V-----T-.........AY---K.......-VI-----N--YKKIR------------DAEE-.AA.....LG.-V-PF...I-G-.INN-...-   82
F1.ES.02.ES_X845_4           MSNLL.....AIS-------L--V-I--T-.........-FL----.......-L------N--YE--R------------DAEE-.AA.....LG.-V-PF...I-GN.INN-...-   82
F1.RO.96.BCI_R07             MSHVL.....AI-----I--L----I--T-.........AY-----.......-LV-----N--YE-IR------------DAEE-.AA.....LG.---PF...I-G-.INN-...-   82
F1.RU.08.D88_845             MTE-I.....-IG----I--L-V-----T-.........-F-----.......---------N--YE-IR----------D-DAEE-.AA.....LG.---PF...V-G-.-NN-...-   82
F2.CM.02.02CM_0016BBY        MSYLI.....-LV---FI--L-A--I--T-.........-Y---K.......-Q---KR-N--YE-IR------------DAEE-.AA.....LG.-V-LF...I-GN.INN-...-   82
F2.CM.10.DEMF210CM001        MPYLI......IL---FIAVL----I--T-.........-Y-----.......----K---NK-YK-IR---------N--DAEE-.AA.....LG.-V-PF...I-G-.INN-...-   81
F2.CM.10.DEMF210CM007        MSLSI.....--V-A-FI-VL----I--T-.........-----K.......-L---KR-NK-YE-IR------------DAEE-.AA.....LG.---PF...I-GN.INN-...-   82
G.BE.96.DRCBL                M--LE.....-S---G-I--S-A-----T-.........-F-----.......--RK-KR-EK-L--IR------------DTEE-.AT.....LM.-L-DF...D--V.G-N-...-   82
G.CM.10.DEMG10CM008          M-ALV.....-A---G-I--FLA--T--T-.........-F-Q---.......E-RK-N---K-L--IR--I---------DTEE-.AA.....L-.DV-DF...DL-I.GNN-...-   82
G.CN.08.GX_2084_08           M--LE.....-A---G-I--F-A-----T-.........-Y-----.......--R--KR--KIL--IR------------DTEE-.AT.....L-.D-VEF...D--V.G-NM...-   82
G.CU.99.Cu74                 MKSLE.....-S---G-II-G-A-----T-.........-F-Q---.......E-RK-K----ILE-IR------------DTEE-.SA.....L-.---DF...DH-V.G-N-...-   82
G.ES.09.X2634_2              M--LE.....-A---G-I--GLG--I--T-.........-F-L-K.......E-RK-QR-Q-IL--IG------------DTEE-.AK.....LM.---DF...D--V.G-N-...-   82
G.GH.03.03GH175G             M-SLE.....-S---G-I--S-A--I----.........-F-----.......R-RK-K--EK----IR------------DTDE-.AA.....L-.---DF...D--V.G-N-...-   82
G.KE.09.DEMG09KE001          M--LE.....-AS--G-I--FLA-TI--T-.........-F-Q---.......E-RK-------Y--IR------------DEEE-.AT.....L-.GR-AF...D--V.G-N-...-   82
G.NG.09.09NG_SC62            M-SLE.....-A---G-I--F-A--I--T-.........-F-----.......--RK-KR--S-LN-IR------------DTEE-.AT.....L-.D-VDF...D--V.GNN-...-   82
G.ZA.01.TV546                M--LE.....-S---G-I--F-A-----T-.........-LL----.......--RKE---EK-L--IR------------DTEE-.AT.....LM.---XF...D--V.G-T-...-   82
H.BE.93.VI991                MNILG......IG-G-----F-------T-.........AY-----.......-L#K--------E-IR----------D-DTEE-.SK.....L-.----L...NLGY.-A--...-   80
H.CF.90.056                  MYILG......LG-G----TF---VI--T-.........-Y---K.......-LV--K------E-IG----------D-DTEE-.SK.....LM.----L...NLGY.-A--...-   81
H.GB.00.00GBAC4001           MYILG......IG-G--I--G-LC-I--T-.........-YL----.......-LVQ-KR----LE-IR----------D-DTEE-.ST.....L-.-R--L...NLGY.-A-V...-   81
J.CM.04.04CMU11421           MNSLQ.....-AS-------FFL-VXA-XV.........AY-----.......-LV---R--K-X--IR--E-----D-D-DTEE-.AK.....L-.---P-...DL-N.------...-   82
J.SE.93.SE9280_7887          MI-LQ.....-A----FI---FL--GM-T-.........-Y---K.......-L-------K---IR----------D-DTEE-.AD.....L-.-R-P-...DL-N.-N--...-   82
J.SE.94.SE9173_7022          MVSLQ.....-------I--FFL--C--T-.........-Y---K.......-L-------K--N-IR----------D-DTDE-.A-.....L-.---P-...DL-N.-N--...-   82
K.CD.97.97ZR_EQTB11          MV-LT......-G-I---A-L-L--I--T-.........AYL----.......-VV--KR-NW-F--IR------------DTEE-.AA.....LG.-T--L...ILG-.INN-...-   81
K.CM.96.96CM_MP535           MVSLA......IS-------L-L--I--T-.........-Y-----.......-LVK-KR-NW----IR------------DAEE-.AD.....IG.-L--L...ILGN.I-N-...-   81
U.CA.01.TV749                MYTLQ.....-------I--TL---C--T-.........-YV----.......-L---K--------IT----------D-DTEE-.SK.....L-.---P-...DL-N.-N--...-   82
U.CA.99.TV721                M-TLQ.....--S----I--AL---C--T-.........-YV----.......RL---K---K----IR----------D-DEEE-.SA.....L-.---P-...DL-N.-N--QD.-   84
U.CD.83.83CD003_Z3           MSILQ.....------II--L-L-----T-.........-Y---K.......RL-------W----IR------------DTEE-.ST.....L-.--EPD...NFRN.DN-M...-   82
U.CD.90.90CD121E12           MWIL-......-G---II--LGL-----T-.........-YL--K.......-L---QR--W----IR------------DTEE-.SL.....L-.--EPD...NFRN.DNNM...-   81
U.CY.05.CY090                M-TLL.....-AS-I-----SL-V-A--T-.........-L---K.......-LR--ER-N--YE-IR--T---------DTEE-.AA.....LG.----L...V-G-.IHN-...-   82
U.CY.08.CY223                MNTPT.....-LS---I--F--VVI--T-.........-F---K.......-LRK--------Q-IR----------D-DTEE-.ST.....L-.-V-DN..VHLLV.N-GM...-   83
U.ES.10.DEURF10DZ001         M--LE.....-A---G----F-A--------.........-Y-----.......#-RK-K---K-L--IR------------DTDE-.AK.....L-.---DF...D--V.G-N-...-   81
U.GR.99.99GR303              M-SLE.....-AS--G-I--F-A-----D-.........-L-V---.......-*EN--R-----V-IR---------I--DTEE-.SK.....LM.---PD...DF-N.-N--...-   81
U.NL.95.U_NL_95_H10986_D1    MNILV......-G-A-----F-LVVI--T-.........-YC--K.......-LK--ER--KI-N-IR------------DAEE-.AA.....LG.---YL...VLR-.I----...-   81
01_AE.AF.07.569M             MT-LE.....-S---G-I--L-S-----T-.........-A--AK.......-L---------VK-IR------------DTEEW.AK.....L-.---DF...D--V.G-N-...-   82
01_AE.CF.90.90CF11697        MSALQ.....-*---G-I--L-L-----T-.........-F---K.......------------E-IR--------G---DTNE-.AT.....L-.-V-DF...D--V.G-N-...-   81
01_AE.CN.10.YNFL03           MT-LE.....-S---G-I--L-L-----T-.........-A--F-.......R-----------K-IR------------DTDE-.AK.....L-.---DF...D--V.G----...-   82
01_AE.HK.04.HK001            MT-LQ.....-S---G-I--L-L-----T-.........-GL-V-.......------------KKIR--E---------DTDE-.AK.....L-.---DF...D--V.G-N-...-   82
01_AE.IR.10.10IR.THR48F      MT-LE.....-S--AG---L-L-----T-.........-A-----.......---N-------K-IS---X--------DTEE-.AK.....L-.---DF...D--V.G-N-...-   82
01_AE.JP.x.JRC77AE           MT-LE.....-S--AG-I--L-L-----T-.........-A---K.......-----K-----LN-IR------------DTDE-.AK.....L-.---DF...D--V.G-N-...-   82
01_AE.TH.09.AA090a_WG11      MT-LE.....-S---G-I--L-L-----T-.........-G--IK.......RL----------K-IS------------DTDD-.AK.....LL.---DI...D--V.G-N-...-   82
01_AE.TH.90.CM240            MT-LE.....-S---G-I--L-L----*T-.........-A--VK.......-----------VK-IR------------DTDE-.AK.....L-.---DF...D--V.G-N-...-   81
01_AE.US.05.306163_FL        MS-LE.....-S---G-I--A-S-----T-.........-A--VK.......--F--------VK-IR------------DTDE-.AK.....L-.---DF...D--V.G-N-...-   82
01_AE.VN.98.98VNND15         MT-LE.....-C---G-I--L-L-----T-.........-G--F-.......---K-------VK-IR------------DTDE-.AK.....L-.---DF...D--V.G-N-...-   82
```

Draft of 30 September 2015, not for public release
hiv1prot_main.tex r8876–10546

```
                               Env start              signal peptide end  gp120 start                                                        glycosylation NVT                                              CD4 binding      glycosylation NDT
                                                                                                                                                                                                                                     V1
B.FR.83.HXB2                   MRVKEKYQH.L.WRW....GW.RWGTM...LLGM.LMIC.SA...T.EKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVL.VNVTENFNMWKNDMVEQMHEDI.ISLWDQSLKPCVKLT.PLCVSLKCT.DL...KNDTNT.......  140
A1.AU.03.PS1044_Day0           --A-GIQMN.-.HC.....LL.K----...I---.IL--.--...A.-QR------------D-E----------------------------------IN-.X----E------N-----Q---.---------------.----T-N-S.HE...VIFNS-LNNSTHS  146
A1.CH.03.HIV_CH_BID_V3538      ---MGMQRN.Y.PH.....W-.K--IV...-FWL.-I--.N-...-.---------------D-V--------------------------------H-IP-.A----E------N------T--.--------------A.----T-E-N.-A...NRTNIN.......  139
A1.CY.08.CY236                 ---RGIQMN.S.HC.....LL.-----...I---.II--.--...V.--T------------D-E------------T--A--I---------------IN-.E----E-----LN------T--.---------------.----T-D-R.HN...VTIH-N...VTNN  143
A1.ES.05.X1608_8               ----GIQMN.S.Q-.....LL.-----...I---.II--.--...A.-N-------------D-E---------------K------------------IR-.D----E------N------T--.------------E--.----T-N--.EA...Y---DV.......  139
A1.KE.11.DEMA11KE001           ---MGTLMN.W.R-.....FV.---I-...V---.II--.--...A.QN-----H-------D-D-I----------E--K----------------R--Y-.E----D------S------T--.---------------.----T-N-S.NA...NVSVSN...VSVT  142
A1.RU.11.11RU6950              -KARGMKRN.Y.QH.....L-.----W...-FW-.IIM-.K-...A.-H------------RD-E-------------K----------------D---IM-.E------D----N-----QT--.------------Q--.----T-H-A.GS...NSTSSN...NSTS  143
A1.RW.11.DEMA111RW002          -K-MGIQRN.Y.GQ.....L-.-----...I---.II--.--...-.-N-------------N-E-------------K---------------------Y-.E----E------N------T--.------------Q--.----T-N--.NN...ITVN.........  137
A1.SN.01.DDI579                ---MGIQRN.Y.LP.....L-.---MI...S-C-.II--.K-...-.-N-------------D-E-------------Q--------------------LD-.E----E------N------T--.------------Q--.----T-N--.-I...PKNFTN.......  139
A1.UG.11.DEMA110UG001          ---MGIQMI.W.QN.....L-.-----...I---.-I--.-V...A.GN------------RD-E---------------------------A------IE-.K----K------N------A--.---------------.----T-----.--...DLNGTQ...SLNG  143
A1.ZA.04.04ZASK162B1           ---M-TQRN.Y.P-.....W-.I--I-...I---.II-G.R-...S.DN-------------D-D------------SK--------------------MN-.E----K------G------T--.--------T------.----T-N--.NS...NFTR-S...NLTD  142
A2.CM.01.01CM_1445MV           -K-MGTQRN.C.QH.....L-.---IL...IF--.-I--.K-.....TN------------RD-E-------------R---------------------D-.K-----------N---------.-N------M------.----T-N-S.EV...-DN-.........  136
A2.CY.94.94CY017_41            ---MGTQRN.Y.QH.....L-.-G-IL...I---.-IM-.K-.....-D-------------D-D-I--------------------------------IN-.E-----------N-----Q---.---------------.----I-N-S.NA...NTS-HS...NS..  140
B.BR.10.10BR_MG029             -----IRKN.Y.QX.....Y-.---IL...----.----.--...-.DR---------------I------------------A----------------R-.E--------V--N-D-------.---------------.----T-N--.---...YANNTA...NGQN  143
B.CA.07.502_1191_03            ----GIRRN.Y.QHSWRGGI-.---I-...----.----.--.AE.DQG--------------K----------G----A--------------S-----M.E----E--------A-------.SN-------------.----T-N--.NA...NVSET-...Y...  146
B.CH.08.M2_0803101_NFLG8       -K--GTKKN.Y.QH.....L-.---M-...----.----S--...A.DQ-----------------------------------------------------.G-----------N---------.---------------.----T-N--.-D...LKNN-N...SSSN  144
B.CN.10.DEMB10CN002            --A-VIRKS.Y.QP.....W-.I----...----.----.--...K.-----------------I-P----------A--A-------------------D-.E----D------N---------.-----E---------.----T-N--.-W...-PNNS-...TNNN  143
B.ES.10.DEMB10ES002            ----GIRKN.Y.QNL....W-.I---L...---I.----.-T...A.-N-----H---------S-------------------------------S----L-.E-----------N--Q---D--.-----E---------.----T-N--.KV...SCTNS-...SSNN  144
B.FR.11.DEMB11FR001            -K--KTR.....KH.....L-.---IL...---I.----.--...A.-Q---------------N-------------------------------------.G----D------N---------.---------L-----.----T-N--.G-...N-A--I.......  136
B.GB.05.MM45d213_GN1           -----IMKS.W.QH.....W-.KG-IL...---I.----.-N...A.-N-------------------------------A------------------M--.E----E-D----N-----Q---.T--------------.----T-N--.-A...DVT-TN...ST..  141
B.HK.06.HK003                  -K--GIRKT.C.Q-.....F-.K----...---I.----.-T...K.-----------------N---------------A---------------------.G----E-DI---N---------.-N-------------.----T-N-S.NN...ITI-TS...TNR.  143
B.HT.05.05HT_129389            ----GIRKN.Y.QH.....W-.T---L...----.----.--...A.--M---------------------------E---------------------IE-.K-----------N---------.-----E---------.----T-N--.KP...V-C-KE...ENGT  142
B.JP.12.DEMB12JP001            -K--GTRKN.C.QH.....L-.K--I-...----.----.-.VEK.DN-------------------------------A-------------------R-.E-------I-E-N--------V.----E----------.----T-N--.-A...G-T-EA...GNT.  144
B.KR.07.HP_18_07JHS10_3909     ----GIRKN.Y.QH.....W-.---I-...----.W---.--.....DEM------------------------------------------I-------E-.------------N---------.---------------.----T-N-S.N-...R-E--I...KNA.  141
B.PE.07.502_0525_wg5           --ARGIMRN.Y.QH.....L-.-----...----.----.N-...A.DQ--------------TA-------------P--------------------IL-.E----D------N---------.---------------.----T-N-S.-K...LK..........  135
B.RU.11.11RU21n                ---M-IR.....KH.....L-.----W...----.----.--...-.QN----------------------------SK-A------Y-------S---IP-.K--------G--N--D------.------------R--.----I-N--.NV...AS..........  132
B.TH.08.MERLBDTRC10            -KA-GIRKN.Y.QHL....W-.K----...----.----.--...-.KQ-------------D-S---------------A------------------IA-.G----K----E-N--D------.------------R--.----T-N--.-V...WRN-TN.......  140
B.US.11.ES38                   ---MGIRKN.Y.QH.....W-.-----...----.----.--...-.-----------------N----------G-KK-A--I-----------S---IK-.E---------------------.------------S--.----T----.NA...TTS--V..T....  140
C.AR.01.ARG4006                --A-GIQRN.W.KQ.....W-.I--IL...GFC-.X---.NV...R.GN---------------K----------G-----------------------L--.K----------E----D--Q--.---------------.----T-N--.-A...NRTISX...TISN  143
C.BR.07.DEMC07BR003            ---MGTQRN.W.KH.....W-.I--IL...GFW-.F-L-.NV...M.GN-------------D-K--------------------------------H-MI-.E--------------D--QQ--.---------------.----T-N--.NA...TSTPLCVTPNCTN  146
C.BW.00.00BW5031_1             ---MGTQRN.C.QQ.....W-.I--IL...GFW-.---Y.NV...G.GN---------------K------------EK----------------S---I--.E----T---------D------.---------------.----T-N-S.AA...N-..........  135
C.CN.10.YNFL19                 ---R-ILRN.Y.QQ.....W-.I--VL...GFW-.----.-V...G.GNR------------D-K----------G-EK--------------------M--.E--------------N-----V.---------------.----T-N--.NV...T-VNSN...STNN  143
C.CY.09.CY260                  ---MGIMRN.C.Q-.....W-.I--IL...GFW-.----.NG...G.D----------------KA------E-R--ER--------------------L--.E--------------D------.-----E---------.----T-N-S.-Y...-GTH-N...TTSN  143
C.ES.08.X2363_2                ---RGILRN.W.QH.....W-.I--IL...GFW-.---X.NV...W.GN---------------K-------------X-A--------------D---L--.D-A--Y-------XD#-L---#.---------------.----T-H--.NYLIPPT..IN...XTIT  143
C.ET.02.02ET_288               ---RGIQRN.C.QQ.....W-.I--LL...GFW-.F---.NV...V.GN-------------D-SP---------------------Y-----------L-M.E------------------Q--.-----E---------.----T-H-SSNV...TS.NSN...VTSN  143
C.IN.09.T125_2139              ---M-ILRN.Y.QQ.....W-.I--IL...GFW-.----.NV...G.GN---------------K-----------HE-----------------K---I--.G--------------D-----V.---------------.----T-E-G.NV...SVSNVS...VSNT  143
C.KE.00.KER2010                ---RGIQRN.W.QQ.....W-.I--IL...GSW-.--T-.-V...E.GN---------------KAP----------EK----------------S---I--.E--------------D------.---------------.----T-N-S.NV...TSNATS...NSTN  143
C.MW.09.703010256_CH256.w96    ---RGIQRN.Y.QQ.....W-.I-SIL...AFW-.--V..-V...R.GN---------------K-------------K--------------------IK-.E--------------D------.---------------.----T-N-S.HN...ITVNG-...MGNG  142
C.TZ.08.707010457_CH457.w8     -K-M-ILRK.W.KH.....W-.I--IL...GFW-.-L-S.NV...M.GN------------R--K------------ER--------------------MR-.E--------------D------.---------------.----T-R-E.NA...TYNDTN.......  139
C.YE.02.02YE511                ---TGILRN.Y.QQ.....W-.I--IL...GFWI.----.NG...G.GN---------------K------------ES--------------------I--.E----S---------D--Q---.---------------.----T-N-N.NN...ITNNV-...ND..  141
C.ZA.08.705010534_CH534.w12    ---MGIQKN.C.QQ.....W-.I--IL...GFW-.-L--.NV.....QSS--------------K-------------E----------------S---MF-.E----D---------D------.---------------.----T-E--.NA...TFN-.........  136
C.ZA.10.DEMC10ZA001            -K-MGIKRT.C.QQ.....W-.I--IL...GFW...--Y.NV...L.GN---------------K------------ER----------------D---I--.G----------E---D------.---------------.----T---G.NV...SRNVTN...TNSS  141
C.ZM.11.DEMC11ZM006            ---RGILRN.Y.PQ.....W-.I--IL...GFW-.V-S-.-V...R.GN---------------K------------EK--------------------IW-.E--------------D--N---.-----E---------.----T-S--.NF...TTNSTI...ITIN  143
D.CM.10.DEMD10CM009            --ARGIKRN.Y.QH.....L-.T--I-...---LL--TY.--...A.DN---------------K-----------TKA-A--I-----------S---II-.E----P----------------.---------------.----T-N--.ET...HLNSTQ...SVN.  143
D.CY.06.CY163                  ---RGIKRN.Y.HN.....L-.K----...----L--TY.--...A.D----------------KA---------VTKA-A--I---------------I-GE------D--------------.-----E---------.----T-D-I.-V...ETNNK-.......  141
D.KE.11.DEMD11KE003            --A-GMKRN.Y.QH.....L-.G--I-...---I.--TY.-V...A.-Q--------------------------S-KE-A--I----------------L-.E-----------N---------.---------------.----T-N--.EW...-DNSTN...SNST  143
D.KR.04.04KBH8                 ---RGMKRN.Y.QH.....L-.K----...----.W---.--...A.GN-------------------------NSLKA-A--I-----------S------.A-----LTC---N--D------.----E----------.----T-N--.-I...R-K-IN...N...  140
D.SN.90.SE365                  --AR-MKRN.Y.QH.....L-.I----...----.--T-.-V...A.----------------------------S-E--K--I---------------IE-.A-----------N---------.---------------.----T-N-R.-I...SS-ATS...NT..  141
D.TZ.01.A280                   ---M-TQRN.Y.QH.....L-.---I-...----.W-TY.-V...A.-Q----I-------R--N----------SF---A--I---------------MD-.---S--------N---------.----EE---------.----T-H-S.-A...NTTNSG...NG..  141
D.UG.10.DEMD10UG004            ---R-MKKN.Y.QN.....S-.---I-...---I.--T-.-V...-.GQ---------------K------------KP-A--I---------------IY-.E-----------N---------.---------------.----T-N--.KW...TSN-TS...NST.  142
D.UG.11.DEMD11UG003            ---RGIEKN.Y.RH.....L-.G---L...---I.-VTY.-V...A.-Q------------R-------------S-E---------------------IH-.E-----------N---------.---------------.----T-N--.EW...EPNATN...ANNA  143
D.YE.02.02YE516                ---RGMKRN.Y.QH.....L-.K---L...----L--TY.--...A.D----------------KA---------SVK--A--I--S------------IE-.K---------------------.---------------.----T-S--.-I...--V-VD...NN..  142
D.ZA.90.R1                     --ARGMERN.C.QN.....L-.K--M-...----.----.--...A.-N--------------------------G-KK-A--I----------------G-.E----------R---N------.---------------.----T-N--.-A...IFNSTS...FS..  141
F1.AO.06.AO_06_ANG32           ---MGMQRN.W.QN.....LG.K--LL...F--V.-I--.N-...A.-N--------------------------G-ER--------------------IL-.K---------Q-N------T--.---------------.----T-H-K.-A...NTSSTE.......  139
F1.AR.02.ARE933                ---RGMQRN.W.QH.....LG.K--PL...F--I.-I--.N-...-.-N-L----------R-------------S-ER----------------D-R----.E------D----N------T--.---------------.---XT-X--.-A...IR...........  135
F1.BR.10.10BR_RJ015            -K-RGMQRN.W.QH.....LG.K--LF...F--I.-I--.N-...E.-----------------N------------KK-A------------------Q--.G----K-D--N--------T--.-----E---------.----T-N--.NA...AR..........  135
F1.CY.08.CY222                 ---RVTQRN.W.QH.....LG.---LL...F--I.-I--.--...A.DN--------------------------G-EK--------------------IH-.E-I--D------N---------.---------------.----T-N--.AK...NSTE.........  137
F1.ES.02.ES_X845_4             ---RGMQRN.W.QH.....LG.K--LL...F--I.-I--.N-...A.DN-----------------------N--G-ER----------------D---IW-.E------D----N------T--.V--------------.----T-N-S.-A...RDANSQ.......  139
F1.RO.96.BCI_R07               -.....QKN.W.QH.....WG.K--LL...F--I.-I--.--...-.D---------------------------G-EK--------------------ID-.------------N---------.-----E---------.----T-N--.NA...NGTLSA...NQSS  138
F1.RU.08.D88_845               ---MGMQRN.W.QH.....WG.K--LL...F---.-I--.N-...-.NN------------R----------------E-A-------------------F-.E----D-D--N-S---------.---------------.----T-N--.NA...TIPNGN...NN..  141
F2.CM.02.02CM_0016BBY          ---RGMQRN.W.QH.....LG.K--FL...F--I.-I--.N-...A.DN----------------------------EK-A--------------D---IF-.D-----------N--D------.--------------I.----T-H-S.-V...NITA-N...T...  140
F2.CM.10.DEMF210CM001          --MR-MQRN.W.QH.....LG.---LL...F--I.-I--.N-...-.-N------------R--N-V--------G-ER--------------------M--.A----S----E-N--N------.--------Q------.----T-N--.NV...TTTN--...TYAK  143
F2.CM.10.DEMF210CM007          ---RGMQRN.W.QH.....LG.K--LL...F--I.-I--.N-...A.DN-----------------V-----E----EK---A--Y-------------L--.E--------------D--N---.T--------------.----T-S--.-A...NDTSAV...NS..  141
G.BE.96.DRCBL                  ----GIQRN.W.QH.....L-.N--IL...I--L.VI--.-....A.--------------ED-NAP---------HS--S--I-----------S---INM.R-----------N---------.-----E---------.----T-N--.EI...N-NSTR...NI..  140
G.CM.10.DEMG10CM008            ----GIQRN.W.QH.....L-.T---L...IF-L.VI--.--...S.-N------------E-TD------------S--R--------------D---LF-.E-------A---N---------.T----E---------.---IT-H-N.NG...NLNYTD...ANGE  143
G.CN.08.GX_2084_08             ---RGIQRN.W.QH.....L-.I-WNL...I--L.VI--.-T...S.-N------------E--D-P-------R--S--K------------------IKM.D----P------N---------.-----E---------.----T-N--.AV...TCANA-...HGT.  142
G.CU.99.Cu74                   ---RGIQRN.C.QH.....L-.K---L...II-L.AI--.--...S.XN----I-------ED-DAP----------S--S------------------IP-.E----K------N---------.-----E---------.----T-N-V.NV...SCTNS-...G...  140
G.ES.09.X2634_2                -K--GTQKS.W.QN.....L-.K---L...I--L.VI--.--...S.NN------------ED-D---------R--S--S------------------IP-.---S-E------N---------.-----E---------.----T-N--.NV...TTSSTS...TN..  141
G.GH.03.03GH175G               ----GIQTN.W.QH.....L-.K---L...I--L.VI--.--...S.D-------------ED-D----------S-SS-K------------------I-M.K----Y------N---------.-----E---------.----T-N--.-I...VYSNC-...RKHP  143
G.KE.09.DEMG09KE001            --A--MKRN.W.Q-.....L-.EG-LL...I--L.VI--.--...S.IN--A---------ED-D------------S--K--I---------------LP-.T-----------N---------.-----E---------.---IT-N--.NV...TSTGIE...NNTE  143
G.NG.09.09NG_SC62              -K--GIQKN.W.PP.....L-.T-W-L...I--L.VI--.N....A.DN------------ED-E-P----------S--K------------------IG-.E----Y------N---------.-----E---------.----T-N-V.T-...NCTDA-...YIGN  142
G.ZA.01.TV546                  ---RGIQRN.W.QH.....L-.K---L...I--L.VIL-.N-...S.NN------------ED-DAP--------S-S--R--I---------------II-.K-----------P---------.-----E---------.----T-----.-Y...NGTVTN...KT..  141
H.BE.93.VI991                  -...-TQRN.Y.PS.....L-.----L...I---.-L--.-V...V.GN---------------K---------------R------------------M--.E----T----V--------T--.---------------.----T-D-S.SV...NATNV-...KS..  138
H.CF.90.056                    -...-TQRN.Y.PS.....L-.----L...I---.-L--.--...A.QN---------------K------------E--K------------------M-M.E----S----E-N------T--.---------------.----T-N--.NV...R-N-SN.......  136
H.GB.00.00GBAC4001             -...-TQRN.Y.PH.....LL.SG-IL...I---.-LM-.-T...G.-N---------------K------------S--K--------------S---M--.E---------E-N--D------.---------------.----T-E-G.EV...G-K-SN...SN..  138
J.CM.04.04CMU11421             -T-M-TQRS.W.QN.....L-.K--L-...IF--.----.K-...A.D-M--------------K----------S-S--S------------------MII.A--------------D------.--V--E-------M-.----T-N-S.HN...ISTNS-...NSRT  143
J.SE.93.SE9280_7887            -...-TQKN.W.QT.....L-.-G-L-...IF--.----.K-...K.-D-------------D-K------------S--K--------------S---MN-.P--------------D--Q---.--V--E-------I-.----T-N-S.NI...TSNS--...TS..  138
J.SE.94.SE9173_7022            -...-TQTS.W.LS.....L-.---L-...IF--.----.--...R.-N------------RD-K------------S--K------------------MS-.P--------------D--Q---.--V--E-------I-.----T-N-S.-V...NSNNS-...DSN.  139
K.CD.97.97ZR_EQTB11            --AR-IQRN.W.QH.....LG.KR-IL...F--I.-I--.--...A.NN----------------------------E------------------------.E-----------N------T--.-----E---------.----T-T--.NV...T-NRTN...ANKN  143
K.CM.96.96CM_MP535             ---RGMQRN.W.QT.....LG.N--IL...F--I.-I--.-N...A.D-----------------P-----------EK---------------------EM.E-----------N------T--.-----E------E--.----T-N--.-Y...-GTNS-...NN..  141
U.CA.01.TV749                  ---MGTQRN.Y.QS.....L-.---L-...IF--.----.NTK..G.-N--A---------R--KA--------QS-S--K------------------IX-.K-------V-------------.-----E---------.----T-N--.-I...TSSNA-...IST.  143
U.CA.99.TV721                  ---M-TKRN.Y.Q-.....L-.---L-...IF--.----.KI.EGK.DN---------------K---------QS-S--K--------------D---ID-.K-------A------D------.-----E---X-----.----T-T--.SV...A-NIT-...TNVT  145
U.CD.83.83CD003_Z3             -----IQRN.Y.QH.....L-.K-SLI...I---.I---.K-...I.--S------------D-E------------EK-S--------------S---L--.G-----------K---------.---------------.F---T-N-I.-V...--S--N...N...  140
U.CD.90.90CD121E12             --A--IQRN.Y.HY.....L-.K-SLI...I---.II--.KG...E.-QS------------D-D----------G-EK-A--------------S---L--.------------N---------.--------------S.S---T-N--.-V...GTNNTA.......  139
U.CY.05.CY090                  --A--TQKN.W.QH.....LG.K--IL...F--I.FI--.-TAEKA.-N---------------K------------SP-A------------------I--.E----Y-------------K-V.---------------.----T-D-V.NI...TANITR.......  142
U.CY.08.CY223                  ---MGTQRN.C.LL.....SL.------.CIFWL.M-V-.N-.....N--------------D-K-------N-Q--N--K------------------ML-.I--------------D------.-----E---------.----T-N--.NA...RANKV-...VNAT  143
U.ES.10.DEURF10DZ001           --ARGIQTN.W.QN.....L-.K---L...I--L.VI--.--...-.NN------------RD-E-------------Q------------------R-IP-.K-------A---N---------.---------------.----T-N-K.PF...NSN-SG.......  139
U.GR.99.99GR303                --L-GIQRN.C.QN.....L-.--AL-...IF--.----.K-...-.DN-------------D-E-----------HSA-K---L----------S---MI-.G-----------N------T--.---------------.----T-N-K.NP...SCVALN.......  139
U.NL.95.U_NL_95_H10986_D1      ---RGMQRN.W.RH.....LG.---IL...F--I.----.--...A.D-----I-------QD-N----------S-SN-R------------------LP-.------------N----V----.-----E---------.----T-N--.-V...T-R--S.......  139
01_AE.AF.07.569M               ----GTQRN.G.PN.....L-.K---L...I--L.VI--.--...S.DN------------RD-D-----------HE-----------------S---IY-.E-----------N-----Q--V.---------------.----T-N-S.NA...NLTNIN...NTIT  143
01_AE.CF.90.90CF11697          -G--GTQMN.W.PH.....L-.K---L...I--L.VI--.--...S.DT------------RD-D---------N-QE----------V----------IH-.E------D----N-----Q--V.---------------.----T-H--.KA...-LNDTY...NGTA  143
01_AE.CN.10.YNFL03             -----TQMN.W.PS.....W-.K---L...I--L.V---.N-...S.NN------------RD-N-----------HE--------IY-----------IH-.------------N-----Q---.-----E---------.----T-T-I.NA...NLTKL-..ANASD  144
01_AE.HK.04.HK001              -----TQMN.W.PN.....L-.----L...I--L.VI--.--...S.DN-------------D-D-----------HE---------------------IK-.GD--------X-K-A---Q--V.---------------.----T-N--.QA...NWKS-N...TTQN  143
01_AE.IR.10.10IR.THR48F        -K---TQRN.W.LN.....W-.K---L...I--L.VT--.--...S.DN------------RD-E-----------QE--A------------------IY-.E-----------P-A---Q--V.---------------.----T-N--.KA...NLTKLA...NASN  143
01_AE.JP.x.JRC77AE             ----GIQMN.W.PN.....L-.----L...I--L.GI--.--...S.D-------------RD-D-----------HE---------------------IH-.E-----------N-----Q--V.---------------.----T-N--.EA...NWTNVS...STSN  143
01_AE.TH.09.AA090a_WG11        ---RGTQMT.W.PN.....F-.---IL...I--L.VI--.--...S.DN-------------D-D-----------HME-A--I---------------IH-.E-----------N-----Q--V.---------------.----T-N--.NA...DLRYAD...NSTN  143
01_AE.TH.90.CM240              -----TQMN.W.PN.....L-.K---L...I--L.VI--.--...S.DN------------RD-D-----------HE---------------------IH-.E-----------N-----Q--V.---------------.----T-N--.NA...NLTNGS...SK..  141
01_AE.US.05.306163_FL          ----GTQMN.W.PN.....L-.K---L...I--L.AI--.--...S.NN------------RD-E-----------HE---------Y-----------MH-.E-----------N-----Q--V.---------------.----T-N--.-F...-VP.........  136
01_AE.VN.98.98VNND15           -----TQMN.W.PN.....L-.K---L...I--L.VI--.--...S.D-------------RD-E-----------HVP--------------------IP-.E-----------N-----Q--V.---------------.----T-N--.SA...TLPNSN...ANL.  142
```

Draft of 30 September 2015, not for public release
hiv1prot_main.tex r8876–10546

glycosylation NIS
glycosylation NCS
CD4 binding
glycosylation NVS
glycosylation NGT
V1
V2
glycosylation NSS
glycosylation NDT
glycosylation NTS
glycosylation NKT

Draft of 30 September 2015, not for public release
hiv1prot_main.tex r8876–10546

HIV-1 Proteins
CD4 binding
glycosylation NNT
glycosylation NCT
V3 tip
V3
glycosylation NFT
V4
glycosylation NGS
glycosylation NTS
glycosylation NNT
glycosylation NKT
CD4 binding
glycosylation NST

Draft of 30 September 2015, not for public release
hiv1prot_main.tex r8876–10546

```
                              glycosylation NST
                        glycosylation NST
                                                     V4                                                                                           V5
                    glycosylation NST                       glycosylation NNT                 CD4 binding          glycosylation NIT  CD4 binding glycosylation NES  CD4 binding
B.FR.83.HXB2                TQLFNSTWFN.......STWS..TE.GS............NNT..EG......SD.....TITLPCRIKQIINMWQKVGKAMYAPPISGQIRCSSNITGLLLTRDGGN.SNN..........ESEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVV....  506
A1.AU.03.PS1044_Day0        SE---R--NA.......TDQL.......N..S.....TE..DS-..AL......NE.....--I-------V-----TP-Q-------R-A---E------I--------.D-T..........STN-T----------------------R-----I---T-------....  513
A1.CH.03.HIV_CH_BID_V3538   -E------ND.......TSNT..................S..QG-..NN......NS.....-----------FV----R--Q-------K-K-Q-I-------------K.GKS..........NIT-----E--N-K---------------------------------....  497
A1.CY.08.CY236              SK---T--NV.......TDS.....................IEE..S-......NT.....I-V---K---------RA-Q-I-P---R-I---E-------------D.TE-..........TT-----------------------------------E-R------....  505
A1.ES.05.X1608_8            -G--DG--NG.......TKS........N..I.........TWE..SI......N-.....------K---VVR---R--Q-V-----T-A-S-N-T----I------D.NSS.........GR---T-------------------------------R-------....  515
A1.KE.11.DEMA1IKE001        SG------SL.......NNTWS.....LN..S....TQS..-D-..GI......N-.....----------------R--Q-------Q-V-Q-V---------------.V-S..........SNGN-------------------------------R-------....  518
A1.RU.11.11RU6950           -E------DG.......NGT......................ATN..NT......E-.....NV--------------R--Q-------R-S-K-V--------I----G.G--..........SNN-T---I--N---------------Q---I-I---T-------....  506
A1.RW.11.DEMA111RW002       -G------SR.......N-TDN.NT.WFRN.N.....AS..SS-..GL......NG.....----Q-K---------RT-R-------P-T-M-N-S-------------.-G-..........NSN-T-------K-------------------------R-------....  514
A1.SN.01.DDI579             SG------ND.......TAIN.......D..T......I..-VQ..NS......TY.....N-------R--VH---R--Q-------Q-V-N-T-------------D.NSS..........TN-T------N-K----------------------------------....  504
A1.UG.11.DEMA110UG001       SG------N-.......--QES...........................NS......TE.....-----------V---RA-Q-I-----K-I-N-R------I------H.N-...........TN-T-----------------------V--------T-R------....  498
A1.ZA.04.04ZASK162B1        SG---R--ND.......TA-N.......E..T.....TW..--I..NS......NG.....S---Q------L----R--Q-------P-E---E------I------L.EL-.........NTNN-T-------------------I----------R-------....  509
A2.CM.01.01CM_1445MV        -G------TK.......EGN-...T.W.K..K...NDSV..SER..-S......N-.....N---Q-K---------R--R-------A-V-K-I-----MI------K.R-D..........SS--T-------------------------I-----R------....  505
A2.CY.94.94CY017_41         -G---G--W-.......NGTW.......N..G.....PY..TPN..NT......NG.....S-I--------------R--R-------A-I-K-T-----II-------.NG............TN-T----------------------L--------R-------....  508

B.BR.10.10BR_MG029          -X--------........---N..XT.X.N..G.........XIE..GS......DA.....N---Q-K---------Q----------R-K----------------G.NGS..........STN-T--------K----N--------------I------------....  518
B.CA.07.502_1191_03         -L------GS.......NETR......................-E-..ID......NS.....---IQ-K---L-----G--R-------R-N-S---------------.NGT..........NGT-T---E--N-K----N--------------I-----R------....  512
B.CH.08.M2_0803101_NFLG8    S-------NA.......TGTQ.......N..T......TGTQ-E..TK......IG.....N-----K------R--E----------K-L-N-V-----I-------H.-K-TN....STDNDT-V-----------------------------------------....  535
B.CN.10.DEMB10CN002         SK------PS.......DGTD.....-.N...........VTG..GS......N-.....-----Q---R----R--E----------R-V-N-T--------V----I.N-S..........TN-T----------------------RL--------E---I---....  508
B.ES.10.DEMB10ES002         -----N--NE.......TA-........N..E......T..--V..NN......-E.....-----K--------R--G----------R-S-----D-------------.T-E..........TT-----------------------Q---I-I------------....  515
B.FR.11.DEMB11FR001         -P--D---NG.......GEP......V-N...........-TG..GN......DT.....I---Q-----------E----------R-------------I-----DL.N-T..........QN-T-----------------------------------------....  490
B.GB.05.MM45d213_GN1        --------NS.......TEG-..I-.--N...........STE..KN......IT.....L----------V-L--G---I-----R-L-S---------I--------.GX-I.......XNNT-T-----------------------------------------....  512
B.HK.06.HK003               -K--D-S-NG.......TN-G.......N...........DTD..AE......N-.....---------R-------E----------K---K---------------T.NTT..........NGT-----------------------R-----I-----R-E----....  503
B.HT.05.05HT_129389         -K---N--NG.......TDF........I...........YPW..NK......TE.....N---------------G------------P---E--------------D.Q-GT.......DATN-T----------------------R---------E--------....  512
B.JP.12.DEMB12JP001         -K------N-.......-N-P.......N...........DTM..PN......NN.....-------------V-----------N---A-L------------V------.TT-..........TTT-T-----------------------------------------....  513
B.KR.07.HP_18_07JHS10_3909  -K------RF.......NSTW.......N...........GTE..-I......N-.....P-------------R--E----------R---K---------------T.E--K.....TGSNRT----------------------R-----I-----R------....  506
B.PE.07.502_0525_wg5        -K----SGD-.......--D.....................SDQ..DK......N-.....----------------E----------Q---S----------------.-S-..........TNGT---------------------------I-------------....  491
B.RU.11.11RU21n             --------DK.......NITE.....FN............KTNELNN......TE.....I---------------A-------------K----------------.............NDT-T----------------------R-----I-------------....  492
B.TH.08.MERLBDTRC10         SA---H-FN-.......TEGE.......N...........ITE.SIN......TT.....I-K-------------E--------------S-------I-------G.N-T..........GN-T-------N-K---------------------I-------------....  512
B.US.11.ES38                -K------EK.......TNDT.....W.S...........G-E..TL......NG..TITP----------------I---------P-P-N---------I-------G.DRS.........NATN-T---------------------------L-----Q-------....  511

C.AR.01.ARG4006             -------YXT.......NGTE.......I.........T..GW-..ND......-X.....P--IS------------E--R-------R-M-T-R------I--------T.NGT..........NNT-----------------------E-N---I---X-------....  505
C.BR.07.DEMC07BR003         SS----ELDS.......NGTF.......K.........I..-G-..-NGTGTE.NS.....N---Q------------E--R-------A-E-T-R---------------D............TSD-----------------------EVK-------E-------....  529
C.BW.00.00BW5031_1          S------YRA..............................-S-..NSTS....NS.....N-----------------G--R-------A-N-I-K------V---Y---E.GV..........K-N-T---A--N-K--------------E-K---I------------....  493
C.CN.10.YNFL19              S----G-YNG.........TY.......M.........S..-G-..-S......NA.....---I------------E--R----N--A-N-T-K---------V----E.GT-..........ATN-T----------------------E-K---L---E-------....  511
C.CY.09.CY260               SK---Y-YPY.......NGTH..........................SH......NG.....-TITQIV-I-LVIR--E--R-IS---TA-ENISG-Y--RV-WPTG--K.N*G.........QNDT-T----------------------E-K---I---E-------....  502
C.ES.08.X2363_2             -K---G-YS-.........GTY.......S.........PG.Y-S..SSHS.....NN.....S-IIQ-----------------Q------A-N-T-N--------------S.--..........DTK-----T--N-K--------------E-K---I------------....  515
C.ET.02.02ET_288            SK---T-GLS.......D..............................NA.....-L-I-----------E--R-------K-I-T-R--------------M.N-T..........NNT-T----------------------E-K-------G---A-------....  501
C.IN.09.T125_2139           SG---R-YTT..DL..FNKTY.......N.........D..-D-..-DNS....GP.....V-II----------E--R-------A-N-T-N-S-------------.D-K.........TSTN-T---L--N------N---------E-K---I---T-------....  525
C.KE.00.KER2010             SG---E-LTR......LNDTY.......S.........SN...............GT.....N------------------R-------A-N-T-T-K------Q------.-T..........NKT-T------N-K--------------E-K---I---E-------....  505
C.MW.09.703010256_CH256.w96 SKI-D-..................................D-A..NKT......-G.....N-----K---------G--R-------A-N-T-T-----------------.VT..........NDT--------N------------------E-K---I-------------....  496
C.TZ.08.707010457_CH457.w8  SK---N...................................-I-..SDN......-S.....N--I--K-------------------------A-N-T-T------I---------.NSD..........SDP-T---T-----------------EV-----I---G-------....  488
C.YE.02.02YE511             SK---R-YMF.......NSTY.......M.........P..-S-..NSTS....NS.....----------------G--R-------A-N-T-T---------------G.KD............N---T---------------------E-K---I---A-------....  505
C.ZA.08.705010534_CH534.w12 SR---G-YMS..................M.........S..----..KNDT....NA.....----------------G--R-------A-D-T-K-T----I------E.ESK.........TNNT-----A-----------------E-K-------E-------....  496
C.ZA.10.DEMC10ZA001         SK-WDN................................Y..TE-..T-......NG.....--I--------------------R-------A-N-T------------------.TTDN.....TTNNNT-V------NIK------------E-K---I-------------....  496
C.ZM.11.DEMC11ZM006         SG--ENRTYT............................L..-D-..GSTSNNT.NS.....---IQ-----------R--Q-------A-N-T-K------------------A.KK-E.......SDPP-T----------------------EVK---L-----Q-------....  506

D.CM.10.DEMD10CM009         SY---G--I-.......G-RN.......D...........TR-..N-......TG.....DTII------------G----------K---K---TV---I--V-EV-.NSS..........PEAT-----------------------R-----I-------------....  494
D.CY.06.CY163               SK---T--SS.......E--E.......N..S.....TG..----..-L......NG.....---I--T-----I-C-G--V---T--PA-I-Q-K----ITG--------.NS-..........CN-T----------------------R-----I---R-------....  507
D.KE.11.DEMD11KE003         SK--TTNITS.................N.................TS......EG.....N---------------G--G------RE-L-T--------------QEG.---..........IN-T-------N-K-------------R-----L---R-------....  500
D.KR.04.04KBH8              SG------MV.......NSTV.......N...........QTG..LN......NS.....----Q-----------E--R-------E-V-K---------I-----A-.N-..........TN-T------N-K-------------QV--I-L-------------....  492
D.SN.90.SE365               SG------KT...................................KS......G-.....N-----------L--R----------E-V-S---------L----I.H-T.........SN-T-T----------------------RL--------T-------....  500
D.TZ.01.A280                SL---T--IK.......G-Q........N...........-TE..TN......NS.....---I-------------G----------A-L---T--------------.V--..........SRE-----------------------R---I-----R-------....  502
D.UG.10.DEMD10UG004         SG------QE.......NSTL.....LN..K......T..-Y-..-E......NK.....----------------RI----------E-L-K---------E---SV.E--...........SQ-TV--V--E--N-------------R-K-I-L---R-------....  497
D.UG.11.DEMD11UG003         SE----I-L-.......N-GN....WTGN.T.....TN..STE..KN......NG.....----------------G----------E-E-K-I---------------.N-A..........TN-T----------------------T-----I-------G-----....  507
D.YE.02.02YE516             ST--TR-INT.......TNDN..T.V.N..S.........TGN..NT......N-.....---I------------G--R-------A-I-K---------------S.N-T..........NST-T----------------------R-------------------....  496
D.ZA.90.R1                  SG------NR.......TGNN.......S...........SS-..GS......NE.....--------R-------G----------E-L---------I------D.NSS..........SHN-T----------------------R---I-L---R-------....  508

F1.AO.06.AO_06_ANG32        SG---NSD...................................TS......NR.....-----------V----G--Q----A-VA-K-T-N-----------------..........NSHT-T------N-K-------------EL----------R----AR..  490
F1.AR.02.ARE933             SG---D-.....................................DD......NG.....-----------V----E--R-NNH--HA-N-T-G-------------T.G-CTD...GGTGNCT-T------N-K-------------E-----------R-Q------....  495
F1.BR.10.10BR_RJ015         ----STK......................................S-......N-.....-----------V----E--R---N-VA-K-T-N-T-------------.DKST.......XNNT-----W--N----------------------I-------P-------....  496
F1.CY.08.CY222              SE---NSN....................................SN......NG.....------K---------G--Q-------A-K-T-N-S-----------Q.--KSK......DNG-------R-NIK-------------E-----I-------Q-------....  494
F1.ES.02.ES_X845_4          SR--TDSRT...................................-F......N-.....------K----V----G--Q----A--A-N-T-N-----------------.N-T..........S-N-T------N-K-------------E-----I-----------GR..  492
F1.RO.96.BCI_R07            SG----STSP.......EGV.....................F--..SH......KG.....N---L-----------G--Q----A--A-S-T-N-------------Q..............NN-T-T------N-K-------------E-----I---R-Q-Q-----....  508
F1.RU.08.D88_845            SK---H-V--.......D-NN.......I.........V..-GS..SL......NG.....-----------V----G--R----A--A-K-V-N-S-----------X.NDT..........RTN-T---A--N-------------------R-------------H-----....  512
F2.CM.02.02CM_0016BBY       -A---T-LL-.......T-M.........................DN......NG.....--II-------V-V--R--R--------A-K-Q-N------------S.KA-..........NTD-L--I--E---------------Q-Q---I--SR---Q-----....  495
F2.CM.10.DEMF210CM001       SA---TEALK............................----..GQ......NG.....-T-IY-----------R----------A-K-Q-N-S----I--I--NR.T--..........K--TL-------K--------------------------R-------....  505
F2.CM.10.DEMF210CM007       D----ISQSE.......KA..........................GN......NG.....S-F-----R--V-------R--------A-V---N-S-------------E.D-S..........STDTL-------K--------------------------R---Q-----....  497

G.BE.96.DRCBL               SG---NSILK.......-NI-.........................-N......N-..........---N-K---VR---R--Q-------A-N-T-R------I------D.N-S..........T----------KN----------T---KS--I---R-R------....  497
G.CM.10.DEMG10CM008         SR--.......................-..........PL..-S-..NS......TG.....N----------VR---R--Q-------A-K-T-I----------------.N-...........SI-T---A----K---------------K---I---R-R------....  497
G.CN.08.GX_2084_08          SG---NSKL-.......--T..........................TN......NT.....-----------VR---R--Q-------A-N-T-R------------A-.NT............GN-T------------------I---K---I-----R------....  501
G.CU.99.Cu74                SK----SXMX.......NS...........................T-......NG.....PFK---K----VR---R--Q-------A-N-T-V-------------T.N-T..........NRT-V---A-----------R------K-------X-R------....  508
G.ES.09.X2634_2             SE----SYL-.......-............................TS......N-.....-----------VRL--R-EQ-------E-N-T-R---------V------.NDT..........K-N-T---A---------------------K---I---R-R------....  502
G.GH.03.03GH175G            SR---NETG-....................................ST......NT.....N-----K----VR---R--Q-------E-E---D-T----I-------.---.........TNNT-T--------------------T---KS--I---R-R------....  505
G.KE.09.DEMG09KE001         SR---M-ELD..................K.................S-......NE.....------K----VR---R--Q-------E-K-T-N--------------E.N-S..........TN-T---A-----------------IR-K----------R------....  501
G.NG.09.09NG_SC62           -G---TSG--.......D-...........................IT......N-.....--M--------VR------Q-------E-S-T-R---------M-----.N-T.........NKT-T---S-----------------I---K--EI--G--R------....  498
G.ZA.01.TV546               SL---RNETS.......ENLT.......V...........-GN..KT......GE.....N-----N----VR------Q-------E-N-T-X---------------K.N-............S-----------------------------K----------R------....  498

H.BE.93.VI991               -K------T-.......-SYT.......N..DT..YNSN..STE..DI......TG.....N---Q-K----V----R--Q-------R-N-T-I------I--F-RN-..............TNNVT----------------------------E-R------....  509
H.CF.90.056                 SG----S-EM.......H-NY.......T.........S..-D-..K-......NE.....N-----------V----R--R-------Q-N-M-V------I--I-E--.ASA...........-NYT-------------------------I----TR------....  495
H.GB.00.00GBAC4001          SR------NS.......NDTG.......N..G...............TD......NS.....------K---------R--Q-------Q-N-T-V------I--L-RYV.D-G..........TNVTL-------------------Q-----I-----R------....  502

J.CM.04.04CMU11421          SS----S-SS.......KNDT.......N..S...............TA......NY......TII--K-R--VR------Q-I-----E-N-T---------I------D.K-ST......SNNST-----T----K-------------E---------R---------....  510
J.SE.93.SE9280_7887         SK----S-DK.......NSIE.......A...........T-D..TS......XA.....---I--K----VR---RT-Q-I-----A-N-T-T---------------.RG-G......SENGT-T---T--N-K-------------E---------------------....  498
J.SE.94.SE9173_7022         ST----S-DE.......NNIK.......D..T.........-S-..ND......NT.....---I--K----VR---RT-Q-I-----A-N-T-K---------------.R-G.......SENGT-T---T--N-K-------------EL--------------------....  500

K.CD.97.97ZR_EQTB11         -DTVDD-EE.....................................-E......DT.....---I---------------Q-I----TA-N-T-R-----MI--------.D--.........TRTE-T----------------------Q-----I---R-R------....  501
K.CM.96.96CM_MP535          -K---E-GE......................................NG.....---------------------I-----A-S-N------MI--------.NT............HN-T----------------------Q-----I---R-R------....  492

U.CA.01.TV749               SK----S-WF.......N-TN.......S..I...GINT..-S-..-I......N-.....--I-------VR---R--Q-------A-N-T-V-------------KY.RP-..........STN-T---T--E---I--N---N---E-----------R------....  511
U.CA.99.TV721               SG---KS-GS.......WEPN.......T..T...EV.N..-V-..KE......NG.....----Q-K----VR---R--Q-V-----A-N-T-T----------------.N-S..........DN-T---T------I------N--I-E-----------R------....  514
U.CD.83.83CD003_Z3          SE--TGI-NG.......TWDK.......N......CTS..TES..NC......TG.....N---------VVRT--G--Q-------E-T-----------------------.G-A..........TQN-T--------K--------------------------R------....  501
U.CD.90.90CD121E12          SG------NG.......TWAE.......N..N....CTNSES-S..SC......NE.....---I-------VRT--G--Q-------K-T-M-N-S-------------D.NE-..........DTK-T--------------------------------R------....  499
U.CY.05.CY090               SK-W-D-VGS.......NGIA.......G.................SN......EN.....------RPFRL-LS----L-L-L-----A-KST-T--S--PI---A-DW...........TRT-T----------------------R-----L---R-R------....  495
U.CY.08.CY223               SK---G----.......G--V.NGT.WYN..N....TLD..S-D..TE......NG.....------K-R--VR---R--Q-------Q-V-H-Q--------------DK............N-T-T---T--------------------------R-R------....  530
U.ES.10.DEURF10DZ001        ----D-S-GI.......NDTV.......S..N......S..-DA..DS......D-.....N---Q-------K---R--Q-I-----R-N-S-H-----------------.NS..........TN-T---------------------I-----------H------....  496
U.GR.99.99GR303             SG------RP.......NDMQ.....W-.............-D-..#AI......*R.....I------#---------R--Q-------Q-E-K-N-T-------------.-S-..........D--T------------X---------------------G------....  504
U.NL.95.U_NL_95_H10986_D1   -R---T-HLY.......NSTG.......N..D....TQD..--S..TN......-K.....N-----Y----VR---R-EQ-------A-R-Q-I-------------K.N-E.........TNGN-T---A--N-K--------------------I-----R------....  511

01_AE.AF.07.569M            -R---N-CIK.......NATV.......R.................GC......N-.....--I-------------EA-Q---------I-N-V------I---------.-S..........TN-T---E--NIK-------------Q-----I---R--------....  501
01_AE.CF.90.90CF11697       -K------TT.......NEIM.......E..E.........FKG..TN......-S.....-----------VV----E-----------EAVN-V------II-------..............ATN-T------N-K-------------Q-----I---R-R------....  506
01_AE.CN.10.YNFL03          -K---K-NIE.......NETV.......K.................SD......N-.....--I-------VR---G--Q----------N-M-V-----I------D.N-...........ST--------N-K-------------Q-----I-------------....  502
01_AE.HK.04.HK001           ----D-S-NT.......T-TN.......R.................-N......RS.....NFI-------------ET---------R-S-Q-V-----I---------.N-G.........SN-T-T------I--------------Q---------R--------....  501
01_AE.IR.10.10IR.THR48F     -K------G-...RTENNGTI...........................--......NG.....--I---K---------ET-Q------V--I-N-L-----I---S-T-G.K--.....GTGENGT-T---A--N-K-------------E-----I-------------....  510
01_AE.JP.x.JRC77AE          SK------G........NGTI.......E.................GH......NG.....------K---------GA-Q-------R-K-N-V-----I----------.-SA..........NN-T------NIK-------------Q-----L---R--------....  499
01_AE.TH.09.AA090a_WG11     -----R-IG........NVTG.......E..G......S..-DE..NN......NK.....--I--------------GA-Q-------R-K-K-V-----I---------.I--..........ATN-T------Q-K-----R-------Q----------R-R------....  507
01_AE.TH.90.CM240           -K---N-CLG.......NETM.......A.................GC......N-.....------K---------GA-Q---------R-N-V-----I-------V.N-T..........DN-T------NIK-------------Q-----I---R--------....  500
01_AE.US.05.306163_FL       -S---N-YIG.......-NG........T...................KG.....--I---K---------G--Q---------R-N-V-----I-------L.N-...........TN-T------NIK---X-------Q-----L-------------....  479
01_AE.VN.98.98VNND15        -K---D-CIK.......NNS........R.................GC......N-.....--I-------------R--Q---------R-N-V-----I-------T.NSRTN...ETNETNE---------IK-------------Q-----I---R-R------....  509
```

Draft of 30 September 2015, not for public release
hiv1prot_main.tex r8876–10546

gp120 end · gp41 start · fusion peptide · immunodominant region · glycosylation NAS · glycosylation NKS · glycosylation NHT · glycosylation NYT

```
B.FR.83.HXB2                ..QREKR..A.VG.IGALFLGFLG.AAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWS.NK.............SLEQI.WNHTTWMEWDREINNYTSLIHSLIEESQNQQEK  655
A1.AU.03.PS1044_Day0        ..-----..-.--.---V-I----.----------I-----------------S-----------M-K------------V--L-----------------------N----NT--.--.............NKSE-.-DKM--LQ--K--S---QI-YN------T---I  662
A1.CH.03.HIV_CH_BID_V3538   ..E----..-.--.L--V------.----------I---A-------------S-------------K------------V-------Q------------------N----S---.--.............-Q-D-.-GNM--LQ-EK--S----I-YE-L---------  646
A1.CY.08.CY236              ..-----..-.--.---V------.----------I-----------------S----------QM-R------------V-------G------------------N----ST--.--.............--DK-.--NM---Q--K--S---Q--YT----------I  654
A1.ES.05.X1608_8            ..E----..-.--M---V------.---------AL-----------------S-----------M-R-----V------V--L----R-------------IV---N----S---.--.............THX--.-ENM--LQ-----S---EI-YN-L-X------X  665
A1.KE.11.DEMA1IKE001        ..G----..-IG-.L--V------.M-----------------N---------S-------------K------------V--M----R-------------V--N-N----S---.--.............-YNE-.-DEM--QQ-EK--S---DI-YN-L---------  668
A1.RU.11.11RU6950           ..E----..-.I-.L--A------.----------I-----------------S-------------K------------V-------R------------------S----S---.--.............-RQE-.-DNM---Q--K-VF---N--YE-L---------  655
A1.RW.11.DEMA111RW002       ..E----..-.--.M----I----.----------V-----------------S---K---------R---------------------------------------S----S---.--.............-QDE-.-QNM--IQ--K------E--YN---Q-------  663
A1.SN.01.DDI579             ..E----..-.--.L--V------.----------------------------S-----------M-K------------V--I-----------L-----------N----SN--.R-.............-Q-E-.-GNM--LQ--K--S---DI-Y------------  653
A1.UG.11.DEMA110UG001       ..-----..-VF-.L--V-I----.----------I-----------------S-------------K------------V--L-S--R------------------N----S---.--.............-QDE-.-QNM--LQ-----S---D--YT------I----  648
A1.ZA.04.04ZASK162B1        ..E----..-.--.L--V------.----------I-----------------S-------------K------------V---------------------I----T----S---.--.............-YDD-.-DNM---Q-EK--G---E--Y--L----V----  658
A2.CM.01.01CM_1445MV        ..E----..-.--.L--A------.----------I-----------------S---K------Q--K------------V--I----Q------L-----------T----S---.--.............NQS--.-DNM--L---K--S---ET-YK-L-D-------  654
A2.CY.94.94CY017_41         ..E----..-.--.L--V------.----------L-----------------S---Q---------K------------V------------------------A-T----T---.--.............-QDE-.-DNM--LQ--K--S---NI-YR-L---------  657

B.BR.10.10BR_MG029          ..-----..-.--.----------.-------------------------------------------------------L-------M---R--------R-----N--------.--.............TYDE-.-QNM--RD-E---D---G--YN-L-K-------  667
B.CA.07.502_1191_03         ..---R-..-.A-.L----I---S.----------L-----------------------------M--------------V-------R------L------I----S-T--ST--.--.............--DE-.-KNM---Q-EK--D---DI-YT-L-K-------  661
B.CH.08.M2_0803101_NFLG8    ..-----..-.-T.L--M------.----------L-----------------------------M--------------V-------R-----------------------TT--.--.............T----.-GNM---Q-EK--D---NV-YT---DA------  684
B.CN.10.DEMB10CN002         ..-----..-.--.L--M------.----------I--M-------T------S-----------M--------------V-------Q--R---------------T--------.--.............--ND-.--NM---Q-EK--D---G--Y-------I----  657
B.ES.10.DEMB10ES002         ..-----..-.A..F---------.----------L-----------------S-----------M--------------V-------R---------------------------.--.............TYN--.-DNM---Q-E---D----I-YD---K-------  663
B.FR.11.DEMB11FR001         ..-----..-.--.----------.----------------------------------------M--------------V-------------------------------T---.--.............T-NE-.-ENM---Q---------N--YT---K------R  639
B.GB.05.MM45d213_GN1        ..-----..-.I-.L--V------.----------V---------------------K----------------------V--I-S---------------------T----T---.--.............--NE-.-DNM---Q--K-------T-YT---KA------  661
B.HK.06.HK003               ..-----..-.--.----------.----------L-----------------------------M--------------V-------R------------------N--------.-R.............T-TE-.-DNL-----EK--D---G--Y------------  652
B.HT.05.05HT_129389         ..-----..-.--.L--M------.----------IA--A----------------------------------------V-------Q------L------R----T----T---.--.............--DY-.-QNM---Q-E---S---G--Y-------I--D-  661
B.JP.12.DEMB12JP001         ..-----..-.--.---V------.----------LM----------------------------M--------------V-------R------------------T----T---K-R.............N-TD-.-DNM--VQ-E---D---KE-YT-L--A------  663
B.KR.07.HP_18_07JHS10_3909  ..-----..-.--T---M------.------------------L------------------------------------V--------------------------N----T---.--.............-QKE-.-DNM---Q----------F-YT-L---------  656
B.PE.07.502_0525_wg5        ..-----..-.--TL--V------.----------I---A---------------M-K----------------------V-------Q----------------A----------.--.............--N--.-ENM---Q-EK--D---E--YT-L-K-------  641
B.RU.11.11RU21n             ..-----..-.--.---V------.----------I----------A-----------------------------------------------------------------T---.--.............--DD-.--N------E---D---EV-Y----QT------  641
B.TH.08.MERLBDTRC10         ..-----..-.--V---M------.----------L--------------------------------------------V-------------------------------D---.--.............--TE-.--NM-----E-------GI-Y---------R--  662
B.US.11.ES38                ..-----..-.--.F---------.----------V--------------------------------------------V--------------L----------------N---K--.............T-GE-.-DNM--IQ-EK--D---DI-Y--L--A-T--DI  661

C.AR.01.ARG4006             ..E----..-I--.L--V------.----------I-----------------S-----------M------------T-V--L----R------------------N----S---.--.............--DN-.-DNM---Q---------DT-YR-L--------Q  655
C.BR.07.DEMC07BR003         ..E----..-.--.---V--I---.----------------------------S-----------M--------------V--I---------------------P------T---.--.............-QQE-.-DNM---Q---------DT-YR-L---------  678
C.BW.00.00BW5031_1          ..E----..-.A..L--V------.----------I-----------------S-----------M--------------V--I----R-------------I---------S---.--.............NE-D-.-GNM-*-Q-----S---DT-YR-L-D--S---Q  640
C.CN.10.YNFL19              ..E----..-.--.---V------.----------I-----------------S------------------------T-V--I-----------L----------------S---.--.............THDE-.-DNM---Q-----S---NT-YQ-L----S---S  660
C.CY.09.CY260               ..E----..-.--.---V------.----------L-----------------S---K----------------------V--I-----------------------T----I---.--.............-KD--.-DNM---Q--K--S---DT-YR-L---------  651
C.ES.08.X2363_2             ..---REKR-V--.L--V------.----------I--------------------------------------------V--I-----------M------F----T----S---.--.............-ETE-.-ENM---Q-----S---YT-YN-L-K--I---T  667
C.ET.02.02ET_288            ..G----..-.A..L---------.----------I-----------------S-----------M-----------RT-V--I----Q----------------P-T----S---.--.............-Q-E-.-ENM---Q-----S----T-YR-L-V-------  649
C.IN.09.T125_2139           ..E----..-.--MM--M------.----------I-----------------S--------------------------V--I----------------------------S---.--.............-YDE-.-ENM---Q--K------DT-YR-L-D--I---R  675
C.KE.00.KER2010             ..E----..-.--.---VL-----.----------I-----------------S-----------M--------------V--I-----------------------N----S---.--.............-QTD-.--NL---Q--K------DT-YR-L----S---R  654
C.MW.09.703010256_CH256.w96 ..E----..-.--.---V------.----------IA----------------------------M--------------V--I----E-----------------------N---.--.............-QDD-.-GNM---Q-----S---DT-YR-L-D--I----  645
C.TZ.08.707010457_CH457.w8  ..E-Q--..-.--.M--M------.----------I-----------------S-----------I------------T----I-----------------------S----S---.--.............-ESE-.--NM---Q--K--S---NT-YR-L---------  637
C.YE.02.02YE511             ..E----..-.A-.M--------S.----------I-----------------S------------------------T-V--I--------------------------------.--.............-QGD-.-DNM---Q-----S---GI-YK-L--------E  654
C.ZA.08.705010534_CH534.w12 ..E----..-.--.----------.T---------I---------------------K-------M--------------V--L----R------M-------------Y--S---.--.............TQ-E-.-ENM---Q-----S---NT-YK-L---------  645
C.ZA.10.DEMC10ZA001         ..E----..-.--.---M------.----------I-----------------S-----------M--------------V--L----Q---------------------------.--.............--DD-.-KNM---Q---------YT-YR-L----T---Q  645
C.ZM.11.DEMC11ZM006         ..G----..-.--.---V------.----------I-----------------S--------------------------V--I---------------N-------N----S---.--.............-Y-D-.-ENM---Q-EK------DT-YR-L-V--T---Q  655

D.CM.10.DEMD10CM009         ..E----..-.I-.L--M------.----------I---------------------------------------L--------------------------H----N----S---.-R.............--ND-.-GNM---Q-E---DK--GV-YN---D--I----  643
D.CY.06.CY163               ..E----..-.I-.L--M------.----------IA--A----V----------------------------------------------------------H----N-A--ST--R--.............--NE-.-DNM---Q-E---E---GV-YT-L----L---R  657
D.KE.11.DEMD11KE003         ..E----..-.I-.L--M------.----------L-------------------------------------------------S---------V------H----N----S---.--.............-QNE-.-GNM--I--E---D---E--Y----S--I---A  649
D.KR.04.04KBH8              ..E----..-.I-.L--M------.----------V--------------------------------------------V--L-----------------------T----S---.-R.............-VDT-.-DNM---Q-E---D---GI-YD-L---------  641
D.SN.90.SE365               ..E----..-.I-.L-VM------.-----------A---------------------------------------------------R------------RH----N----S---.--.............--TE-.-QNM-----E---D---G--Y-------T----  649
D.TZ.01.A280                ..E----..-.I-.L--M------.----------L--------------------------------------------V----S----------------H---------S---.--.............--DD-.--NM-----EK--D---GV-Y-------V----  651
D.UG.10.DEMD10UG004         ..EKQ--..-.I-.L--M------.-----------A-------V-----------------------------------V----S--R------------RH---------ST--.--.............T--D-.-GNM-----EK--D---E--YN---V--T----  646
D.UG.11.DEMD11UG003         ..E----..-.I-.L--M------.--------V-L-----V-----------------------M-------V-----------S----------------H----N----S---.--.............-V-Y-.-ENM---Q-EK--E---GV-Y-------I----  656
D.YE.02.02YE516             ..E----..-.I-.L--M------.----------V-------------------------------------------------Q----------------H----S----S---.-R.............--DA-.-QNM-----EK--D---G--YA-L----I--D-  645
D.ZA.90.R1                  ..A----..-.I-.LASW------.----------V--------------------------------------------V-------R-------------H----T---DPE--.--.............TYDE-.-GNM-----E---D---GY-YT------V----  657

F1.AO.06.AO_06_ANG32        ..G----..-.--.-----F----.----------I-----------------------------M--------------L----------R-------------P-N----S---.--.............-QAE-.-ENM-----EK--S--SNT-YE---K------I  639
F1.AR.02.ARE933             ..K----..-.--.M----I---A.---G------I-----------------------------M--------------L-------R------L-----------T----S---.--.............TQKE-.-ENM-----EK--S--SNE-YR----A------  644
F1.BR.10.10BR_RJ015         ..-----..-.I-GL--V------.----------L-----------------S--------------------------V-----------I--L-----------S----S---.--.............-YTD-.-ENM---Q-EK--S--SEE-YR---Q-------  646
F1.CY.08.CY222              ..K----..-.-E.---V--T---.----------L---------------------K----------------------V--------------------------N----S---.-R.............-Q-D-.-KNM---D-EK--D---DT-YR-L-Q------Q  643
F1.ES.02.ES_X845_4          ..E----..-.--.----------.----------I--------------------------------------------V--------------------------N----S---.--.............--DY-.-ENM-----EK--S--SNT-YR---Q------I  641
F1.RO.96.BCI_R07            ..K----..-.A-.L-L-L-----.----------I---------------------K----------------------V--------------------------N----S---.--.............-QD--.-DNM------K--S--SNI-YN-L-Q--I----  657
F1.RU.08.D88_845            ..E-Q--..-.L-.M---------.----------I-----------------------------M--------------V-------Q--R-----------F---N----S---.--.............TYDE-.--NM---Q-EK--S-HSET-YR-L----S--DR  661
F2.CM.02.02CM_0016BBY       ..K--R-..-.--.---V------.----------I---------------------K-------------------------------------------------N----S---.--.............-Q-E-.-GNM---Q-EK--D---DT-YR----A------  644
F2.CM.10.DEMF210CM001       ..S----..-.--.L--M------.----------I-----------------S---K----------------------L--------------------------T----S---.S-.............PHDD-.-QNM---Q-EK--A---GT-YQ----A-D----  654
F2.CM.10.DEMF210CM007       ..N----..-.--.---V------.T---------I-----V-----------S-----------M--------------L-----F--------L-----------S----S---.--.............-Y-E-.-DNM---Q-EK--S---ET-Y----NA------  646

G.BE.96.DRCBL               ..E----..-.--.V--I------.T---------I-----V-----------S-----------------------R--V--L-----------------------N----T---.--.............-YNE-.-ENM--I--E---D---YH-Y----Q--I----  646
G.CM.10.DEMG10CM008         ..G----..-.--.L--V------.----------I-----V--------------------------------------V------------------------P-N----T---.--.............GYDE-.-ENM--------VS---QQ-Y----K-------  646
G.CN.08.GX_2084_08          ..-----..-.--TM---------.T---------L-----V-----------S--------------------------V--------------------------N----T---.--.............-YSD-.-DKM--IQ-E--VS---QH-Y------------  651
G.CU.99.Cu74                ..E----..-.--.MR-V------.----------I-----V-----------S--------------------------V--------------L-----------N--------.--.............-VDE-.-GNM--I--E-------AQ-Y------------  657
G.ES.09.X2634_2             ..G----..-.A..L--I------.----------I-----V-----------S-----------M--------------V--L----R------------------N----T---.--.............-YDD-.-DNM--VQ-E-------QQ-YT-L---------  650
G.GH.03.03GH175G            ..E----..-.--.L--V------.----------I-----V-----------S-----------M--------------V-------R------------------N----T---.--.............TYNE-.-DNM--IQ-EK------KQ-Y------------  654
G.KE.09.DEMG09KE001         ..G----..-V--.V--V-----S.----------I-----V-----------S--------------------------V--I-GF------------T-----P-N----T---.--.............TY---.-ENM--TQ-----S---QE-YR---N--T---I  651
G.NG.09.09NG_SC62           ..-----..-A--.M--------S.----------L--------------------------------------------V--M------------------------N----V---.-R.............-YD--.-DNM--IQ----VR---QQ-YT--K-A------  648
G.ZA.01.TV546               ..G----..-.--.L--V------.----------I-----V-----------S--------------------------V--------------------------N--------.--.............-XSE-.-DNM--I--E-------QX-Y----X--X----  647

H.BE.93.VI991               ..E----..-.--.M--F------.----------I-----------------S------Q----M--------------V--------------------------N----S---.--.............--DE-.-DNM------KQ-----DE-YR-L-V-------  658
H.CF.90.056                 ..E----..-.--.M--S------.----------I-----------------S------Q-R--M--------------V-------R------------------N----S---.--.............-QSE-.-DNM------KQ-S---EE-YR-L-V--T----  644
H.GB.00.00GBAC4001          ..E----..-.--.M--F------.----------I-----------------S------Q----M-------V------V--------------------------N----S---.--.............-YXX-.-XNM-----E-Q-D---XE-YX-L-I-------  651

J.CM.04.04CMU11421          ..E----..-.L-.---V------.T---------I-----V-----------S---K---------K------------V--------------------------T----TT--.-R.............-HDD-.-GNM--IQ-E-------GI-YT----A------  659
J.SE.93.SE9280_7887         ..E----..-.--.---V------.T---------I-----V-----------S---K---------K------------V--------------------------N--------.--.............-Y-D-.-ENM--IQ-E-------GI-Y-----A-----N  647
J.SE.94.SE9173_7022         ..E----..-.--.---V------.T---------I-----V-----------S---K--X------K------------V--------------------------N--------.--.............-Y-D-.-ENM--IQ-E-------GI-Y-----A-----T  649

K.CD.97.97ZR_EQTB11         ..-----..-.--.----------.----------I----------------------------QM-----------R--V-------R------------------N----S---.--.............-QSE-.-ENM---Q-EK--S-H--T-YR------I----  650
K.CM.96.96CM_MP535          ..-----..-.--.L--V-F----.----------I-----------------S-----------------------R-----------------------------N----S---.--.............-W-E-.--NM-----EK--G--SDT-YK------T----  641

U.CA.01.TV749               ..-----..-.--.----------.----------I-----------------S-------------K------------V--------------------------N----T---.--.............--DE-.--NM---D----V----AI-YG---Q-------  660
U.CA.99.TV721               ..-----..-.--.----------.T---------I-----------------S-------------K------------V--------------------------N----T---.--.............-YNE-.--NM-------------GI-YN---Q-------  663
U.CD.83.83CD003_Z3          ..E----..-.--.M---------.-----------A--A----------------------------------------V-------ES-----L-----------T----S---.--.............--DN-.-DNL---------S---QV-YG-L-D--K----  650
U.CD.90.90CD121E12          ..E----..-.--.M---------.----------IA--A----------------------------------------V-------EG-----L-----------T----S---.-R.............--ND-.-DNM------K--S---QI-YG---D--K----  648
U.CY.05.CY090               ..E----..-.--.L---------.----------L-----------------------------M-----------R--V--L----Q------------------N----S---.--.............TYND-.-ENL---Q--K--S---GI-YG------I----  644
U.CY.08.CY223               ..E----..-.--.M--I------.----------I----------A------S---K------Q--R------------V--L-----------------------N----S---.--.............-Y-E-.-DNM--LQ-E---D---VQ-YT----A-Y----  679
U.ES.10.DEURF10DZ001        ..G----..-.I-.M--F------.----------I---A-------------S---K------N--R------------V--L-A--R------------------N----S---.--.............-Y-D-.-ENM--QQ--K------GI-YE---K-------  645
U.GR.99.99GR303             ..E----..-.--.MA-II-----A----------V--------------R--S---Q---------R------------V-----------I-E--A---------T----S---.--.............-Q-LTYGKNM--L---K--S---QQ-YE------I----  655
U.NL.95.U_NL_95_H10986_D1   ..-----..-.--.L----F----.----------I-----------------------------------------------M-----------------------N----S---.--.............-Q-E-.-DNM--I--E---SK--GT-YR-----------  660

01_AE.AF.07.569M            ..E-Q--..-.--.---MIF----.----------I-----------------S-----------M--------------V----------KF--L------I---------ST--.--.............-Y-D-.--NL--T--E---S---NQ-YA-LT---D--DR  650
01_AE.CF.90.90CF11697       ..E----..-.--.---MIF----.----------L-----------------S--------------------------V----------KF--L------I----N----T---.--.............-YAE-.-DNM--I---K------NQ-YE-LTK-------  655
01_AE.CN.10.YNFL03          ..E----..-.--.---MIF----.----------I-----------------S-----------M----------------------Q--KF--L------I---------ST--.--.............TY---.--NM--I--E---G---NQ-YDILIK--D---R  651
01_AE.HK.04.HK001           ..E----..-V--.---MIF----.----------L-----------------S-----------M--------------V----------RF--L------I---------NT--.--.............-Y-E-.--NM--TQ-E---S---DQ-YAILT------D-  651
01_AE.IR.10.10IR.THR48F     ..E-Q--..-.--.---MIF----.----------I-----------------S--------------------------V----------KF--L------I--P------ST--.--.............T----.--NM--TQ-E---S---EQ-YEILT--------  659
01_AE.JP.x.JRC77AE          ..E-Q--..-.--.---MIF----.----------I-----------------S--------------------------V----------KF--L------I---------ST--.-R.............-Y-E-.-DNM--IQ-E---S---NQ-YG-LT-A-D--D-  648
01_AE.TH.09.AA090a_WG11     ..E----..-.--.---MIF----.----------I-----------------S-----------M--------------V-----------F--L------I----N----ST--.--.............-Y-K-.--NM--I--E---S---NQ-YDILT------D-  656
01_AE.TH.90.CM240           ..E----..-.--.---MIF----.----------I-----------------S--------------------------V----------KF--L------I---------ST--.-R.............-F-E-.--NM--I--E---S---NQ-YEILT------DR  649
01_AE.US.05.306163_FL       ..E-Q--..-.--.--TMIF----.----------I-----------------S----------RM--------------L----------RF--L------I---------ST--.--.............-Y-E-.-DNM--R--E---S---DQ-YG-LT-------R  628
01_AE.VN.98.98VNND15        ..E----..-V--.---MIF----.----------I-----------------S--------------------------V----------TF--L------I----N----S---.--.............-F-E-.-DNM--I--E---S---NQ-YAILT------D-  659
```

Draft of 30 September 2015, not for public release
hiv1prot_main.tex r8876–10546

```
                                                                               transmembrane domain   gp41 cytoplasmic tail start                glycosylation NGS
B.FR.83.HXB2                NEQELLELDKWASLWNWFNITNWLWYIKLFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTHLPTP..RG.PDRPEGIEEEGGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELL.GRR..............GWEALKYWWNLLQY.WSQE 806
A1.AU.03.PS1044_Day0        ------A-----N-----D-SQ------I--------I-----------IS--------------T-N-..E-.L---GRT------QG-------S-F---A-------------------S----T----.-HSSLK.......GLRL---G---L----V-.---- 820
A1.CH.03.HIV_CH_BID_V3538   ---D--A------------S-------I--------I------T----I----------------T-N-..E-.L---GR------QG-G------S-F---A-------L---------FA--AA-T----.-HNSLK.......GLRL---G---LG---L-.-GR- 804
A1.CY.08.CY236              ---D--A-----N-----D-SH------I--------I-----------IR--------------T-N-..G-.L---GR--G---Q--S------S-F---A-----N-----------FI--AA-T----.-HSSLK.......GLRL---G---LE---V-.--R- 812
A1.ES.05.X1608_8            --KD--A-----N-----D-SK------I--------I------T-F--I--------------LI-SA..GD.----R----G---Q--G------S-F---A-------------------FI--A--T----.-HNSLK.......GLRL---G---L----V-.-GR- 823
A1.KE.11.DEMA1IKE001        ---D--AF----N--S--K---------I--------I----------VI-K--------------N-..G-.L---GR------QG-G------S-F---A-------------------FT--AA-T----.-HSSLK.......GLRL---G---L----L-.-GR- 826
A1.RU.11.11RU6950           ---D--A----T------D-S-------I--------I----I-T----I--A--------L--LT-H-..E-.----GR-S-----QG-------S-F---A-------------------FI--AA-T----.--SSLK.......GLRLV--G---LG---G-.-G-- 813
A1.RW.11.DEMA111RW002       ------A--------T--D-S-------I--------I-----------I--------------I-T-NQ..GE.----R----G---QG------AS-F-----E---N--I------------AA-T----.--SSLK.......GLRL---G---LG--T--.-I-- 821
A1.SN.01.DDI579             ---D--A--------S--D-SK------I--------I-----------I--------------LT-N-..--.-----R------Q--------S-F-----------------------FT--AA-T----.-................QG---L--V-L-.-GR- 800
A1.UG.11.DEMA110UG001       ---D--A-----N-----D-S-------I--------I----IC---A-I--------------I-T-N-..--.---H-R------Q--N--T-SAS-F---A------F-----L---FI--AA-T----.-HSSLK.......GLRL---G---L----L-.-GR- 806
A1.ZA.04.04ZASK162B1        ------A-----------D-SR------I--------I----I-----ML--------------I-N-..E-.L---GR------QG-------S-F---A-----N-----------FI---A-T---..............GLRL--KGIQ-L----L-.---- 809
A2.CM.01.01CM_1445MV        ---DV-A-----N-----S---------I--------I--K--IVIF-VI------------L-IPT-R-..E-.L----R-K-G---Q---T-V---S-F---A-------------------CIS-AA-T-T--.-HSSLK.......GLRL---GI--L----L-.-GR- 812
A2.CY.94.94CY017_41         ---D--A-----D--S----SH-----RI--------I-------IITV---------V---IPT-S-..E-.----R-T--G---QG----------FF--A-------------------CI--AA-T----.-HCSLK.......GLRL---G--NL----L-.-GR- 815

B.BR.10.10BR_MG029          -----------------D--H------I---X----I------T-------------------RF-A-..--.T---G-------G----RI--D-FF-IF-V--------I-----------------.---...............-------------I---.---- 818
B.CA.07.502_1191_03         ---------------S--D---------I--------I------T-I---S--------------R--AQ..--.-------T--G-------T-GW-AT-F-E---I---N-WI-L-----------LA------.---...............-R-V----------.-G-- 812
B.CH.08.M2_0803101_NFLG8    ------------N--S--D-S-------I--------IV---I-------R---------------V-..--.-------T---------T----D-F-------------------------T-G--.---...............------L---------.---- 835
B.CN.10.DEMB10CN002         --L---K--Q--------D--R------I--------I--K-----------------------R----..--.----------------------GP---RFFI-----WN-----Q--------------I-.---...............-------G--I--.-T-- 808
B.ES.10.DEMB10ES002         ---D--------------E--S------I-------I-----I-T--------------------A-..--.L-----------D---N--RP--H-F--I--V--------------------A-T----.---...............---------------.---- 814
B.FR.11.DEMB11FR001         -----------S------D--S------I--------------CT--------------------AQ..--.---------G----KS--T-GC--D-F-TI--V---N--R----------S-AA------.---...............---------------.---- 790
B.GB.05.MM45d213_GN1        -----------------D---------R-----------------L-----------L----------..--.---------G----------GG----F-T---V--QN-------L-----T---------.---...............-----I---------.-I-- 812
B.HK.06.HK003               ---D--A-----------D---------I-----A--------T-I--------------------V-..--.-----G------N--S---SP--G-L--I--V---------------------I-SA---.-H-...............----I---C-------.---- 803
B.HT.05.05HT_129389         ------Q-----------D-S-------I--------------T--------------------RF-A-..--.-E----T-----------T-----F-P-------------H---F---AA-V--I-.---...............---------------.---- 812
B.JP.12.DEMB12JP001         -----------------D--S------I----I--------------------------------RF-A-..--.---------G----------S---S-F-P--------------------------I-.---...............----I----------.---- 814
B.KR.07.HP_18_07JHS10_3909  -----------S--------SQ-----RI--------I----I-S----------------------F-S-..--.---------G----N--G-GP----F-T---V--------I-RH--------A-S----.-L-...............---V-----------.---- 807
B.PE.07.502_0525_wg5        -----------------D--Q------V----I---I-----CV-IN---------------------..--.-G-----------D-N--T-A---T-F-P-------------------------A---L---.---...............---I-----------.---- 792
B.RU.11.11RU21n             ------A--------S--D---------I--------I----I-----V-------------R---Q..--.-------------------T-V-----LF-------------L-----------AA------.-L-...............----------------.-I-- 792
B.TH.08.MERLBDTRC10         --L---------------D--K------I--------------------------------L--RF-A-..--.-GG-------------------GT----F-T------QN------R------------V--------.---...............----------------.-G-- 813
B.US.11.ES38                -------------------S-------I--I----II------T-------------------R--AQ..--.-------T---------E--VP----F--IL-V--------------------LA-T--F-.-H-...............----I------------.---- 812

C.AR.01.ARG4006             --KD--A----QN-------S-------I----I---I----I------------------LI-N-..--.-GGLG--------Q-K-------S-F-S-A------------RQ----I--AA-AA-R-.--S...............---I---LGS-V---.-GL- 806
C.BR.07.DEMC07BR003         ---D--A----QN-----G-S------RI--------I----I------------------LI-S-..--.----GE-------Q--G-----S-F-P-A--------------------F---AA-A----.--SSLR.......GLQR---I---LGS-V---.-GL- 836
C.BW.00.00BW5031_1          --KD--A--S-KN--S--D-S-------I--------I----I------------------LT-SQ..-E.L--LGR-------Q--------A--F---A-E----------RH---FI---V-A----.--SSLR.......GIQK--D----LGS-V---.-GL- 798
C.CN.10.YNFL19              --KD--A--S-KN--T--D---------RI--------I----I------------------LI-NQ..G-.L--LGR-------Q-KN--V-----F---A-----N-------L---FI-VTA-A----.--NSLR.......GLQT--R----LGG-V---.-GL- 818
C.CY.09.CY260               ---D--A--S-N----------K------I--------I----I-----------------L--LI-N-..--.-----G--V--E--QQ-RD--S--I-F-VVVAW----S---FC----SYF-LIA-T--V-V-HILLG.......GLRR-GVIMYFVS-VVL-.---L- 810
C.ES.08.X2363_2             --KD-------NN--T--D---------I--------I------I----------------L--LT-N-..--.----LG--------Q--S--------F---A-------------------FI--AA-T--Y-.-H-...............-------LGS-V---.-GL- 818
C.ET.02.02ET_288            ---D--A----KD--S--S--Q------I--------I----I----L--------------LT-N-..-E.----LR---------Q------------F--IA---I------------------I--AA-A----.--SSLK.......GLQR---I---LGSIV---.-GL- 807
C.IN.09.T125_2139           --KD--A--N-KN-------S-------I--------I----I------------------LI-S-..G-.----L-R--------Q-------------F---A-----N-------------FI-VTA-A----.-HHSLR.......GLRK-------LGS-V---.-GL- 833
C.KE.00.KER2010             --KD--A--S-NN-----D--R---------I--VS-I----I-V-F---------------LT-N-..--.----LGR-------Q-K-G-------F---A-------------------CI---A-V----.--SSLR.......GLQK-------LG-CV---.-G-- 812
C.MW.09.703010256_CH256.w96 --KD--A--S-QN--------K------I--------I----I------------------LI-NQ..-E.----LGR-------P--N-------F---T-------------------FI-VTV-V----.--SSFR.......GLQR-------LGS-V---.-GL- 803
C.TZ.08.707010457_CH457.w8  ---D--A--S-KN--S--S---------I--------I----I--I---------------LT-N-..--.----LGR---G---Q-N---T---S-F---A------F-----------FT---A-V----.--SSLR.......GLQK---T---LGS-V---.-GL- 795
C.YE.02.02YE511             --KD--AS-S-NN---------------I--------I-V--I-T----I----------L---LT-NS..--.----LG--------Q--N---Q-A--F--VV-N--------------FT---A-A----.--SSLK.......GIQR-------LGS-V---.-GL- 812
C.ZA.08.705010534_CH534.w12 ---D--T--S-NN-----S--K------I--------I----I------------------LT-NQ..-E.L--L-R--G---Q----------F-S-A--------------------FI---V-A----.--SSIK.......GLQR---I---LGS-V---.-IL- 803
C.ZA.10.DEMC10ZA001         --KD--A----QN--S------------I--------I----I------------------LN-N-..--.----L----------Q--G-----S-F-P-A----------H---FI---V-V--R-.-SSSLR.......GLQR-------LGS-V---.-GL- 803
C.ZM.11.DEMC11ZM006         --RD--A----N---T--D-S-------I--------I----I------------------LT-SS..-E.----LGR-------Q-KG-----S-F---A----------H---FI--AA-AA---.--SSLR.......GLQR-------LGS-I---.-GL- 813

D.CM.10.DEMD10CM009         --K---------------S--Q------I---V----I--------------K--------L--A-..--.---------G---Q--G---------FS--------N-----------I--AA-T-N--.---...............------L--------.--R- 794
D.CY.06.CY163               --KD----------R---D-SK------I---------IR-K-------PIH----R---P--H-L--AS..-P.-A---P---G---Q--G---------FS--------N-----------I--AA-T-G--.---...............------L-------.-R-- 808
D.KE.11.DEMD11KE003         ------K--Q--------S--K------I--I-----I----------V--K-------------L--A-..--.----------------QG---------FS--------N------RH--------A-------.---...............-----I--L-------.-I-- 800
D.KR.04.04KBH8              --KD--------------S--R------I--I-----I----I-----L-----------------PF-A-..--.---------T--G---Q--GK-----S-LS--------N--------------------A-------.---...............----V--LG------.-I-- 792
D.SN.90.SE365               ------A-----------D---------I---V----I----I-----L-------------------L--A-..--.----------------QG---------FS--------N-----------I--AA-------.---...............----V---L-S-----.-I-- 800
D.TZ.01.A280                --K---------------S--K-------------I--K---T-F-L------------L--L--AS..--.----------------QG-G--------FS--------N------RH----I--A---G--.-H-...............-----I--L-------.-I-- 802
D.UG.10.DEMD10UG004         ------Q--Q----------S------RI--------I----------L---------------L--A-..--.----------------QG-------LG-LST-------N-I----------I--AL--L---.---...............-----SR-L-------.-I-- 797
D.UG.11.DEMD11UG003         -----WQ-----------T---------I--------I----------------------------L--A-..--.-----------------G----V-----FFP-------N------------------AA-------.---...............-----I--L-------.-I-- 807
D.YE.02.02YE516             --K---------------S-SQ------I-------S-I----------M-----------L--L--A-..--.----------------PGKGT-------FS--F-E---N---Y---Q----I---A-T-D--.---...............------L--------.--R- 796
D.ZA.90.R1                  --K-------------------------I--------I----------V---------------L--AQ..--.-----------------G--K--------FYS-------N------R----------AA-------.---...............------L-------.---- 808

F1.AO.06.AO_06_ANG32        ------A--------S--D---------I--------I-----------------K------L--LI-S-..--.---------GD--Q---K-V---S-F-P-V-----N------RH---FI--AA-T-NRG.L-S...............--------L---V--.--R- 790
F1.AR.02.ARE933             ------A-----------S-----------I----FI-----------------K------L--LI-S-..-E.---------G---QGK---V---T-F---A-----N------RH---FI--AA-T-DRG.L--...............------LLG--TL-.---- 795
F1.BR.10.10BR_RJ015         --L---------------D--H------I--------I------T---------K---------L--S-..--.---------G---Q-----V-----FF--A-----N------RL---FI-T-A---GKG.L--...............S------LG-IIL-.--K- 797
F1.CY.08.CY222              --K---A-----------D-S-------I----I---I----I------------K------L--LI-S-..--.---------G---Q-----V-----FFS-V-----N------RH---FI--AA-T-NRG.L--...............---I---L---I--.-I-- 794
F1.ES.02.ES_X845_4          --K---A-----------D-S-------I--------I-----------------K--------LI-S-..--.---------G---Q-----V-----F---V-----N------RH---FI--AA-T-NRG.L--...............--------L---A--.--R- 792
F1.RO.96.BCI_R07            ------A--------T--D--K-----RI--------I-----------------K------L--LI-S-..-E.---------G---Q--G--V-----F---V-----N------RH---FI--AA-T-NRG.L-G...............--------L---T--.---- 808
F1.RU.08.D88_845            --K-----------S--D-S-----RI--I-----I--------F---S---K------L--L---Q..--.----------------QG----V---S-F-TIV-----N----N-RLW--FI--AA-V-DRG.LK-...............--------L---A--.-G-- 812
F2.CM.02.02CM_0016BBY       ---D--A----D---S--T--------RI---V----I-----------I------------L--LI-S-..--.-----G----------Q-K---V---S-F---A----------------RH---FI--AA-T-DRG.LKG...............---V---L---A--.-GR- 795
F2.CM.10.DEMF210CM001       ---D--S-SQ-D------S--Q------I--------I------------R--------V-L--LI-S-..--.----------------QGK---V---S-F-S-V-----I------RH---FI--AA-T-TRG.LQL...............---T---L---TLH.-G-- 805
F2.CM.10.DEMF210CM007       ---D--A----D------A--H------I--------I---------A-IKK---------L--LI-S-..--.-----G---------Q-K---V---S-FF--V-----N------RH---FI--AA-T-NRG.LE-...............--K----L---A-H.---- 797

G.BE.96.DRCBL               ---D--A--Q-----S--S-S------RI-V------I---------------------------L-HHQ..-E.----A----G---Q---------S-F---A-------------------FI--AA-T----.--NSLK.......GLRL-------L----L-.-AR- 804
G.CM.10.DEMG10CM008         ------A-----------D-S------R--VI-----I---------------------------LTHQQ..-E.-----R---G---QG--------S-F---A-----N---WC------FT--AA-T--I-.--SSLQ.......GLSL.......V----L-.-G-- 797
G.CN.08.GX_2084_08          ---D--A-----------D--K------I--------I----I-----------------------THHQ..-E.-----R-------Q-K---V---S-F---A-------------L------FI--AA-T----.K-SSLQ.......GLRLV--G---L----L-.-G-- 809
G.CU.99.Cu74                ---D--A--------T--D-S-------I--------I------T---LI----K---------PTHHQ..-E.-----R---G---Q--N------S-F---A-------------------V--AA-T----.-HSSLK.......GLRL---G---L----V-.-G-- 815
G.ES.09.X2634_2             ---D--A--T--------Q-SQ------I--------I-----------I----K---------LTHHQ..-E.----GR--------Q---K-----S-F---V--------------H---FI--A--T----.--NSLK.......GLRL---G---L----L-.-GR- 808
G.GH.03.03GH175G            ------A--------S--D--K------I--------I------T-------------------LTHQQ..-E.---L-RT--G---Q-----V---S-F---A-------------------FA--AA-T----.--SSLK.......GLRL---G---L----L-.-V-- 812
G.KE.09.DEMG09KE001         ------A--------S-LS--------RI--------I----I---------------------LTHHQ..-E.-G---R---G-----K-------S-F---A-----N---L------FI-LAA-T---...............GLRLS--G---L----V-.-G-- 802
G.NG.09.09NG_SC62           ---D--A--S-----D--A-SK------I--------I-FK-----F-----------------LPHHQ..-E.-----R---G---Q---------S-L---A---V---------------FIS-AA-TA---.--SSLQ.......GLRL--AG---L----L-.-G-- 806
G.ZA.01.TV546               --HD--A-NN-----S--D--K-----RI-V------I--X----------------V---IPTHHQ..-D.----R--------Q--GK-V---S-F---A--------------------S--V-TL---.-HNSLK.......GLRL---G---LKS--E-.-G-- 805

H.BE.93.VI991               ---D--A-----N-----S--------RI-------II-------------------L--LI-NQ..--.----RE--------Q-----------F-P-V-E---N------R-------S--A-T----.---...............--------LLG---L-.-G-- 809
H.CF.90.056                 ---D--A--------T--D-SH------I--------I----I-------------------LV-N-..--.----------T--G---Q-----V-----F-PVV-------S----RL--------V-T----.---...............-R------L-------.-G-- 795
H.GB.00.00GBAC4001          ---D--S----S------D-SX------I---VI---I------T----I-S-------------LT-NA..-R.-----------------QG----TP--S-FFI-V-Q--WN-----R----------A-T----.---...............-------L-------.-G-- 802

J.CM.04.04CMU11421          ------A--Q-GN-----D-SK-----RI--------I----I---FA----------------LT-HL..AE.A-----T--G---QG-TX-------FF--A-E---N--------------FI---A-T--T-.---...............---I---LG--AW-.-G-- 810
J.SE.93.SE9280_7887         --KD--A----TN-------S-------I----I---I----I----A----------------LI-N-..TE.A---G----G---QG-T---------F---A-----N-------------FV--AA-T-GT-.-L-...............---I---LV--VW-.-G-- 798
J.SE.94.SE9173_7022         --KD--A----TN-------S-------I----I---I----I----A----------------LI-N-..TE.A---G----G---QG-T---------F---A-------------------FV--AA-T-GT-.-L-...............---I---LV--VW-.-G-- 800

K.CD.97.97ZR_EQTB11         ---D--A-----------D-S-------I--------I------T---V----------------LT-S-..--.---------G---Q-K---V---S-F---A-----N------RH----V--A---LDRG.LKG...............S------L---IL-.-G-- 801
K.CM.96.96CM_MP535          ---D--A-----------D--K------I----I---I---A-----V-----------------LI--S.--.A-------------Q-KN--V---S-F---A-----N------RQ--N-I------L-RG.L-G...............-------L---V--.---- 792

U.CA.01.TV749               ---D--A--Q-----S--D--------RI--------I----I-T-I-M------------L-I-P-N-..E-.L--L-R-AG----P--T---------F---A-N-----F-----------V--AG-T----.-HSSLK.......GLRL---G---LG---L-.-G-- 818
U.CA.99.TV721               ---D--A----T------D--K------I---V---XI----I---I--I-------------I-T-S-..E-.L--L-R-AG----PG-T-----AD-F---A-----------F---------V--AA-T----.--SSLK.......GLRL---G---L----A-.-G-- 821
U.CD.83.83CD003_Z3          S-KD--------------D---------I--------I------T-F--I---------------L-----..--.----GRT------E-NN--V-----F---A-E-------I---------I---VK...................GLRR-------LG--VL-.-G-- 797
U.CD.90.90CD121E12          ------------------D-S-------I--I-----I------T----I---------------LI-----..--.----G-T------K-NT--V-----F---A-------------------FIS--AK...................GLRR-------LG--V--.---- 795
U.CY.05.CY090               --KD----SN--------D-S-------I--------I------S----K---------------LI-R-..-E.---------GG---Q-----V-----F---A-----N----L--H--T-V---A-T-N........K.......GLQGS--I---LG--AL-.-G-- 795
U.CY.08.CY223               --KD-----N--------T-SK-----RI--------I-----------I--------------IPT-N-..E-.L--LGR--G-----Q--T------G-F--IA-----N-------Q---FI---V----T-.---...............-------LGS-V--.---- 830
U.ES.10.DEURF10DZ001        ---D--A----S---S--D-S-------I---V----I--------FNLI--------------PIHLQ..-E.L--L-RT--G---Q---------S-F---A-------------Q--N-V--AA-A----.-HSSLK.......GLRL-------L--I-S-.-G-- 803
U.GR.99.99GR303             ---D--A--TX-------K-S------RI--------I----------V-K--------------IPI-S-..--.L-------------Q-K-------S--FS-A-----N------------FI---A-T--I-GTQ#...............--QX---LGEPFC-#-V-- 806
U.NL.95.U_NL_95_H10986_D1   ---D--A--T-----S----S-------I--------I------T----I---------------LF-SS..--.-------T--R---Q------------F---A-----N------RH--------V--L-RG.L-G...............S------L-S-V--.---- 811

01_AE.AF.07.569M            --MD--K--Q--------D---------I--------I----I---------------------PSHHQ..-E.---------G---QG----V---S-F---A-------------------FI--A--T----.-HSSLK.......GLRR---G---LG---L-.-G-- 808
01_AE.CF.90.90CF11697       ---D--A-----N-----S---------I---V----I--------------------------LTHHQ..-E.-----R---G---EGK-------S-F-S-A-------------------FI--AA-T----.-HSSLK.......GLRR---G---LG---S-.-G-- 813
01_AE.CN.10.YNFL03          --KD--------------D--K------I--------I----I---------------------PTHHQ..-EP----R---G---QG---V-----FF---------------L---FIS-AA-T----.-HSSLK.......GLRR---G---LG---L-.-G-- 810
01_AE.HK.04.HK001           --KD-----Q--------S--K------T--------I----I--------------------IP-HQQ..-E.---GR------GQ-----V---S-F---L-----N-------------FI-----T----.-HSSLK.......GLRL---G---LG---L-.-G-- 809
01_AE.IR.10.10IR.THR48F     --KD--------------D--K------I--------I----I------------------L-IPTHQQ..-E.-----R---G---Q-----V---T-F--IA---------------K--A--AA-T----.-HSSLK.......GLRR---G---LG---A-.---- 817
01_AE.JP.x.JRC77AE          --KD--------------S--H------I-------II----I---------------------PTHHQ..-E.---------G---QG----V---S-F---V-----------------------AA-T----.-HSSLK.......GLRR---G---LG---I-.-G-- 806
01_AE.TH.09.AA090a_WG11     --KD-------T------S---------I--------I----I------------------L--LSHQQ..-E.L--------G---QG---------S-F---A-------------------FI--AA-TLG--.-...............RQ---G---LG---L-.-G-- 807
01_AE.TH.90.CM240           --KD--------------D---------I--------I----I---------------------PSHHQ..KE.---------G---QG----V---S-F---A----------------------T--AA-T----.-HSSLK.......GLRR---G---LG---L-.-G-- 807
01_AE.US.05.306163_FL       --KD--------------D---------I--------I----I-------X-------------PIHHQ..-E.-----R---G---Q-----V---S-F---A-------------------FI---A-T----.-HSSLK.......GLRR---G---LG---V-.-G-- 786
01_AE.VN.98.98VNND15        --KD------------------------I--I---S-I----IC--HT-------D---S-L--PTHHQ..-E.-----R---GRC-QG----V---S-F---A--------------L---FI--AA-T----.-HSSLK.......GLRR---G---LG---L-.---- 817
```

Draft of 30 September 2015, not for public release

hiv1prot_main.tex r8876–10546

```
                                                                      gp41 cytoplasmic tail start
                                            transmembrane domain                                        glycosylation NGS  *
B.FR.83.HXB2                   NEQELLELDKWASLWNWFNITNWLWYIKLFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTHLPTP..RG.PDRPEGIEEEGGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELL.GRR.............GWEALKYWWNLLQY.WSQE  806
02_AG.CM.10.DE00210CM013       ---D--A----D---S------------I--------I--------FA-I--------------LTHHQ..GE.----GRT--G---Q--N--V---S-F---F-----N--I--------FV-----T----.-HNSLK.......GLRL--G----L----S-.-G--  834
02_AG.CY.09.CY256              ---D--A-----N-----D--R------I--------I------T-FALI--A-----------LTHHQ..-E.-----R---G---Q--S--V---S-F-T-V-----------------FA--AA-T----.--SSLK.......GLRL-------LG---S-.-G--  807
02_AG.ES.06.P1423              ---D--A---------------------I--------I---------N-I--------------LTHHQ..-E.-----R---G---Q-----V---S-F---A-E---N----C-------V--AA-G--I-.-HSSRR.......GLEL------FLG---S-.-GR-  812
02_AG.GW.05.CC_0048            ---D--A--------S--D--K------IV-I-----I------T--T-I--------------LYHHQ..-D.-G---R---G---Q-----V---S-FF--A------------------F---A--T----.-HSSLK.......GLRL-------L----L-.-G--  817
02_AG.KR.12.12MHR9             ---D--A--R--------D---------I--------I---------T-I--------------LTHHQ..-E.-----R---G---QG--K-V---S-F----------------------V---A-T----.-HGSLK.......GLRL------HL-S--A-.-G--  819
02_AG.LR.x.POC44951            ---D--A-----------D--K------I--------I---------T-I--------------LTHRQ..-E.-----R---G---Q---K-V--LS-F---A-----------------FA--AA-T----.-HSSLK.......GLRL-------L----S-.-G--  802
02_AG.NG.09.09NG_SC61          ---D--A-----------D--K-----RI--------I---------ALI--------------LARHQ..-E.---------G---Q--N--V---S-F---A-----------------F---AA-T----.-HSCLK.......GLGL-------L----S-.----  811
02_AG.NG.x.IBNG                ---D--A-----------D--------RI--------I---------T-I--------------LTHHQ..-E.-----R---G---Q-K---V---S-F---A------------------I--AA-T----.-HNCLK.......GLRL--G----L---IS-.-V--  804
02_AG.SE.94.SE7812             ---D--A--------D--S-SS-----RI--I-----I---------A-I--------------LTHHQ..-E.-G---R---G-GQ-----V---S-F---A-----------------FVS--A-T----.-H-...............-------L----S-.-G--  809
02_AG.SN.98.98SE_MP1211        ---D--A-----------D--------LI-L-V----I---------A-I--------------LTHHQ..--.A--------G---Q-----V---S-F---A---------L------FV--AA-T----.-HSSLK.......GLRL------HL----S-.-G--  811
03_AB.RU.97.KAL153_2           ----I-A-----------D-SK------I------------I---------G-----------R--AQ..--.----------D------T------RF------------F-I--H-------AA------.---...............----------------.-I--  789
04_cpx.CY.94.94CY032_3         ---D--A-----------S--K------I--------I----I------------------L--LI--TQ..--.L--G-T-------Q--S--------F-P-------N------RH--N-----A-T----.-I-..................-------L--F-L-.-G--  805
05_DF.BE.x.VI1310              ------S--Q--------D--K------I--------I------T---------K---------PF-A-..--.-----ET--G---Q--G---------FS-------N-------H----T---V-------.---..................-------L-S-P--.--R-  811
06_cpx.AU.96.BFP90             ---D--A----N--S--D----------I--------I--------I--------------L--LI-N-..T-.A--GE---G---QG-T---------F---A-----------------FG--AA-T--I-.---...............---I---LG--IC-.-G--  809
07_BC.CN.98.98CN009            --KD--A--S-KN--S--D---------I--I-----I---I-------------------LT-N-..G-.---LGR-------Q-K-----------F---A-----N-----------FI-VAA-V----.--NSLR.......GLQR-------LGS-V--.-G--  809
08_BC.CN.97.97CNGX_6F          --KD--A--S-KN--S--D---------I--------I---I------------------ILT-N-..G-.-G-LGR-------Q-KTS---------F---A-----N-----------FI-LTA-G----.--NSLR.......GLQR-------LGS-V--.-GL-  814
09_cpx.GH.96.96GH2911          ---D--A--S-----S--D--------R---------I---I----V--------------LT-N-..-E.---------G---Q-KV-----S-FF--A-----N-----------FI--AA-T--I-.-H-..............V-QI---LG--A--.-G--  792
10_CD.TZ.96.96TZ_BF061         ------Q--------S--S---------I--------I----C----L-----------LH-L--A-..--.-----------------G-G-----D-F-T--------N-------------I--A-----------.--M.................I---L-------.-I--  790
11_cpx.CM.95.95CM_1816         ---D--A--------T--D---------I--------I-----------C------V-L--LI-SH..KE.A-S-G---G---Q--T-----S-F---A-----N---Y---X----I------LGT-.---.................---T---LG--T--.----  815
12_BF.AR.99.ARMA159            ------A-----------D-SK-----RI--------I-----------K-----L---I-S-..-E.---------G---QGKV--V-----L-T----------------------------V-------.---.....................-------------R-  792
13_cpx.CM.96.96CM_1849         ---D--A--Q--------S-SK------I--------I-----------KS-----------LTHPK..-E.-----R---G---Q---------S-F---A-----------------FI--A--T----.-HSSLK.......GLRL---G---L----V-.-GR-  817
14_BG.ES.05.X1870              ---------------S--D---------I----I---I-----------I----K---------PTHHQ..-E.----G---------Q-K------S-F---A-------------X--FI-VAA-T----.--NSLK.......GLRL---G---L----L-.-GR-  815
15_01B.TH.99.99TH_MU2079       ------A-----------D---------I--------I----------------------L--PIHQQ..-E.A---E---G---QG-NG-V---S-F---A-----------------K--VS-A--T----.--SSLK.......GLRL---G---LG---L-.-G--  804
16_A2D.KR.97.97KR004           --KD--A-----------D-S-------I--------I-----M-II-V-----------V---IPP---..ED.---HGR--DG---QG----V---S-F---V-E--------------CIS-AA-T----.-HSSLK.......GLRL---G---L----L-.-GR-  812
17_BF.AR.99.ARMA038            ------A-ET--------D-S-------I------------V--T---------K---V-L--RP-S-..-E.---------G---QGK---------F-P---------------C-------W--A------.---................-R---R-------------  798
18_cpx.CU.99.CU76              ---D--A--------T--D--------RI--------I-----------K------------LI-NQ..L-.------T-G---QG-G--V---S-F-T-A-E-----------------W-AA-T----.--SSLK.......GLRL-------L----V-.-G--  813
19_cpx.CU.99.CU7               ------A----GK-----D---------I--------I-----------------------LTRH..---.---H-R---G---Q-K---------F---A-----------------FIS-AA-T----.-HS...............---G---L----L-.-G--  751
20_BG.CU.99.Cu103              ---D--A-------------S----------------I------T----I---------------THHQ..-E.-------GG---Q-N---V---S-F-P-A---------------------AA-T----.--SSLK.......GLRL---G---L----G-.-G--  799
21_A2D.KE.99.KER2003           --K------Q---------SQ-----RI---V----I----------L-------------L--A-..--.------------------G------L--LS-I------N-------------I--AG------.---..................-E-G---L-------.-GR-  802
22_01A1.CM.01.01CM_0001BBY     ---D--A---SN--T--D-S--------I--------I-----------I-------------LT-NQ..E-.L---GR---G---Q--NK-----S-F---A-----------------FI--AA-T----.-HSSLK.......GLRL---G---L----I-.-G--  798
23_BG.CU.03.CB118              --RD--A-----------D-S-------I--I-----I-----M-----I-------------LTHHQ..-E.---------G---Q--G-----AS-F-PIV--------------------AA-TL---.--SSLK.......GLRL---G---L----L-.-G--  805
24_BG.ES.08.X2456_2            ---D--A--------T--D-S------RI--------I------T-I-LI-------------THHQ..-E.----GE--GG---Q-KG-----S-F-T-A--------------------AA-.....................GLRL---G---L----L-.-GR-  799
25_cpx.CM.02.1918LE            ---D--A--------S--S-S------RI--------I------T-I------------------PHHQ..-E.----G--------Q------------F-P-V-E---N-----------FI--AA-T----.-HSCLK.......GLRL---G---L-------.-GR-  807
26_AU.CD.02.02CD_MBTB047       ---D--A--E--------S-SK------I--------I---AL--IT-----------------LF-S-..E-.VA--G--------Q--T---------F---A-E---N---C---H---F---AA-TL-I-.-HSSLK.......GLRR---G---L----A-.-G--  808
27_cpx.FR.04.04CD_FR_KZS       --KD--A-----------D-S-------I--------I------T-F--I-------------I-T-N-..E-.L--LGR-------QGKS-----S-F---A-----N----L----K-CA--AA-T----.-HSSLK.......GLRL---G---L-------.-G--  813
28_BF.BR.99.BREPM12609         ---D--Q--T--------D---------I------------C--------------------R--VQ..-E.----G---------N------AD-F--I--E-------------------------------.---.....................------.-G--  818
29_BF.BR.01.BREPM16704         ------Q-----N--S--D--K-----R----------------------------------R--V-..--.-----------D---G------D-F--I-------------------------A------.---.....................--------L-  820
31_BC.BR.04.04BR142            ---D--A----QN--S--D--K------I--------I----I-------------------RT-I-..--.---LGE-------Q-------V---S-F---A-----------C------T--AA-A----.--SILR.......GLQR---I---LKS-V--.-GL-  799
32_06A1.EE.01.EE0369           ------A--------S--D-S-------I-----------------------V-L--RI-N-..T-.A--GE---G---QGNTK----AT-F---V-----------------FG---A-T--T-.-H-...............---I---LK--VC-.-G--  800
33_01B.ID.07.JKT189_C          --KD--------T---S----S------I--------I---I--------------------PIHQQ..-E.----GR-------Q-----V---S-F---A------------------V--AA-T----.-HSSLK.......GLRL---G---LG---S-.-I--  809
34_01B.TH.99.OUR2478P          --RD--------T------D-S------I--------I---I--------------------PTHHQ..-E.-----R---G---QG-N--V---T-F-S-A---------------------A------.-H-...............-----R---------.-I--  793
35_AD.AF.07.169H               ---D--T-----N-----S-S------RI--------I---------A-I--------------LT-SQ..GD.LG--GR-------Q--T------S-F---A-----------------FS---R-T----.-HSSLK.......GLRL---G---L----L-.-G--  804
36_cpx.CM.00.00CMNYU830        ---D--A--Q--------D--K------I--------I---------NVI----------P---LTRHQ..-EE-----R-K-G---Q---K-V-----F---A-----------------FI--AA-T----.-HNSLK.......GLRL---G---L----V-.-G--  807
37_cpx.CM.00.00CMNYU926        ------A----TN--S--S--------RI--------I-------------------------LTHHQ..-E.-----R---G---QG----V---S-F---A-----------------FI--AA-T--I-.-HSSLK.......GLRL---G---L----V-.-G--  802
38_BF1.UY.03.UY03_3389         --K---------------D-SK------I--------I---I--------------------RF-S-..-E.---------G---Q-K---V---S-F---A-----N------RH---FT--AA---GRG.LN-...............---I-Q---------.-I--  794
39_BF.BR.04.04BRRJ179          ------A--Q------------------I--------I---------A------K---------I-S-..PE.---------G---QGK---V---S-F---V---V-N------RH---FI-VAA---NRG.L--...............------LLG---L-.----  794
40_BF.BR.05.05BRRJ055          ---------------S--S---------I-------------T------------------IRP-AA..--.----G---------------G----L--I--------------------------AA------.---.....................-----------K-.----  801
42_BF.LU.06.luBF_18_06         ---D--A-----------D---------I-------------S------------------RH-AQ..--.------T------------RP--D-F-T---V-----F-----H-------AA------.---...................---V----------.-R--  790
43_02G.SA.03.J11223            ------A-----X-----D-S------RI--------I-----------X-------------LTHHQ..-E.---------G---Q------------F---A-E-V-N-----------FI--AA-T----.--SSLQ.......GLRL---G---L----L-.--R-  796
44_BF.CL.00.CH80               ---D--A-----------S--------RI--------I------T--------K-----L--PI-N-..--.---------G---QG----V-----F---V-----N-----RL---FI-TAA---DRG.LKG..................-------LG------.-G--  794
45_cpx.FR.04.04FR_AUK          ---D-------TN-----D-SH------I--------I--------V--------------LI-N-..--.---------G---Q-K-------S-F---V-----N-----------FI--AA-T--T-.---................--QI--HLG--A--.-G--  795
46_BF.BR.07.07BR_FPS625        ---D--AF----------D-S------RI--------I---I-----------K-----L---I-S-..EE.----R----G---QG----VG---FF--V-----N-----RH---FI--AA---DRG.LK-..................-----LLG--AW-.-G--  785
47_BF.ES.08.P1942              ------------------D-S-------I-----------------I----------------H-A-..--.----G---------------G-SAS-F-T---I-----------H-------AA------.---..................L--------------  798
48_01B.MY.07.07MYKT021         --KD-----------------------RI--------I-----...A-----------------P-HQQ..-E.-----R-------QG----V---S-F---A------------------AA-T----.-HSSLT.......GLRR---G---LG---I-.-A--  814
49_cpx.GM.03.N26677            ---D--S----TN--S--D-S------RI--------I---I---A-I--------V-L--LI-N-..TE.A---G---G---QG-T---------F-SIV--------C--RH---FV--AA-T--T-.-H-...............---I---LG--VL-.-GK-  805
50_A1D.GB.10.12792             --K------------------K------I--------I---I--L-----------------L---..--.-E------------QG-G------F-T---E----V----------I--AA---I-.-H-...............---I--L-------.-I--  806
51_01B.SG.11.11SG_HM021        ------------N------S--------I--------I-----T--N-I--------------R----..--.A------G--------TFESSAT-F----V----F---R-----------A-V----.---.....................----------------  801
52_01B.MY.03.03MYKL018_1       -------------------D--------I--------I---I---------------------LSHHQ..-E.-----R---G---QGS---V---S-FF--A----------L------------AA------.---.................-G---R---------.-I--  810
53_01B.MY.11.11FIR164          --K--------TN-----D-S------RI--------I---I---------------------LTHQQ..-EQL----R---G---QG----V---S-F---A-X----------------I--AA-A----.-HSSLK.......GLRR---G---LGX--S-.-G--  806
54_01B.MY.09.09MYSB023         --KD-------T---S----S------RI---V----I--------------------------LTHHQ..-E.---------G---QG----V-----F-EI--E-----------------FI--AA-T----.-HSSLK.......GLRL---G---LG---V-.-G--  814
55_01B.CN.10.HNCS102056        -------------------D--K------I-------II----IL------K-------------PSHHQ..-E.-----R---G---QG----V---S-F-P-A-----------L------FI--AA-T-K--.-....Q.......GLRR---G---LG---L-.-G--  812
56_cpx.FR.10.URF5_patient_A    ---D--A--------T--D-S------RI--I-----I-------------------------LTHHQ..-E.----GR-------Q-----V---S-F---A---------I-L------FV--A--G----.-HSCLK.......GLRL------HL-S--S-.-G--  802
57_BC.CN.09.09YNLX19sg         --KD--A--S-Q--------S-------I--------I---I---A-----------------LA-N-..G-.---LGR-------Q-KT------Q-F---A-----N-----------FI-VTA-V----.--ISLR.......GLQK-------LGSIV--.--LK  820
58_01B.MY.09.09MYPR37          --L---------------D--K------I--------I-------------------------LTHHQ..-E.--K-------G---P-----V-----F---A-----------------FI--AA-T----.--TSLK.......GLRR---G---LG---L-.-GK-  814
59_01B.CN.09.09LNA423          -------------------D--K------I-------------I-------------------ILSHHQ..-E.---------G---QG---------S-F---A-----------------FISVTA-T----.-HSSLK.......GLRR---G---LG---A-.----  809
60_BC.IT.11.BAV499             ---D--A----QN--X--D--S------I--------I---I-------IF-------L--LI-G-..--.---LRE--K--K-Q-K-------S-F---A-----------------I--TA-A----.--SSLK.......GLQRR--I---LGS-V--.----  809
61_BC.CN.10.JL100010           --KD--A--S-KN--S--D--R------I-------------I-----V--------------PT-N-..G-.H--LRR-------Q--T--V-----F---A-----N-----------FISVTA-V-V--.--NSLR.......GLQK-------LGS-V--.-GL-  798
62_BC.CN.10.YNFL13             --KD--A--S-KN--S--S-S-------I--------I---I---------------------LT-N-..--.---LG--------Q-K---------F---A-----------------FI-VAA-AGK--.-HSSLK.......GLQR-------LGSIV--.-GL-  808
63_02A1.RU.10.10RU6637         ------A----TN--S--D-S------RI--------I------T----I---K---------LTHHQ..-E.----G----G---Q-----V-----F---F-------------Q----A--AA-T-Q--.-HSSLK.......GLRL-------L----A-.-G--  805
64_BC.CN.09.YNFL31             --KD--A--S-KN-----D---------I--------I---I--------------------L-IPTQS-..G-.---LGR-------Q-KS---------FF--A-----N-----------FI-VTA-V----.--SSLR.......GLQT--R----LGS-V--.-GL-  809
65_cpx.CN.10.YNFL01            --LD-M---R--------D--K------I--------I---I---------------------PFHHQ..-EQ---------G---QG---------T-FF--A-----N-----------FI-VT--V----.--SSLK.......GLRQ---G---LG---L-.-G--  820
67_01B.CN.11.2011.ANHUI.HF115--KD--------------D--K------I--------I---I---------------------PXHQQ..-E.----G--------QG----V---S-F---F-----------------FI--A--T----.-HSSLK.......GLRR---G---LG---L-.-G--  797
68_01B.CN.11.2011.ANHUI.WH73--KD--------------D---------I--------I---I-----------K---------PFHHQ..-E.----GE-------Q-----R---S-F-P---------------R--I--V--A--T----.--SSLK.......GLRL---G---LG---L-.-G--  798
70_BF1.BR.10.10BR_PE004        ------A-----------D--K------I--------I-----------I---------------F-A-..--.------TG-----------GQ--T-F--I--V--------------------A------.---.....................----------------  800
71_BF1.BR.10.10BR_PE008        ---D--A-----------G---------I--------------V---I----------------R--AQ..--.--G---GG---G---NT-GQS-D-F--I--V-----S--I--S------T-A------.---...................-S-------.----  806
72_BF1.BR.10.10BR_MG002        --K------------S--D-S------RI--------I-----S------------------H-A-..--.----G---------S---ET-AG-F-T---I---------FRH---------T---.---..................I---G--I---.----  797
O.BE.87.ANT70                  --KK-----E---I---LD--K------IA-I---A---V-VIMI--N--KNI----Q---L-IPNHHQ..EE.AGT-GRTGGG---EG-P-W-PSPQ-F-P-LYT---TII-WT--L-SN-ASGIQKVISY-.RLGLWI.......LGQKIINVCRICAAVT--.-L--  809
O.CM.98.98CMA104               --KK----SE---I----D--K------IA-I---A---V---MI--NL-RNI----Q-----IPIHHQ..SE.VGV-G-TG-G---G--P-L-P-QQ-F-Q-LYT---TII-WT--L-SN-ASGTQ-VISH-.-IGLWI.......LGQKIIDGC-LCKAVI--.-L--  830
O.CM.98.98CMABB141             --KK-----E---I---LD--K------IA-I---A-I-V---MV--NL-RNI----Q---L-IPIHHX..SE.AGT-GRTG-G---EG-P-L-ASPQ-F-P-LYT---TII-W---L-SN-ASGIQKVISH-.RLGLWT.......LGQKIIS-CRICAAVT--.-L--  833
O.CM.98.98CMABB212             -VKQ-----E-------D--K------IA-I---A-I-I-VIMI--N--KNI----Q---L-IPIQHR..QE.AEA-GRTG----DG--P-WTP-PP-F-QQLYT---TII-WI--L-SN-ASGIQ-LISY-.RLGLWI.......LGQKII--CRLCGAVI--.-L--  836
O.CM.98.98CMU5337              --RK-----E---I---LD--K------IA-I---A-I-V-VIMI--NL-RNI----Q---L-IPIHHQ..PE.AGT-G-TG-G---EG-PKL-PSPQ-F-X-LYT---TII-W---L-SN-ASGIQTVISH-.-LGLRI.......LGQKIID-CRICTAVT--.-L--  816
O.CM.99.99CMU4122              --KK----NE----D-LD--K------IA-I---A-I-I---MI--NL-RNI----Q-----IPNHHQ..SE.AGA-G-TG----G--P-L-P-PP-F-H-LYT---TII-W---L-SN-ASGIQTVISH-.-LGLWI.......LGQKIIN-CRICAAVV--.-L--  821
O.FR.92.VAU                    --K------E---I---LD--K------IA-I---A-I-V-V-ML--NL-KNI----Q---L-IPIQQQ..AE.VGT-G-TG-G-DE--R-WTP-PQ-F-H-LYT---TII-WI--L-SN-ASEIQKLIRH-.-LGLWI.......IGQRTI--CRLFKAII--.-L--  823
O.GA.11.11Gab6352              --KK-----E---I----D--K------IA-I---A---V--IMVI-NL-RNI----Q-------SHHR..SE.AET-GRTG-G---EG-S-LTASPQ-F-P-LYT---TIVVW---L-SI-ASGIQKAISH-.RLGLWI.......LGEKIIS-CRICTAVV--.-L--  831
O.SN.99.99SE_MP1299            --KK-----E---I---LD--K------IA-I---A-I-V---MI--NL-KNI----Q----K--IHHQ..PE.AEA-G-TG-G----GMPTL-PWPQ-F-P-LYT---TII-W---L-SN-ASGIQTVISH-.-LGLWT.......LGQKIIS-CRLCIAVI--.-L--  828
O.US.10.LTNP                   --KS-----E-V-I---LD--K------TA-I---A-I-V-VIMI--NL-RNI----Q---L-IPIHHQ..VE.AGT-G-TG-G---EG-P-WRP-QQ-F---LYT---TII-W---L-SN-ASGIQKVISH-.RLGLWI.......LGQKLIN-CRICTAVT--.-L--  812

N.CM.02.DJO0131                --KS-----Q--------S---------IA--V-A-II-I--ISVIIT-IA----------L--LI--TTT--.-----ET--GV-GQG----V---S-F-T-V-E-F-N-LI-L----T-S---LQ-TL---.--SLIR.......GLQLLN-LRTRL-GIIA-.-GK-  787
N.CM.02.SJGddd                 --RS-----Q----T--DL--------IA---A-II-I--ISVIIT-IA----------L--LI--T.---.X-----T--GV--Q--G--V---S-F---V-E---N-SI-L----T-S---LR-TL---.RQSLSR.......GLQLLN-LITRL-GI-A-.-GG-  786
N.CM.04.04CM_1015_04           --KS-----Q-T-----D-S--------IA--V-A-II-I--ISVIIT-IA----------L--LI--T.---.------T--GV--Q--G--V-----F---V-E-C-N-VTLL----------TQ-TL-I-.-QSLNR.......GLQLLN-LR-RL-GI-A-.-GK-  783
N.CM.06.U14296                 --KS-----Q--------D-S--------IX----A-II-I-XISVIIT-IA----------L--LI--T.---.----GETG-DV--Q--G--V---S-FFS---E---N-VI-L-R--T-S---LX-TL---.-QSLSR.......GLQLLN-LRIHL-GIIA-.-GK-  808
N.CM.06.U14842                 --KS----XH--------D--K------IA----A-II-I--IS-IIT-IA----------L--LI--T.---.-----ETG-G---Q-----V-----F---V-E---N-LT-L----T-S---LR-TL---.-QSLSK.......GLQLLN-LRIHL-GIIA-.-GK-  805
N.FR.11.N1_FR_2011             --KS-----Q-T-------S--------IA----A-II-I--IS-IVT-IA----------L--LI--A..-D.-----T--GV--Q--G--V-----FS--V-E---N-LI-L----T-S---LR-TL---.-QSLSK.......GLQLLN-FRI-L-GI-A-.-GK-  794

P.CM.06.U14788                 --RK--------Q--S-MD--K------I-------II--K-LM-IGNV-RK-------V-L--LRLHG..GP.AGIAP-TD-----AGNG-----LE-L-P-------N-VVQI-QI-VSCI-RIKDLLTI-.WIHLGQ.......LLNS..NC-RDCFAACG-.-T--  835
P.FR.09.RBF168                 --KK-----Q--Q----LD--Q------I-------II---LL-IINV-R-I-------V-L--LGLNG..DP.AGIAP-TN-----AGNG-----LD-F-P-V---KN-VVQI-QI-VGCI-GIKDLLTI-.WIHLGQ.......LLTR-LNC-RDCFAACG-.-T--  829

CPZ.CD.06.BF1167               -HKQ-------S------D---------IA-I--AS-------MFIINL-GKL-----L--P-IPIQ-A..--.-GQ---T--G---SG---WRGYPV-F-Q---E-C-Q-L-WI-QTSVN--WHLWTLIK-I.HQWSLLLWETVRQLLRRLIHL-TQFLH--ET.-GIL  834
CPZ.CM.05.SIVcpzMT145          --K------S--N--S--D---------M-LIV---II----C--IG-LI----K-------L--LI-AN..Q-.--GLGET-KG---NV-G------S-F-P-V-E---N-LS-L--Q---CASLIW-LL-I-.-QYSLR.......GVQQIGTL-HQLRGT---.-TT-  805
CPZ.GA.88.GAB1                 --RD-----Q--------D--K------I-L-A---II---IMT-F-V-R-----------L--LI-VQ..-E.QG-LGE-D-G---Q--S--V---E-C-P-------N-GIW--QS-TS-ACN-W-QLKT-.-HLILH.......SLRLLR-R-CLLGGII--.-GK-  804
CPZ.TZ.06.SIVcpzTAN13          -KK---------ST--D--D--Q-------A-I-IA--I----LMFIVNVLRQ-----M--FS-IPTQAE..QD.-EQ-G--A-G--D-G-I-WTPSPR-FFSIV-E---N-L-WI-QTCLNFVWLLGT-LQA-.KQGTISFIQTIIIVHRNLVLKWRLFRDW-AN.AYVI  820
CPZ.US.85.US_Marilyn           --KS-----Q-S------D----------I-LIV-AS---I---GVIF-L-AK---------L--LF--T..-E.------T--GA-KT-NV--T---S-F---V-E---N-LI-L----E-----LR-T-QI-.-QNINK.......GLQLLN-LRARC-GVIA-.-AR-  786

GOR.CM.07.SIVgor2139_287       --KA--------D--S-LD--Q------I--I-IA-------LM-IINMCH------------ILGRNQ..VP.AGIVPE-G----RA-N------LA-F-P-----I-T-VVWI-RT-VI--SGIKE-TLS-.IEYLTR.......LLRRVNNLTRDCFAFIA-.-G--  819
GOR.CM.07.SIVgorCP2135con      --KA--------D--S-LD--Q------I----IA-------LM-IINMCH------------ILARNQ..VP.AGIVPE-G----RA-S------LA-F-P-------T-IVWI-RT-VI--SGIKE-TLS-.IEHLTR.......LLRRVNNLTRDCFAFIA-.-G--  814
```

	Env end gp41 end glycosylation NAT cytoplasmic tail end	
B.FR.83.HXB2	LKNSAVSLLNATAIAVAEGTDRVIEVVQGAC*RAIRHIPRRIRQGL.ERILL*	856
A1.AU.03.PS1044_Day0	--I---N-YDTI-----GW------IG--I----LN---------.--A---	871
A1.CH.03.HIV_CH_BID_V3538	-RV--IT--D----T--GW------IG-RF----L----------.--A-I-	855
A1.CY.08.CY236	--I--IN-FDTI-----GW------I---IG---LN---------.--A---	863
A1.ES.05.X1608_8	--L---N-YDTI-----GW------IG-RL----LN---------.--A---	874
A1.KE.11.DEMA11KE001	--I--T---DTV-----GW------IR-RIG---L----------F.--A---	877
A1.RU.11.11RU6950	-----IN-VDTI--V--GW------IG------CN---------A.--A-Q-	864
A1.RW.11.DEMA111RW002	--I--I---FDTI--T--GW------I---RI----LN--T-----F.--A---	872
A1.SN.01.DDI579	--S--I--------V---W------IG-R-G--LL-----V-------.--A---	851
A1.UG.11.DEMA110UG001	-----IN--DTI-----GW---I--IG-SIG---L---------A.--A---	857
A1.ZA.04.04ZASK162B1	-R---I---DTI--T--GW-G----IG-RIV---L----------.--A-Q-	860
A2.CM.01.01CM_1445MV	------N-VD-I-VT----------LG-RI----LN--T-----F.--A---	863
A2.CY.94.94CY017_41	-----I--FDTI-V----W------IG-R-F---LN---------.--A---	866
B.BR.10.10BR_MG029	----------------------------RI#---LN--T-MG---#--A---	868
B.CA.07.502_1191_03	--------F----------------L-R-G--VL-V-------F.--S---	863
B.CH.08.M2_0803101_NFLG8	-----TN---TI----------I--LA-RGF--FL----------.--A-Q-	886
B.CN.10.DEMB10CN002	-QK-----------------I--A--RIG---I----------.----V-	859
B.ES.10.DEMB10ES002	------N---T---V-------I---A-R-W--FL---T-----A.--V---	865
B.FR.11.DEMB11FR001	------------------------I--RI----I---V-------.--A---	841
B.GB.05.MM45d213_GN1	--K--I----------------A-DII-RXF---L----------.--S-V-	863
B.HK.06.HK003	--K-------FI----------I--A--RGV---L----------.--A-V-	854
B.HT.05.05HT_129389	---------D---V--G-----I--I-RRVV---L-V-T------.--A---	863
B.JP.12.DEMB12JP001	--K--I----------------I---I-RL----L---V-------.--A---	865
B.KR.07.HP_18_07JHS10_3909	--K------H---R-I------A--IL-R-Y---L----------F.--A---	858
B.PE.07.502_0525_wg5	----------I----------I---GLNIF---FI---A-------.--L---	843
B.RU.11.11RU21n	---------F-------G----------RTF--VL----------F.--A---	843
B.TH.08.MERLBDTRC10	----------V----------------R-Y---I----------.---T-V-	864
B.US.11.ES38	--S--I--F--I------------GL-R-W--FI----------.------	863
C.AR.01.ARG4006	--K--I---DTI----------I---I-RIWEI-CN--------F.-AA-Q-	857
C.BR.07.DEMC07BR003	--KR-I--FDTV----------I--II--IW---CN--------F.-AA-Q-	887
C.BW.00.00BW5031_1	--K--I----DTI----------I-----RLY---LN--------F.-AA-Q-	849
C.CN.10.YNFL19	--K--I---DT----------I-AIX-R-W-------.T--G-DP..VA-Q-	865
C.CY.09.CY260	--K-VS--DDT-TT-A-GAA--IV-LLRRISGI-CS-R----RC-.AAA-Q-	861
C.ES.08.X2363_2	--K--I---DTI----------I--FI-RI----SNL-------F.-AA---	869
C.ET.02.02ET_288	--K--I----T---V-------I--L--RFW---CNA-------F.-AA---	858
C.IN.09.T125_2139	--KG-I--FDTI----------I--FL-RL----YN---------..AA-Q-	884
C.KE.00.KER2010	--R--I---DTI----------I--LI-RI----LN--T-------..AA---	863
C.MW.09.703010256_CH256.w96	--K--I---DT-----------I--FL-RI----Y---------F.-AA-Q-	854
C.TZ.08.707010457_CH457.w8	--K--I---DTI---T------I--L--RI----CNV-T-------..AA-Q-	846
C.YE.02.02YE511	--K-----VDTI----------I--I--RS---LCN--------F.-AA-Q-	863
C.ZA.08.705010534_CH534.w12	--K--I---DT------T----I---I-RG----CN--------F.-TA---	854
C.ZA.10.DEMC10ZA001	--K--I---DTI----------IF-A-LRF-----N--T-----F.-AA---	854
C.ZM.11.DEMC11ZM006	--K--I---DTI--RT------I--L-LNI----LS--------F.-AA-Q-	864
D.CM.10.DEMD10CM009	-----I--FD--------------A-LR-F---L----------.---A---	845
D.CY.06.CY163	-----IN--D-------------T---RT---VLR--------F.--A---	859
D.KE.11.DEMD11KE003	----------DTI--VI------L--IGLRIG---IN---------..--A---	851
D.KR.04.04KBH8	-----IR-FD----------------RR-F--VL----------F.--A-I-	843
D.SN.90.SE365	-----I---DTI----------I-D---R----L---T-------.-GA---	851
D.TZ.01.A280	-----I---V-------------I--II-R-F--VL-----V-------.-A---	853
D.UG.10.DEMD10UG004	V----I--FGT---V-------I---L-RIG---LN--T-----A.--AF--	848
D.UG.11.DEMD11UG003	-----I--F-T---V----------L--R-G--VLN--------F.--A---	858
D.YE.02.02YE516	--I--I--F-------------I--I--R-G---LN---------.--A---	847
D.ZA.90.R1	-----I---DTI---T------------R----VLNV-V-------.--L---	859
F1.AO.06.AO_06_ANG32	-----IN--DT---V----------L-R-G--VL----------.------	841
F1.AR.02.ARE933	-RI--I--F-T---V----------A--R-V--VCN---------.--AF--	846
F1.BR.10.10BR_RJ015	--------------V-G-----I--AL-RFG---LN--------F.--A-I-	848
F1.CY.08.CY222	-----I---DTI--V----------AL-R-G--VLNV--K-----.------	845
F1.ES.02.ES_X845_4	-----I---DT---V-------I--AL-RTG---L-----V----.--A---	843
F1.RO.96.BCI_R07	-----I----TI--V----------AL--VLT--LN---------.--A---	859
F1.RU.08.D88_845	-----I----T---V---W----L-AL-R-G--VL----------.--L---	863
F2.CM.02.02CM_0016BBY	--I--I----T---V----------IL-R-G---L---------F.--A---	846
F2.CM.10.DEMF210CM001	--K--I--FDT---V-------I---L-R-G---L---------A.--A---	856
F2.CM.10.DEMF210CM007	-----I---DT---V-------I-D-L-RIG--VL---------F.--A---	848
G.BE.96.DRCBL	-----IN--DTI-----NW-------A-R-G--VLN---------.--A---	855
G.CM.10.DEMG10CM008	-----IN--DT------DW-------A-RVG--FLN---------.------	848
G.CN.08.GX_2084_08	-----IN--DTI-----NW-------AHR-----CN---------.--A-V-	860
G.CU.99.Cu74	-----IN-VDTV-V---NW------I--RT---FLN--------F.--A-Q-	866
G.ES.09.X2634_2	--S--IN--DTV-----NW------I--RVG---LN---------.--A---	859
G.GH.03.03GH175G	-----I---DTV-----NW-------I-R-G---LN--T-----F.--A---	863
G.KE.09.DEMG09KE001	-----IN--DTV-----N-------I--R----VL---------F.-SA---	853
G.NG.09.09NG_SC62	-----I--VDTL--V-GNW------A-R-G--FL-----V-----.--A---	857
G.ZA.01.TV546	-----IN--DTX-V---NW------IA-RTG---CN---------.--A---	856
H.BE.93.VI991	-----I----T-----------I--L--R-W---L----------F.--A---	860
H.CF.90.056	-----ID---T----------GI-VI--R-W---L----------F.--S---	846
H.GB.00.00GBAC4001	------N---T---V-------I--L--R-G---I----------F.--A---	853
J.CM.04.04CMU11421	-----I---DT---VX------I--IXRR-F--VL----------F.--A---	861
J.SE.93.SE9280_7887	-----I----T-----------I--IA-R-F---L----------.--A---	849
J.SE.94.SE9173_7022	-----I----T-----------I--IA-R-F---L----------.--A---	851
K.CD.97.97ZR_EQTB11	I----IN---T-----------I--I-YR-F--LL----------F.--L---	852
K.CM.96.96CM_MP535	-----I----T------G----I--IG-R-F--LL----------.--A---	843
U.CA.01.TV749	--T--IN---TV--------------G-RI-----N---------F.--A---	869
U.CA.99.TV721	--T--IN---TV--V-------I---G-RI----CN--------F.--A---	872
U.CD.83.83CD003_Z3	-----I--------V-------I---G-RI----LN--------F.--A---	848
U.CD.90.90CD121E12	-----I--------V-------I--L--RVG---L---------F.--A---	846
U.CY.05.CY090	-----I---DT---V-------I--IG-RIF--FIN---------.--A---	846
U.CY.08.CY223	-----I----TI----------I--IG---I---LN---------.--A---	881
U.ES.10.DEURF10DZ001	-----I---DTI---T-NWI--A--GI-RIG--LLN---------.--A---	854
U.GR.99.99GR303	-QK--T---DI---V--GW------TG-RIG---L---------F.--A---	857
U.NL.95.U_NL_95_H10986_D1	----IN---T-----------I--IG-R-F--VL---------F.--A---	862
01_AE.AF.07.569M	-RI--I---D-I-----GW------I----W--FI----------.------	859
01_AE.CF.90.90CF11697	--T--IT-FD-I-V---GW---------R-W--LI----------.---A---	864
01_AE.CN.10.YNFL03	--T--I--FD-------G----------R-W--LL---------F.--A---	861
01_AE.HK.04.HK001	--I--I----T-------W--------A-R-W---L----------.---T-V-	860
01_AE.IR.10.10IR.THR48F	--I--S--D-------GW-------A-R-W--LL----------.--A---	868
01_AE.JP.x.JRC77AE	--T--I---D-A------W-------L-R-----L----------.---T---	857
01_AE.TH.09.AA090a_WG11	--I--I---DTI-----GW-------A-R-W---IQ---------.--A-V-	858
01_AE.TH.90.CM240	--I--I---D-----A-GW-------A---W---L----------.---T---	858
01_AE.US.05.306163_FL	--T--I---D-----T--W-------A-RTG---LN---------.--A-I-	837
01_AE.VN.98.98VNND15	--I--I--IDTA-----GW-------A-R-W--FL----------.--A-Q-	868

Draft of 30 September 2015, not for public release
hiv1prot_main.tex r8876–10546

```
                                                                     Env end
                                                                    gp41 end
                            glycosylation NAT              cytoplasmic tail end
                                                         *
B.FR.83.HXB2                LKNSAVSLLNATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQGL.ERILL*   856
02_AG.CM.10.DE00210CM013    -----TN--DTI-----NW------I--R-----YN---------.--A---   885
02_AG.CY.09.CY256           -----IN--DTI-V--GNW--G----G-R-----CN--T------.--A-Q-   858
02_AG.ES.06.P1423           -----IN--DTI-----NW------IA-R-G--FL---------F.--A---   863
02_AG.GW.05.CC_0048         -----IN---TI--V--NW------IG-S-----CN--T------.--L---   868
02_AG.KR.12.12MHR9          -----L--FDTI-----NW------IG-R-G---LN--------F.--A---   870
02_AG.LR.x.POC44951         -----IN--DTI-----NW------IG-R-G----N--T------.-KT---   853
02_AG.NG.09.09NG_SC61       -----------I-----NW---I--IG-R-G-...................-   844
02_AG.NG.x.IBNG             -----IN---TI--V--NW---A--IG-RVG----N--------F.--A---   855
02_AG.SE.94.SE7812          -----I---DTI--V--NW------L--R-G---LN--------F.--A---   860
02_AG.SN.98.98SE_MP1211     -----IN--DT--V---NW------I--RTG---CN---------.--A-Q-   862
03_AB.RU.97.KAL153_2        --S--IN-IGTI-----GW------IG-RF---M-N--------A.-KA-Q-   840
04_cpx.CY.94.94CY032_3      -----IN-F-T-----------I--A--R-----CN---------.--A---   856
05_DF.BE.x.VI1310           -----I----T--VV----------AL-R-G---LN---------.--A---   862
06_cpx.AU.96.BFP90          -Q---I--FD-A-----NW----A----RIF--FLNV-------F.--A---   860
07_BC.CN.98.98CN009         --K-TI--VDTI----------I--L---L----YN--------F.-AA-Q-   860
08_BC.CN.97.97CNGX_6F       --K-TI--VD-I----------I-NR---I----HNV-------F.-AA-Q-   865
09_cpx.GH.96.96GH2911       -----I----T---V-----------L-R-G---L---------F.--A---   843
10_CD.TZ.96.96TZ_BF061      -----I---DTI--E--GW------I--R-V---LN--T------.--A---   841
11_cpx.CM.95.95CM_1816      -----I----T---V-------I---IHRXL--VL----------.--A---   866
12_BF.AR.99.ARMA159         -----I----T---V--G----IL-AL-RIG--VLNA---V----.--A---   843
13_cpx.CM.96.96CM_1849      --I--IT--DS-------W---I--TG-RIG-G-L---------F.--A-Q-   868
14_BG.ES.05.X1870           -----I---DTI-----NW---A-X----VG---LNV-V------.--F---   866
15_01B.TH.99.99TH_MU2079    --I--I---DT------GW-------A-R-W---LR--------F.--A---   855
16_A2D.KR.97.97KR004        -----I--F---------W------I--R-----IN---------.--A---   863
17_BF.AR.99.ARMA038         -----I----T---V----------AL-R-G--VLN--------F.--A---   849
18_cpx.CU.99.CU76           I-S--IN--DTV---T-NW------I--RTG---LN--------A.--A-Q-   864
19_cpx.CU.99.CU7            --T--I----T---V--GW---A--II-R-W--VL----------.--A---   802
20_BG.CU.99.Cu103           -----IN--DTV-----NW---I--I--RGF--FL---------F.--A---   850
21_A2D.KE.99.KER2003        -R---N--Y----------------I-RR-----LN--------F.--A---   853
22_01A1.CM.01.01CM_0001BBY  -----IT--D----T--GW------IT-R-W--FL---T------.--A---   849
23_BG.CU.03.CB118           -----IN--DTV-----N----I--I--RGF--FL---------F.--A---   856
24_BG.ES.08.X2456_2         -----IN-YDTV-----N----I--I--RGF--FL---------F.--A---   850
25_cpx.CM.02.1918LE         -----IN--DTI---T-DW------IA-R-G---LNV-V------.--A---   858
26_AU.CD.02.02CD_MBTB047    -----IN---TI-V----C------IA-RGG---L----------.--A-Q-   859
27_cpx.FR.04.04CD_FR_KZS    -----IN--DTV---T-NW-------A-R-G---IN--------A.--A-Q-   864
28_BF.BR.99.BREPM12609      -------------------------IL-RI-----N--T------.--A---   869
29_BF.BR.01.BREPM16704      ----------------------------RIW---L---T------.--A---   871
31_BC.BR.04.04BR142         --K--I---DTI----------I---I--IG---YN--------F.-AA-Q-   850
32_06A1.EE.01.EE0369        -----IN-IDTI-----DW-------G-RVF--FLN---------.-IA---   851
33_01B.ID.07.JKT189_C       --I--I-----------GW-------I-IVW---L---------F.--S---   860
34_01B.TH.99.OUR2478P       ----------------------I-----R-Y---L----------.--A-I-   844
35_AD.AF.07.169H            --T--TN-IDTI-----GW-----DIG-RFG---L----------.-GA---   855
36_cpx.CM.00.00CMNYU830     --T--IT--D---VT--GW-------A-RTG---LNV-T------.--A---   858
37_cpx.CM.00.00CMNYU926     --I--IN--DT---V--GW------IA-R-----L---------F#--A-Q-   853
38_BF1.UY.03.UY03_3389      -----I--F-T---V-------I--AL-R-G--VLN---------.--A-I-   845
39_BF.BR.04.04BRRJ179       -----I----T---V-------IL-AL-TIG-G-L----------.--S---   845
40_BF.BR.05.05BRRJ055       ----------TL----------I--A--RIG-G-L----------.--S---   852
42_BF.LU.06.luBF_18_06      -----I----------------I---L-RIG--VL---T------.--A---   841
43_02G.SA.03.J11223         -----IN-VDTI-----GW-------A-R-----IN--------F.--A---   847
44_BF.CL.00.CH80            -----I----T---V--GW-------L-R-G--VL---------F.--A---   845
45_cpx.FR.04.04FR_AUK       ---------D-------GW-----T--SL---FLN---------.--A---   846
46_BF.BR.07.07BR_FPS625     -----I----T---V-------I--AL-R-G--VLN---------.--A---   836
47_BF.ES.08.P1942           -----I----T---V-------I--AL-S-G--VL----------.--A---   849
48_01B.MY.07.07MYKT021      --I--I--------T-------I---ARR-W---I----------.--A---   865
49_cpx.GM.03.N26677         -----I--FD-I----------I---ARR-F--LLN----T----.--A---   856
50_A1D.GB.10.12792          -----I--FDTI--V--SW---A---G-RIG---L---------A.--A---   857
51_01B.SG.11.11SG_HM021     --------F----V-----------I--RVW--FL----------.--S---   852
52_01B.MY.03.03MYKL018_1    ---------DT-----------I-----R-Y---L---T------.--A---   861
53_01B.MY.11.11FIR164       --T--I---D-A-----GW---------R-W---L----------.--A-Q-   857
54_01B.MY.09.09MYSB023      --I--I---DT--V---GW---------R-W---L----------.--A-Q-   865
55_01B.CN.10.HNCS102056     --I--IA---TI--V--GW------AA---G---L----------.--S---   863
56_cpx.FR.10.URF5_patient_A --K--I---DTI---I-NW------LG-A-G--VL-V-------F.--A---   853
57_BC.CN.09.09YNLX19sg      --K-T---YDSI----------I--FI-RI--S-LN---K----F.KAA-Q-   871
58_01B.MY.09.09MYPR37       --I--I---D-------GW---------SIW---L----------.--A---   865
59_01B.CN.09.09LNA423       --T--I---DTL-----GW------LA-R-W--LL-V--------.--A-Q-   860
60_BC.IT.11.BAV499          -----I----T-------R--KI--IL-RF-K--L---AK---D-.------   860
61_BC.CN.10.JL100010        --K-TI--YDTI----------I--L--RLY----NV-T-----F.--A---   849
62_BC.CN.10.YNFL13          --K--IN--DTI----------I--TI-R-Y---CN--------F.-AA-Q-   859
63_02A1.RU.10.10RU6637      -----IN-FDTI-----NW-------G-RVG--LL----------.--A---   856
64_BC.CN.09.YNFL31          --K--I---DTI--------------I-R-----L---------F.-AA-Q-   860
65_cpx.CN.10.YNFL01         --I--I---D---VT--GW-------A---W---L-V--------.--A---   871
67_01B.CN.11.2011.ANHUI.HF115 --I--I--FD----XI-GW------AA-R-W--FL---------F.--A-V-   848
68_01B.CN.11.2011.ANHUI.WH73 --I--I---D---V-T-GW------AA-R-W---L---------F.--A---   849
70_BF1.BR.10.10BR_PE004     ----------T---V--------V-AL-R-G--VLNV-------F.--A---   851
71_BF1.BR.10.10BR_PE008     ----------T---V---------GAL-SIG--FL----------.--A---   857
72_BF1.BR.10.10BR_MG002     -R--------TI--V--------L-AL-TLG---I----------.--A-Q-   848
O.BE.87.ANT70               -Q---T---DTL-V---NW--GI-AGI-RIGTG--N---------.--S---   860
O.CM.98.98CMA104            -Q---T---DTI-V--GKW--DI-LGI-RIG-G-LN---------.--S---   881
O.CM.98.98CMABB141          -Q--VI--VDTI-V---NW--NI-LGI-RLG-G-WN--T-----F.--S---   884
O.CM.98.98CMABB212          -Q---I--FDTI-V-A-NW--NI-LGI-RIG-GXLN--------A.--X-X-   887
O.CM.98.98CMU5337           -Q---T---DTI-V-T-NW--SI-LGI-RIG-G-YN---------.--S---   867
O.CM.99.99CMU4122           -Q---TN--DTL-V---NR--SIFLGI-RIG-G-LN---------.-QA-V-   872
O.FR.92.VAU                 -QT--TN--DTV-V---NW--SI-LGI-SIG-G-LN---------.--L---   874
O.GA.11.11Gab6352           -Q---TN-VDTL-V---NW--GI-LGI-RIG-G-LN--T------.--S---   882
O.SN.99.99SE_MP1299         -Q---T---DTI-V---NW-VTI-LGI-RIG-G-LN---------.--S---   879
O.US.10.LTNP                -Q---T---DTF-V---NW--GI-SGI-RIG-G--N---------.--S---   863

N.CM.02.DJO0131             --D--I----TI--V-------L--LA-RIG-G-L----------.--A---   838
N.CM.02.SJGddd              -RD--I----T---V-------L--LA-RIG-G-L---------A.--A-V-   837
N.CM.04.04CM_1015_04        --D--I----T------G----L--LA-RIG-G-L----------.--T---   834
N.CM.06.U14296              --D--I----T---V-------L--LA--IG-G-L----------.--A---   859
N.CM.06.U14842              -XD--I----T---V-------L--LA-RIG-G-L----------.----V-   856
N.FR.11.N1_FR_2011          --D--I----T---V-------L--LA-RVG-G-L----------.--A-I-   845

P.CM.06.U14788              --Q--T----TV--S--NW--Q--A-G-QIG-GFLN----L----.--S---   886
P.FR.09.RBF168              --Q--T---DTV--S--GW--Q--I-G-QIG-GFLN--------I.--S--X   880

CPZ.CD.06.BF1167            IRQHCIVCIDTL-EFTGWW--G---ALRV-VDI-----T-------.-IA-N-   885
CPZ.CM.05.SIVcpzMT145       I----I---DT-----G-----IL-AITRLG-G-L----------.--A---   856
CPZ.GA.88.GAB1              --I--I---D-----------I--AF-VTL-I--N---------.--A---   855
CPZ.TZ.06.SIVcpzTAN13       V-T-FIRN-DRL-NFT-WW--IL--GAVNIF-G--N--T------.-IA-N-   871
CPZ.US.85.US_Marilyn        --V--T---DT-----------I--LTRRLFLG-I----------.--S--X   837

GOR.CM.07.SIVgor2139_287    --Q--I---DCV-VWT-NW--Q--AIA-RIG-G-LN---------.--S---   870
GOR.CM.07.SIVgorCP2135con   --Q--I---DCV-VWT-DW--Q--AIA-RIG-G-LN---------.--S---   865
```

```
                              myristoylation
                              Nef start                                                                                   acidic cluster  poly-P helix  phosphorylation      HXB2 premature Nef end
B.FR.83.HXB2                  MGGKWSKSSVIGWPTVRERMRR.......AE.................PAADRVGAASRDLEKHGAITSSNTAATNAACAWLEA...QEE..E..EVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQ.RRQDILDLWIYHTQGYFPD*QNYTPGPGVRYPL  136
A1.AU.03.PS1044_Day0          -------R-SK--SEI--KI-Q.......TP.................T--EG---V-K--D----V----V..NHPS-T----.....-..G.-------R---------F-G-F---F--------D---Y-X.K--E-----V-N-------W--------I----  135
A1.CH.03.HIV_CH_BID_V3538     ---------IV---EI-------.......TP.................A--EG---V-----R-------------S----------E.G.-------R---------F-G--------------D---Y-R.K--E-----V-N-------W--------T-F--  138
A1.ES.05.X1608_8              ---------IV---EI-D-I--.......XP.PA...............---SG---V-Q--D-R--V----I..NHPS----N-.....X.-..D-----R-----------G-F-----------------Y-R.K--E-----V-N-------W-------------  137
A1.KE.11.DEMA11KE001          ---------IV---E-------.......-P.................S--PG---V-Q--A----V----I..NHPS-----......D.-.-------R-------------I------------D---Y-R.K--EL----V---------WH------------  135
A1.RU.11.11RU6950             ---------IV---Q----I--.......-P.AP..............A---SG--PV-Q--DR---V----------------------.-.-------R-----------G-F------------D---Y-K.----------V---------W--------I----  140
A1.RW.11.DEMA111RW002         --S------IV---E------Q.......-R.RAT............PP---AG---V-------------V..NHPS-------......D.-.-------R-------------F---F-------------Y-K.---E-----V-----F---W-------------  140
A1.SN.01.DDI579               --------RQE--HK--D---Q.......TP.PT.............---EG---V-Q--S----V--T-Q-..S-S-T-----.....-.-.-------R-----------G-F------------D---Y--.Q--------V---------W--------T----  137
A1.UG.11.DEMA110UG001         ---------IV---E----I-N.......TP..............AP---KG---V-Q--D---------L..NHPSNT----.....-.-.-------R----T-------------------------Y-R.K--E-----V-----F---W--------I-F--  137
A1.ZA.04.04ZASK162B1          ---------IV---QI---I-Q.......TP.TA...............A--EG---V-----R--------E-N-ESL--VR-......E.-.-------R-----------G-F------------D---W-R.K--E-----V-N-------W-------------  140
A2.CM.01.01CM_1445MV          --S-L--.....---A-------.......TP.................---EG---V-Q--GTR--V-IR------SD-----......S.D.-------R-----------G-F------------D---Y--.K--E-----V---------W----------F--  134
A2.CY.94.94CY017_41           -------R-IP---AI-------.......TP.PTAQRTE.......AVS---PG---V-Q--AT---V---------PD---V-......E.S.-------R---------F-G-F---F--------D---Y--.K------M-V---------W--------I----  148
B.BR.10.10BR_MG029            -------R-GK-----------.......--.................---EG---V-----TR---------E--D----------.....-.-.-------R---------------------------Y--.---------V---------W--------I----  137
B.CA.07.502_1191_03           -------GGMG--NA--D----.......T..................---EG---------R-----------N-PD-----......N.-.-------R----------------D-------TW--.K--------V-N-------W--------T----  137
B.CH.08.M2_0803101_NFLG8      --S-L--C-PV--S-I------.......T-.PRNE..........QNEQ--VG---------R------------------D-----.....-.-.-------R-----------G-L-------------------.K---------------W-------------  144
B.CN.10.DEMB10CN002           -------PRLG--SAI------.......T-.P.............VE---EG--------ER--L-T-----N-----------......D.-..D-----R-----------G-L-------------------.K--E-----V---------WH-------T-F--  140
B.ES.10.DEMB10ES002           ---------V---AI-D-----.......--.PA............AE---AG--------DR---L-------N--DY------......DD.G.-------K-----------G-L----------------Y--.Q--------V-N-------W----Q--------  142
B.FR.11.DEMB11FR001           ---IF--R-GG---A--------.......TG.TEPKAERLGKERVRRTE---EG---------R-----T----DN--DS------......D.-.-------R-----------G-L------R---------Y--.K--------V---------W--------I----  154
B.GB.05.MM45d213_GN1          ---------V---X--------.......-R..............AE----G---V-Q--GR-----------N------V--.....-T-G.-------R------------F----------------W--.Q--------V-N---F---W--------T-F--  140
B.HK.06.HK003                 -------RGRD--DAI---I--.......--.PR............AE---AG---V----AQR-------Q..N--DRV-----......EE-.-------R-----------D-F---------------VY--.K-K-------V-----F---W-C------I----  141
B.JP.12.DEMB12JP001           ---------IV---A----I--.......--.PR............AE---EG---V--------------I..N--D-V------.....-.-.-------R---------F-E-F----------------Y--.K-K---N--V-----I---W-E------T-F--  139
B.KR.07.HP_18_07JHS10_3909    -------R--P--DAI------.......T-.................---EG---V-----QR---T----SN------------.....-.-.-------R-----------G----------------VY--.K--------V---------W--------T-F--  137
B.PE.07.502_0525_wg5          -------C-PV----I-D----.......--.................---EG--------D----L-A-----N------------.....-.-.-------K---------F-G----------------VW--#K--------V---------W--------T-F--  137
B.RU.11.11RU21n               ---------L----RI----Q-.......T-.................---EG---V----------------------------.....-.-.-------R-----------G------------------Y--.K----------------W--------T----  137
B.TH.08.MERLBDTRC10           ---------LV---AI----K-.......--.................----G---V-----R-----NN--P-N--D------......E.-.-------K-----------G-L------R--------VY--.K-------------F---W--------T-F--  138
B.US.11.ES38                  -------I-AD--SR-------.......-G.T.............AE---EG---------R---------VTN------VA-.....DD-.-.-------R---------F---L------N--------VW--.K--------V-N-------W--------P----  142
C.AR.01.ARG4006               --N----...L---A--------.......-P.PGVGDXR......RQDA---EG--P--Q--DRY--L-----P-N--D-------...Q-E.-.-------R---------F-G-F---F-------------Y--.Q-K-----V-N-------W-------------  146
C.BR.07.DEMC07BR003           --N----C-PV--SAI------.......T-.................---EG--P--Q--DRY--L-----PGN----------...Q-E.-.-------RX-------------F---F-------------Y-K.K----I---V-N-------W--------I----  138
C.BW.00.00BW5031_1            ---------LV---E--D-I--.......TD.................---EG----Q--D----L-N-------KD-------.....E.G.-------R-----------G-F--GW--------D---Y-K.K--E-----V---------W-------------  138
C.CN.10.YNFL19                --A------IV---AI------.......T-.................----G---------R--------I..N-D--T--Q-.....E.-.-------R----V------G-F---F-------------Y-K.K--E-----V-N---F---WS------------  136
C.ES.08.X2363_2               --------RRV--TE-------.......TK.PAAEP.........AVE---VG--P--Q--D-Y--L-----PSN--D----A-.....E.DE.-------R-----------G-F--GF--------D----XK.K--------V---------W--------I----  147
C.ET.02.02ET_288              --SA---C-PR--S-I-----Q.......TP.................---EG---------Q---L-T---PEN-------Q-.....E.-.-------R-------------F--GF-------------Y--.K-KE-----V-N---F---W-------------  138
C.IN.09.T125_2139             --S----C-IV---A--D----.......T-.................---EG----Q--D-Y--------GT---DV----T.....A.G.-------R-----------G-F---F-------------Y-K.K--E-----V---------WH------------  138
C.KE.00.KER2010               ---------IV---A-------.......--.................---EG----Q--D-Y--L-----GN--D----Q-.....E.-.-------R---------R-G-I---F--------D---Y-K.K--------V-N---F---W-------------  138
C.MW.09.703010256_CH256.w96   ---------IV---N----I--.......TD.................---EG---------R---L-T---VSN------------.....E.K.-------R-----------G-F---F-------------Y-K.K--------V-N---F---W-------------  138
C.TZ.08.707010457_CH457.w8    -------C-MV---E----I--.......TA.................---EG--------DR---------T------------.....A.-.-------R-----------G-----F-------------Y-K.K--------V---------W-------------  138
C.YE.02.02YE511               ---------IV---V--D-I--.......TN.PAAAGVG...A..ASQDS--AG----Q--A----L-----PTN------------.....QD.-.-------R-----------G-----W-------------Y-K.K--E-----M-N-------W--------T----  150
C.ZA.08.705010534_CH534.w12   ---------IV---E----I--.......TR.PAAEGE.........KT--EG--T--Q--D----L-T----RN-D-------.....D.G..D-----R---------F---F---F-------------Y-R.K--E-----VHN---F---W-------------  145
C.ZA.10.DEMC10ZA001           --N------IV---A-------.......TR.PTAKRER.......RAK---VG----Q--D---------Q..-------Q-.....D.D.-------R-----------G-F--GF--------D---W--.K--------V---------W-------------  146
C.ZM.11.DEMC11ZM006           --N----.....---AI---I--.......T-.................---EG----Q--D-F--L-T---KN--DV---Q-.....EGDG.-------R-----------G-F--GF--------D---Y-K.K--K-----V-----F---W-------------  136
D.CM.10.DEMD10CM009           ---------IV---AI-------.......TQ.PT.............E----G---V----------------Q---D-------.....D.G.-------R--------------------------D--VY--.K--E-----V---------W--------T----  141
D.KE.11.DEMD11KE003           ---------IV---AI---I--.......T-.................---AG---V-----R--------SQ---D-------.....-.D.-------R--------------------------VW-R.K--E-----V-----F---W--------I----  137
D.KR.04.04KBH8                -------.....--AI-------.......T-.................----G---------R-----R------D-------.....E.-.-------R----T--------L-------D-------VW--.K--------V-N---F---WH-------I----  133
D.SN.90.SE365                 ---------IVE--AI---I--.......T-.................----G---V--------V-----KD---D-------.....D.G..K-----R--------------------------VW-K.K--E-----V-----F---WH-------T----  138
D.TZ.01.A280                  ---------IV---AI---I-K.......-D.................----G---V-----R---------H---D-------.....A.D.-------R--A----------L-------D-------VW--.---------V-----F---WH-------T----  138
D.UG.10.DEMD10UG004           ---------IV---AI---I--.......TD.................---EG-----K-------------Q--P-------.....-.-.-.G-----R-----------G-L------------D-I-W--.K--------V---------W--------T----  137
D.UG.11.DEMD11UG003           ---------IV---AI---I--.......TA.................---EG------------------Q--S-------.....E.G.-----IR-----------G-L-----------------VW-P.K--E-----V---------W--------I----  138
D.YE.02.02YE516               ---------IV----I--KI-K.......TD.................----G---V--------------..N-P-------.....D.-.-------R--------------------------V---.---------V---------W--------I-H--  135
D.ZA.90.R1                    ---------IV---AI---I--.......TD.................----G---V--------------S---D-------.....S.-.-------K---------F-E------------------W-K.K--E-----V-N---I---W--------I----  138
F1.ES.02.ES_X845_4            ---------IV---A--D----.......TP................A--EG---V-Q--DRR-----------SDL------.....-.A.-------R-------------L------------D-----R.K--------V-----F---W--------I----  137
F1.RO.96.BCI_R07              ---------LA---A----I--.......TP.PAAE...........RG--EG---V----DRR--------G---PDL--V-.....D.G.-------R----P------G-L------R---------Y--.K-RE-----V-N-------W--------T----  143
F1.RU.08.D88_845              --A------IV---AI-D----.......-P.PAAPP.......AAEGX--EG---V----DQR--------R---PDL----------.....-.-.-------R-------------F-X-------------Y-K.K--E-----V---------W-----------L--  146
F2.CM.02.02CM_0016BBY         ---------IV---K------Q.......TP.................V--EG--KV-Q--D----------G---NDI-----.....D.-.-------R---------F-G--------------D---Y-R.K--E-----V---------W-------X.....  132
F2.CM.10.DEMF210CM001         ---------IV----I---L--.......TP.AAAAE..........GG--EG---V----D-Y--------R---DL---E.....-.-.-------R---------F-G-F--GF-------------Y-K.K--E-----V---------W--------I----  143
F2.CM.10.DEMF210CM007         ---------IV----I---IKQ.......TP.................V--EG---V-Q--D-Y--------KE--PDL-------...-D-.-.-------R---------F-G-F---F-------------Y-K.K--E-----V---------W--------T-F--  137
G.BE.96.DRCBL                 --N----RK-A---E----L-Q.......HP.................A--EG---V-Q--DR------R---G--PD------.....-..DS.------R-------------F---F--------D--VY-K.K--E-----V-----F---W--------T-V--  138
G.CM.10.DEMG10CM008           ---------IV---E----I--.......-P.PG...........QTPQ--AG---V-Q--DR-----------N-PD------.....-..DS.D-----R-----------G-F---F--------D---Y-K.K--E-----V-N---F---W--------T----  143
G.CN.08.GX_2084_08            #-------CIV---EI---I-Q.......TP.................---EG----Q--DR-------------PD------.....-..-S.------R---------F-G-F---F--------D---W-K.K--------V-N-------W--------T-F--  137
G.CU.99.Cu74                  ---------IV---A----I-Q.......TP.................---AG---V-Q--AR-----------N-PD---Q-.....-..DS.------R-------------F---F--------D---Y-K.K--------V-N---F---W--------T----  138
G.ES.09.X2634_2               -------R-IV---EI---L--.......TP.S..............A--EG---V-Q--AR-----------N-PDR------.....-..DS.------R-----------G-F---F--------D---Y-K.Q--------V-N-------W--------T-F--  139
G.GH.03.03GH175G              I------R-IKE--K-K-QI-Q.......-P.................A--EG----Q--D----L-------N-PD------.....-..DS.------K-------------F---F--------D---W-K.K--E-----V-N---F---W--------T-L--  138
G.KE.09.DEMG09KE001           --N-------V---Q----I--.......TP.PAAAGEG......AASP---AG----Q--A-----------TN-PD------...QD..-P.------R-------------F--GF--------D---Y--.K--------V-N---F---W--------I----  149
G.NG.09.09NG_SC62             -EN-----.....--A----I--.......TP.................A--EG----Q--TR---------SDR-PDV------...-Q-..DS.------R-------------F--GF--------D--TY-K.Q--------V-N-------W--------I----  134
H.BE.93.VI991                 -------GCIS---A----I-Q.......T-.................---EG---V-Q--DRR--V-IN-I-SN--DS------......E.-.-------R-----------G-F----------------Y-K.K--E-----V-N-------WH-------E----  138
H.CF.90.056                   --------RMG--S-I------.......--.................-V-EG---V----DRR--V-IN---S--RDA------......DG.-.-------R-----------G-F------------D---Y-K.Q--------V-N-------W--------E-F--  138
H.GB.00.00GBAC4001            ---R-----IV---RI---I--.......--.................---EG--PV----DXX--V-IN---DR-PDS------......E.-.-------R-----------G-L---------------VY-K.K--E-----V-N-------WH------------  138
J.CM.04.04CMU11421            ---------IV---Q--------.......-A..............APV----G---V-Q--AT---------T---D-------...-T-.-.-.----------------G-F--GF--------D---Y-K.K--E-----V-----F---W--------X----  140
J.SE.93.SE9280_7887           --N-----.....--Q--------.......-P.................A----G---V-Q--A---------------D-------...-T-.-.-.-------K--I-------G-----F--------D---Y-K.K--E-----VHN-------W--------I----  134
J.SE.94.SE9173_7022           --N-----.....--Q--D-----.......-A.P...............A----G---V-Q--A--------------DD-------...-T-.-.-.-------R-----------G-F---F--------D---Y-K.K--E-----VHN-------W--------TX---  135
K.CD.97.97ZR_EQTB11           ---------IV--S-------K.......TP.................----G---V-Q--D----V------FN-PD-------.....D.-..D-----R---------F-G-F--GF--------D---Y-K.---E-----V-----F---W--------I----  137
K.CM.96.96CM_MP535            ---------IV---AI------.......-R.PAADRV........GTQ----G---V-Q--AR---V-----SHN-PD------.....-.-.-.-------R-------------F--GF--------D---Y-K.K--E-----V-----F---W--------I----  146
U.CA.01.TV749                 -------R-IV---AX------.......-X.PXP..............A--XG---V-K---------T-.--SN-PDL--V-......E.G.-------R--------------------------X---Y--.X--------V-N---F---W--------I----  140
U.CA.99.TV721                 --A----.....---AI------.......TP.PTP.............T---EG---V-KS-Q---V-T-.A--N-PDL--V-......G.-.-------R---T-------F--GF--------D---Y--.Q--------V-N---FX--W--------I----  137
U.CD.83.83CD003_Z3            --N-----....Q--AI------.......-R.PAA.............E----G----------Y--------E--E---Q-----AQ-D.G.-------R-------------F--GF--------D---Y-K.K--E-----V-N-------W--------I----  138
U.CD.90.90CD121E12            --N-----....Q--AI-----Q.......-R.PAA.............E----G---V----DR------------S---Q-.....E.G..D-----K------------------F-----------Y-K.K--E-----V-N-------W--------I----  138
U.ES.10.DEURF10DZ001          ---------IV---EI---I--.......TP.TAAE............GA--EGA--V-Q--AR---V------T--PDL------.....-.-.-------R-------------F---F--------D---Y-K.K--------V-N---F---WH-------I----  142
U.GR.99.99GR303               ---------IV---EI---I--.......TP.PKDR..........QPL---EG---V-Q--DR---V-----R-N--NL------.....-.-.-.-EA--R--L--------E-L--------------W-----.Q--------F-N-------W--------I----  144
U.NL.95.U_NL_95_H10986_D1     ---------IV---A----I--.......TP.AAEGR..........RA----G---V----DR--------------D-------...-T-.-.-.-------R-------------F---F--------D------.K--------V---------W--------I----  144
01_AE.AF.07.569M              --S------IV---Q---KIKK.......TP.PAAEGVGAVSQDLDKH.G--EG---V----D----V----M..N--DNV--R-....EGDDEG-----R---------F-G-F---F--------D---Y-K.K--E-----V-----F---W--------T-F--  155
01_AE.CF.90.90CF11697         -------NRIV---Q----I--.......TP................A--EG----Q--D---------M..N--D-------.....-.-.-------R-----------G-F---F--------D---Y-K.K--E-----V---------W--------I----  135
01_AE.CN.10.YNFL03            --N----.....---R----IQQ.......TP.................---EG---V-Q--D----V----M..N--DST--R-.....D.-.-------K-----------G-F---F--R-----D---Y-K.K--E-----V-----F---W-------------  132
01_AE.HK.04.HK001             ---------IV---Q----I-Q.......TP.................--TEG---V-Q--D----V--N-M..N-DDSV--R-.....DE-..G-----R-----------D-F---F--------D---Y-K.K--E-----V-N---F---W--------I----  137
01_AE.IR.10.10IR.THR48F       --S------IV---Q----IEQ.......TP.................---EG---V-Q--D----V----M..N--D-V------.....X.-.-------R-----------G-F---F--------D---Y-K.K--------VHN---F---W--------I----  135
01_AE.JP.x.JRC77AE            ---------IV---QI---I--.......TP.................---EG---V-Q--D----V----M..N--ESV--R-.....DE-E.G-----R---------F-E-F---F-------------Y-K.K--E-----V-N---F---W--------T----  138
01_AE.TH.09.AA090a_WG11       --S-----.....--Q----I-K.......TP.PAA...........EEV.--KG---V-Q--A----V----M..N-PDSV--R-...H--DE-.-------R-------------E-F---F--------D---Y-K.K--E-----V-N---F---W-------------  138
01_AE.TH.90.CM240             --S------IV---Q---KIKQ.......TP.................--TEG---V-Q--D---------I..D--D-V--R-.....D.-.-------M-----------G-F---F--------D---Y-K.K--E-----V-N---F---W--------I-F--  135
01_AE.US.05.306163_FL         --------.....--Q----I--.......TP.................--TEG---V-Q--D---------#..N--D-T--R-.......D.-------K---------F-G-F---F--------D---Y-K.K--E-----V-N---F---W--------T-F--  130
01_AE.VN.98.98VNND15          ---------IV---QI---I--.......TP.................--TEG---V-Q--D----V----M..N--DSV--R-...H--ED-..G-----R---------F-G-F---F--------D---Y-K.K--------V-N---F---WH-------I----  137
```

myristoylation Nef start — acidic cluster — phosphorylation poly-P helix — HXB2 premature Nef end

```
B.FR.83.HXB2      MGGKWSKSSVIGWPTVRERMRR.......AE.....................PAADRVGAASRDLEKHGAITSSNTAATNAACAWLEA...QEE..E..EVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQ.RRQDILDLWIYHTQGYFPD*QNYTPGPGVRYPL  136
02_AG.CM.10.DE00210CM013  ---------IV---KI---I-H.......TP.ARGERAGG.......EEAP-EG-----Q--GRR---------S--PDV-------...-D.....---------R-----------H---F-----------Y-K.K--E-----V-----F---W-------T-F--  147
02_AG.ES.06.P1423         --A------IV---QI---------.....TP.....................E---X--Q-KD--E-S--N--T-------GA---.....-.....---------V-----------I-----------D--VW---K--E-----V-N-------W-------T-F--  137
02_AG.GW.05.CC_0048       ---------IV---KI-D-IKQ.....TP.....................EG-----Q--D----L-------E--PD-------...-.....---D-----R-----------S-F---F-----------Y-R.K--E-----V-----F---W-------I----  137
02_AG.KR.12.12MHR9        ---------IV---R----L--.....TP.PAAEGVGA.....ASTP---EG-----Q---RR--------------------...-.....EEG-----R----------G-F-F----R----------Y-K.K--E-----V-----F---W-------I---  151
02_AG.LR.x.POC44951       ---------LV---K----I-Q.....TP.VS.............ER--KG---V-Q--DRR-------GS---D-------...-.....---------R-----------F---F-----------D-----K--E-----V-----F---W-------T-F--  140
02_AG.NG.09.09NG_SC61     ---------IVE--K----I-Q.....TP.....................T-TG-----Q--D-----------E--PD-------...-D.....---K---------R-----------F---F-----------D--Y-K.K--E-V-N--F---W-------T-F--  137
02_AG.NG.x.IBNG           ---------IV---K-MK---Q.....TP.....................T--TG-----Q--DR---------Q--PD-------...-D...N---R---------G-----------D---F-----------D--Y-K.K--E-----V-N--F---W-------T-F--  137
02_AG.SE.94.SE7812        ---------IV---QI-D-I-Q.....TP.....................---RG-----Q--DR--------------------...-.....---R---------R-----------F---F-----------D--Y-K.K--E-----V-----F---W-------T-F--  137
02_AG.SN.98.98SE_MP1211   ---------LV---Q----I-.....TQPTP.....................S--IG---V-Q--DR------R---H---D-------...-.....---R---------R-----------------------Q-------Y-K.K--E-----V-----F---W-------I----  140
03_AB.RU.97.KAL153_2      ---------IV---Q----I-.....TP.....................P.....A---RG--PV-Q--D-Y--V-------N--D-------...-.....---R---------R-----------G-F---F-----------D--Y-K.K--E-----V-----F---W-------I-F--  133
04_cpx.CY.94.94CY032_3    ---------IV---EI---------.....R.AEPERMR...R..AQAE---AG---V-Q--D------IN------PDKT---------E.....---R---------F-G-L---------D--Y-K.K--E-----V-----F---WD------E-F--  150
05_DF.BE.x.VI1310         ---------V---AI--K---------.....TP.....................EG---VFQ--DNR--V-IR--V---PDM--V-------E.....---K---------R-----------F---F-----------Y-K.K--------V-N---S--W-------I----  138
06_cpx.AU.96.BFP90        --N---------SQ---------.....TP.PT.....................ER--EG---V-Q--D----------T-------V----T-------...-.....---R---V---------F---F-----------D---Y-K.K--E-----V-N---F---W-------T-F--  136
07_BC.CN.98.98CN009       ---------IV---DI-----T.....EQ.....................G---V-----------P---DD-------...-.....---K---------R-----------F---F-----------D---Y-R.K--E-----V---------WH------F--  137
08_BC.CN.97.97CNGX_6F     ---------IV---AI---I-.....T.....................G---V-----------D---D-----T.....-.....---R---------F-G-F---F-----------D---Y-K.K--E-----V---------WH------F--  137
09_cpx.GH.96.96GH2911     --------XIV---EI---I-.....TS.....................PSP-EG-----Q--GR---L-----S--HD-------...D.....D-----R---------G-----F-----------D---Y-K.K--E-----V-----F---W-------F--  138
10_CD.TZ.96.96TZ_BF061    ---------IV---Q----------.....TS.....................EG-----Q--A-F--L-----PGN--D-------...AE.....---R---------F-G-F--GF--------D---Y---K-K-----V-N---F---W-------T----  138
11_cpx.CM.95.95CM_1816    --------RIV---EI---I-.....TS.PT.....................A--PG---V-K--------V----GE-------...-.....E.....---R---------G-F--GF--------D---Y-K.K--E-----V-----F---W-------I----  140
12_BF.AR.99.ARMA159       ---------IV---EI---L-.....P.....................EG---V-Q---NR--------R-N-P-L-------...-.....---R---------F-G-F---------------D---Y-K.K--E-----V---------W-------L--  137
13_cpx.CM.96.96CM_1849    ---------R-TV---A----I-.....T.....................TG---V-Q--DR---V----L---.KDYV-------...G.....---K---------G-F--GF--------D---Y-K.K--------V---------W-------I----  130
14_BG.ES.05.X1870         ---------R-IV---EI---L-.....SP.....................EG---V-Q--AR---V--------N-PD-------...Q..DS.....---R---------G-F---F-----------D---Y-K.Q--------V-N---F---W-------I----  138
15_01B.TH.99.99TH_MU2079  --S-------.---Q---KIKQ.....TP.....................EG---V-Q--DN---V----M..N--DSV--R---.E-D.....---R---------F-E-F---F-----------D---Y-K.K--E-----V-N---F---W-------I----  131
16_A2D.KR.97.97KR004      -------R-LP---AI---------.....TP.PAAERT.....PPAA--EG---V-Q--ATR--V-IN-------DS-------...-.....E.....---R---------G-F------R----D---Y-K.K--------V-----F---W-G-----I----  147
17_BF.AR.99.ARMA038       ----------.....A---------.....P.....................E---EG---V-----RR--V-----S-N-PDL-------...-.....---R---------G-L---------------Y-K.---E-----V-N-------W-------I----  135
18_cpx.CU.99.CU76         ---------LV---EI---I-.....TQ.....................ELPRTPP-E-QEER---SG-N-S-----PTHGR--E...E.....---R---------G-F---F-----------D---FF-.K-PE----VFN---FL--W-------F--  137
19_cpx.CU.99.CU7          #--------I----A----Q-.....-.....................-----G---V-----------R----D-------...A.D.....---R---------G-L---------IVW---K-------V---------WH------I----  137
20_BG.CU.99.Cu103         --N-------.---Aa---L-.....TP.....................-----AG---V-Q--A----------N-PD-------...DS.....---R---------F---F-----------D---Y-K.Q--------V-N-------W-------T-F--  134
21_A2D.KE.99.KER2003      ---------IV----I---I-.....TD.....................G--P------R---------T--PD-------...D...G---R---I-----G-F---------------D---Y---K--E-----V-----F---W-------T-F--  138
22_01A1.CM.01.01CM_0001BBY ---------IV---Q----I-Q.....TP.....................A--EG---V-Q--D----V----M..NH-S-V-------...D..N.....---R---------G-F---F-----------D---Y-K.K--E-----V-N-------W------X.....  130
23_BG.CU.03.CB118         --N-------.--AA---------.....P.PAXN.....P---AG---V-Q--AR---------------DR--V---.....DS.....---R---------F---F-----------D---W-K.Q--------V-N---F---W-------I----  139
24_BG.ES.08.X2456_2       --N-------.SE--R----I-.....TP.LTTTRGGASQNSARQES---VG-----Q--AR----------N-PD-------...-Q..DS.....---R---------F---F-----------D---Y-K.Q--------V-N-------W-------T-F--  154
25_cpx.CM.02.1918LE       -------NR.V---EI---I-.....TP.PAE.....AAA--EG-----Q--A----------S--PD-------...S.....---R---------G-F---F-----------VW---K--E-----------W-------X.....  137
26_AU.CD.02.02CD_MBTB047  ----------V---E----------.....TP.....................TG---V---D----V----SF--PD-------...E.G.....---R---------F--GF--------D---Y---Q-K-----V-N---F---W-------I----  138
27_cpx.FR.04.04CD_FR_KZS  -----Q--PV---A---K---------.....P.....................G--V-Q--A------NN----N-PD--V-------.D.S.....---R---------F---F-----------D---W-K.K--E-----V-N-------W-------E-F--  138
28_BF.BR.99.BREPM12609    ---------RGTG--EA---------.....TP.PA.....AE.....R-----V-----R----------HN--D-------...-.....---R---------G-L-----R---------P.K--E-----V---------W-------T-F--  141
29_BF.BR.01.BREPM16704    ---------C-MS---AA---------.....T-PGP.....AE---EG---------Y-----------------------...-.....---R---------I-I---------------K--E-----V---------W-------I----  142
31_BC.BR.04.04BR142       --N----CIPV---AI---L-.....-H.PAAERIG..SERVRGAE---EG--P--Q--DR---L-----RDN--D-------...Q-E.....---R---------F---F-----------Y-K.Q--------V-N-------W-------T-F--  153
32_06A1.EE.01.EE0369      ---------IV---EI---L-.....TP.....................EG---V-Q--D-------H-PT-------...-D.....---R---------G-F--GF--------D---Y-K.K--E-----V-----F---W-------I----  137
33_01B.ID.07.JKT189_C     --S-C-E-CIV---Q----I-Q.....TP.....................EG---V-Q--D----V--T-M..N--DSV-------...-.....---R---------F-G-F---F--R------D---Y-K.K--------V-N---F---W-------I----  135
34_01B.TH.99.OUR2478P     ---------IV---Q----I-.....TP.....................EG---V-Q--D----V----M..N--D-V--R------NE.....---R---------F---F-----------Y-K.---E-----V-N---F---W-------T-F--  137
35_AD.AF.07.169H          ---------KKE--E-K-K--Q.....TP.....................A--EG---V-Q-----------I..NHPS-T--TQ...-D.....---R---------F-G-L--------Q---D---Y-R.K--E-----V-N-------W-------I----  136
36_cpx.CM.00.00CMNYU830   ---------I----QI-D---Q.....TP.....................IG---V----DR---------TS--DY-------...G.....---R---------G-F---F-----------D---K-------V-----F---W-------I----  131
37_cpx.CM.00.00CMNYU926   ---------LV---Q----IKQ.....TP.....................EG---V-Q--A------R--ST--PD-------...G--.DP.D---R---------F---F-----------D-------K---------N-------W-------.....  132
38_BF1.UY.03.UY03_3389    -------RRIV--ADI---------.....-.....-----------EG---------Q--DRR-------KDN-PDV-------...-.....---R---------G-----------------Y-K.---E-----V---------W-------I----  137
39_BF.BR.04.04BRRJ179     ---------K----AI-D---------.....TP.PTAER.....GAAK--EG---------SR--L-----KDN-PDV-------...E.....---R---------L---------------VW---K-------V---------W-------V--  146
40_BF.BR.05.05BRRJ055     --------R-P---EI---------.....---PAADG.....ER---G-----Q--A----L-----VSN--D-------...N.....---R---------R-----------HR--I------RKE---------K.---------V-----F---W-------I----  143
42_BF.LU.06.luBF_18_06    ---------LV---A-----K.....T.....................VG--P---------L-T---TK---V-------...-.....---R---------G-L---------------Y-----------------------W-------I----  137
43_02G.SA.03.J11223       -------R-MG--SEI---I-.....TP.PTAAE.....GAGT--EG-----Q--AR---------PQ---D-------...-D.....---R-R---------F--GF--------VY-K.K--E-----V-----F---WH------I----  145
44_BF.CL.00.CH80          ---------IV---DI---------.....TP.....................EG--KV-Q---RR----S-N-PDV-------...D.....---R---------F---------------K-------V-----F---W-------I----  137
45_cpx.FR.04.04FR_AUK     -------C-IV---EI---I-.....TP.....................TG--V-Q--D----V----------S-------...-N.....---R---------G-I---F-----------VY---K-------V---------W-------I----  138
46_BF.BR.07.07BR_FPS625   --NI--R-RIV---A----L-.....TP.....................EG-----Q---RR----N---RDN-PDV-------...D.....---R---------G-----------------TY-K.K-HE----V-----F---W-------I----  137
47_BF.ES.08.P1942         --N-----IV----I---I-.....TP.....................I-EG---V----G----------N-PDL-------...G.....---R---------G-F---F-----------Y-R.K--E-----V---------W-------I----  137
48_01B.MY.07.07MYKT021    ---------KV---Q----I-Q.....TP.....................EG---V-Q--D----V--T-M..N--DSV--R------...D.....---R---------G-F---F-----------G---Y-K.K--------V-N---F---W-------T-F--  136
49_cpx.GM.03.N26677       --SS----RIV---EI---------.....R.ATP.....TAPA---EG---V-Q--D----------P----------E...-T-.....---R---------F--G---X.....  99
50_A1D.GB.10.12792        -------R..V---ETEK---M.....GP.....................AP---PG---V-Q--A----------V..NHPS---------...D.....---R---------F--GF--------D---Y-K.K--E-----V-----F---W-------I----  135
51_01B.SG.11.11SG_HM021   --A------PV---KI----K-.....T-.PAA.....................EA--EG--P---------L------DN--D-------...H-D.....---R---------F-G-F---------------W---K-------V---------W-------E-F--  141
52_01B.MY.03.03MYKL018_1  -------C-----R----E-.....-G.....................G---------R---------------D-------...-.....---R---------G-F---F--R----D---Y-K.K--E-----V-----F---WH------I----  135
53_01B.MY.11.11FIR164     ---------IV---Q----I-.....TP.....................EG---V-Q--D----V-VT-M..N--DST--R------...E.D..G.....---R---------F-D-F---F-----------D---Y-R.K--E-----V-N---F---WS------T----  136
54_01B.MY.09.09MYSB023    ---------GRLG--TQI---------.....TP.AAAEE.....VGA--EG---V-K--D----V--T-M..N--DKV-------...D.....---R---------F--GF--------D---W-K.K--------V-N-------W-------T-F--  143
55_01B.CN.10.HNCS102056   --S------IV---Q----IKQ.....TP.....................S--EG---V-Q--D---------I..N--DSV--R------...E...G.....---R---------F-G-F---F-----------D---Y-T.K--E-----V-N---F---WH------------  136
56_cpx.FR.10.URF5_patient_A -------RIV---Q----I-Q.....TT.PX.....................VG---------D----L-T----QN-PD-------...-.....---R---------G-----------------D-------V---------W-------I----  84
57_BC.CN.09.09YNLX19sg    ---------IV---AI---------.....TQ.....................EG---------Q--D-Y--L-----DTN-TD---R------...E...G.....---R---------F-G-F---F-----------Y-K.K--E-----V---------W-------I----  138
58_01B.MY.09.09MYPR37     ---------IV---QI-KI-Q.....TP.....................EG---V-Q--D----V--T-L..N--D-S-RT-------...E.....---R---------F---F-----------D---Y-R.K--E-----V-N---F---W-------I----  136
59_01B.CN.09.09LNA423     ---------KA---Q----IEQ.....TP.....................V-TEG---V-Q--D-----V--I..N-TDSV--R------...DE..G.....---R---------F-G-F---F-----------D---Y-K.K--E-V-N--F---WH--X.....  129
60_BC.IT.11.BAV499        ---------IVR--AIK--L-.....T.....................EG---V-------K------L.N-PDRT--K----.K-D.K...-S---R---------G-L---------------Y---Q-R------V-N-------W------RI----  136
61_BC.CN.10.JL100010      -------C--V---S-----------.....-.....-----------G---V-------Y----------D-------...K.....---R---------G-L---------------Y-K.K--E-----V---------W-------I-F--  137
62_BC.CN.10.YNFL13        --S----R-IV---DI----------.....-.....-----------VG----Q--D-Y--L-----NT---D----S-------...E.G.....---R---------F-G-F---F-----------Y-K.K--E-----V-----F---WD------F--  138
63_02A1.RU.10.10RU6637    -------R-IV---Q----I-.....-P.....................AS--TG-PV-Q--D----V--------N--D-------...-.....---R---------G-L---------------Y-K.K--E-----V---------W-------T-F--  138
64_BC.CN.09.YNFL31        ---------IV---AI---I-Q.....TD.....................-----G---V-----------------DD-------...E.G..G.....---R---------F-G-F---F-----------Y-K.---E-----V---------W-------I-F--  138
65_cpx.CN.10.YNFL01       -N----R-IV---EI---L-.....TR.TD.....RTE----G---V----AR----------D-------...-.....---R---------F-G-F---F-----------Y-K.K--E-----V---------W-------F--  142
67_01B.CN.11.ANHUI_HF115  -------X-IV---Q--D-IKQ.....TP.....................A-TEG---V-Q--D--------I..N--DNX--R------...EG-E.....---R---------F---F-----------D--LY-K.K--E-----V-N---F---W-------I-F--  138
68_01B.CN.11.ANHUI_WH73   ---------IV---Q----IKQ.....TP.....................A-TEG---T-Q--D--------I..N--DNV--R------...EG-E.G.....---R---------F-G-F---F-----------D--VY-K.K--E-----V-N---F---W-------I-F--  138
70_BF1.BR.10.10BR_PE004   ---------IV---EI---L-.....NP.....................-----VG--------QR--L-IN----N-D---F----.H--E.....---R---------G-L---------------------------------------W-------I----  97
71_BF1.BR.10.10BR_PE008   ---------V----I--I-.....TP.....................-----G---V-----R----------.N--D-V-------...-.....---R---------G-L---------------W---K-------V---------W-------T-W--  135
72_BF1.BR.10.10BR_MG002   ---------IV---QI---------.....TP.TAAERRERIRQ..TPTAA--EG---V--------------..FN--D-V-------...-K-DE.....---R---------G-L-----R---------Y---K-------V---------W-------I----  152
O.BE.87.ANT70             --NALR-GKFE--AA-------TR.....TF.PES.....................E-C-PG--QI--E-AAR-G-P--H-PQN---L-F-S...HQ.....---R---------A---------G-F---F-----------Y-H.K-AE----V-N---F---W-------T-F--  143
O.CM.98.98CMA104          --NVLG--KFA--S-----------.....-PXPEP.....................E-C-PG--XL--E-AAR-G-P--Y-PQN---L-F--S...HQD..-D.....---R---V----F-G-F---F-----------D---Y-D.A-AE----VHN---FX--W-------T-F--  143
O.CM.98.98CMABB141        --NALR--KFE---A-------K.....TS.PEP.....................KDC-PG--QI-KE-AAR-G-P--Y-PQN---L-F--S...HQDE.....---R---V----F-G-F---F-----------D---Y-H.K-AE----VHN---F---W-------T-F--  142
O.CM.98.98CMABB212        --NA----KFA--SE--N---Q.....SS.PDP.....................Q-C-PG---V--E-ADR-G-P--Y-PQN---L-F-T-...H-D.....D-----R---------F-G-F---F-----------Y-H.K-AE----VH---F---W-C------I----  141
O.CM.98.98CMU5337         --NALR-HRDQR--A------N.....-P.S.....................A-C-PG--QV-QE-AAR-G-STAH-PQN---L-F--S...HQD.....D-----R---------F-G-F---F-----------Y---A-AE----V-N---F---W-------T-F--  139
O.CM.99.99CMU4122         --NALR--RFP--SAI-----K.....TS.PEP.....................E-C-XG--PV--E-AAR-G-P--Y-PQN---L-F--S...HSD.....---R---V----F---F---V----N-------Y-H.G-AE----N---F---W-------T-F--  141
O.FR.92.VAU               --NVLG--KWA--SE--R--Q.....TS.PDP.....QSSDTQ.P-PG-R-V--A-ADRRG-P--Y-PQN-Q-L-F-GS...HQG..K..GI----R---------G-F---F-------D---Y-H.K-AE----VH----F---W-------T-F--  145
O.GA.11.11Gab6352         --NVLG-DKFG---A-------TR.....TS.PES.....................E-P-AG---V--AIAEK-G-PN-H-PQN-E-L-F-DS...HQ.....---R---------F-G-H---V---------Y-H.A-AE----V-----F---W-H------T-F--  143
O.SN.99.99SE_MP1299       --NVLG-DIFK--SA------G.....TS.PDP.....................E-C-PG--QI--E-AAR-G-P--Y-PQN---L-F--S...HQD.....---R---------G-F---F-------D---Y-H.K-AE----V-----F---W-G------P-F--  141
N.CM.02.DJO0131           --KI-----IV---EI---------...QRP.H.EPA.....VEP-VG-----Q--ANR--L-T---RTN-PTV--V-------...G.....---R---------F-----------------W-R.K---V----V-----F---W-------T-F--  144
N.CM.02.SJGddd            --KA-----IV---A--------.ARQ..TQ.EXA.....AGAE--VG---V-Q--ANR-GV-T---KDN-ETV-------...-.....---C---------F---F-----------W-R.K--E-----V-----F---W-.....------  133
N.CM.04.04CM_1015_04      --KS-----V---EI---------....Q..RQ.TQA.....AA--VG-----Q--ANR----T--RDN-ETV--V-------...E.....---R---------F---F-----------VW-R.K--E-----------N-------W-------  143
N.FR.11.N1_FR_2011        --KI-----IV---EI---------....Q..TP.GAA.....VEP-AG-----Q--ANR----IN--KDN-QSV-------...H-D.....---R---------F-MGF---F---D-------VW-R.K--------V----I---W-------T----  142
P.CM.06.U14788            --NA-K---LV---A---KIKQ.....TT.PTPDPTTP.....VTP-PG--EI-KE-AQGKG-P-KFSSKN---L-F-D---.H-.....---R-----C------F---F-----------D--VY-P.K-AA---T-MFN---I--W-------E---  147
P.FR.09.RBF168            --NA-K---LV---A-----NQ.....TA...P.....-----ETP-SG--EI-KE-AQGKG-P-KYNSKN---L-F-----.H-D.....---R----V---------F---F-----------Y-P.G-AE----V-----F---W-H------E-F--  139
CPZ.CD.06.BF1167          --AVLG....TRF-GIDNMLN-.....IR.....................NTPAADAPG-Q--HNR-GL-TNTIGTEHDVLQHSGD...HT.-...A-----R----M----E-L-----W----------YW-P.K-AA---T-MFN---I--W-------------  133
CPZ.CM.05.SIVcpzMT145     --A------LV---EI---IK-.....-P.....................T---G--EV-Q-IASR--V-IR--PQ--ETL----E...MQD.....-K--------F-G-F---------------W-R.K--E-----V---------W-------------  137
CPZ.GA.88.GAB1            --T------LV---E--R-I-E.....-P.....................T--EG--EV-K---R----R--PE--QTL----E...MDN.....---R----T---------F---------------VY-R.---E-----V-----F---W----T---T-F--  137
CPZ.TZ.06.SIVcpzTAN13     --NI.....FGR-AGAT-VL-.....LH.....................A-PAHD-Q--KE-HER-GL-NNTIGTEKDVAHYSQD...HT.----------R-A--M----E-L-T---W-------D---Y-T.--AA---T-M-N---V---W-------------  132
CPZ.US.85.US_Marilyn      --N-----IV---E--N-L-Q.....TQ.T.....................TA--EG--PV-Q--AE-----TR--PQN-QTL---DEMTNH---.S.-----R---------Q-F-GF--------VY-R.---E-----V-----I---W-------------  134
GOR.CM.07.SIVgor2139_287  --NT-K---F----A------Q.....TN.PEP.....K.P-PG--EV--A-AQR-G-PAKY-PQN-Q-L-F----..H-D.-.-------R---------F-G-F---F--N--------VF---A-S------V---------W-------I----  140
GOR.CM.07.SIVgorCP2135con --NT-K---F----A------Q.....TS.PEP.....K.P-PG--EV--V-AQR-G-PAKY-PRN-Q-L-F----..H-D.-.-------X---------F-G-F---F-----------VF---A-SE-----V---------W-------I----  140
```

```
                                                                                                                           normal Nef end
B.FR.83.HXB2                 TFGWCYKLVPVEPDKI.EE.ANKGENTS.LLHPVSLHGMDD..PEREVLEWRFDSRLAFHHVARE.LHPEYF.K............NC*   205
A1.AU.03.PS1044_Day0         -----F-----D--EV.-K.-TEA--N-.----ICQ------..DD--T-M-K---K--LK-----.----FY.-............D-.   204
A1.CH.03.HIV_CH_BID_V3538    -----F-----DQTEV.--.-TV---N-.----ICQ------..E-----M-K---H---R-F-K-.----FY.-............D--   208
A1.ES.05.X1608_8             -----F-----D--EV.-K.--E--DNC.----M-Q--V--..E---T-M-K------LK-R---.I----Y.-............D--   207
A1.KE.11.DEMA11KE001         C----F-----D-EEV.-K.--E---N-.----MCQ------..D-K-T-V---------K-T---.M----Y.E............---   205
A1.RU.11.11RU6950            -----------D-ADV.--.-TE---N-.----ICQ------..E-K---K-Q------LT-R---.M----Y.-............D--   210
A1.RW.11.DEMA111RW002        -----F-----N-EEV.-R.--E---N-.----ICQ------..E-G-T-M--------LR-----.E---FY.-............D--   210
A1.SN.01.DDI579              -----F-----DQASV.--.-TE--...-----CQ------..K-K---M--------LK-T---.M---FY.-............D--   204
A1.UG.11.DEMA110UG001        -----F-----D-EEV.-K.-TE---N-.----MCQ------..D---T-M-K-------Q-T---.M---FY.-............D--   207
A1.ZA.04.04ZASK162B1         -----------QAEV.-Q.DTQ---N-.----ICQ--I---..Q-K---M-K------LK-L---.M---FY.-............D--   210
A2.CM.01.01CM_1445MV         -----F-----N-SEV.--.-TE---N-.----ICQ--I---..------K-----H--LR-R---.-Y-DFY.-............D--   204
A2.CY.94.94CY017_41          -----F-------SEV.--.-TQ---N-.----ICQ--V---..------R-E--RS--RR-R---.-----Y.-............D--   218

B.BR.10.10BR_MG029           -----------D---V.--.--E---NC.----MG----E-..------M-K-------------.M----Y.-TQELLTASNQ#.---   217
B.CA.07.502_1191_03          -----F-------EQV.--.-TV--TNN.----G-----E-..--K---V-----K-------K-.-K--F-.RDR............-   207
B.CH.08.M2_0803101_NFLG8     C----F-----D---V.--.--T---IN.----M-Q---E-..--K---V-K---------M---.----FY.-............D--   214
B.CN.10.DEMB10CN002          C----F----L--EE-.-K.-TE--D-C.------Q---E-..A-----V-K---------M---.-------.R............D--   210
B.ES.10.DEMB10ES002          C----F-----DREQV.--.--E---N-.----M-----E-..--K---Q-K-------------.-------.-............D--   212
B.FR.11.DEMB11FR001          -----F-----DKE--.--.--Q---NC.----M-Q-----..------V-K---L--VRYM---.K---FY.-............D--   224
B.GB.05.MM45d213_GN1         -----F-------E-V.--.--A----R.----I-Q-----..------Q-K---H---X-----#-----Y.-.............N-   209
B.HK.06.HK003                -----F----L--EEV.-K.--E---NC.----IN------..--G---V-----H--LR-M---.-----Y.-............---   211
B.JP.12.DEMB12JP001          -----F-----D--QV.-K.D-E---IN.----M-Q-----..-----*M-K---H--LQ-----.K----Y.-............D--   208
B.KR.07.HP_18_07JHS10_3909   -----F-------E-V.--.-TV---NC.----MN------..--G---V-K-------------.K----Y.-............D--   207
B.PE.07.502_0525_wg5         C----F-----DQEQV.-K.D-ER-DKC.----M-Q-----..--K---I-K---S-----K---.M---FY.-D.............-   206
B.RU.11.11RU21n              -----F-------EE-.-K.-TT---NC.----ANM-----..------V-K------Y------.K---FY.-............D--   207
B.TH.08.MERLBDTRC10          -----F-------EEV.-K.--E---NC.----M-Q---E-..------K-K---S--Y--R---.----FY.-............D--   208
B.US.11.ES38                 -----F-----D---V.--.--T---NC.----M-Q-----..T-K---V-K---H---N-----.-----Y.-............---   212

C.AR.01.ARG4006              -----F-----D-REV.--.--T---N-.----M-----E-..SDG---Q-K---L--RR-M---.-----Y.-............D--   216
C.BR.07.DEMC07BR003          -----F-----D-REV.--.--T---N-.----M-----E-..T-G---Q-K--RL--RR-M---.-----Y.-............D--   208
C.BW.00.00BW5031_1           -----------D-KEV.--.-TE---NC.----ICQ---E-..ED----R-K---E--RR-I---.R---FY.-............D--   208
C.CN.10.YNFL19               -----F-----D-KEV.--.--E---NC.-----CQ---E-..EH----K-Q--IQ--HR-M---.----FY.-............D--   206
C.ES.08.X2363_2              ----PF-----D-KEV.---.D-----AC.----A--------..AD--I-K-Q---S-VRI-K---.-----Y.-............D--   217
C.ET.02.02ET_288             -----F-----D-SEV.--.--E---NC.----I-------..EDK---M-K--IS--HR-L---.-----Y.-............D*-   207
C.IN.09.T125_2139            -----F--I--D-KEV.--.--E---NC.-----CQ---E-..EH----M-K---Q--RR-L---.Q---FY.-............D--   208
C.KE.00.KER2010              -----F-----D-REV.--.---E--NC.----M-Q--IE-..E-----R-Q---L--HR-L---.Q----Y.-............D--   208
C.MW.09.703010256_CH256.w96  -----F-----D-REV.--.--E---NC.----M-Q-----..DH----K-K---Q--HK-M---.----FY.-............D--   208
C.TZ.08.707010457_CH457.w8   -L---------D-KEV.---.ETE---N-.----M-----E-..EH----K-K---M--RR-M---.K----Y.-............D--   208
C.YE.02.02YE511              -----F-----D-REV.--.--E---N-.----I-----E-..GDG---M-K------HR------.----FY.-............D--   220
C.ZA.08.705010534_CH534.w12  -----F-----D-REV.--.--E---NC.----M-Q---E-..ED----M-K--IS--HR-M---.I----Y.-............D--   215
C.ZA.10.DEMC10ZA001          -----------D-REV.--.-----DNY.----I-----E-..E-G---K-Q---S--RR-L---.-----Y.-............D--   216
C.ZM.11.DEMC11ZM006          ----PF-----D-QEA.---.H--RDDSC.-M--M-Q--IE-..EH--I-M-K--TQ--RR-L---.----WY.-............D--   206

D.CM.10.DEMD10CM009          -----FE----D-RVV.-G.ETE----C.------Q---E-..T-----V-K-N-----E-K--Q.K---F-.-............D--   211
D.KE.11.DEMD11KE003          -----FE---L--EEV.-K.-TE---NC.------Q------..Q-----I---N-----E-K--Q.-----Y.-............---   207
D.KR.04.04KBH8               -----FE-----SEEV.--.--E---NC.----M-Q--IE-..---Q--K---N-K---E-K--M.----FY.-............D--   203
D.SN.90.SE365                C----FQ----D-EEV.--.-TE--DNC.----LCQ--IE-..E--Q--V---N-K---E-K--K.----FY.-............D--   208
D.TZ.01.A280                 -----FE----D-KEV.---.TTE---SC.----I-Q---E-..T-K---V---N-----E-K--S.-----Y.-............D--   208
D.UG.10.DEMD10UG004          -----FE----D-KEV.---.DTE---NC.----MCQ---E-..Q-----K---N-----E-K-KI.Q----Y.-............D--   207
D.UG.11.DEMD11UG003          -----F----IDAKEA.---.DTE---NC.----ACQ---E-..T-----M---N-----E-K--V.---DFY.-............D--   208
D.YE.02.02YE516              -----FE----D-KMV.---.E-E--DKC.----MHQ---G-..--K---V---N-K---E-K--K.T---F-.-............D--   205
D.ZA.90.R1                   -----F-----D-QEV.-Q.-TE--TNC.----MNQ---E-..------K-K-------E-K---.-----Y.-............D--   208

F1.ES.02.ES_X845_4           -L---------D-EEV.-K.--E---NC.----M-Q------..ADG---K-K---T--LR-I---.RF---Y.-D.............-   206
F1.RO.96.BCI_R07             --E--F-----D-EEV.-K.--E---NC.----M-Q---E-..ED----K-Q-------K-E---.R---FY.QD.............-   212
F1.RU.08.D88_845             -----F-----D-EEV.-R.--E----C.----M-Q------..EDG---R-K------LR-L---.K---FY.-D.............-   215
F2.CM.02.02CM_0016BBY        ...........................................................................................   132
F2.CM.10.DEMF210CM001        -----F-----S-EEV.-K.--E---NC.----M-Q---E-..EDG---M-K---G--RR-I---.K---FY.-D.............-   212
F2.CM.10.DEMF210CM007        -----F-----D-EEV.-K.--E--TNC.----I-Q--IE-..ED----M-K---T--MR-I-K-.K---FY.-D.............-   206

G.BE.96.DRCBL                -----F----M--SEV.--.------N-.----ICQ---E-..EDG---V-----S--RR-L---.-----Y.-............D--   208
G.CM.10.DEMG10CM008          ----------MD-AE-.--.-T----NC.----ICQ---E-..ED----V-----S--RR-I---.K---FY.-............D--   213
G.CN.08.GX_2084_08           -----F----MD-AEV.--.---E--N-.----ICQ--Q--..ED----V-----S--RR-R---.----FY.-............D--   207
G.CU.99.Cu74                 -----F----LD-AE-.--.-T----N-.----ICQ---E-..DD----I-----S--RR-L---.---DFY.-............D--   208
G.ES.09.X2634_2              -----F--E-MD-AEV.--.-TR---N-.----IC---AE-..GDK---R-K---S--RR-L---.----FY.-............D--   209
G.GH.03.03GH175G             -----F----ID-AEV.--.-----TN-.----ICQ--IE-..EDK---V-K---S--RR-L---.-----Y.-............---   208
G.KE.09.DEMG09KE001          -----F-----D-AE-.--.-TE--RNT.----ICQ---E-..SD----K-K---S--RR-I---.-Y---Y.-............---   219
G.NG.09.09NG_SC62            ----PF----MN-AE-.--.------N-.----I-Q---E-..DDK---V-K---S--RR-I---.---D-Y.-.............-*-   203

H.BE.93.VI991                -----F-----D-QDV.-K.--E---N-.----MCQ--IE-..------M-K------LR-R-K-.----FY.-............D--   208
H.CF.90.056                  -----F-----N-QEV.-Q.--E---N-.----M-----E-..DG----M-K------LT-L--V.K---.Y.-............D--   207
H.GB.00.00GBAC4001           C----F-----D-QEV.-K.--E---N-.----MCQ---E-..E-K---V-----H---R-R---.Q---FY.-............D--   208

J.CM.04.04CMU11421           -----------D-REV.--.--E---NC.----X-Q--IX-..X-X---M-K--XS--RR-X---.----FY.-............D--   210
J.SE.93.SE9280_7887          -----------D-SEV.--.--E---NC.----ICQ--IE-..E-----Q-K---S--RR-I---.----FY.-............D--   204
J.SE.94.SE9173_7022          -----F-----D-SEV.--.--E---NC.----ACQ--IE-..E-----K-K---S--RR-I---.----FY.-............D--   205

K.CD.97.97ZR_EQTB11          -----------D-REV.--.-TE---NC.-----NQ---E-..EH----K-K---S--RK-----.M----Y.-............D--   207
K.CM.96.96CM_MP535           -----------D-AEV.---.TTE--DNC.----INQ---E-..EH--I-M-K---S--RR-----.---D-Y.-............D--   216

U.CA.01.TV749                -----------D-XEV.--.-TE---N-.----MCQ--AE-..S-----R-K---K--LK-I-K-.---DFY.-............D--   210
U.CA.99.TV721                -----F-----D--EV.--.-TE---N-.----ICQ--V---..--K---K-Q---Q--RR-I---.-Y-DFY.-............D--   207
U.CD.83.83CD003_Z3           ----PF-----D-KEV.--.--E---NC.----ICQ------..E-K---M-K---S--RR-L---.----FY.-............D--   208
U.CD.90.90CD121E12           -L---F-----D-VEV.--.--E---N-.----ICQ---E-..E-K---M-K---H--RS-K---.-------.-............D--   208
U.ES.10.DEURF10DZ001         -----------D-REV.---.ET----NC.----LCQ--AE-..E-----R-K---S--RR-L---.K---FY.-............D--   212
U.GR.99.99GR303              -----------D-REV#--.--E---NC#----I-Q---E-..T-----I-K------LR-R--G.E----Y.-............D--   214
U.NL.95.U_NL_95_H10986_D1    -----------D-REV.--.--E---NC.----M-Q------..AD----R-E---S--RV-Q--V.-----Y.-............D--   214

01_AE.AF.07.569M             C----F-----D-STV.---.D-----NC.----M-Q--IE-..E-----I-K---A--RK-I---.-----Y.-............D--   225
01_AE.CF.90.90CF11697        C----------D-KEV.---.D--D-SNC.----M-Q--V---..E-----M-K---S--RR-I---.-R---Y.-............D--   205
01_AE.CN.10.YNFL03           C----F-----D-RAV.---.ET-E-.SC.----INQ--V---..ED----M-K---A--RK-I---.Q---FY.-............---   201
01_AE.HK.04.HK001            C----F-----D-REV.---.D-----NC.----A-Q--I---..E-----M-K---A--RK-L---.----FY.-............D--   207
01_AE.IR.10.10IR.THR48F      C----F-----DRQEV.--.------NC.----M-Q--QE-..E-----R-K---X--RK-----.-----Y.-............D--   205
01_AE.JP.x.JRC77AE           C----F-----D-TEV.---.D----DNC.----M-Q--I---..ED----M-K---A--RK-I---.-----Y.-............D--   208
01_AE.TH.09.AA090a_WG11      C----F-----D-REV.---.DT----NC.----M-Q--V---..E-----M-K---A--RK-I---.-R---Y.-............D--   208
01_AE.TH.90.CM240            C----F-----DQREV.---.D-----NC.----M-Q--IE-..E-----M-K---A--RK-----.Q----Y.-............D--   205
01_AE.US.05.306163_FL        C----F-----D-REV.--.------NC.----L-Q--IE-..E-----M-K---A--RK----D.-----Y.-............DX.   199
01_AE.VN.98.98VNND15         C----F-----D-REV.---.DT----NC.------Q--LE-..G-----M-K---A--RK-I---.---D-Y.-............D--   207
```

normal Nef end

```
B.FR.83.HXB2                TFGWCYKLVPVEPDKI.EE.ANKGENTS.LLHPVSLHGMDD..PEREVLEWRFDSRLAFHHVARE.LHPEYF.K.............NC*  205
02_AG.CM.10.DE00210CM013    -----F----LD-VE-..R.--E---N.-----ICQ------..ED----V-----T---K-K---.---FY.-D.............-  216
02_AG.ES.06.P1423           ----------M--SEV.-K.--E---NC.----ICQ------..ED----V-----L---K-T---.------Y.............D--  207
02_AG.GW.05.CC_0048         -----F----MD-TEV.---L--E--N.------Q---E-..ED----V-----X---X.............................  192
02_AG.KR.12.12MHR9          -----F----LD-AEV.---.---E--N.------CQ---E-..ED----V---G-----T-K-K-.-----FY.............D--  221
02_AG.LR.x.POC44951         -----F----MD-AEV.---.------N.-----ICQ--IE-..TD-Q--V-----S---K-R-Q-.-----FY.............D--  210
02_AG.NG.09.09NG_SC61       -----F----MD-AD-..---.---N.-----ICQ---E-..E-----V-----S--RR-I---.-----Y.............D--  207
02_AG.NG.x.IBNG             -----F----MD-AEV.---.--E---N.-----ICQ---E-..DD----I---------R-T---.-----Y.............D--  207
02_AG.SE.94.SE7812          -----F----MD-AD-..K.DTE---N.-----ICQ---E-..ED----V---------T-K---.M---FY.............D--  207
02_AG.SN.98.98SE_MP1211     ----------MD-AEV.---.--Q---N.-----ICQ---E-..EDK---V---------R-T---.-----Y.............D--  210
03_AB.RU.97.KAL153_2        ----------D-AEV.----TE---N.-----ICQ------..E-K---M-K------LT-R-----.---FY.............D--  203
04_cpx.CY.94.94CY032_3      C----F-----D-QEV.---.TE----C.-----I-Q---E-..E-----K-K------YK------.----FY.............D--  220
05_DF.BE.x.VI1310           -L---F-----N-EEV..K.--E--DNC.----M------E-..DD----Q-K---S--LR-I---.R---FY.QD..........-.  207
06_cpx.AU.96.BFP90          ----------D-EEV.---LT----NC.----ICQ--AE-..E-----K-K---S--RR-I---.K---FY.............D--  206
07_BC.CN.98.98CN009         -----F-----D-REV.---.--E---NC.------CQ---E-..EH----K-K---Q--HR-R-----.-----FY.............D--  207
08_BC.CN.97.97CNGX_6F       -----F-----D-REV.---.--E--DNC.------CQ---E-..EH----K-K---Q--HR-R-----.-----FY.............D--  207
09_cpx.GH.96.96GH2911       -----F-----D-KEV.---.TG---N.-----MCQ--V-...SGG---M-----T--LK-I---.K---FY.............D--  207
10_CD.TZ.96.96TZ_BF061      -----F-----D-REV.---.------NC.----M-----E-..KHG---V-K---T--HK-I---.-----Y.............D--  208
11_cpx.CM.95.95CM_1816      C----F-----D-REV.---.--A---NC.----L-Q--T-...E-----M-K---S--RR-I---.---DFY.-D..........-.  209
12_BF.AR.99.ARMA159         -----F-----D-VEV.-K.--E---NC.----M-Q---E-..ED----I-K------LR-L---.K---WY.-D..........-.  206
13_cpx.CM.96.96CM_1849      C-E--F-----ID-KEV.---.--E---NC.----M-Q--IE-..A-----M-K---S--RK-I---.----FY.............D--  200
14_BG.ES.05.X1870           ------F--E--D-AEV.---.TG---N.-----ICQ--IE-..ADK---R-K---G--RR-I---.R-DFY.............D--  208
15_01B.TH.99.99TH_MU2079    C----F-----D-SDV.---N-----NC.----M-Q---E-..E-----K-Q---A--RK-I-------.-----Y.............D--  201
16_A2D.KR.97.97KR004        -----F-----S-AEV.---.TE--TN-.-----ICQ--IE-..------K-V---H--LV-K-----.---FY.............D--  217
17_BF.AR.99.ARMA038         -L-----------EE-..K.--E---NC.----M-Q-E-----.ED----K-K---S---R-----.R---FY.QN..........-.  204
18_cpx.CU.99.CU76           -----F-----DSNEV.---.--E---N.-----ICQ------..E-----K-K---A--RR-I---.Q---FY.............D--  207
19_cpx.CU.99.CU7            -----FE----D-KEA..K.DTE---NC.----ACQ------..--K---V---N-----E-K--I.K---FY.............D--  207
20_BG.CU.99.Cu103           -----F----MD-AEV.---.--N---N.-----ICQ---E-..N-----I-----S--RR-L-----.----FY.............D--  204
21_A2D.KE.99.KER2003        -----F-----D-SEV.---.DT----N.----MCQQ-I---.------R-K--PH--LK-R-Q-.K---FY.............D--  208
22_01A1.CM.01.01CM_0001BBY  ..........................................................................................  130
23_BG.CU.03.CB118           -----F-----D-ADV.---.SN---N.-----ICQ---E-..DDK---M-----S--RR-I---.R--DFY.............D--  209
24_BG.ES.08.X2456_2         -----F-----D-TEV.---.--S--TN-.-----ICQ---E-..SDK---I-K---S--RR-I---.---DFY.............D--  224
25_cpx.CM.02.1918LE         ..........................................................................................  137
26_AU.CD.02.02CD_MBTB047    -----F-----D-AEV.---.------N.----AM-----E-..S-G---M-K---Q--RK-T---.-----Y.............D--  208
27_cpx.FR.04.04CD_FR_KZS    -L---F-----D-KEV.---.--E---N.-----M-----E-..E-----M-K------LR-----.K---FY.-.............D--  208
28_BF.BR.99.BREPM12609      -----F-----D-QEV.-A.--E---N.-----M-Q------..---K---M-K--------M-----.----Y.-D..........-.  210
29_BF.BR.01.BREPM16704      -----F-----D--QV.---.E---N.-----M-Q------..---K---M-K--------M-----.----F-.............D--  212
31_BC.BR.04.04BR142         -----F-----D-REV.---.--A---N.-----M-----E-..THG---M-K---L--RR-M-----.----FY.............D--  223
32_06A1.EE.01.EE0369        ----------D-REV.----.T----NC.----IC---AE-..E-G---M-----S--RR-I---.K---FY.............D--  207
33_01B.ID.07.JKT189_C       C----F-----D-TEV.---.E-----NC.----M-Q---E-..ED----K-Q---S--RR-I---.Q---Y.............D--  205
34_01B.TH.99.OUR2478P       C----F-----D-GEV.K-.-----SC.----M-Q--IE-..E-----M-K---A--RK-I-------.----FY.............D--  203
35_AD.AF.07.169H            ----------D-AEV..K.D-E---N.-----MCQ------..E---T-M-K-------T-K---.----FY.............D--  206
36_cpx.CM.00.00CMNYU830     ..........................................................................................  131
37_cpx.CM.00.00CMNYU926     ..........................................................................................  132
38_BF1.UY.03.UY03_3389      -M---F-----D--EV.-K.--E---NC.----M-Q--A---.EDK---V--------LR-I---.R---FY.QD..........-.  206
39_BF.BR.04.04BRRJ179       -----F-----S-EEV.-R.--E--DNC.----M-Q------..ED----V-K---K---------.-----Y.............D--  216
40_BF.BR.05.05BRRJ055       -----F-----D-EEV.-K.--E---NC.----M-----E-..E-K---Q----AT---------.----FH.............---  213
42_BF.LU.06.luBF_18_06      C----F----L-Q-QV.-K.-TE-----.----I-----E-..T-----V-K---K-------H-.----FY.............D--  207
43_02G.SA.03.J11223         -----F----MD-ADV.--.----E--N.-----ICQ---E-..ED----V-----S--RR-I---.----Y.............D--  215
44_BF.CL.00.CH80            -L---F-----D-EEV.-K.--E----G.----M-Q------..ADK---M-K---S--LR-I----.RF--FYQDD..........--  207
45_cpx.FR.04.04FR_AUK       -----F-----D-KEV.---.TE----C.----I-Q---IE-..T-----I-K---H--QR-L---.K---FY.-.............D--  208
46_BF.BR.07.07BR_FPS625     -----F-----N-EEV.-R.--E---NC.----I-Q---E-..E-----Q-K---S----R-----.---DFY.-NCCI.....EL---  213
47_BF.ES.08.P1942           -----F-----D-EEV.-K.--E---NC.----M-Q---E-..ED----M-K--RH--SK------.-----Y.-.............D--  207
48_01B.MY.07.07MYKT021      C----F-----D-QEV.---.D-----SC.----M-Q---E-..E-----M-K---S--RR-I----.R---Y.-.............D--  206
49_cpx.GM.03.N26677         ..........................................................................................  99
50_A1D.GB.10.12792          -----------QEEV.-K.-TE-G-NC.----ICQ------..E-K---Q-K---S--RK-I-------.----FY.............D--  205
51_01B.SG.11.11SG_HM021     -----FX----D--QV.-K.--EK----.----M-Q--N-...--K---M-K------LR-M-K-.K--D-Y.-.............D--  211
52_01B.MY.03.03MYKL018_1    C----F-----D-KEV.---.D-----SC.----I-Q--QE-..E-G---M-K---A--RT-T-------.----FY.-.............D--  205
53_01B.MY.11.11FIR164       C----F-----D-AEV.---.---E--NC.----M-Q---E-..D-----I-K---S--RR-I---.-----Y.-.............D--  206
54_01B.MY.09.09MYSB023      C----F-----D-REV.---.------NC.----M-Q--AE-..E-K---M-K---A--RT-----.M-----.-.-----------D--  213
55_01B.CN.10.HNCS102056     C----F-----D-REV.---.TT-E--NC.----M-Q------..E-K---M-K---A--RK-R---.----FY.-.............D--  206
56_cpx.FR.10.URF5_patient_A ..........................................................................................  84
57_BC.CN.09.09YNLX19sg      -----F-----D-RE-.---.--E---NC.----CQ---E-..EH----R-K---Q--HR-R---.----FY.R.-.........D--  208
58_01B.MY.09.09MYPR37       C----F-----D-REV.---.------NC.----M-Q--IE-..N-----M-K---A--RK-I---.K---Y.-.-.........D--  206
59_01B.CN.09.09LNA423       ..........................................................................................  129
60_BC.IT.11.BAV499          C----F------TKD.....IEE-SSC.----MQQ-RVE..------M-K---S-----M---.K-----.-.-.........D--  202
61_BC.CN.10.JL100010        -----F-------KEV.---.TE---NC.------CQ---E-..EH----M-K---H---S-----.K---WY.-.-.........D--  207
62_BC.CN.10.YNFL13          -----F-----D-REV.---.--E---NC.----M-----E-..EHG---R-K---Q--HK-R---.----FY.-.-.........D--  208
63_02A1.RU.10.10RU6637      -----F----MD-AE-.---.TE---N.-----ICQ--LE-..ED----V-----K---D-----.R---FY.-.-.........D--  208
64_BC.CN.09.YNFL31          -----F-----D-REV.---.--E---NC.------CQ---E-..EH----M-K---Q--HR-R---.----FY.-.-.........D--  208
65_cpx.CN.10.YNFL01         -----F-----D-REV.---.S-E---NC.----M-Q------..EHG---V-K---M--HR-M---.Q---FY.-.-.........D--  212
67_01B.CN.11.ANHUI_HF115    C----F-----D-REV.---.D-----KC.----M-Q---E-..E-----M-K---A--RK-I---.---X-Y.-.-.........D--  208
68_01B.CN.11.ANHUI_WH73     C----F-----D-RDV.---.D--R-DSC.----I-Q--I---.E-----M-K---A--RK-I---.---D-Y.-.-.........D--  208
70_BF1.BR.10.10BR_PE004     ..........................................................................................  97
71_BF1.BR.10.10BR_PE008     -----F-----D-E-V.---.---K--S-.----M-----E-..A-K---V-K---S-G-------.-----Y.-.-.........-.  205
72_BF1.BR.10.10BR_MG002     -----F----L--EEV.-K.-SE#--NC.----AN------P.G-G---V----------M---.------Y.-N.........-.  221
O.BE.87.ANT70               ----LF-----SEEEA.-RLG-TC-RAN.----ACA--FE-..THK-I-M-K--RS-GNT---MI.T---L-Q-D.........-.  214
O.CM.98.98CMA104            ----MF-----SEAEA.--LG--H-RAK.----ACA--FG-..AHG-I-K-Q--RS-GLT---KI.T---L-P-D.........-.  214
O.CM.98.98CMABB141          -Y--L------TREEA.-RLG-EC-RA-.----ACN--FE-..HG-I-K-Q--RS-GNT---KI.T---X-X-D.........-.  213
O.CM.98.98CMABB212          ----L------TAEEA.-RLG-TR-DA-.----ACN--AE-..AHG-I-K-Q--RT-GLT---LQ.K---L-.PK.........-.  211
O.CM.98.98CMU5337           ----MF-----SEAEA.DXLG--H-RAK.----ACN--FE-..-HG-I-K-Q--RS-GGT---MI.T--HL-N-D.........-.  210
O.CM.99.99CMU4122           ----LF-----SEAEA.--LG--CVTAK.----ACN--SE-..-HG-I-K-Q--RS-GNI---KV.T---L-L-D.........-.  212
O.FR.92.VAU                 ----LF-----SEAEA.-ALG--C-RA-.----ACN--YE-..QHK-I-K-Q--RS--NT---LI.T---L-S-D.........-.  216
O.GA.11.11Gab6352           ----LF-----TE-EA.KRLG-ECVRAH.----ACA--FE-..-HG-I-K-K---S-GLT--SKV.I---L-T-D.........-.  214
O.SN.99.99SE_MP1299         ----LF-----SEAEA.--LG--C-RA-.----ACN--FE-..NHGQI-K-Q--RS-GST---MV.TN--L-N-D.........-.  212
N.CM.02.DJO0131             -----F----LSAEAV.---.-E-D-NA.----LCQ--V---.GHK---Q-Q---S--RR-I---.---DFY.-.............  214
N.CM.02.SJGddd              ..........................................................................................  133
N.CM.04.04CM_1015_04        -----F----LSAEAV.---.-E-D-NA.----ICQ--V---.DHK---V-----S--RR-I---.---DFY.-.............  213
N.FR.11.N1_FR_2011          -----F----LSAEAV.---.TE---N.-----ICQ--V---.DHK---V-----S--RR---K-.-----F-.-.............D--  212
P.CM.06.U14788              ----LF-----SEVEA.--MGD-Q-KAK.----ACTY-FS-..-HK-I-V-K---S-GRE---LQ.K---L-I-D.........-.  218
P.FR.09.RBF168              ----LF-----SEVEV.--MGDEQ-KAK.----ACTY-DS-..HH--I-M-K--RS-GER---LQ.K---L-T-D.........-.  210
CPZ.CD.06.BF1167            CR--LF----AD-PE.......ED-RNL.-M--ACSY-R----.-NG-L-V-K---E--RR-I---.R-----.-.........-.  196
CPZ.CM.05.SIVcpzMT145       -----F----LP-EEV.-K.--E---NC.----MCQ---E-..E-----V-T---K--LR-L---.K----Y.R.........D.-  206
CPZ.GA.88.GAB1              C----F----LTEEQV.-Q.--E-D-NC.----ICQ---E-..EDK---V--------LR-I---.Q----Y.-.........D--  206
CPZ.TZ.06.SIVcpzTAN13       CR--LF-----T-PE.......DD-RNI.----ANS--RE-PDGDG-T-M-Q--TS--RR-I---.R-----.-.........-.  197
CPZ.US.85.US_Marilyn        -Y---F----LTEEEV.-Q.-----TNI.----MCQ---E-..EHG---I-Q--TE--RR-R-K-.------.R.........-.-  210
GOR.CM.07.SIVgor2139_287    ----LF-----SEAEA.--MGTSN-KAM.----ACS--R----.YG-I-V----RA-GTT---LQ.K---L-KRD.........-.  211
GOR.CM.07.SIVgorCP2135con   ----LF-----SEAEA.--MGTAS-RAI.----ACS--R----.HG-I-V----RA-GTT---LQ.K---L-KRE.........-.  211
```

VI

HIV-2/SIV Proteins

Contents

VI-1	Introduction	**371**
VI-2	Annotated features	**372**
VI-3	Sequences	**373**
VI-4	Alignments	**378**
VI-4.1	Gag	378
VI-4.2	Pol	382
VI-4.3	Vif	389
VI-4.4	Vpx	391
VI-4.5	Vpr	392
VI-4.6	Tat	393
VI-4.7	Rev	394
VI-4.8	Env	396
VI-4.9	Nef	402

VI-1 Introduction

The number of HIV-2 and SIV sequences is growing much more slowly than for HIV-1, particularly for complete or near-complete genomes. In selecting the set of sequences to present here, we eliminated some sequences derived from the same sample, patient, or isolate. We further omitted a few that were very closely related.

The selection of the HIV-2 protein alignments was slighly different than in previous years. In addition to the proteins from the full genome alignment, we have included all of the gene-length fragments available for each protein. There are many additional gene-length sequences included for Nef, and a few for the other genes. The differing numbers of sequences in each protein alignment represent the differing availability of gene-length HIV-2 sequences.

In these alignments, as in the database, we label sequences according to their native host. Thus, transfers of SIVsmm into humans have produced HIV-2 groups A through G. Unintentional transfers of SIVsmm into captive macaques are named according to the receptive host (Rhesus macaques = mac, Stump-tailed macaques = stm, and Pig-tailed macaques = mne). However, intentional, experimental cross-species transfers of virus remain labeled as being from the original host (e.g., HIV-1 injected into a chimpanzee is labeled HIV-1, not SIVcpz; and SIVsmm injected into a Rhesus macaque is labeled SIVsmm, not SIVmac).

VI-2 Annotated features

Features of HIV-2 annotated in the alignment that follows.

Feature	Protein	Location	Page
Gag p15 start	Gag	1	378
p15 end	Gag	135	378
p27 start	Gag	136	378
p27 end	Gag	364	380
p2 start	Gag	365	380
p2 end	Gag	381	380
p8 start	Gag	382	380
p8 end	Gag	433	380
p1 start	Gag	434	380
p1 end	Gag	447	380
p6 start	Gag	448	380
PTAP motif	Gag	458-461	380
PSAP in HIV-2 B, U	Gag	476-479	380
p6 end	Gag	511	381
Gag end	Gag	511	381
Pol p15 start	Pol	1	382
p15 end	Pol	67	382
protease start	Pol	68	382
protease end	Pol	166	383
p51 RT start	Pol	167	383
D catalytic site	Pol	276	383
DD catalytic site	Pol	351	384
p51 RT end	Pol	605	385
p15 RNase H start	Pol	606	385
p15 RNase H end	Pol	725	386
p31 Integrase start	Pol	726	386
p31 Integrase end	Pol	1020	388
Pol end	Pol	1020	388
Vif start	Vif	1	389
Vif end	Vif	215	390
Vpx start	Vpx	1	391
Vpx end	Vpx	113	391
Vpr start	Vpr	1	392
Vpr end	Vpr	102	392
Tat start	Tat	1	393
exon 1 end	Tat	99	393
exon 2 start	Tat	100	393
Tat end	Tat	131	393
Rev start	Rev	1	394
exon 1 end	Rev	24	394
exon 2 start	Rev	25	394
Rev end	Rev	108	395
Env start	Env	1	396
signal peptide end	Env	22	396
gp120 start	Env	23	396
V1	Env	113-167	396
V2	Env	169-211	397
V3	Env	311-344	398

Feature	Protein	Location	Page
V3 tip	Env	328-334	398
V4	Env	402-432	398
V5	Env	472-477	399
gp120 end	Env	525	399
gp41 start	Env	526	399
Env end	Env	880	401
gp41 end	Env	880	401
Nef start	Nef	1	402
R17Y mutation	Nef	17	402
max HIV-1 similarity	Nef	153-182	402
premature stop in Mac239	Nef	93	402
normal Nef end	Nef	264	403

VI-3 Sequences

Sequences included in the HIV-2/SIV protein alignments.

Name	Accession	Proteins	Author	Reference
MAC.US.x.239	M33262	All	Kestler, H.	*Science* **248**(4959):1109-1112 (1990)
A.CI.88.UC2	U38293	All	Barnett, S.W.	*Virology* **222**(1); 257-61 (1996)
A.DE.x.BEN	M30502	All	Kirchhoff, F.	*Virology* **177**(1):305-311 (1990)
A.DE.x.PEI2_KR_KRCG	U22047	All	Kraus, G.	*ARHR* **14**(1); 65-77 (1998)
A.GH.x.GH1	M30895	All	Hasegawa, A.	*ARHR* **5**(6):593-604 (1989)
A.GM.87.D194	J04542	All	Kuehnel, H.	*Proc Natl Acad Sci USA* **86**(7):2383-2387 (1989)
A.GM.x.ISY_SBL_6669_85	J04498	All	Franchini, G.	*Proc Natl Acad Sci USA* **86**(7):2433-2437 (1989)
A.GM.x.MCN13	AY509259	All	Schmitz, C.	*J Virol* **78**(4):2006-2016 (2004)
A.GM.x.MCR35	AY509260	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Nef	Schmitz, C.	*J Virol* **78**(4):2006-2016 (2004)
A.GW.86.FG_clone_NIHZ	J03654	All	Zagury, J.F.	*Proc Natl Acad Sci USA* **85**(16):5941-5945 (1988)
A.GW.87.CAM2CG	D00835	All	Tristem, M.	*J Gen Virol* **72**(PT 3):721-724 (1991)
A.GW.x.MDS	Z48731	All	Becker	Unpublished
A.IN.07.NNVA	EU980602	All	Gurjar, S.R.	*JAIDS* **52**(3); 329-35 (2009)
A.IN.95.CRIK_147	DQ307022	All	Santhosh, C.V.	*ARHR* **24**(10); 1315-7 (2008)
A.IN.x.NIM_8	DQ973520	Gag	Gurjar, R.S.	*ARHR* **25**(3); 363-72 (2009)
A.JP.08.NMC786_clone_41	AB731742	All	Umeki-Sakamoto	Unpublished
A.PT.x.ALI	AF082339	All	Reeves, J.D.	*J Virol* **73**(9); 7795-804 (1999)
A.SN.85.ROD	M15390	All	Clavel, F.	*Nature* **324**(6098):691-695 (1986)
A.SN.86.ST_JSP4_27	M31113	All	Kumar, P.	*J Virol* **64**(2):890-901 (1990)
B.CI.88.UC1	L07625	All	Barnett, S.W.	*J Virol* **67**(2):1006-1014 (1993)
B.CI.x.20_56	AB485670	All	Takekawa	Unpublished
B.CI.x.EHO	U27200	All	Rey-Cuille, M.A.	*Virology* **202**(1):471-476 (1994)
B.GH.86.D205_ALT	X61240	All	Dietrich, U.	*Nature* **342**(6252):948-950 (1989)
B.JP.01.IMCJ_KR020_1	AB100245	All	Kusagawa, S.	*ARHR* **19**(11):1045-1049 (2003)
G.CI.92.Abt96	AF208027	All	Brennan, C.A.	*ARHR* **13**(5):401-404 (1997)

Name	Accession	Proteins	Author	Reference
AB.CM.03.03CM_510_03	EU028345	All	Yamaguchi, J.	*ARHR* **24**(1):86-91 (2008)
H2_01_AB.CI.90.7312A	L36874	All	Gao	Unpublished
H2_01_AB.JP.04.NMC307_20	AB731738	All	Umeki-Sakamoto	Unpublished
H2_01_AB.JP.07.NMC716_01	AB731740	All	Umeki-Sakamoto	Unpublished
H2_01_AB.JP.08.NMC842_10	AB731744	All	Umeki-Sakamoto	Unpublished
U.CI.07.07IC_TNP3	KC693505	All	Ayouba, A.	*AIDS* **27**(15):2488-2491 (2013)
U.FR.96.12034	AY530889	All	Damond, F.	*ARHR* **20**(6):666-672 (2004)
MAC.US.x.17EC1	AY033233	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Nef	Anderson, M.G.	*Virology* **195**(2):616-626 (1993)
MAC.US.x.251_1A11	M76764	All	Marthas, M.L.	*J Med Primatol* **18**(3-4):311-9 (1989)
MAC.US.x.251_32H_PJ5	D01065	All	Rud, E.W.	*(in) Brown,F, Chanock,RM and Ginsberg,HS(Eds); VACCINES* **92**: MODERN APPROACHES TO NEW VACCINES INCLUDING PREVENTION OF AIDS: 229-235; Cold Spring Harbor Laboratory Press, New York (1992)
MAC.US.x.251_BK28	M19499	All	Hirsch, V.	*Cell* **49**(3):307-319 (1987)
MAC.US.x.MM142_IVMXX	Y00277	All	Chakrabarti, L.	*Nature* **328**(6130):543-547 (1987)
MNE.US.82.MNE_8	M32741	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Nef	Kimata, J.T.	*J Virol* **72**(1):245-256 (1998)
MNE.US.x.MNE027	U79412	All	Kimata, J.T.	*J Virol* **72**(1):245-256 (1998)
SMM.CI.79.SIVsmCI2	JX860430	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.LR.89.SIVsmLIB1	JX860431	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.SL.92.SIVsmSL92A	JX860432	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.SL.92.SL92B	AF334679	All	Chen, Z.	*J Virol* **70**(6):3617-3627 (1996)
SMM.US.04.G078	JX860415	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.G932	JX860416	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M919	JX860417	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M922	JX860418	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M923	JX860419	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M926	JX860420	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M934	JX860421	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M935	JX860422	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M940	JX860423	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M946	JX860424	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M947	JX860425	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M949	JX860426	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M950	JX860427	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M951	JX860428	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.04.M952	JX860429	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.05.D215	JX860413	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.06.FTq	JX860414	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.11.SIVsmE660_FL10	JQ864084	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Nef	Wu, F.	*J Virol* **86**(16); 8835-47 (2012)
SMM.US.86.CFU212	JX860407	All	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)

Name	Accession	Proteins	Author	Reference
SMM.US.x.F236_H4	X14307	All	Hirsch, V.M.	*Nature* **339**(6223); 389-92 (1989)
SMM.US.x.H9	M80194	Gag, Pol, Vif, Tat, Rev	Courgnaud, V.	*J Virol* **66**(1):414-419 (1992)
SMM.US.x.PBJA	M31325	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Nef	Dewhurst, S.	*Nature* **345**(6276):636-640 (1990)
SMM.US.x.PGM53	AF077017	All	Novembre, F.J.	*J Virol* **72**(11):8841-8851 (1998)
SMM.US.x.SME543	U72748	All	Hirsch, V.	*J Virol* **71**(2):1608-1620 (1997)
SMM.x.x.pE660.CG7G	JX648292	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Nef	Lopker	Unpublished
STM.US.89.STM_37_16	M83293	All	Novembre, F.J.	*Virology* **186**(2):783-787 (1992)
A.PT.93.JAU1	L28935	Vif	Ribeiro, A.C.	*ARHR* **14**(5):465-469 (1998)
A.SN.x.A2057	U81849	Vpx	Stivahtis, G.L.	*J Virol* **71**(6):4331-4338 (1997)
A.SN.x.A640	U81845	Vpx	Stivahtis, G.L.	*J Virol* **71**(6):4331-4338 (1997)
A.SN.x.A640	U81837	Vpr	Stivahtis, G.L.	*J Virol* **71**(6):4331-4338 (1997)
A.US.93.7924A	U81835	Vpr	Stivahtis, G.L.	*J Virol* **71**(6):4331-4338 (1997)
A.GM.90.CBL24	AJ238995	Env	Vella, C.	*ARHR* **15**(15); 1399-402 (1999)
A.GM.x.CBL21	U05350	Env	Breuer, J.	*J Gen Virol* **1995** Feb;76(Pt 2):333-45
A.GM.x.CBL22	U05351	Env	Breuer, J.	*J Gen Virol* **1995** Feb;76(Pt 2):333-45
A.GM.x.CBL23	U05352	Env	Breuer, J.	*J Gen Virol* **1995** Feb;76(Pt 2):333-45
A.GW.06.CA65316_9	JN863892	Env	de Silva, T.I.	*J Virol* **86**(2); 930-46 (2012)
A.GW.06.CA65319_7	JN863893	Env	de Silva, T.I.	*J Virol* **86**(2); 930-46 (2012)
A.GW.06.CA65330_5	JN863894	Env	de Silva, T.I.	*J Virol* **86**(2); 930-46 (2012)
A.GW.06.CA65409_14	JN863896	Env	de Silva, T.I.	*J Virol* **86**(2); 930-46 (2012)
A.GW.87.CAM2env	U05354	Env	Breuer, J.	*J Gen Virol* **1995** Feb;76(Pt 2):333-45
A.GW.x.CA7205_8	JN863897	Env	de Silva, T.I.	*J Virol* **86**(2); 930-46 (2012)
A.GW.x.CA7253	JN863898	Env	de Silva, T.I.	*J Virol* **86**(2); 930-46 (2012)
A.GW.x.CAM1	U05359	Env	Breuer, J.	*J Gen Virol* **1995** Feb;76(Pt 2):333-45
A.GW.x.CAM3	U05355	Env	Breuer, J.	*J Gen Virol* **1995** Feb;76(Pt 2):333-45
A.GW.x.CAM4	U05356	Env	Breuer, J.	*J Gen Virol* **1995** Feb;76(Pt 2):333-45
A.GW.x.CAM5	U05357	Env	Breuer, J.	*J Gen Virol* **1995** Feb;76(Pt 2):333-45
A.GW.x.CAM6	U05358	Env	Breuer, J.	*J Gen Virol* **1995** Feb;76(Pt 2):333-45
A.PT.00.00PTHCC20_1	GU983949	Env	Rocha, C.	*Retrovirology* **2013** Oct 24;10:110 doi: 101186/1742-4690-10-110
A.PT.03.ABG	JF683340	Env	Dias-Rito	Unpublished
A.PT.92.93PTHDESC_13	JX219596	Env	Rocha, C.	*Retrovirology* **2013** Oct 24;10:110 doi: 101186/1742-4690-10-110
A.PT.98.98PTHDECT_13	GU983928	Env	Rocha, C.	*Retrovirology* **2013** Oct 24;10:110 doi: 101186/1742-4690-10-110
B.x.06.8704A_06_01	JX235884	Env	Kong, R.	*J Virol* **86**(2); 947-60 (2012)
MAC.US.x.750.p4i3	KJ201187	Env	Sandler, N.G.	*Nature* **511**(7511); 601-5 (2014)

Name	Accession	Proteins	Author	Reference
SMM.US.07.FBr_365wpi	KF478030	Env	Schmokel, J.	*Cell Rep* **5**(4):997-1009 (2013)
SMM.US.07.FFr_365wpi	KF478067	Env	Schmokel, J.	*Cell Rep* **5**(4):997-1009 (2013)
SMM.US.08.FTv_32wpi	KF478169	Env	Schmokel, J.	*Cell Rep* **5**(4):997-1009 (2013)
SMM.US.10.FJV_154wpi	KF478191	Env	Schmokel, J.	*Cell Rep* **5**(4):997-1009 (2013)
SMM.US.10.FPY_154wpi	KF478124	Env	Schmokel, J.	*Cell Rep* **5**(4):997-1009 (2013)
A.CI.x.IC763124	U76641	Nef	Switzer, W.M.	*J Infect Dis* **177**(1):65-71 (1998)
A.ES.x.S1084	U76642	Nef	Switzer, W.M.	*J Infect Dis* **177**(1):65-71 (1998)
A.PT.x.1096	AJ344398	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.1139	AJ344392	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.1147	AJ344390	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.1148	AJ344377	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.1215	AJ344393	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.1227	AJ344391	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.1268a	AJ344395	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.1320	AJ344394	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.1378	AJ344414	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.138	AJ344378	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.1395	AJ344397	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.1428	AJ344408	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.1543	AJ344405	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.1544	AJ344407	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.1567	AJ344409	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.268	AJ344410	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.293a	AJ344399	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.293b	AJ344400	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.427d	AJ344415	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.483	AJ344401	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.511	AJ344402	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.546	AJ344403	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.741	AJ344384	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)

Name	Accession	Proteins	Author	Reference
A.PT.x.794	AJ344388	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.956	AJ344369	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.B1_1	AJ344406	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.EP	AJ344387	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.LF3	AJ344383	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.MP2	AJ344386	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
A.PT.x.P1	AJ344381	Nef	Padua, E.	*J Gen Virol* **84**(PT 5):1287-1299 (2003)
B.CI.x.IC762993	U76639	Nef	Switzer, W.M.	*J Infect Dis* **177**(1):65-71 (1998)

```
                              Gag p15 start                                                                                                                    p15 end p27 start
MAC.US.x.239                  MGVRNSVLSGKKADELEKIRLRPNGKKKYMLKHVVWAANELDRFGLAESLLENKEGCQKILSVLAPLVPTGSENLKSLYNTVCVIWCIHAEEKVKHTEEAKQIVQRHLVVETGTTETMPKTSRPTAPSSGRGGNYPVQQ.IGGNYVHLPLSPRTLNAWVKLIEEKKFGAE  169
A.CI.88.UC2                   --A-----R--------------G-----R---I----------R---N---S--------T--D-------------F--------L-------D-----KL------A----A-K--N-------P--K---F----.A------V-------------V--------  169
A.DE.x.BEN                    --A-----R---------V----G-----R---I--------K---------S--------R--D-------------F--------L-------D-----KLA-----A----A-K--N-------P--KR-------.A------V-------------V--------  169
A.DE.x.PEI2_KR_KRCG           --A-S---R---V----------G-----R---I-------GK---------S-------IT--D-------------F--------L-------D--G----------A----ADK--S----A--P-----------.-A---S-V-------------V--------  169
A.GH.x.GH1                    --A-----R--------------S-----R---I--------K---------S--------T--D-------------F--------L-------D-----KL-----GA----A-K--S-------P----R-F----TG----I-V-------------V-D------  170
A.GM.87.D194                  --A-----R---------V--------R-R----------------------S--------K--E-------------F--------L-------D-----KLA-----A----A-K--NI------P--K---F----.A----I-V-------------V--------  169
A.GM.x.ISY_SBL_6669_85        --AK----R--------------G-----R---I------------T-----S-------I---E-----------------T-----L-------D-----R--G----A--E-A-K--NI------P--K---F----.---------------------V--------  169
A.GM.x.MCNI3                  --A-----K--------T-----G-----R---I------------------S-----R--T--G-------------F-----------------D--G---------AA----A-K--N-------P--K---F----.V----T-V-------------V--------  169
A.GM.x.MCR35                  --A-----K--------T-----G-----R---I------------------S-----R--T--D-------------F-----------------D--G---------AA----A-K--N-------P--K---F----.V----T-V-------------V--------  169
A.GW.86.FG_clone_NIHZ         --A-----R--------------G-----K---I------------------S--------T--D-------------F-----------------D--G----------A----A-K--N-------P-..-K-F----.VA---T-I----G--------V--------  167
A.GW.87.CAM2CG                --A-----R---------V----G-----K---I------------------S-----R--K--D-------------F-----------------D-----R-AL---AA----A-K--D-------P--K-------S.-----T-V-------------V--------  169
A.GW.x.MDS                    --A-S---R---------V----G-----R---I------------------S--------T--D-------------F------F----------D-----KFAR---A----A-K--N-------P----------N.-----T-V-------------V--------  169
A.IN.07.NNVA                  --AK----R--------R-----G-----RV---------------T-----S--------T--D-------------F-----------------D-----KL------A----A-K--N-------P-EG---F---R.-----T-V--------V----V--------  169
A.IN.95.CRIK_147              --AKS---R--------R-----G-----R------------K---T-----S--------T--D-------------F-----------------D-----K-------A----A-K--N-------P--G---F---H.-----T-V-------------V--------  169
A.IN.x.NIM_8                  --A-----R--------R-----G-----R----------------------S-----------D---S------R--F-------------------------K-------A----A-K--N-------P--KA--F---S.-----S-V-------------V--------  169
A.JP.08.NMC786_clone_41       --A-----R--------------G-----K---I------------------S-----R--T--N--I----------F--------L-------D-----KL------A----A-K--E-------P--K---F----.V------V-------------V--------  169
A.PT.x.ALI                    --A-----R-R------R-----G-----Q---I------------------S-----R--K--E-------------F-----V--V-------D--G----I----AA-IE-A-K--S-------P-EQ---F----.VA---T-V-------D-----V--------  169
A.SN.85.ROD                   --A-----R--------R-----G-----R---I-----K------------S--------T--D-M-----------F-----------------D--G-----R----A----A-K--S---------EK-------H.V----T-I-------------V--------  169
A.SN.86.ST_JSP4_27            --A-----R--------------G-----R---I------------------S--------T--D-------------F---------------A-D------K------A--K---K--S-------P--N---F----.VA---T-V-------------V--------  169
B.CI.88.UC1                   --A-S-------T-----V----G---R-C---II--V--------------S----H---T----------------F------Y-L-------D-----K-A----AAD-...-K--A-------P-..----------.-A-----M-------------V--------  164
B.CI.x.20_56                  --A-G-------T-----V----G-----C-------V-----Y----T---S-Q--H---T----M-----------F------Y-L-------D-----K-A-S--A-DK...-K--A-------P-..----------.-A-----V-------------V-D------  164
B.CI.x.EHO                    --A-G-------T-----V----G---R-----I---V---E--------R-GS----R--RK--G-------------------F-L-------D-----K-A----AAD-...-K--AM-K-SK-T-..RLA------.-A---S----------------V--------  164
B.GH.86.D205_ALT              --A-G-------T-----V----G-------------V--------------S--------K----------------F-I----F-L-------D-----K-A----AAD-...-K--A-NK----P-..----------.LA--------------------V--------  164
B.JP.01.IMCJ_KR020_1          --A-G-------T-----V----G-----C---I---V-----Y--------S----H---T--E-------------F------Y-L---V---D--D--K-A-S--AAN-...-KA-A--K----P-..-R-------.VA---T----------------V--------  164
G.CI.92.Abt96                 --A-S-------------V----G----------------------------S-----------------------------------L---Q---------V--------------KV-A----I--P-----------.V---------------------V--------  169
AB.CM.03.03CM_510_03          --A-S-------T-----V----G---R------I--V----------T---S--------K-------------F------F-L-------D-----K-A----A.DA...-K-TA-NK----P-..----------.-A-----V-------------V--------  163
H2_01_AB.CI.90.7312A          --A-G-------T-----V----G----------I-------------N---S--------T-----M----------F--------L---K---D-----KLA-S----D-...-KT-TADK-A-T--..-R-S----.VA-----V-------------V----L---  164
H2_01_AB.JP.04.NMC307_20      --A-G-------T-----V----G----------I-----------------T--------TA----M------V---F--------L---K---D-----KLA-S--AAA-...-KA-T..ADKPAATSS---------.VA-----V-------------V--------  164
H2_01_AB.JP.07.NMC716_01      --A-S------ET-----V----G---R------I-----------------S--------K------------V---F--------L---Q---D-----KLA-S--AAA-...-KA-A..ADKPAATSS---------.VA-----V-------------V--------  164
H2_01_AB.JP.08.NMC842_10      --A-G-------T-----V----G-----V---II-----------------T--------TA----M------V---F--------L---K---D-----KLA-S--AAA-...-KT-TPTADKPAATSS---------.VA-----V-------------V--------  166
U.CI.07.07IC_TNP3             --A-H-------T-----V---------R-----------------------S-------------------------------T-----L--------------L-----------ADK--A-------P---R--------.V---------------------V--------  169
U.FR.96.12034                 --A--#------V-----V---------------------------------A-----------------------------------L-------------------A---E-A-K--A-----X-P--GR-------.-----A---------------V--------  168
MAC.US.x.17EC1                ----------------------------------------------------------------------------------------------------------------------------------------------.------------------------------  169
MAC.US.x.251_1A11             --A-------------------------------------------------------------------------------------------------------------------A-I------------------.------------------------------  169
MAC.US.x.251_32H_PJ5          --A--------------------G-------------------------------------------------------------------------------------------A--------------------------.------------------------------  169
MAC.US.x.251_BK28             --A--------------------G-------------------------------------------------------------------------------------------A--------------------------.------------------------------  169
MAC.US.x.MM142_IVMXX          --A--------------------G-------------------------------------------------------------------------------------M----A-----------F-----------.-----T------------------------  169
MNE.US.82.MNE_8               --A--------------------G-------------------------------------------------------------------------------------------A--------------------------.V----T------------------------  169
MNE.US.x.MNE027               --A--------------------G---------------------D-----------A---------------------------------------------------------A-------------KR--------.-----T------------------------  169
SMM.CI.79.SIVsmCI2            --A---------------V----G---------L-------------D----S----H--------------------F-----------------------------------ADK--A-------PG-K---------.V----T------------------------  169
SMM.LR.89.SIVsmLIB1           --A-----------Q---V------R---------------------D----D-----------V---------------------------------------------------AD-V---------P-D-R--------.M---------------*---V--------  168
SMM.SL.92.SIVsmSL92A          --A---------------V----G----------I-----------------S-----R-----V--------------------------M---N----------K--A--NE-ADQL-A-------P--KR--------.V---------------------V--------  169
SMM.SL.92.SL92B               --A-G-------------V----G-R--------II---R------S-----S-----R--A-----M----------FS----V--L---M---D-----KT--S-----S--A-KL-AQ------P-..----------.V-N----T-------------V--------  167
SMM.US.04.G078                --A-S--------------C-------R-------------------D-------------------------------------------------------R--------------ADK--T-------P--K--------.V----T---------------V---R----  169
SMM.US.04.G932                --A---------------V------R-R------------------------T-------I---------------------------L------------------------------A-A---I------P-----------.V----T---------------V--------  169
SMM.US.04.M919                --A-H--------------G--------------------------------T-----R-I------I-----------------------------------------I---K----------ADK--I-------P-----------.V----T---------------V--------  169
SMM.US.04.M922                --A---------------V---------------------------------T-------I-----------D------------------------------------T------A----ADK--A---LV--P--K--------.V----T--S------------V--------  169
SMM.US.04.M923                --A--------------------G-------------------------------------------------------------------------------------------ADK--A-------P-----------.V---------------------V--------  169
SMM.US.04.M926                --A---------------V------R--------------------------S-----------------------------------L---------------------P---------DNL----T----P-----------.V---------------------V--------  169
SMM.US.04.M934                --A------#-----------------------------------G---------R--------------------------------V--------------V------------ADN-----------P-----------.V---------------------V--------  168
SMM.US.04.M935                --A----------------------R--------------------------T-------IA--------------------------------------------------------ADK--A-------P-----------.V----T---------------V--------  169
SMM.US.04.M940                --A-S-------------V------R-----------------------------------------------------------------------------R----K-------AD-----------P--AR--------.V----T---------------V---R----  169
SMM.US.04.M946                --A-S-------------V--------------------------------------R------------------------------L--------------V------------ADN-----------P-----------.V---------------------V--------  169
SMM.US.04.M947                --A--------------------G----------------------------T-----------N-----------------------------------------------------ADK--A--------------------.V----T---------------V--------  169
SMM.US.04.M949                --A-S-------------V------R-------------------------------------------------------------------------------------------AD------------P--AR--------.V---------------------V---R----  169
SMM.US.04.M950                --A---------T-----V----G----------I-----------------T-----R-----------------------------L--------------V------------ADN-----------P-----------.V---------------------V--------  169
SMM.US.04.M951                --A---------------VK---G-R-R------I--------------------------------------------------------------------------------------AD------------P--IR--------.V----T---------------V---R----  169
SMM.US.04.M952                --A---------T-----V---------------------------------T-D---R-----------------------------L------R-------V------------ADN-----------P-----------.V---------------------V--------  169
SMM.US.05.D215                --A-C-------------V----G-----------------------------------R-I-------------------------------------------------------E-AG-----------P--MR--------.V----I-Q-------------V--------  169
SMM.US.06.FTq                 --A-S-------T-----V-----------------------------------------------------------------------------Q------------------AD---A-------P-----------.----------------------V--------  169
SMM.US.11.SIVsmE660_FL10      --A----------------------------------------------D----------------------------------------------------------------------ADK--A-------P-----------.----------------------V--------  169
SMM.US.86.CFU212              --A-S-------------V----G----------------------------X-----------T-------------------------------------------------------ADP---S-------P--MK--------.V---------------------V--------  169
SMM.US.x.F236_H4              --A--------E------V--------------------------------D-------------------------------------------------------------------ADR--A-------P-----------.V---------------------V--------  169
SMM.US.x.H9                   -----------------------G----------I-------------------------------------------R-L----------------------------------------ADK--A-------P-----------.V----T---------------V--------  169
SMM.US.x.PBJA                 --A--------------------G---R-Q---I-----------------------------------------------L----------------------------------------ADK--A-------P--K--------.-----T------------------------  169
SMM.US.x.PGM53                --A-------------------------------------------------T-------I--------A------------------------------------------------------ADK--I-------P-----------.-----T---------------V--------  169
SMM.US.x.SME543               --A------------------------------------------------D-------------------------------------------------------------------ADK--A-------P-----------.V---------------------V--------  169
SMM.x.x.pE660.CG7G            --A-------------------------------I----------------D-------------------------------------------------------------------ADK--A-------P-----------.----------------------V--------  169
STM.US.89.STM_37_16           --A-S-------------V----G----------------------------S-------IT--E-------------F----------------------V-K----------ANK--A-------P-----------.V---------------------V--------  169
```

Draft of 30 September 2015, not for public release

hiv2prot_main.tex r8876–10546


```
MAC.US.x.239                   VVPGFQALSEGCTPYDINQMLNCVGDHQAAMQIIRDIINEEAADWDLQHPQ..PAP.QQGQLREPSGSDIAGTTSSVDEQIQWMYRQQNPIPVGNIYRRWIQLGLQKCVRMYNPTNILDVKQGPKEPFQSYVDRFYKSLRAEQTDAAVKNWMTQTLLIQNANPDCKLVLK 336
A.CI.88.UC2                    --------------------------Q--------E----------A---I..-G-LPA----D-R---------T--------------V-----------I-------------V---------S----------------A-P-----------V------------ 337
A.DE.x.BEN                     -----------------------------------E----------S---I..-G-LPA----D-R---------T----------P---V-----------I-------K--------I-------------------------P------------------------ 337
A.DE.x.PEI2_KR_KRCG            -----------------------------------E----------V---I..-G-LPA------R---------T-E------F-A---------------I------------------------------------------P-----------V------------ 337
A.GH.x.GH1                     -----------------------------------E---D------A---I..-G-LPA----D-R---------T-E--------P---V-----------I------------------------------------------P------------------------ 338
A.GM.87.D194                   -----------------------------------E----------A---I..-G-LPA----D-R---------T-----------P--V-----------I-----------------------S------------------P------------------------ 337
A.GM.x.ISY_SBL_6669_85         -----------------------------------E----------V---I..-G-LPA----D-R---------T-E---E-----E--V-----------I----------------I------S------------------------------V-.S--------- 336
A.GM.x.MCNI3                   -----------------------------------E-V--------V---I..-G-LPA------R---------T--------F-P---V-----------I----------------I-------------------------P-----------V------------ 337
A.GM.x.MCR35                   -----------------------------------E----------V---I..-G-LPA------R---------T--------F-P---V-----------I----------------I-------------------------PT----------V------------ 337
A.GW.86.FG_clone_NIHZ          -----------------------------------E----------VA--I..-G-LPA------R---------T-E------F-P---V-----------I----------------IN------------------------P------------------------ 335
A.GW.87.CAM2CG                 -----------------------------------E----------AN--I..-G-LPA----D-R---------T-E------F-A---V-----------I----------------I------S------------------P-----------V------------ 337
A.GW.x.MDS                     -----------------------------------E----------A---I..-G-LPA------R---------T--------F-A---V-----------I------------------------------------------------------------------- 337
A.IN.07.NNVA                   -----------------------------------E----------V---I..-G-LPA------R---------T-E---E--F-A---V-----------I----------------I-------------------------P--------------------II-- 337
A.IN.95.CRIK_147               ---R-------------------------------E----------A---I..-G-LPA------R---------T-E------F-A---V-----------I----------------I-------------------------P-----------V------------ 337
A.IN.x.NIM_8                   -----------------------------------E----------V---I..-G-LPA------R---------T-E---E--F-A---V-----------I----------------I-------------------------P--------------S--------- 337
A.JP.08.NMC786_clone_41        -----------------------------------E-------E--A---I..-G-LPA----D-R---------T-E--L------P--V-----------I----------------I-------------------------P------------------------ 337
A.PT.x.ALI                     -----------------------------------E----------VA--I..-G-LPA------R---------T-E------F-PR--V-----------I----------------I-------------------------P-----------V------------ 337
A.SN.85.ROD                    -----------------------------------E-------E--V---I..-G-LPA------R---------T-E------F-P---V-----------I----------------I-------------------------P-----------V------------ 337
A.SN.86.ST_JSP4_27             -----------------------------------E----------A---I..-G-LPA------R---------T-E------F-P---V---S-------I----------------I-------------------------P-----------V------------ 337
B.CI.88.UC1                    -----------------------------------E----------Q---I..-G-LPA----D-R---------T-E--------A---V-----------I----------------I-------------------------P------------------------ 332
B.CI.x.20_56                   -----------------------------------E----------Q---I..-G-LPA----D-R---------T-E--------A---V-----------I-----------------------S------------------P------------------------ 332
B.CI.x.EHO                     -------------------------E---------E----------Q---S..-G-MPA------R---------T-E--------P---V----------------------------I-------------------------P------------------------ 332
B.GH.86.D205_ALT               -------------------------E---------E----------Q---S..-G-MPA----D-R---------T-E--------A---V----------------------------I-------------------------P------------------------ 332
B.JP.01.IMCJ_KR020_1           -------------------------E---------E---D------Q---S..-G-MPA----D-R---------T-E--------P--------------------------------I-------------------------P-----V------------------ 332
G.CI.92.Abt96                  ------------L------------E---------E-------X-------..QG-PPA--X---T---------TIE-----TH------------------------------------------------------------P------------------------ 337
AB.CM.03.03CM_510_03           -----------------------------------E---D------Q---S..-G-LPA----D-R---------T-E--------G---------------I----------------I-------------------------P---S-------V------------ 331
H2_01_AB.CI.90.7312A           -------------------------E------V--E----------Q---V..-G-LPA----D-R---------T-E--------G-S------------------------------------------------------A-T--RA---E---V------------ 332
H2_01_AB.JP.04.NMC307_20       I------------------------E------V--E----------Q---V..-G-LPA----D-R---------T-E--------G--S-----------------------------------------------------A-T--RA---E---------------- 332
H2_01_AB.JP.07.NMC716_01       I------------------------E------V--E----------Q---I..-G-LPA----D-R---------T-E--------G--SV------------------------------------------------------Q--RA---E---------------- 332
H2_01_AB.JP.08.NMC842_10       I------------------------E------V--E----------Q---I..-G-LPA----D-R---------T-E--------G--SV----------------------------------------------------A-T--RA---E---V------------ 334
U.CI.07.07IC_TNP3              -------------------------E---------E----------I-N--..-G-LPA------R---------T-E------------------------------------------------S------------------P------------------------ 337
U.FR.96.12034                  -------------------------E---------E---D------T---N..QG-PPA------R---------T-E--------P--------------------------------I-------------------------P------------------------ 336
MAC.US.x.17EC1                 ---------------------------------------------------..---.----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.251_1A11              ---------------------------------------------------..---.----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.251_32H_PJ5           ---------------------------------------------------..---.----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.251_BK28              ---------------------------------------------------..---.----------------------------------------------------------------------------------------------------------------- 336
MAC.US.x.MM142_IVMXX           --S---------L--------------------------------------..Q--.------------------T-E-------------------------------------------------------------------P------------------------ 336
MNE.US.82.MNE_8                -------------------------E---------E---------------..Q--.------------------T---------------------------------------S-----------------------------P------------------------ 336
MNE.US.x.MNE027                ------------I------------E---------E---------------..Q--.------------------T---------------------------------K---------IR------------------------P------------------------ 336
SMM.CI.79.SIVsmCI2             -------------------------E--------KE--------------N..VG-LPA----------------T--------------V------------------------------------------------------------------------------- 337
SMM.LR.89.SIVsmLIB1            -------------------------E---------E--------*------..-S-IPA------R---------T---------------------------R-----------------------------------------P------------------------ 335
SMM.SL.92.SIVsmSL92A           -------------------------E---------E---------------..-G-LPA------R---------T-E--------AA--------------------------V----I-------------------------P------------------------ 337
SMM.SL.92.SL92B                -------------------------E---------E--------------RGQQPAQPA-G--------------TPS---E----A---V---D--------------------------------------------------P------------------------ 337
SMM.US.04.G078                 -------------------------E---------E---------------..-G-LPA------R---------T--------------------------------------M----------------------C-------P------------------------ 337
SMM.US.04.G932                 -------------------------E---------E---------------..-G-IPA----D-R---------T-E-------------------------------------------------------------------P------------------------ 337
SMM.US.04.M919                 -------------------------E---------E---------------..-G-IPA----D-R---------T-E------------------------------------V----I-------------------------P------------------------ 337
SMM.US.04.M922                 -------------------------E-------V-E---------------..-G-LPA------R---------T-E-------------------------------------------------------------------P------------------------ 337
SMM.US.04.M923                 -------------------------E---------E---------------..-G-LPA------R---------T-E------------------------------------------------S------------------P------------------------ 337
SMM.US.04.M926                 -------------------------E-------V-E---------------..-G-LPA------R---------T--------------------------------------V----I-----------------G-------P------------------------ 337
SMM.US.04.M934                 -------------------------E-------V-E---------------..-G-LPA------R---------T-E-----------------------------------------I------T----------G-------P------------------------ 336
SMM.US.04.M935                 -------------------------E-------V-E---------------..-G-IPA----D-R---------T-E-----------------------------------------I-------------------------P------------------------ 337
SMM.US.04.M940                 -------------------------E---------E---------------..QG-LPA------R---------T--------------------------------------V----------------------C-------P------------------------ 337
SMM.US.04.M946                 -------------------------E-------V-E---------------..-G-LPA------R---------T-E-----------------------------------------I------T----------G-------P------------------------ 337
SMM.US.04.M947                 -------------------------E---------E---------------..-G-IPA----D-R---------T---------------------------------------------------------------------P------------------------ 337
SMM.US.04.M949                 -------------------------E-------V-E---------------..-G-LPA------R---------T---------------------------------------------------------------------P------------------------ 337
SMM.US.04.M950                 -------------------------E---------E---------------..-G-LPA------R---------T-E-----------------------------------------I-------------------------P------------------------ 337
SMM.US.04.M951                 -------------------------E---------E---------------..VG-LPA------R---------T--------------V-----------------------M----I-----------------C-------P------------------------ 337
SMM.US.04.M952                 -------------------------E-------V-E---------------..-G-LPA------R---------T-E-----------------------------------------I---------N---------------P------------------------ 337
SMM.US.05.D215                 ----------------V--------E---------E---------------..-G-LPA------T---------T--------------V----------------------------I-----K-------------------P------------------------ 337
SMM.US.06.FTq                  -----------------------------------E---------------P.MG-IPA----D-R---------T-E---E-----------------------------------------------A---------------P--------------------VI-- 338
SMM.US.11.SIVsmE660_FL10       -------------------------E---------E--------------P..-G-PPA------R---------T---------------------------------------------------------------------P------------------------ 337
SMM.US.86.CFU212               -------------------------E---------E---------------..-G-IPA----D-R---------T-E-------------------------------------------------------------------P------------------------ 337
SMM.US.x.F236_H4               -------------------------E---------E---------------..-G-LPA------R---------T---------------------------------------------------------------------P------------------------ 337
SMM.US.x.H9                    -------------------------E---------E---------------..-G-IPA------R---------TX-------------------------------------------------S------------------P-X---------------------- 337
SMM.US.x.PBJA                  -------------------------E---------E---------------..-G-IPP------R---------T---------------------------------------------------------------------P------------------------ 337
SMM.US.x.PGM53                 -------------------------E---------E---------------..-G-IPA----D-R---------T-E-------------------------------------------------------------------P------------------------ 337
SMM.US.x.SME543                -------------------------E---------E---------------..-G-LPA------R---------T-E-------------------------------------------------------------------P------------------------ 337
SMM.x.x.pE660.CG7G             -------------------------E---------E---------------..-G-LPV------R---------T---------------------------------------------------------------------P------------------------ 337
STM.US.89.STM_37_16            -------------------L-----E--------KE----------M----P.-G-LPA-----------------PE------------------------------------V-----------T----------------A-PS------R---------------- 338
```

```
                            p27 end p2 start   p2 end p8 start                                               p8 end p1 start   p1 end p6 start  PTAP motif                        PSAP in HIV-2 B, U
MAC.US.x.239                GLGVNPTLEEMLTACQGVGGPGQKARLMAEALKEALAPVPIPFAAAQQR.....GPRK.PIKCWNCGKEGHSARQCRAPRRQGCWKCGKMDHVMAKCPDRQAGFLGLPWGKKPRNFPMAQVHQGLMPTAPPE..................DPAVDLLKNYMQLGKQQRE 482
A.CI.88.UC2                 ---M--------------------------------T-P---------........-R.T-R-----------K---------------PG-I--N---------M-SR--------V--AP---I-----A..................---A---EK-L-Q-R.... 476
A.DE.x.BEN                  ---M-------------------------------MG-S---------........--.A-RY--------------------------PG-I--N--E---------R--------VT-AP---I-----A..................---AE--ER---Q-R.... 476
A.DE.x.PEI2_KR_KRCG         ---M-------------I--------------------A---------........-R.T--------D--------------------SG----N--E------I-----------VTR-P---T-----A..................---A---EK-L-Q-R.... 476
A.GH.x.GH1                  ---M--------------------------------T-P---------........--.V-R---------------------------TG-------E------M-----------V--APP--I-----A..................-------ER---Q-R.... 477
A.GM.87.D194                ---M-----------------S--------------T-A---------........-R.A-R-----------K---------------SG-I--N--E------M--R--Q-----A--AP---I-----I..................-------EK---Q-R.... 476
A.GM.x.ISY_SBL_6669_85      ---M-------------I-----------------MR-A---------........KR.A-----------------------------SG-I--N---------------------VVPSS---T-----M..................-------EK---Q-R.... 475
A.GM.x.MCN13                ---M--------------------------------T-A---------........KR.T-----------------------------PG--I-N----V----M--R--------V---P---T-----V..................-------E----Q--.... 476
A.GM.x.MCR35                ---M--------------------------------T-A---------........KR.T-----------------------------PG--I-N----V----M--R--------V---P---T-----V..................-------EK---Q--.... 476
A.GW.86.FG_clone_NIHZ       ---M---------T--------------------VM--A---------........--.TF--------------WS------------SG----N---------I-H----A----A---P---T-----L..................-------EK---Q--.... 474
A.GW.87.CAM2CG              ---M-------------------------------MG-P---------........-R.T-----------------------------PG-I-TN---------I-----------V---P---T-----L..................-------EK---Q--.... 476
A.GW.x.MDS                  ---M-------------------------------MT-A---------........-R.T-----------------------------TG-I------------I-----------V---P---T-----A..................-------EK---Q-R.... 476
A.IN.07.NNVA                ---M-----G-------------------------MG-T----V----........--.T-R------------------K--------PG-I--N---------M--R--------V--IP---T-----M..................-------EK---Q-R.... 476
A.IN.95.CRIK_147            ---M----------------------------------P---------........--.T-R---------------------------TG----N--E------IS-R--------V---P---T-----M..................-------EK---Q-R.... 476
A.IN.x.NIM_8                ---M-------------I------------------S-P---------........--.---------D----K---------------TG----N---------------------V---P---T-----M..................-------EK---Q-R.... 476
A.JP.08.NMC786_clone_41     ---MG-------------------------------T-A---------........--.AF----------------------------QG-I-SN--E------IS-R--------V--AP---I-----A..................---M---EK---Q-R.... 476
A.PT.x.ALI                  ---M-------------------------------MT-A---------........-R.T--------------------K--------PG-L--N--E------------------VTR-P---T-----A..................E--A---EQ---Q-R.... 476
A.SN.85.ROD                 ---M------------------------------VIG-A---------........--.AF----------------------------PG-I-TN---------------------V---P---T-----V..................-------EK---Q--.... 476
A.SN.86.ST_JSP4_27          ---I-------------------------------M--A---------........-R.T-----------------------------AG-I-----E------------------V--IP---T-----I..................--VE---EK---Q--.... 476
B.CI.88.UC1                 ---M-------------I------------------T-A---------K....A-.KRGTVT------V--T-K---------------QG-I-S---E-------------------T--P--VT-S---MDPAEGMTPRGATPSAPPA----EM--S--KM-R.... 493
B.CI.x.20_56                ---I--------------------------------T-A---------K....A-.KR.T-T------G--T----K------------PG---SQ----V----F---------------P-RVT-S---MDPTQDMTPQGATPSAPPA----EM---------.... 492
B.CI.x.EHO                  ---M-------------I------------------T-STN------P-....A-.KR.TVT------A--T----K-----------QQG-I-S---E------F-----------V.-AP--IV-S---MNPAFGMTPQGAIPSAPPA---EEM---------.... 491
B.GH.86.D205_ALT            ---M-------------I------------------T-A------V--K....A-.KRGTVT------Q--T-----------------TG-I-S---E-------------------T--P--VT-S---MNPAEGMTPRGATPSAPPA----EM--S---M-R.... 493
B.JP.01.IMCJ_KR020_1        ---M-------------I------------------T-P-L----V---....A-.KRGAVT------I--T----K------------QG-I-S---E------F-----------V--AP--IL-S---MNPAENMTPQGAMPSAPPA----EM--D-----R.... 493
G.CI.92.Abt96               ---M-------------X------------------N-TAL-------K....T-GKRST-----------TV----------------PG-I-----E------F------------T--P---T-S---M..................--------------R.... 481
AB.CM.03.03CM_510_03        ---M-------------I-------K----------T-A------F--K....T-.KRGTVT------G--T----K------------QG-I-S---E--V----------------T--P--VT-S---MNPAEXMTPQGVTPSAPPA----EM--S---M--.... 492
H2_01_AB.CI.90.7312A        ---PH------------I------------------T-A---------K....A-.-RGTVT------G--T-K---------------PG-N--N--E----------------------P-EIV-S---MNTAEGKTHQGAIPSAPPA----EM--S------.... 493
H2_01_AB.JP.04.NMC307_20    ---PH------------I------------------T-A---------K....A-.KRGTLT------G--T-----------------TG-N--N---------------------A---P-EVV-S---MDIAESKTHQRVTPSAPPA----EM--------R.... 493
H2_01_AB.JP.07.NMC716_01    ---PH------------I------------------T-A---------K....A-.KRGTLT------G--TX----------------TG-N--N----X----F-----------A---P-EVV-S---MDTAESKTYQRVTPSAPPA----EM----T---R.... 493
H2_01_AB.JP.08.NMC842_10    ---PH------------I------------------T-A---------K....A-.KRGTLT------V--T-K---------------TG-N--N---------------------A---P-EVV-S---MDTAESKIHRRMTPSAPPA----EM--------R.... 495
U.CI.07.07IC_TNP3           ---M----------------------------------G-L----V---....T-GK-SM-----------------------------PG-I-----E------F------------T-IP---T-----A..................------------M-R.... 481
U.FR.96.12034               ---M-------------I------------------Q-M------V---....G---R.-VR---------T----K----------E-PG-N--S--ETKV-------E-----------P---T-----AEPAVDLL....TPTAPPA--------S---Q--.... 493
MAC.US.x.17EC1              -------------------------------------------------.....----.-------------------------------------------------------------------------..................------------------- 482
MAC.US.x.251_1A11           -------------------------------------------------.....----.------------------------------------------------------------------T------..................------------------- 482
MAC.US.x.251_32H_PJ5        -------------------------------------------------.....----.------------------------------------------------------------------T------..................------------------- 482
MAC.US.x.251_BK28           -----------------------------------------------K-.....----.------------------------------------------------------------------T------..................------------------- 478
MAC.US.x.MM142_IVMXX        ---T----------------------------------A---------K.....----.---------------------------------------N--------------------------T------..................E----------H---.... 478
MNE.US.82.MNE_8             ---M----------------------------------G-L-------K.....----.------------------T----------Q-G--------------F--------------M----T------..................---------------.... 478
MNE.US.x.MNE027             ---M---------------------K------------G-L-------K.....----.-----------------------------Q-G--------------F--------------M----T------..................---------------.... 478
SMM.CI.79.SIVsmCI2          ---M-------------I--------------R--FQ-G-L--------...GG-V-R.-----------------K------------SG-------E------F------------T-LQ---T-----MDPA.........VAAPPM-----M-R--LK--Q.... 490
SMM.LR.89.SIVsmLIB1         ---M-------------------------------FQ-ASL----V---....G-TKR.TVR---R----------K------------TG-I-----E------F--------------M----T-----A..................---------------.... 478
SMM.SL.92.SIVsmSL92A        ---M--------------------------------G-TQL-------K......-S-..VR--------------K-------------G-I-----E------F--------------MP---T-----A..................--------D--K--R.... 478
SMM.SL.92.SL92B             ---M---------------------------M-D--TG...SLV---F-GAAKGQGN-PI-R-F----T-------------K-----EEGRIQ-N--NQK-----------------.-T.TS-T-S----..................---ARIV-E-LEKAQ.... 479
SMM.US.04.G078              ---M----------------------------------G-L--------....G----.TV----------T----K------------TG--------------F---------------P-R-V------..................--------------R.... 480
SMM.US.04.G932              ---M-------------I--------------------G-L----V--K.....-Q--.------------------------------TG-------E------F--------------M----T-----A..................--------------R.... 479
SMM.US.04.M919              ---M--------------------------------T-GQL----V--K.....----.T-----------------------------PG------TE---------------------MP---T------..................-----------KM--.... 479
SMM.US.04.M922              ---M--S-----------------------------T-GQL----V---......QKR.T-----------------------------TG-------E---------------------MP---T------..................--------S--KI-R.... 479
SMM.US.04.M923              ---M-----------------------------D--TQG-L----V--......---R.T-------------K--K------------TG-------E---------------------MP---T------..................------------M-R.... 479
SMM.US.04.M926              ---M----------------------------R---G-ASL-------K......-K-.T-Q----------------------------G-------E------F--------------M----T-----A..................-----------KM-Q.... 479
SMM.US.04.M934              ---M----------------------------------A-L----V--K.....-Q--.T-T-----------K---------------TG-------E------F--------------M----T-----A..................--------S--KM-Q.... 478
SMM.US.04.M935              ---M-------------I------------------T-GQF----V---......QK-.T-----------------------------TG-------E---------------------IP---T------..................--------S--KI-R.... 479
SMM.US.04.M940              S--M---------------------------I--V---G-L----V---....G---R.TV-------D--T----K------------TG--------------F-----------------R-I------..................--------------R.... 480
SMM.US.04.M946              ---M----------------------------------A-L----V--K.....-Q--.T-T---------------------------TG-------E------F--------------M----T-----A..................--------S-LKM-H.... 479
SMM.US.04.M947              ---M--------------------------------T-GQL----V---.....-Q-R.T-------------K---------------AG-------E---------------------MP---T------..................-----------KM-R.... 479
SMM.US.04.M949              ---M--------S-----------------------TTG-L--------....G---R.TV----------T----K------------TG-------E------F---------------P-R--------..................--------------R.... 480
SMM.US.04.M950              ---M--S----------I------------------T-A-L----V--K.....-Q--.T-T---------------------------TG-------E------F--------------M----T-----A..................--------S--KM-Q.... 479
SMM.US.04.M951              ---M----------------------------------G-L--------....G----.T-----------T----K--------R--QTG-I-----E------F---------------P-R-T------..................-----------R--R.... 480
SMM.US.04.M952              ---M--------------------------------V-ASL----V--K.....-Q--.T-T---------T-----------------TG-------E------F--------------M----T-----A..................-----------KM-R.... 479
SMM.US.05.D215              ---M-----------------------------D--Q-GHL-------K....G----.M------------------------------G-T--N---------F-------------------T-----A..................--------------R.... 480
SMM.US.06.FTq               ---M-------MS--------------------S----GVL-------K....G-L-H.TV------------K----------------G-I-----E--V---F------------T-MP---T------DLAR..............---E----R---Q-R.... 485
SMM.US.11.SIVsmE660_FL10    ---M-------------I------------------R-DQL----V--K.....-Q-R.T-------------------------G---TG-------E---------------------MP---I------..................-----------KM-R.... 479
SMM.US.86.CFU212            ---M--------------------------------T-G-L------L-....G--KR.TV-------D--T----K-------------G------TE------F-------------------T-----A..................--------E--KQ-R.... 480
SMM.US.x.F236_H4            ---M--------------------------------R-DQL----V--K.....-Q--.T-------------K---------------TG-------E---------------------MP---I------..................-----------KM-R.... 479
SMM.US.x.H9                 ---I--------------------X--------D--TQG-L----V--K.....-QX-.I-X-X--X--------F-------------AG-------E---------------------MP---T------..................-----------KV-R.... 479
SMM.US.x.PBJA               ---I-----------------------------D--TQG-L----V--K.....-Q--.I-----------------------------AG-------E---------------------MP---T------..................-----------KM-R.... 479
SMM.US.x.PGM53              ---M--------------------------------T-GQL----V--K.....----.T-------------K---------------TG-------E---------------------MP---T------..................-------R---KM-R.... 479
SMM.US.x.SME543             ---M-------------I------------------R-DQL----V--K.....-Q-R.T-------------------------G---TG-------E------F--------------MP---T------..................-----------KM-R.... 479
SMM.x.x.pE660.CG7G          ---M-------------I------------------R-DQL----V--K.....-Q-R.T-------------------------G---TG-------E---------------------MP---I------..................-----------KM-R.... 479
STM.US.89.STM_37_16         ---M------------------------------VFQ-D-L-------Q.....-.-R.TV-------A--T-K--KG----------QQG-Q-----E--V---F-------H------IP---T------..................---A---RS------.... 479
```

p6 end
Gag end

```
MAC.US.x.239                KQRESREKPYKE..................VTEDLLHLNSLFGGDQ*  510
A.CI.88.UC2                 ----Q--R----VTEDLLHLEQGETPRRE.A-----------K---  522
A.DE.x.BEN                  ----Q--R----VTEDLLHLEQRETPHRE.E-----------K---  522
A.DE.x.PEI2_KR_KRCG         --K-QKMR----VTEDLLHLEQGETPHKE.A-----------K---  522
A.GH.x.GH1                  E---Q--R----VTEDLLHLEQGKAPHRE.A-----------K---  523
A.GM.87.D194                ----Q--R----VTEDLLHLEQGETPHRG.A-----------K---  522
A.GM.x.ISY_SBL_6669_85      ----Q-QR----VTEDLLHLEQGETPHRE.T-----------N---  521
A.GM.x.MCN13                R---Q--R----VTEDLLHLEQGEAPCRE.T----------*K---  521
A.GM.x.MCR35                R---Q--R----VTEDLLHLEQGEAPYRE.T----------*K---  521
A.GW.86.FG_clone_NIHZ       R---Q--R----VTEDLLRFEQAETPCRE.T-----------K---  520
A.GW.87.CAM2CG              R---Q-QR----VTEDLLHLEQGEAPCRE.T-----------K---  522
A.GW.x.MDS                  ----Q--R----VTEDLLHLEQGKTPCKE.T-----------T---  522
A.IN.07.NNVA                ----Q--R----VTEDLLHLEQGEAPCGE.T-----------K-L-  522
A.IN.95.CRIK_147            ----Q--R----VTEDLLHLEQGGAPCGE.M-----------K---  522
A.IN.x.NIM_8                --K-Q-QR----VTEDLLHLEQGGTPCGE.A-----------K-L-  522
A.JP.08.NMC786_clone_41     R---Q--R----VTEDLLHLEQGETPHRE.A-K----------E---  522
A.PT.x.ALI                  ----Q--R----VTEDLLHLEQGETPHKE.------------K---  522
A.SN.85.ROD                 R---Q--R----VTEDLLHLEQGETPYREPP-----------K---  523
A.SN.86.ST_JSP4_27          R---Q--R----VTEDFLQLEKQETPCRE.T-----------K---  522
B.CI.88.UC1                 Q------R----..................------------E---  522
B.CI.x.20_56                Q------R----..................------------E---  521
B.CI.x.EHO                  --K-N--R----..................------------E---  520
B.GH.86.D205_ALT            Q------R----..................------------E---  522
B.JP.01.IMCJ_KR020_1        --KGG------..................------------E---  522
G.CI.92.Abt96               --K-Q-N-----..................---X----S----D---  510
AB.CM.03.03CM_510_03        Q------R----..................------------E---  521
H2_01_AB.CI.90.7312A        Q---KQGR----..................------------E---  522
H2_01_AB.JP.04.NMC307_20    Q---KQGR----..................------------E---  522
H2_01_AB.JP.07.NMC716_01    Q---KQGR----..................------------E---  522
H2_01_AB.JP.08.NMC842_10    Q---KQGR----..................------------E---  524
U.CI.07.07IC_TNP3           --K----R----..................------------D---  510
U.FR.96.12034               --K-N--R----..................------------N---  522
MAC.US.x.17EC1              ------------..................----------------  511
MAC.US.x.251_1A11           ------------..................M---------------  511
MAC.US.x.251_32H_PJ5        ------------..................----------------  511
MAC.US.x.251_BK28           Q-----------..................----------------  507
MAC.US.x.MM142_IVMXX        Q-----G-----..................----------------  507
MNE.US.82.MNE_8             Q----KR-----..................------------E---  507
MNE.US.x.MNE027             Q----KR-----..................-A----------E---  507
SMM.CI.79.SIVsmCI2          Q---RK-R----..................------------D---  519
SMM.LR.89.SIVsmLIB1         ----N--R----..................------------D---  507
SMM.SL.92.SIVsmSL92A        --K----R----..................------------E---  507
SMM.SL.92.SL92B             REKTR-SR----..................------------E---  508
SMM.US.04.G078              ----N--R----..................--------S---E---  509
SMM.US.04.G932              Q-----KT----..................------------E---  508
SMM.US.04.M919              ----N--R----..................------------E---  508
SMM.US.04.M922              ----N--R----..................------------E---  508
SMM.US.04.M923              R---N--R----..................------------D---  508
SMM.US.04.M926              Q---N-R-----..................------------E---  508
SMM.US.04.M934              Q-K-N-R-----..................------------D---  507
SMM.US.04.M935              ----N--R----..................------------E---  508
SMM.US.04.M940              -------R----..................------------E---  509
SMM.US.04.M946              Q-K---RR----..................------------D---  508
SMM.US.04.M947              R---N--R----..................------------E---  508
SMM.US.04.M949              -------R----..................------------D---  509
SMM.US.04.M950              Q-----R-----..................------------D---  508
SMM.US.04.M951              ----N--R----..................------------E---  509
SMM.US.04.M952              E---N-RR----..................------------D---  508
SMM.US.05.D215              ----N--R----..................------------E---  509
SMM.US.06.FTq               -------R----..................-R------S----E---  514
SMM.US.11.SIVsmE660_FL10    ----N--R----..................------------E---  508
SMM.US.86.CFU212            ----N--R----..................------------D---  509
SMM.US.x.F236_H4            ----N--R----..................------------E---  508
SMM.US.x.H9                 R---N--R----..................------------E---  508
SMM.US.x.PBJA               R---N--R----..................------------E---  508
SMM.US.x.PGM53              ----N--R----..................----------------  508
SMM.US.x.SME543             ----N--R----..................------------E---  508
SMM.x.x.pE660.CG7G          ----N--R----..................------------E---  508
STM.US.89.STM_37_16         ------KT----..................------------E---  508
```

```
                              Pol p15 start                                                                                   p15 end protease start
MAC.US.x.239                  FFRPWSMGKEAPQFPHGSSASGADANCSP.................RGPSCGSAKELHAVGQAAERKAERKQREALQG.................GDRGFAAPQFSLWRRPVVTAHIEGQPVEVLLDTGADDSIVTGIELGPHYTPKIVGGIGGFINTKEYKNVEIEV  134
A.CI.88.UC2                   ---DGLT------L-R-P-S----T-ST-.................SRSGS-PVR-IF....--GE---GAEG-TI--GDGRLTAPRAGRDTSQR.----L--------K------Y--D---------------A-----DN----V-----------------K-  147
A.DE.x.BEN                    ---VGPT----S-L-RDP-P----T-STS.................GRS-S-TVG-IY....--RE---GAEG-TI-RGDGGLAAPRAERDTSQR.----L--------K------Y--D---------------A-----DN----------------------K-  147
A.DE.x.PEI2_KR_KRCG           ---D-P-----S-L-RDP-PA---T-ST-.................SR--SRP-R-VL....--REE---AEN-TI--GDRGLTAPRTRRDTTQR.-------------K------YV-----------------A-----SN-S--------------------K-  147
A.GH.x.GH1                    ---DG--------L-R-P-S----T-ST-.................SRS-S--IGKIY....--GER--GAEG-TI-RGDGRLTAPRAGKSTSQR.----L--------K------Y--V---------------A--Q--DN-V--------------I--I--K-  147
A.GM.87.D194                  ---DGPT--A---L-R-P-S----T-ST-.................NRS-S-PVG-IY....--RE----AEG-TI--GDGGLTAPRAGRDAPQR.----L-T------K------F--D---------------A-----DN----------------------K-  147
A.GM.x.ISY_SBL_6669_85        ---A-T-------L-R-PKFA--NT-ST-.................N-S-S-PTG-V-....--RE-T--AETKTI-RSDRGLAASRARRDTTQR.D---L--------K------Y--D---------------A-----SN-S-----------------D---R-  147
A.GM.x.MCN13                  ---DGPP------L-R-P-PA--NT-ST-.................SRS-S-PTG--Y....--RK-TK-TE--TI-RSDRGLTAPRAGRGTMQR.D---L-------LK------YV-----------------A-----SN-S-------------------K---  147
A.GM.x.MCR35                  ---DGPP------L-R-P-PA--NT-ST-.................SRS-S-PTG-IY....--RK-TK-TE--TI-RSDRGLTAPRAGRSTIQR.D---L-------LK------YV-----------------A-----SN-S-------------------K---  147
A.GW.86.FG_clone_NIHZ         ---D-PL---G--L-R-P-PA--NT-ST-.................I-S-S-PTG-IY....--RK--KGAE--TV--SDRGLTAFRAGRDTMQG.D---L--------K-----------------R-N---A-----SN-S------------------------  147
A.GW.87.CAM2CG                ---D-PL--------R-P-ST--NT-ST-.................I-S-S--TG-IY....--RE---GAET-TI-RGDRGLTAPRTRRGPMQG.DN--L--------K---------------------------A-----SN-S------------------------  147
A.GW.x.MDS                    ---D-PL------L-R-PGSA--NT-ST-.................SRS-S-PTG-IY....--RE---GAE--TI-RGDRGLAAPRAGKDTMQG.DN-----------N---------------------------A-----SN-S------------------------  147
A.IN.07.NNVA                  ---DGPP------L-R-PNPA--NT-ST-.................N-S-S-PIG-IY....--RE---GAE--TI--GDRGPTAPRAGRSTLRG.N--RL--------K------Y-------------------A-----SD-N------------------------  147
A.IN.95.CRIK_147              ---DKPP----S-L-R-P-ST--NT-STS.................N-S-S-PTG-IY....--RE---GAE--TI--GDRGPTAPRAGRGTLRG.N---L--------K------Y-------------------A-----SN-S---------------N------  147
A.JP.08.NMC786_clone_41       ---DK-T----S-L-R-P-S--T-T-ST-.................SRS-N-PVG-VY....--RE---GAEG-TI-RGDRGLTAPRARRDTSQR.--K-L---------------Y--D---------------A-----NN--------------------------  147
A.PT.x.ALI                    ---A-P-----S-L-RNP-SA-INT-ST-.................SRA-S-P-GAVY....--GE--K-AE---I-RGDGGLTAPRAGRDTTQR.----L--------K------Y-------------------A-----SN----------------ED---K-  147
A.SN.85.ROD                   ---TGPL------L-R-P-SA---T-ST-.................S-S-S--TG-IY....--RE-T--AE--TI--SDRGLTAPRAGGDTIQGATN--L--------K------Y-------------------A-----NN-S------------------------  148
A.SN.86.ST_JSP4_27            ---VGP---------C-PNPA---T-ST-.................DR--R-PTR-V-....--RE----AE---I-RSDRGLPAARETRDTMQR.D---L--------K-------V-------------------A-V---SN-S----------------------R-  147
B.CI.88.UC1                   ---VRTL----S-L--DP----S-TI-T-DGPSRGHDTSGGDTICAPCRS-S-D-EK--.....EDGET---EP--T---..................-----------------R-C------------------A-----SN-----------------D-----  148
B.CI.x.20_56                  ---IRTL----S-L---P---EG-TI-T-DGPNTGHDTSGCDSICTPCRS-S-D-EK--.....--REA---EP--T---..................----L------------K-S------------------A-----SN-----------------D-----  148
B.CI.x.EHO                    ---VRPL----S---R.PGTP-DS-I-A-DEPSIRHDTSGCDSICTPCRS-R-D-------.....-TREE--GE---T---..................-----------------K-T----S------------A-----SN--------------N--------  147
B.GH.86.D205_ALT              ---VRTL----S-L--DP----S-TI-T-DEPSRGHDTSGGDTICAPCRS-S-D-EK--.....-DGETT--EP--T---..................-----------------K-C----S-------V----A-----SN-----------------D-----  148
B.JP.01.IMCJ_KR020_1          ---I-PL----S-L-CSP-TP-DT-I-TTNEPSREHDTSGCNAICAPCRS-S-DVEG--.....--RE----G-G-T---..................-----------------K-Y----S------------A-----SD-------------------------  148
G.CI.92.Abt96                 ---V-TL----S-L--DP-P--S-SISTT.................D---SRPTE-------.....-XGEE-KGAEKQT---..................---XX----L---------Y--S---------------A-----SN-S--------------------K-  130
AB.CM.03.03CM_510_03          ---VRTL----S-L--DPG---S-TI-T-DEPSRXHDTSGGDTICAPCRS-S-D-EK--.....-DGETT--EP--T---..................----L------------R-C------------------A-----SN-----------------D-----  148
H2_01_AB.CI.90.7312A          ---A-TL----S-L---P----DS-I-A-DEHSRGQDTSGSDTICAPCRSGS-D-EK--.....-TREA---EAG-----..................----L------------Q-Y------------------A-----SN------------------I----  148
H2_01_AB.JP.04.NMC307_20      ---A-TL----S-L-R-P----GG-I-A-DGHSREQDTSEGDTICAPCRS-S-D--K--.....-TGEA---EA--T---..................--G-L------------Q-Y------------------A-----SN-----------------DI----  148
H2_01_AB.JP.07.NMC716_01      ---V-TL----S-L-R-P----GG-I-T-DGHSREQDISEGNTICAPCRS-S-D--K--.....-TGEA---EA--T---..................--G-L------------Q-Y------------------A-----CN-----------------DI----  148
H2_01_AB.JP.08.NMC842_10      ---A-TL----S-L-R-P----GG-I-A-DGHSREQDTSEDDTICAPCRS-S-D--K--.....-TGEA---EA--T---..................--G-L------------Q--------------------A-----SN-----------------DIK---  148
U.CI.07.07IC_TNP3             ---V-P-----S-L--DPDP--V-T-SA-.................SR--S---E-------.....-DGQETKGE-GKTI--..................--G------------------------------------A-----VN-H---------------T-----  130
U.FR.96.12034                 ---A-PL--K-S-L---P-PA-T-P-SPS....SRTSSGSADPDSPSS---S---------.....-TGQETKGE---TI--..................--G---------K--I---Y------K---N--------A--K-E-Y----T--E-E------KFH--N-K-  144
MAC.US.x.17EC1                -----------------------------.................--------------------------------..................---------------------------------------------------------------------  134
MAC.US.x.251_1A11             -----------------------------.................--------------------T-----------..................D--------------------------------------------------------------------  134
MAC.US.x.251_32H_PJ5          -----------------------------.................-------------A------------------..................------------------------------------------------------L---------------------  134
MAC.US.x.251_BK28             -----------------------------.................--------------------------------..................-------------------------------------------------------------------K---  130
MAC.US.x.MM142_IVMXX          -----PL----------------------.................-RT----------L------------------..................---------------------------------------------------------------------  130
MNE.US.82.MNE_8               ---L-P-----------PN----------.................---------K------T-------.--K-----..................--------------------------A-------------A-------------------------K---  130
MNE.US.x.MNE027               ---L-P-----------PN-----T----.................---------K------T-------.--K-----..................-G--------------------------A-------------A-----------------------------  130
SMM.CI.79.SIVsmCI2            ---I-P-----S-L--DPTPA---T-GA-.........NGSSCGGAPN-S---HVE--SETRS-T--....EE-K---R..................-NGRL---------------Y-------------------A-----LG--------------R--T--N---  139
SMM.LR.89.SIVsmLIB1           ---L--L----SK----PN-----T----.................SR------E----I--E--G....E--KT---..................--------------------------------------A------N-----------------E----T-  130
SMM.SL.92.SIVsmSL92A          ---I--L----S-L---PN---P-TDSP-.................S---S---EG--E--KET-----.E---TI--..................--G-L---------------------------------A-----LN-K---------------QK-K-Q-  130
SMM.SL.92.SL92B               -----P-----S-L--ADNI#.#-TVST-.................RS-SPDSQGVSRESTEGKD....-EEQ----..................--G---------R-----R-L-----L------------A------G---R--------------RT-K---  126
SMM.US.04.G078                ---V-P-----S-----P-T-E-G--G--.................--SGS--------AREE--G....E---T---..................--G-----------------Y-------------------A------Q-----------------QD---K-  130
SMM.US.04.G932                ---L-P-------L---PN----------.................S---S--------AR--T-G....E-EDT---..................----L---------------Y-----A-------------A------D-----------Y-----------K-  130
SMM.US.04.M919                ---A---------L---PD----------.................-RS----TE---ED--K--G....E---T---..................--G-----------------Y--E----------------A------N------------------D-K-K-  130
SMM.US.04.M922                ---A-P-----------PD-----T----.................--S----T----EDR-ET-G....E---T---..................--G-----------------Y--E----------------A------N-----------------------K-  130
SMM.US.04.M923                ---A-PV----------PD-----T----.................--S----TE----D--K--G....E---T---..................--G-L---------------Y--E----------------A------N------------------D-K-K-  130
SMM.US.04.M926                ---V-P-----------PN-------S--.................S---GR--E---ED-S-T--....E-E-T---..................---RL---------------Y-----A-------------A------N----V------------------K-  130
SMM.US.04.M934                ---V-PL----------PN-------S--.................SR--G---E---ED-S-T--....E-E-----..................-N------------------Y-------------------A------N-----------------------K-  130
SMM.US.04.M935                ---A-P-----------PDT----T----.................--S----T----EDR-ET-G....E---T---..................--G-----------------Y--E----------------A------N-----------------------K-  130
SMM.US.04.M940                ---V-P---K-S-----P-T-E--T----.................--SGS--------A-EE----....E---T---..................--------------------Y-------------------A------Q-----------------QD---K-  130
SMM.US.04.M946                ---V-PL----------PN------YS--.................SR--G---E--LED-SST--....E-E-----..................----L---------------Y-------------------A------N-----------------------K-  130
SMM.US.04.M947                ---A-P-----------PD-----T----.................--S----TEK--ED--K--G....E---T---..................--G-----------------Y--E----------------A------N------------------D-K-K-  130
SMM.US.04.M949                ---V-P-----S-----P-T-E-------.................--S-S---E----A-KETKG....E---T---..................--G-----------------YV------------------A------N-----------------QD---K-  130
SMM.US.04.M950                ---I-PL----------PN-------S--.................SR--G---E---ED-S-T--....E-E-----..................--------------------Y-------------------A------D-----------------------K-  130
SMM.US.04.M951                ---V-P-----S-----PGT-E-------.................--SGS----------EE--G....E---T---..................--G-----------------Y-------------------A------Q---------------E--QD---K-  130
SMM.US.04.M952                ---V-PL----------PN-------S--.................S---G---E---ED-SGT--....E-E-----..................----L---------------Y-------------------A------N-----------------------K-  130
SMM.US.05.D215                ---I-PL----------P-T-----D-PS.................G-S-GRP-E----AR-E--G....E---T---..................----------------I---Y-------------------A------N-----------------TE---K-  130
SMM.US.06.FTq                 ---L-------S-L--DPD----Y--GPS.............RRPSK-SGG--LE-V--A--E--G....E---TI--..................-EG-----------------------------------A-----LN----V---------------A---  134
SMM.US.11.SIVsmE660_FL10      ---A-P-------L---PD----NT----.................--S----TE---ED--K--G....E---T---..................--------------------Y--E----------------A------N----V--------------D-K-K-  130
SMM.US.86.CFU212              ---X-P-----------P-T------ST-.................SRS--RXV-GI-ET--ET--....E---T---..................--G-----------------QV------------------A-V---TN--------------------------  130
SMM.US.x.F236_H4              ---A-P-----------PD-----T----.................--S----TE---ED--K--G....E---T---..................--------------------Y--E----------------A------N------------------D-K-K-  130
SMM.US.x.H9                   ---A-P-----------PD-----T---X.................--S----TE---EG--K--G....E---T---..................-NG-----------------Y--E--I-------------A------N------------------D-K-K-  130
SMM.US.x.PBJA                 ---A-P-----------PD-----T----.................--S----TE---ED--K--G....E---T---..................-NG-------------I---Y--E----------------A------N------------------D-K-K-  130
SMM.US.x.PGM53                ---A-P-----------PD----------.................--S----TE---ED--K--G....E---T---..................--G---------------A-Y--E----------------A------N------------------D-K-K-  130
SMM.US.x.SME543               ---V-P-----------PD-----T----.................--S----TE---ED--K--G....E---T---..................--------------------Y--E-----------------------N------------------D-K-K-  130
SMM.x.x.pE660.CG7G            ---A-P-------L---PD-----T----.................--S----TE---ED--K--G....E---T---..................--------------------Y--E----------------A------N----V--------------D-K-K-  130
STM.US.89.STM_37_16           ---L-P-------L---PNT-------PS.................-R------EK---A--E---....E-EDT---..................---------------------------------------A-----LQ----V------------F---N---  130
```

```
                                            protease end p51 RT start                                                                                                                          D catalytic site
MAC.US.x.239              LGKRIKGTIMTGDTPINIFGRNLLTALGMSLNFPIAKVEPVKVALKPGKDGPKLKQWPLSKEKIVALREIC.EKMEKDGQLEEAPPTNPYNTPTFAIKKKDKNKWRMLIDFRELNRVTQDFTEVQLGIPHPAGLAKRKRITVLDIGDAYFSIPLDEEFRQYTAFTLPSVN 303
A.CI.88.UC2               -N--VRA--------------I---------L-V-RI--I-IT--------R------T---VE--K---.-----E--------------------------------------------I-----------K-------V---------Y-D--P---------- 316
A.DE.x.BEN                -N--VRA--------------I---------L-V--I--I--T--------R------T----E--K---.-----E-----------------------------------K-------I-----------K---SI--V---------H-D----------A-- 316
A.DE.x.PEI2_KR_KRCG       -N-KV-A--------------I---------L-V---D-I--I---------VR----T----E--K---.----RE-----------------------------------K---E---I-----------KR----------------H-D----------T-- 316
A.GH.x.GH1                -N--VRA--------------I---------L----I--I--T--------R-R----T----E------.-----E--------------------------------------------I-----------K-------V---------H-D-------------- 316
A.GM.87.D194              -N--VRA--------------I-AT------L-V--LD-I--T--------R------T----E--K---.----RE--------------------------------------------I-----------K-------V---------H-D-------------- 316
A.GM.x.ISY_SBL_6669_85    -N-KVRA--------------I---------L-V--I-----T---------QR----TR---E------.----RE-----------------------------------K--------------------KR------V---------Y-D-------------- 316
A.GM.x.MCN13              ---KVRA--------------I---------L-V--I--I-IM----------R----T----E--K---.----RE-----------------------------------K-------I-----------K------------------H-D-----------I- 316
A.GM.x.MCR35              ---KVRA--------------I---------L-V--I--I-IM----------R----T----E--K---.----RE-----------------------------------K-------I-----------K------------------H-D-----------I- 316
A.GW.86.FG_clone_NIHZ     ----V-A--------------V---------L-V--I--I-IM--------R------T----E--K---.-----E-----------------------------------K-------I-----------KR------V---------H-D-------------- 316
A.GW.87.CAM2CG            ----VRA--------------I---------L-V--I--I-IM--------R-R----T----E--K---.-----E-------------------R---------------K-------I-----------KR------V---------H-D-------------- 316
A.GW.x.MDS                ----VRA--------------I---------L-V--I--I-IM----------R----T----E--K---.----RE-------------------R---------------K-------I-----------KR------V---------H-D-------------- 316
A.IN.07.NNVA              ---KVRA--------------I---------L-V--I--I--K----------R----T----E--K---.----RE-------------------R---------------K---E---I-----------KR------V---------H-D-------------- 316
A.IN.95.CRIK_147          ---KVRA--------------I--V------L-V--I--I-IM----------R----T----E--K---.-----E-----------------------------------K---E---I-----------KR------V---------H-D-------------- 316
A.JP.08.NMC786_clone_41   -N--VRA--------------I-A-------L----I--I--T--------R------T----E--K---.-----E--------------------------------------------I-----------K-------V---------Y-D-----------T- 316
A.PT.x.ALI                -N--V-A--------------I---------L-V--I--IE-R----------R----T----E--K---.--T-RE-----------------------------------K-------I-----------KR------V---------H-S-------------- 316
A.SN.85.ROD               -N-KVRA--------------I---------L-V-----I-IM----------R----T----E--K---.-----E-----------------------------------K-------I-----------KR------V---------H-D--P----------- 317
A.SN.86.ST_JSP4_27        -N--VRA--------------I---------L-V--I--I-IM----------R----T----E--K---.----RE-----------------------------------K-------I-----------K-------V---------H-D-----------I- 316
B.CI.88.UC1               V---VRA--------------I-NT---T----V--I-----K---------IR---------L--K---.-----E---------------------R-------------K-------------------EKR-----V----------PN-----------I- 317
B.CI.x.20_56              ----VRA--------------I-NT---T----V--I-----K---------IR---------L--K---.-----E-----------------------------------K-------------------EKR-----V-----V---PD----------P-- 317
B.CI.x.EHO                V---VRA-V------------I-NS---T----V-RI-----Q---E-----IR---------L--K---.-----E-----------S-----------------------K---E---------------SK------V-----V---PD----------A-- 316
B.GH.86.D205_ALT          V---VRA--------------I-NT---T----V--------E---------IR----R----L--K---.-----E-----------------------------------K-------#-------##-EKR---I-V----------PN-------------- 314
B.JP.01.IMCJ_KR020_1      V--K-RA--------------I-NT---T----V--I-----K---------IR---------L--K---#K----E---------------------R-------------K-------------------E------------------PD-------------- 317
G.CI.92.Abt96             V----QA-V------------I-VK----------------R----M----IR---------Q------.D---QE-----------------T---------------------------------------EKR------------------VD-------------- 299
AB.CM.03.03CM_510_03      V--TVRA--------------I-NT---T----V-XX-----E---------IR----R---L--K---.-----E-----------------------------------K-------------------E-R----I-V----------PX-----------T- 317
H2_01_AB.CI.90.7312A      V--KVRS--------------I-NT---T----V--I-----Q---------IR---------L------.-----E------S----------------------------K---E---------------EKR----I-V-----V---PD-----------I- 317
H2_01_AB.JP.04.NMC307_20  V---VRA--------------I-S---------M--L-----Q---------IR---------L------.-----E-----------------VR---R------------K-------------------EKR--S---V-----V---PD-----------I- 317
H2_01_AB.JP.07.NMC716_01  V---VRA--------------I-S---------M--L-----Q---------IR---------L------.-----E------------------R----------------K-------------------E-R--S---V-----V---PD-----------I- 317
H2_01_AB.JP.08.NMC842_10  V---VRA--------------I-S---------M--L-----Q---------IR---------L------.-----E------------------R----------------K-------------------E-R--S---V-----V---PD-----------I- 317
U.CI.07.07IC_TNP3         ---KVRA--------------I-VK--------V----T---T----M----IR---------E--K---.----QE---------------------R----------------------------------EKR------V-----------Q--P---------- 299
U.FR.96.12034             ---T-R-------------------E---------G-EAL-IGG----------R--------LE--K---.-----E--------------------------------------------S-------------EK-------V---------C-D-----------I- 313
MAC.US.x.17EC1            ----------------------------------------------------------------------.--------------------------------------------------------------------------------------------------- 303
MAC.US.x.251_1A11         -----R-------------------------L----------------V---------------------.--------------------------------------------------------------------------------------------------- 303
MAC.US.x.251_32H_PJ5      -------------------------------L--------------------------------------.--------------------------------------------------------------------------------------------------- 303
MAC.US.x.251_BK28         -------------------------------L---------T------V---------------------.--------------------------------------------------------------------------------------------------- 299
MAC.US.x.MM142_IVMXX      -------------------------------L---------SP---------------------------.--------------------------------------------------------------------------------------------------- 299
MNE.US.82.MNE_8           ---------------------------E--------------T----------R----------------.------------------------------------I-------K------------------------------V------------------------ 299
MNE.US.x.MNE027           ------------------------------------------T----------R----------------.--------------------------------------------K------------------------------V------------------------ 299
SMM.CI.79.SIVsmCI2        ----V----------------I---------C-L--------T-----E--------------E------.-----E-----------------------------------S----K-----N-------------EKR-----------------W---------A----- 308
SMM.LR.89.SIVsmLIB1       -----R-------------------E-------V--------K----------R---------E--K---.-----E--------------------------------------K-------I--------------R------V------------------------ 299
SMM.SL.92.SIVsmSL92A      ---EVRA-L------------I-----I---Y-V---Q----T---------QR---------E--K---.-----E--------------------------------------K-------I----------QQ-R------V----------PD-------------- 299
SMM.SL.92.SL92B           ---V-RE-L------------I-A---V---Y-V----YT--K--E-M---R-----------Q--T---.-----E----R----------------N-D------------KM--E---------------KEME--------------V---P---------I---- 295
SMM.US.04.G078            ---K---------------------M-----L-V--------K-----E--------------E--K---.-----E----------------------S-----------------I--------------R------V------------------------ 299
SMM.US.04.G932            ---K--------------------T-----------------T----------R---------I------.-----E--------------------------------------K------------------------V---------H-------------- 299
SMM.US.04.M919            ---V---------------------M-----L-V--------T----------R----------------.--------------------------------------------K--------------------R------V------------------------ 299
SMM.US.04.M922            ---V---------------------M-----L-V--------T----------R----------------.--------------------------------------------K--------------------R------V------------------------ 299
SMM.US.04.M923            ---V---------------------M-----L-------I--T----------R---------I------.--------------------------------------------K--------------------R------V------------------------ 299
SMM.US.04.M926            ---K--------------------TM----------------T----------R---------I------.G----E--------------------------------------K--------------------R------V----------D------------- 299
SMM.US.04.M934            ------------------------TM-----C-V--------T----------R---------I------.-----E--------------------------------------K--------------------R------V------------------------ 299
SMM.US.04.M935            ---V---------------------M-----L-------I--T----------R----------------.--------------------------------------------K--------------------R------V------------------------ 299
SMM.US.04.M940            ---K-R-------------------------L-V-R------K-----E--------------E--K---.D----E------------------------------------------I--------------R------V------------------------ 299
SMM.US.04.M946            ---K-R------------------TM-----C----------T----------R---------I------.-----E--------------------------------------K--------------------R------V----------D------------- 299
SMM.US.04.M947            ---V--------------------A-M------------I--T----------R---------I------.-----E--------------------------------------K--------------------R------V------------------------ 299
SMM.US.04.M949            ---K---------------------------L-V--------K-----E--------------E------.D----E--------------------------------------K-------I------------R------V----------D------------- 299
SMM.US.04.M950            ---K--------------------TM-----C----------T----------R---------I------.-----E--------------------------------------K--------------------R------V--------E--------------- 299
SMM.US.04.M951            -------------------------------L-V-R------R-----E--------------E------.D----E--------------------------------------K-------I------------R------------------------------- 299
SMM.US.04.M952            ---K-------N------------TM-----C----------T----------R---------I------.-----E------------------R-------------------K--------------------R------V----------N------------- 299
SMM.US.05.D215            ---K--------------------TM-------V--------T----------R---------E--K---.--------------------------------------------K--------------------R------V----------N------------- 299
SMM.US.06.FTq             ---T-R--V---------------R--------V-R------T----Q-----R---------E--K---.-----E-----------------------------------------------------------R------V---------P-------------- 303
SMM.US.11.SIVsmE660_FL10  ---V---------------------M-----L-------I--T-----E---R---------I------.--------------------------------------------K--------------------R------V------------------------ 299
SMM.US.86.CFU212          -----R--L----------------------M-Q--------T---------VR---------Q--K---.-Q------------------------------------------K---------------------------V----------D------------- 299
SMM.US.x.F236_H4          ---V---------------------M-----L-------I--T-----E---R---------I------.--------------------------------------------K--------------------R------V------------------------ 299
SMM.US.x.H9               ---V---------------------M-----L-------I--T----------R---------I------.---------X-------------XX-G--------------K--------------X-----R------V------------------------ 299
SMM.US.x.PBJA             ---V---------------------M-----L-------I--T----------R---------I------.--------------------------------------------K--------------------R------V------------------------ 299
SMM.US.x.PGM53            ---V---------------------M-----L-V-----I--T----------R----------------.-----------------------------------------------------------------R------V------------------------ 299
SMM.US.x.SME543           ---V---------------------M-------------I--T-----E---R---------I------.--------------------------------------------K--------------------R------V------------------------ 299
SMM.x.x.pE660.CG7G        ---V---------------------M-----L-------I--T-----E---R---------I------.--------------------------------------------K--------------------R------V------------------------ 299
STM.US.89.STM_37_16       ---K---------------------------V--------T---------I----------E--K---.--------------------------------------------K-------I------------R------V----------G------------- 299
```

DD catalytic site

```
                                                                                                                                              p51 RT end  p15 RNase H start
MAC.US.x.239                KIILSQEQEGCYYQEGKPLEATVIKSQDNQWSYKIHQEDKI.LKVGKFAKIKNTHTNGVRLLAHVIQKIGKEAIVIWGQVPKFHLPVEKDVWEQWWTDYWQVTWIPEWDFISTPPLVRLVFNLVKDPIEGEETYYTDGSCNKQSKEGKAGYITDRGKDKVKVLEQTTNQQ  642
A.CI.88.UC2                 R---D-----H----E-E----IQ-------T------E--.-----Y--------------Q-V-------L----RI--------RET-----DN-----------V--------T----G---P-T--F-------R---------V----R---R---------  655
A.DE.x.BEN                  ----------Y----E-E----IQ---GH--T------E--.-----Y--------------Q-V-------L----RI--------RET-----DN-----------V--------T----G---P-A--F-------R---------V-------------------  655
A.DE.x.PEI2_KR_KRCG         ---------------E-E-----Q-D-----T-----GE--.-----Y----------------V-------L----RI--------RET-----DN--------D---V--------A--------P----F-------R--------------R---RI---------  655
A.GH.x.GH1                  ----------Y----E-E----IQ-N-----T------E--.-----Y--------------Q-V-------L----RI--------RET-----DN-----------V--------T----G---P-A--F-------R-------R-V----R---R---R------  655
A.GM.87.D194                ----------S----EEE-------------A-----GERV.-----Y--------------Q-V-------L----R---------R-T-----DN------V----V--------T----G---P-T--F-------R---------V----R-R-R-----S---  655
A.GM.x.ISY_SBL_6669_85      ----------H----E-E-----Q-D-----T--V--GE--.-----Y--------------Q-V-------L----RI--------RET-----DN-------D---V--------A--------P-A--F-------R------------------RI---------  655
A.GM.x.MCN13                ----------H----E-E-----Q-D-----T--V--GE--.-----Y--------------Q-V-------LI---RI--------RET-----D--------D---V--------A--------PRT--F-------R---------V----R---RM--------  655
A.GM.x.MCR35                ----------H----E-G-----Q-D-----T--V--GE--.-----Y--------------Q-V-------L----RI--------RET-----D--------D---V--------A--------PRT--F-------R---------V----R---RM--------  655
A.GW.86.FG_clone_NIHZ       R-----K--H----E-K-----Q-D-----T--V--GE--#S----Y---------#-----Q-V-------L----RI--------R-T-----DN-------D---V--------A----GE-VP-A--F-------R--------------R-R-----------  654
A.GW.87.CAM2CG              R--------H----E-E-----Q-D-----T------E--.-----Y----H------K---Q-V-------L--*-RI--------RE------DN-------D---V--------A----G---P-T--F-------R---------V----R----I--------  654
A.GW.x.MDS                  RV-------H----E-E-----Q-D-----T--V--GE-T.-----Y--V------------Q-V-------L----RI--------REI-----D--------D---V-----------------P-A--F-------R---------V--K-R-------------  655
A.IN.07.NNVA                R--------H--K-E-E-----Q-D-----T--V--GE--.-----Y--V--A---------Q-V-------L----RI--------RET-----D--------D---V--------A--------P-A--I-------------------V----R-R-----------  655
A.IN.95.CRIK_147            R--------H----E-E-----Q-D-----T--V--GE--.-----Y--V--A---------Q-V-------L----RI--------RET-----D--------D---V--------A--------P-T--F-------R---------V----R--------------  655
A.JP.08.NMC786_clone_41     ---------------E-E-----Q-------T------E--.-----Y--------------Q-V-------L----RI--------RET-----DN-----------V--------T----G---PDA--F-------R---------V--K-R-R-----------  655
A.PT.x.ALI                  R--------H----E-E-----Q-D-----T-----GE--.---E-Y--M------------Q-V-------L----RI-R------RET-----D--------D---V--------A--------L-A--F----P--R--------------R--------------  655
A.SN.85.ROD                 R--------H----E-E-----Q-D-E---T------E--.-----Y--V-------I----Q-V-------L----RI--------REI-----DN-------D---V--------A----G---P-A--F-------R---------V---------K--------  656
A.SN.86.ST_JSP4_27          ---------------E-E-----Q-D-----T-----GG--.-----Y--V-----------Q-V-------L----RI--------R-T-----DN-------D-----------------------L-A--F----------R-----------R---RL--------  655
B.CI.88.UC1                 ----E-----S--K--V------Q-NLA---T-----GNR-.-----Y--V---------------V-------L----EI-V------RET-D------------------V--------AY------L-KV----------RA--------V-------------------  656
B.CI.x.20_56                ----E-----S--K--V------Q-NLA---T-----GNRT.-----Y--V---------------V-------L----EI-M------RET-D---------------------------AY------L-KV----------RA--------V------R-----------  656
B.CI.x.EHO                  ----E-----S--K--V------Q-NLA---T-----G---.-----Y--V---------------V-------L----EI-M------RET-D------------------V-----I--AY------L--V-----------A--------V--------P---------  655
B.GH.86.D205_ALT            ----E-----S--K-RV------Q-NLA---T-----GN-V.-----Y--V---------------V-------L----EI-V------RET-D------------------V-----I--AY------L--R----------RT--------V-------------------  653
B.JP.01.IMCJ_KR020_1        ----E-----S--K--I------Q-NLA---T-----G-R-.-R---Y--V-----------------------L-T--EI-M------RET-D------------------V-----I--AY------L--V----------RT--------V--------A---------  656
G.CI.92.Abt96               ----X-----A--K-XE------Q-NL----T--V--GNR-.-----Y------------------V-------L----KL-F------R-T---------------D---V--H-----Y----E-L-QT-----------T--------V-------------------  638
AB.CM.03.03CM_510_03        ----E-----S--K--V------Q-NLA---T-----GNRM.-----Y--V---------X-----V-------L----EI-I------RET-D------------------V-----I--AY---------K-----------A--------V-------------------  656
H2_01_AB.CI.90.7312A        ----G-----S--K--V------Q-NLA---T-----GN-V.-----Y--V---------I-----V-------L-----I-M------RET-D------------------V--------AY------L-KV----------RT--------V----R-----P---------  656
H2_01_AB.JP.04.NMC307_20    ----E-----S--K--V------Q-NLA---T-----G--V.-----Y--V---------I-------------L-----I-M------RET-D------------------V--------AY------L-KV--F-------RV--------V----R-----P---------  656
H2_01_AB.JP.07.NMC716_01    ----E-----S--K--V------Q-NLA---T-----G--V.-----Y--V---------I-------------L-----I-M------RET-D------------------V--------AY------L-KV--F-------RV--------V----R-----P---------  656
H2_01_AB.JP.08.NMC842_10    ----E-----S--K--V------Q-NLA---T-----G--V.-----Y--V---------I-------------L-----I-M------RET-D------------------V--------AY------L-KV--F-------RV--------V----R-----P---------  656
U.CI.07.07IC_TNP3           ----E-----T--R--E------Q-NL----T-----GN--.-----------------------------------K--I------RET----------------------V--------AY----E-L-RV----------RN--------V-------------------  638
U.FR.96.12034               ----------A--R-EEN-----L-N-----T-----G-RT.---------V--------------V-----------KI-M------RE---------------VLD---V--------AY----E-LL.-----------RA--Q------V------R--P---------  651
MAC.US.x.17EC1              -----------------------------------------.-------------------------------------------------------------------------------------------------------------------------------  642
MAC.US.x.251_1A11           -------------------------N---------------.------------------------------------------------------------N----------------------------------------------------------------  642
MAC.US.x.251_32H_PJ5        -----------------------------------------.-----------------------------------------------------------------------------------K------------------------------------------  642
MAC.US.x.251_BK28           -----------------------------------------.-----------------------------------------------R-------------------------------------------------------------------------------  638
MAC.US.x.MM142_IVMXX        ---------------S-------------------------.---------------------------------------------------------------------------------------------V----S---------------------------  638
MNE.US.82.MNE_8             -------------------------N---------------.----------------------------------------------------------K------------------------K------V------------------------------------  638
MNE.US.x.MNE027             -------------------------N---------------.-------------------------------------------------------------------------------------------V--------------------------R---------  638
SMM.CI.79.SIVsmCI2          R---E-----AF-R-EEN-----Q-N-----T-----GSRV.--------------------V-------L----KI-F------RE----------------D---V---------------------P----F-------RV---------V----RE----P-------R  647
SMM.LR.89.SIVsmLIB1         ----------------D-------V-N--S-----------.-R------V-----------V-------L--*--I--------REI*---------------*-------------P--E-L--V--F-V-R---RE--I-------N--R-----A-------K  634
SMM.SL.92.SIVsmSL92A        ----------S-----E-----A--N-E---------GE--.-----Y--------------Q-----------K----------REL-----------V-D---V-----I-MA---------P-------V---------A-----V-------------------  638
SMM.SL.92.SL92B             R---N-----R--R-DE------L-N-----------G-R-.----------------I----N-V------SL----KT-F------RE--D--------A-------------I-----------K--V--I-----RN---------V------E--LP---A----  634
SMM.US.04.G078              ----------------D------IV------------GE-V.-----Y------------------V------------------M---RET-------------------V-----------------A--F-V---------------------RN-I-S---------  638
SMM.US.04.G932              ----------------------M-N----------------.-------E----------------V-----------------------REI-------------------V----------------------V-----R---V----------R---------------  638
SMM.US.04.M919              -----------------------------------------.------------------------V-----------------------REI-------------------V-------------E--Q-A--F-V-----R---------V----R-----P---------  638
SMM.US.04.M922              -----------------------------------------.------------------------V-----------------------REI-------------------V-------------E--Q-A--F-V-----R---------V----R-----P---------  638
SMM.US.04.M923              -----------------------------------------.------------------------V--------------R------REI-------------------V-------------E--Q-A--F-V-----R--R------V----R--A-L---------  638
SMM.US.04.M926              -------------------------N-------------V.---------------------------------------I--------R---------------------V-----------------------V-----R---V---------------I---------  638
SMM.US.04.M934              -------------------------N---------------.--------------------------------------I--------R---------------------V-----------------------V-----R---V---------------I---------  638
SMM.US.04.M935              -----------------------------------------.------------------------V-----------------------REI-------------------V-------------E--Q-A--F-V-----R---------V----R-----P---------  638
SMM.US.04.M940              ----------------D------IM------------GERV.-----Y------------------V------------------M---RET-------------------V-------------------RA--F-V--------------------RS-I-S---------  638
SMM.US.04.M946              -------------------------N---------------.--------------------------------------I--------R---------------------V-----------------------V-----R---V---------------I---------  638
SMM.US.04.M947              -----------------------------------------.------------------------V-----------------------REI-------------------V-------------E--Q-A--F-V-----R--R------V----R--AEL---------  638
SMM.US.04.M949              ----------------D------IV------------GE--.-----Y--V---------------V------------------M---REI-------------------V-------------------RA--F-V--------------------RS-IES---------  638
SMM.US.04.M950              -------------------------N---------------.----R-----------------------------------------I--------R---------------------V-----------------------V-----R---V---------------I---------  638
SMM.US.04.M951              ---------------AD------IV------------GE-V.-----Y--V---------------V--V---------------M---REI-------------------V-----------------A--F-V--------------------RS-I-S---------  638
SMM.US.04.M952              -------------------------N---------------.--------------------------------------I--------R---------------------V---------------K------V-----R---V-----------R--I---------  638
SMM.US.05.D215              R---------------D-Q-----V-N-----------G-T.------------------------------------------------RE-----------A-------V---------------------F-V-----R---------------R-R--T---------  638
SMM.US.06.FTq               R---N-----------D------IV-------------N--.-----Y------------------V----------K-----M---REI-----------------------------------A--F-V-----R-T-V---------------L---------  642
SMM.US.11.SIVsmE660_FL10    ------------------I--------------------V.---------V---------------V-------L----E---------REI---------------D---V-------------E--Q-A--F-V-----R--R------V----R--A-L---------  638
SMM.US.86.CFU212            ----------------------V-N--------------V.-------------------------V-----------RI--------REI---------------D---V--------------L-----F-------R---------V--G---R-----------  638
SMM.US.x.F236_H4            ------------------I--------------------V.---------V---------------V-------L----E---------REI---------------D---V-------------E--Q-A--F-V-----R--R------V----R--A-L---------  638
SMM.US.x.H9                 ---P---P------------P--------------------.-----------------S-----V--------I------R------XEI-------------------V-------------E--Q-A--F-V-R---R--R------V----RG-T-L---------  638
SMM.US.x.PBJA               -----------------------------------------.------------------------V--------------R------REI-------------------V-------------E--Q-A--F-V-----R--R------V----R--A-L---------  638
SMM.US.x.PGM53              -----------------------------------------.------------------------VE----------------------REI-------------------V-------------E--Q-A--F-V-----R--R------V----R--A-P---------  638
SMM.US.x.SME543             ------------------I--------------------V.---------V---------------V-------L----E---------REI---------------D---V-------------E--Q-A--F-V-----R--R------V----R--A-L---------  638
SMM.x.x.pE660.CG7G          ------------------I--------------------V.---------V---------------V-------L----E---------REI---------------D---V-------------E--Q-A--F-V-----R--R------V----R--A-L---------  638
STM.US.89.STM_37_16         ----------Q--R-D-------V-D---------------.--------------------------------------I--------RE------A-------------V--------------E-L--T--F-V-----R---------------N---A---------  638
```

p15 RNase H end p31 Integrase start

```
MAC.US.x.239          AELEAFLMALTDSGPKANIIVDSQYVMGIITGCPTESESRLVNQIIEEMIKKSEIYVAWVPAHKGIGGNQEIDHLVSQGIRQVLF.LEKIEPAQEEHDKYHSNVKELVFKFGLPRIVARQIVDTCDKCHQKGEAIHGQANSDLGTWQMDCTHLEGKIIIVAVHVASGFIE  811
A.CI.88.UC2           ------A---A-----V-----------VASQ-------I------D----EAV----I-----------V-------------.-----------E-----I---IH---I-QL------N--AQ-Q---------V-AEI-V-------------------------  824
A.DE.x.BEN            ----V-R---A-----V-----------VA-Q-----N-I----------EAV----------V-------------.-----------E----II---TH---I-LL------NS-AQ-Q---------V-AEI-V----Y-------------------------  824
A.DE.x.PEI2_KR_KRCG   ------A---------------------VA-Q------K-----------ETL----------V-------------.-----------E---------SH-----KL------N--AQ-Q---------VDAE------------------------------  824
A.GH.x.GH1            ------A-T-------V-----------VV-Q-------I------D----EAV----------V-------------.--R--------E-----M---TH---I-QL------N--AQ-Q---------V-AEI-V-------------------------  824
A.GM.87.D194          ------A---A-----V-----------VA-Q-----N-I------D----EAV----------V-------------.-----------E-----I---TH---I-QL------N--AQ-Q---------V-AEI-V-------------------------  824
A.GM.x.ISY_SBL_6669_85 ------A--V------V--V--------V--Q-A-----I--K--------EA----------V-------------.--R--------G---------AH-----NL------N--AQ-Q---------V-AE--------------------------------  824
A.GM.x.MCN13          ------A---------------------VA-Q------I----------EA----------V-Q-------------.-----------E---------SH-----KL------N--AQ-Q---------VDAE------------------------------  824
A.GM.x.MCR35          ------A---------------------VA-Q------I----------EA----------V-Q-------------.-----------E---------SH-----KL------N--AQ-Q---------VDAE------------------------------  824
A.GW.86.FG_clone_NIHZ ------A---------------------VA-Q-----N-I----------EA----------V-------------.-----------E-----I---SH---I-KL------N--AHVQ--------V-AE----------V-------------  823
A.GW.87.CAM2CG        ------A---------------------VA-Q-----N-I----------EA----------V-------------.-----------E---T-----CH--DI-QL------N--AQYQ--------V-AEV-----------------------------  823
A.GW.x.MDS            ------A---------------------VV-Q-----N-I----------EA----------V-------------.-----------E---------SH-----QL------N--AQ-Q---------V-A-------------V-------------  824
A.IN.07.NNVA          ------A---------V-----------VA-Q-----N-----------EA----------V-------------.--N-----N----SH-----KL------N--AQ-Q---------V-AE--------------V-------------  824
A.IN.95.CRIK_147      S-----A---------------------VA-Q-----NKI----------EA----------V-------------.-----------E---------SH--R--KL------N--AQ-Q---------V-AE--------------V-------------  824
A.JP.08.NMC786_clone_41 ------A---A---T-V-----------VASQ-------I------D----EAV----------V---------I--.I-----------E-----I---TH---I-QL------N--PQ-Q---------V-AEI-V-------------------------  824
A.PT.x.ALI            ------AL-V------------------VA-Q-----N-I----------EA----------V-------------.-----------E---------SH-----NL------N--AQ-Q---------V-AE----------V--I-------------  824
A.SN.85.ROD           ------A---------V-----------SASQ-------KI---------EA----------V-------------.-----------E---------SH---I-NL------NS-AQ-Q---------V-AE------------------------------  825
A.SN.86.ST_JSP4_27    ------A--V------------------VA-Q------KI---------EA----------V-------------.-----------E---------SH-----KL------N--TQ-Q---------V-AE------------------------------  824
B.CI.88.UC1           ------AL--Q----QV-----------VA-Q--T--P-----------EA---G----R-L----V-------------.-----------E--G------H-----QL--K---NS----Q----V---V-AE--------------V-------------  825
B.CI.x.20_56          ------AL--Q----QV-----------A-Q--TD-PI--K--------EA---G------L----V-------------.-----------E--G------H---I-QL--K---NS----Q---------V-AE--------------V-------------  825
B.CI.x.EHO            ------AL--Q----QV-----------VAAQ--T--PI-RE-------EK---G------L----V---------I-.-----------E--N------H---I-QL------NS----Q---------V--E--------------V-------------  824
B.GH.86.D205_ALT      ------AL-----E-QV-----------AAQ--T--PI-AK-------EAV--G------L----V-------------.-----------E--G------H---I-QL--K---NS----Q---------V-A-------------------------  822
B.JP.01.IMCJ_KR020_1  ------AL--Q----QV-----------AAQ--T--PI-KT-V-----EA---G------L----V---------HI---.N-----------E--G------H---I-QL--K---NS----Q---------V-AE------------------------------  825
G.CI.92.Abt96         ------A---Q-----V-----------A-Q--T--P---K----X---EXL--G-----------V-X-----------.-X---------E-F-------TH---I-QL--K---NSYHS-Q---------V-AE----------V----------V-  807
AB.CM.03.03CM_510_03  ------AL--Q----QV-----------AAQ--T--PI-TX-------EA---G------L----V-------------.-----------E--G---X--H---X-QL--K---NS----Q---------V-AE------------------------------  825
H2_01_AB.CI.90.7312A  ------AL--Q----QV-----------VA-Q--T--PI--K--------EA---G------L----V-------------.-----------E--G------H---I-QL--K---NS----Q-----V---V-AE----------V-------------  825
H2_01_AB.JP.04.NMC307_20 ------AL--Q----QV-----------A-Q--T--SI--R--------EAV--G------L----V-------------.-----------E--G------H-----QL--K---NS----Q---------V---V-AE------------------------  825
H2_01_AB.JP.07.NMC716_01 ------AL--Q----QV-----------A-Q--T--SI--R--------EAV--G------L----V-------------.-----------G------H---I-Q---K---NS----Q-----V---V-AE------------------------------  825
H2_01_AB.JP.08.NMC842_10 ------AL--Q----RV-----------A-Q--T--SV--R--------EAV--G------L----V-------------.-----------E--G------H-----HL--K---NS----Q---------V---V-AE------------------------  825
U.CI.07.07IC_TNP3     ----------Q---SQ---V----------A-Q--T--P---K------M-EAL-XG----------L-V-------------.-----------F-----------I--L--K-------------------V-TE----------V-------------  807
U.FR.96.12034         ------A---E-----V--VT----I---A-Q--T--P--SK----L-R-EAV--G-----------V-------------.-----------E---------QH---I-QL--K---N--HV-Q-----M---V-TEV-----------VV-----------  820
MAC.US.x.17EC1        ------------------------------------------------------------------------------------.-------------------------------------------------------------------------------------  811
MAC.US.x.251_1A11     ----------------------------------------------K---------------V-------------.-----------------------------------------V-------------------------  811
MAC.US.x.251_32H_PJ5  ------------------------------------------------------------------------------------.-----------------------------------------V-------------------------  811
MAC.US.x.251_BK28     ---------------T----------------------------------------------------------------.-----------------------------------------V----------------V----------  807
MAC.US.x.MM142_IVMXX  ------------------------------------------T---------------------------.-----------S-----I----------L--K-----------------V----------------V----------  807
MNE.US.82.MNE_8       ----------A-----------------------------------------T---------------------------.-----------------------L--K-----------------V-------------------------  807
MNE.US.x.MNE027       ------------------------------------------T---------------------------.-----------------------L--K-----------------V-------------------------  807
SMM.CI.79.SIVsmCI2    -----IAP--R-A-----VVT--------VA-H---TD-P---E----L---EA---G-------L----V-------------.--Q--------E-------Q--H---I-QL--K---N--TQ-Q---------V-AE----------V-------------  816
SMM.LR.89.SIVsmLIB1   --------MQ----E-----------VA-Q-----NN-----------TA---*--------R-V-----R----S.#-----------E-----I-----Y-I---L--K-----------------V-AE-----I-------------------------  801
SMM.SL.92.SIVsmSL92A  ------A---A--------VT-----L--VL-Q-----PI--E--NH-ML-E-L----------L----V-------------.--N----------I----Y---I-#L--K---N-----Q---------V-A-I-------------------------  806
SMM.SL.92.SL92B       ---Q-L-L--K--PS-V-VVT-----LN----Q-S--D-DI-A----QLVQ-EAV-IG-------L----N-V-R-------.--S-----D----------AQ-YNI-QL--K---NA-N--Q---------T-AEV----------V-------------  803
SMM.US.04.G078        ---------V---TE----------V--Q-----KI----------TA----------L-----------------------.-----------E----------I--L--K---------Q---------V--E--V----------V-------------  807
SMM.US.04.G932        ----------I---------V-----------------------------I----------------------------.-----------E----------L--K---------Q---------V-----------------------------V-  807
SMM.US.04.M919        ------CL--A-----------------A-Q----------H--------EA----------V-------------.-----------E----------I--L--K---------Q---------V-AE------------------------------  807
SMM.US.04.M922        ------YL--A-------V----------A-Q-----N----------EA----------V---------S---.-----------E----------I--L--K--I---------------V-AE------------------------------  807
SMM.US.04.M923        ------YL--A-----------------VA-Q----------------EA----------V-------------.-----------E----------L--K-----------------V-AE------------------------------  807
SMM.US.04.M926        ----------I-----------------V---------K----------E-V----------------------.-----------ER---------Y-I--L--K-----------------V----------V----------  807
SMM.US.04.M934        ----------Q-----------------V-----------------------EG----------------------.-----------ER---------Y-I--L--K-----------------V----------V----------  807
SMM.US.04.M935        ------YL--A-------V----------A-Q-----------------------EA----------V-------------.-----------E----------I--L--K--I---------------V-AE------------------------------  807
SMM.US.04.M940        ---------V----TE----------V--Q-----KI----------TAV--T------L-----------------------.-----------E----------I--L--K---------Q---------V-TE--V----------VV-----------  807
SMM.US.04.M946        ----------Q-----------------V--------N----------EG----------------------.-----------ER---------Y-I--L--K-----------------V----------V--A---------  807
SMM.US.04.M947        ------YL--A-----------------A-Q----------H--------EA----------V-------------.-----------E----------L--K-----------------V-AE------------------------------  807
SMM.US.04.M949        ---------V----AE----------V--Q-----KI----------TAV----------L-----------------.-----------E----------I--L--K-----N--Q---------V-TE--V----------V-----------  807
SMM.US.04.M950        ---------Q-------V----------V-----------S--------G----------------------------.-----------ER---------Y-I--L--K--I---------------V----------V----------  807
SMM.US.04.M951        ---------V----TE---V--------V--Q-----KI----------TA----------L-----------------.-----------E----------I--L--K---------Q---------V-TE--V----------V---------E---  807
SMM.US.04.M952        ---------Q-----------------V----------S---------E----------------------------.-----------ER---------Y-I--L--K---------Q---------V----------V----------  807
SMM.US.05.D215        ---------VK---------V--------VA-Q-----NSI----------TA----------V-------------.---------D-E----------Y-I--L--K-----------------V-AE------------------------------  807
SMM.US.06.FTq         ---------V-----E-----------VA-Q-------V----------TA----------------------------.-----------E----------RI--L--K--I---------------V-AE----------V-------------  811
SMM.US.11.SIVsmE660_FL10 ------YL--A--R-----------A-Q----------------EA----------V-------------.-----------E----------L--K-----------------V-AE------------------------------  807
SMM.US.86.CFU212      ----------------------------VA-Q-----NKI-S--------TA---T------L----V-------------.-----------E----------Y-I--L--K-----------------V-A----------V-------------  807
SMM.US.x.F236_H4      ------YL--A-----------------A-Q----------------EA----------V-------------#*K-----------E----------L--K-----------------V-AE------------------------------  806
SMM.US.x.H9           ------YL--A-----------------VA-Q-------------K----EA---------X--E---V-----E----------.-----------E------K-----X---L--K---------L-------V-AX----------------G-----  807
SMM.US.x.PBJA         ------YL--A-----------------VA-Q----------------EA----------V-------------.-----------E----------L--K-----------------V-AE------------------------------  807
SMM.US.x.PGM53        ------YL--A-----------------A-Q----------------EA----------V-------------.-----------E----------I--L--K-----H-----------V-AE------------------------------  807
SMM.US.x.SME543       ------YL--A-----------------VA-Q----------------EA----------V-------------.-----------E----------I--L--K-------R---------V-AE------------------------------  807
SMM.x.x.pE660.CG7G    ------YL--A-----------------A-Q----------------EA----------V-------------.-----------E----------S---L--K-----------------V-AE------------------------------  807
STM.US.89.STM_37_16   ------A---A--------V-----------Q-----K----------EA----------V-------------.-----------E----------I--L--K-----------------V-AE------------------------------  807
```

Draft of 30 September 2015, not for public release
hiv2prot_main.tex r8876–10546


```
MAC.US.x.239          AEVIPQETGRQTALFLLKLAGRWPITHLHTDNGANFASQEVKMVAWWAGIEHTFGVPYNPQSQGVVEAMNHHLKNQIDRIREQANSVETIVLMAVHCMNFKRRGGIGDMTPAERLINMITTEQEIQFQQSKNSKFKNFRVYYREGRDQLWKGPGELLWKGEGAVILKVGT  981
A.CI.88.UC2           -------S----------S----------P--T---------V--QS-------------------S------T-------------------------------L-R--N--K-Q-------------------------V---A  994
A.DE.x.BEN            -------S----------S----------P--T---------V--QS-------------------S------TI------------------------------L-R--N---Q-------------------------V----  994
A.DE.x.PEI2_KR_KRCG   -------S----------S----------T--T---------T--QS-------------------S------TM------------------------------LHA----L------F-----------------D----V----  994
A.GH.x.GH1            -------S----------S----------S--T---------I--QS-------------------S------TI------------------------------L-R--N---Q-------------------D----V---A  994
A.GM.87.D194          -------S----------S----------P--T---------I--QS-------------------S------TI------------------------------L-R--N--K-Q-------------------D----V---A  994
A.GM.x.ISY_SBL_6669_85 -------S----------S----------T---------V--QS-------------------E------TM----------V---V----------L-A----L------F----N---Q-------D----V----  994
A.GM.x.MCN13          -------S----------S----------T---------V--QS-------------------S------T-------T------------------------LHA----L------F-----------------D----V----  994
A.GM.x.MCR35          -------S----------S----------T---------V--QS-------------------S------T-------T------------------------LHA----L------F-----------------D----V----  994
A.GW.86.FG_clone_NIHZ -------S----------S----------T---------V--Q--------------------T-------------------I----------L-A----L------F-----------------D----V----  993
A.GW.87.CAM2CG        -------S----------S----------T---------V--Q--------------------S------T-------------S----------L-A----L------F-----------------D----V----  993
A.GW.x.MDS            ----------S----------T---------I--QS-------------------S--D---T-------------------S----------L-A----L------F-----------------D----V----  994
A.IN.07.NNVA          -------S----------S----------T---------V--QS-------------------S------T-------------S------A-----L-A----L------F-----------------D----V----  994
A.IN.95.CRIK_147      ----------S----------T---------I--QS-------------------S------T-------------S----------L-A----L------F-----------------D----V----  994
A.JP.08.NMC786_clone_41 -------S----------S----------S--T---------V--QS-------------------S------T-------------------------L-R--N----Q-------------------V----  994
A.PT.x.ALI            -------S----------S---------S-V--T---------V--QS-------------------S------T-------------------S------L-T--L-----P-----------------D----V----  994
A.SN.85.ROD           -------S----------S----------T---------I--QS-------------------S------TI------I-------------S----------L-A----L-D----F-----------------LV----  995
A.SN.86.ST_JSP4_27    -------S----------S----------T---------I--QS-------------------S------T-------------------V-A------L-A----LQ-----F-----------------D----V---A  994
B.CI.88.UC1           ----------S----------T--D---A---I--Q--------------------------D--V-I--V----A----------V----------F-A--L--Q--Q-------------------------LI----  995
B.CI.x.20_56          ----------S----------T--D------V--QS-------------------------D--V-I--V----A----------------A------F-A--L--Q--Q-------------------------LV----  995
B.CI.x.EHO            ----------S----------T--D---A---I--Q-------E----------------D--V-I--V----T----------IV----------L-T--L--Q-----------------D---------I----  994
B.GH.86.D205_ALT      ----------S----------T-PS------V--Q--------------------L-D--V-I--V----T----------V----------F-A--L--Q--Q-------------------------I----  992
B.JP.01.IMCJ_KR020_1  ----------S----------T--D---A---I-V-Q--------------------D--V-I--V----A----------IV----A-----L-T-KL--Q-----F-----------------I----  995
G.CI.92.Abt96         ----------T---------V--QA----------S-----K--K--D----I----------I----------T-------------------------I----  977
AB.CM.03.03CM_510_03  --X----S----------S----------T--D------V--Q--------------------S-L-D--V-I--V----T----------V----------F-A--L--Q--Q-------------------------X----  995
H2_01_AB.CI.90.7312A  ----------S----------T--D------V--Q--------------------D--V-I--V----A----------V----------F-A--L--Q--Q-------------------------I----  995
H2_01_AB.JP.04.NMC307_20 ----------S----------T--D------V--QS-------------------D--V-L--V----A----------V----------F-A--L--Q--Q-------------------------V----  995
H2_01_AB.JP.07.NMC716_01 ----------S----------T--D------V--QS-------------------D--V-L--V----A----------V----------F-A--L--Q--Q-------------------------V----  995
H2_01_AB.JP.08.NMC842_10 ----------S----------T--D------V--QS-------------------D--V-L--V----A----------V----------YF-T--L--Q--Q-------------------------V----  995
U.CI.07.07IC_TNP3     ----------T----S----------T--K------V--QS-------------------D----M-------T-------------------------T-------------------------I----  977
U.FR.96.12034         ----------T---------V--------------------E-------TM------A----------I----S------L-----------------------AD---------LI----  990
MAC.US.x.17EC1        ----------------------------------------------------------------------------------------------------------------------------------  981
MAC.US.x.251_1A11     ----------------------------------------------------------------S--------T-----------------------------------------------------  981
MAC.US.x.251_32H_PJ5  ----------V-----------------------------------------------------------------------------------------------------------------------  981
MAC.US.x.251_BK28     ----------------------------------------------------------------------------------------------------------------------------------  977
MAC.US.x.MM142_IVMXX  ----------S----------------------------------------------------------------------------------------------------------------------  977
MNE.US.82.MNE_8       ----------S----------------------------------------M----------L-----------------------------------------V----  977
MNE.US.x.MNE027       ----------S----------T----------------------------------------M----------L-----------------------------------------V----  977
SMM.CI.79.SIVsmCI2    ----------S----------T---------V--Q----------S----------I------C----------IV---S------L-T----------AN---------I----  986
SMM.LR.89.SIVsmLIB1   ----------VS---------T---------N--Q--R----------T--K-K----I------T----------V----------L-T----------I----  971
SMM.SL.92.SIVsmSL92A  ----------A---VK---------T---------Q----------D--V-L--V----T----------A----YH-T----K----------I---S  976
SMM.SL.92.SL92B       -----R---------I-S----K---------T---------L-V-QS----------DL--KN--K-----E----L----A----------IV-----L-T-YLN-Q----Q----------A---------VI----  973
SMM.US.04.G078        ----------S----------T----------------------------E----------F-----------------------------  977
SMM.US.04.G932        ----------S----------T----------  977
SMM.US.04.M919        ----------S----------T---------Q----------T----------I----------V----------  977
SMM.US.04.M922        ----------S----------T---------Q----------T----------I----------V----------  977
SMM.US.04.M923        ----------S---I------T---------Q----------T----------I----------V----------  977
SMM.US.04.M926        ----------S----------T---------Q----------M----------  977
SMM.US.04.M934        ----------S----------T---------Q----------M----------  977
SMM.US.04.M935        ----------T----------T----------T----------I----------V----------  977
SMM.US.04.M940        ----------S----------T----------E----------F----------  977
SMM.US.04.M946        ----------S----------T---------Q----------S-I----------  977
SMM.US.04.M947        ----------S----------T---------Q----------V-----I----------V----------N----------  977
SMM.US.04.M949        ----------S----------T----------E----------F----------  977
SMM.US.04.M950        ----------S----------T---------Q----------G-----M----------  977
SMM.US.04.M951        ----------S----------T----------E----------I----------F----------  977
SMM.US.04.M952        ----------S----------T---------Q----------  977
SMM.US.05.D215        ----------S----------T---------Q----------TH-----D----I--V----T----------V----------  977
SMM.US.06.FTq         ----------S----------T---------Q----------T----------I----------  981
SMM.US.11.SIVsmE660_FL10 ----------S----------T---------Q----------T----------I----------V----------  977
SMM.US.86.CFU212      ----------S----------T---------TN---------T----------I----------V----------  977
SMM.US.x.F236_H4      ----------S----------T---------Q----------T----------I----------V----------  976
SMM.US.x.H9           -----X----------S----------T---------Q--X-------X---------T----------I------I--I----------V----------  977
SMM.US.x.PBJA         ----------S----------T---------Q----------T----------I----------V----------  977
SMM.US.x.PGM53        ----------S----------T---------Q----------T----------I----------V----------  977
SMM.US.x.SME543       ----------T----------T---------Q----------T----------I----------V----------R----------  977
SMM.x.x.pE660.CG7G    ----------T----------T---------Q----------T----------I----------V----------  977
STM.US.89.STM_37_16   ----------S---V----------T---------Q----------T------D---T---V----------K---L--------V----------V----  977
```

p31 Integrase end
Pol end

```
MAC.US.x.239                DIKVVPRRKAKIIKDYGGGKEVDSSSHMEDT.GEAREVA.........*  1019
A.CI.88.UC2                 ------------R----RQ-L---P-L-GA.R-DG-------------  1033
A.DE.x.BEN                  ------------R----RQ-L---P-L-GA.R-DG-M-CPCQVPEIQ  1041
A.DE.x.PEI2_KR_KRCG         ---I--------R----RR--------L-G-.R-DG-----------  1033
A.GH.x.GH1                  ----I--------R----RQ-L##---L-GARE-DG-----------  1032
A.GM.87.D194                ------------R----RQ-L------L-GA.R-DG-----------  1033
A.GM.x.ISY_SBL_6669_85      ----I--------R---PRQ-M--G--L-GA.R-DG-M---------  1033
A.GM.x.MCN13                ----------------RQ-L--G--L-GA.R-DG-M---------  1033
A.GM.x.MCR35                ----------------RQ-L--G--L-GA.R-DG-M---------  1033
A.GW.86.FG_clone_NIHZ       E---------------RQ-M--G--L-GA.R-DG-M---------  1032
A.GW.87.CAM2CG              ---II--------R----RQ-L-----L-GA.R-NG-----------  1032
A.GW.x.MDS                  ---I---------R----RQDM--GP-L-G-.R-DG-----------  1033
A.IN.07.NNVA                ---II--------R----RQ----GP-L-G-.R-DG-----------  1033
A.IN.95.CRIK_147            ---II--------R----RQ-M--GP-L-G-.R-DG-----------  1033
A.JP.08.NMC786_clone_41     ----I--------R----RQ-L-----L-GA.R-DG-M-*PCQVPEI-  1040
A.PT.x.ALI                  ------------R----RQ-L--GP-L-GA.R-DG-----------  1033
A.SN.85.ROD                 ---II--------R----RQ-M--G--L-GA.R-DG-M---------  1034
A.SN.86.ST_JSP4_27          ---II------------RQ-M--G-NL-GA.R-DG---*PYQVSKIQ  1040
B.CI.88.UC1                 E---I--------RH------L-CGTD-----.RQ---M-QSGQVPEA-  1042
B.CI.x.20_56                E---I--------RH------L-CGADV----.RQ---M-QSGQVPEA-  1042
B.CI.x.EHO                  E---I--------RN------L-C-ADV----.MQ-----QSN*IPEI-  1040
B.GH.86.D205_ALT            E------------RH-----GL-C-AD-----.RQ---M-QSD*VS*V-  1037
B.JP.01.IMCJ_KR020_1        E---I---------N------L-C-AN-----.-Q---M-QPSQISET-  1042
G.CI.92.Abt96               E------------------------TN-----.RQTG-----------  1016
AB.CM.03.03CM_510_03        E------------RH-----GL-C-AD-----.RQ-S-M-QSDQIP*V-  1041
H2_01_AB.CI.90.7312A        E---I--------RH------L-C-TDV----.RQ---M-QSSQVSEA-  1042
H2_01_AB.JP.04.NMC307_20    E------------RH------L-CGTDV----.RQ---M-QSGQVPEA-  1042
H2_01_AB.JP.07.NMC716_01    E------------RH------L-CGTDV----.RQ---M-QSSQAPEA-  1042
H2_01_AB.JP.08.NMC842_10    E------------RH-----GL-CGTD-----.RQ---M-QSGQVPEA-  1042
U.CI.07.07IC_TNP3           ----I-------------------G-NL-GA.RQDG-M-QLGEIP*A-  1023
U.FR.96.12034               ----I--------R-------L----NL-GA.EKV--M-LPDQTPEYN  1037
MAC.US.x.17EC1              ------------------------------.---------------.  1020
MAC.US.x.251_1A11           --------------------M---------.---------------.  1020
MAC.US.x.251_32H_PJ5        ------------------------------.---------------.  1020
MAC.US.x.251_BK28           ------------------------------.---------------.  1016
MAC.US.x.MM142_IVMXX        --------------------M---------.---------------.  1016
MNE.US.82.MNE_8             ------------------------------.---------------.  1016
MNE.US.x.MNE027             ------------------------------.---------------.  1016
SMM.CI.79.SIVsmCI2          ----I----------------LGN-PYL-NP.E-DGKM-QPD.....-  1028
SMM.LR.89.SIVsmLIB1         ---------------------------G--L----.R-I--M-----.  1010
SMM.SL.92.SIVsmSL92A        ----I----------------M---AD--N-.TQV--T-QLSEISKV-  1023
SMM.SL.92.SL92B             E---------------RQ-MG--AS---Q*KNSGAL-----------  1012
SMM.US.04.G078              E--------------A-----G-G--L--P.-----M-----------  1016
SMM.US.04.G932              ------------------------P-L-----V--M-----------  1016
SMM.US.04.M919              E--------------------L--G--L-----K-------------  1016
SMM.US.04.M922              E--------------------L--G--L-----T-------------  1016
SMM.US.04.M923              ---------V----------L--G--L----.---------------  1016
SMM.US.04.M926              ------------R-------M--GP----P.---------------  1016
SMM.US.04.M934              ------------V----------GP-V--P.---------------  1016
SMM.US.04.M935              E--------------------L--G--L----.---------------  1016
SMM.US.04.M940              ---------------A-----G-G--L--P.-----M-----------  1016
SMM.US.04.M946              ------------V----------GP-V--P.---------------  1016
SMM.US.04.M947              E--------------------------G--L----.---------------  1016
SMM.US.04.M949              ---------------A-----GNGP-L--P.-----M-----------  1016
SMM.US.04.M950              ------------------------GP-V--P.---------------  1016
SMM.US.04.M951              ---------------A------G-G--L--P.-----M-----------  1016
SMM.US.04.M952              --------------------M--GP-V--S.---------------  1016
SMM.US.05.D215              ---------V--V---A----------L----.R-TG-----------  1016
SMM.US.06.FTq               ---------------A----L--GT-V----.R--G-M---------  1020
SMM.US.11.SIVsmE660_FL10    E--------------------L--G--L----.---------------  1016
SMM.US.86.CFU212            ----------------------G--L----.R--G-M---------  1016
SMM.US.x.F236_H4            E--------------------L--G--L----.---------------  1015
SMM.US.x.H9                 E--------------------L--G--L----.---------------  1016
SMM.US.x.PBJA               E--------------------L--G--L----.---------------  1016
SMM.US.x.PGM53              E--------------------L--GP-L----.---------------  1016
SMM.US.x.SME543             E--------------------L--G--L----.---------------  1016
SMM.x.x.pE660.CG7G          E--------------------L--G--L----.---------------  1016
STM.US.89.STM_37_16         ----------------------G--L----.R--G-----------.  1016
```

```
                          Vif start
MAC.US.x.239              MEEEKRWIAVPTWRIPER.LERWHSLIKYLKYKTKDLQKVCYVPHFKVGWAWWTCSRVIFPLQEGSHLEVQGYWHLTPEKGWLSTYAVRITWYSKNFWTDVTPNYADILLHSTYFPCFTAGEVRRAIRGEQLLSCCRFPRAHKYQVPSLQYLALKVV..SDVRSQGENPT  167
A.CI.88.UC2               ---G-S--V-----V-G-.M------V-----R----EG-R----H---------------KGE----I-A--N---------SHS--L---TEK-------DC----I-----C--A-----------K-----NY-Q---A---------V--Q.QND-P-RKG-A  168
A.DE.x.BEN                ---DRN--V-----V-G-.M-K--A-V-----R----EE-R----H----------------GK----I-A--N---------SH---L---TEK-------DC----I-----S-------------K-----NY-Q---A---------V--Q.QND-P-RKGTA  168
A.DE.x.PEI2_KR_KRCG       ---GE---V-----V-G-.M-K----V----HR----EG-----H----------------GN----I-A--N---------S--------TER-------DC--S-I-----S-------------K-----NY-Q--RSK--L--F---V--Q.QNG-P-KNST-  168
A.GH.x.GH1                ---G-N--.-----V-G-RM------V-----R-R--EE-R----H---------------KGE----I-A--N---------SHS------TER-------D-----I-----S-------------K-----NY-Q---V---------V--Q.QND-P-RKGTA  168
A.GM.87.D194              ---G-N--V-----V-G-.M------V-H---R----EE-R----H---------------EGE----I-A--N---------SHS--L---TEK-------DC--S-I-----S-------------K-----NY-Q---A---------V--Q.QNG-P-RKGAA  168
A.GM.x.ISY_SBL_6669_85    -DQG----------V-G-.M-K----------R----EQ-R----H---------------KGN----I-A--N---------S-S--M----EG-------DC--T-I-----S-------------KS----NY-Q---SK-----F---V--Q.QNDKP-RD-T-  168
A.GM.x.MCN13              ---G-N--V-----V-G-.M-K----V-----R----ED------H---------------KGN----I-A--N---------S-S--M---TER-------DC--S-I-----S-------------K-----NY-Q--RA------F---V--Q.QND-P-RNGTP  168
A.GM.x.MCR35              ---G-N--V-----V-G-.M-K----V-----R----ED------H---------------KGN----I-A--N---------S-S--M---TER-------DC--S-I-----S-------------K-----NY-Q--RA------F---V--Q.QND-P-RNGTP  168
A.GW.86.FG_clone_NIHZ     ---G----V--I--V-G-.M------V-----R----E-------H---------------K-N----I-A--N---------SHS------TEK-------DC--T-I-----S-------------K-----KY----RS------F---V--Q.QND----NSA-  168
A.GW.87.CAM2CG            ---G-S--------V-G-.M-K----V--------G--EQ------H---------------RGD-R--I-A--N---------S----M---TEK-------DC--T-I-----S-------------K-----KY----RS------F---V--Q.QND-P-RDRT-  168
A.GW.x.MDS                ---G-T--V-----V-G-.M-K----V-----------EE-R----H---------------KGN----I-A--N---------S----L---TEK-------DC--S-I-----T------A-------K-----NY-Q---S---T--F---V--Q.QNG-P-RDST-  168
A.IN.07.NNVA              ---G----V-----V-G-.M-K----V-----------ER------H---------------NNN-N--I-A--N---------S--------TEK-------D---S-I-G---A-------------KV----NY-Q---T---T--F---V--Q.QND---RDST-  168
A.IN.95.CRIK_147          ---D----V----KV-G-.M-K----V-----------E-------H---------------KDN--I-I-A--N----E----S--------TE--------DCT-S-I-----A-------------KI----NY-Q--RT---T--F---V--Q.QND---RDGT-  168
A.JP.08.NMC786_clone_41   ---D-N--V-----V-G-.M-K----V-----R----E--R----H---------------EGK----I-A--N---------S--------TER-------D-----I-----S-------------K-----NY-Q---V---------V--Q.QNG-P-RKGTS  168
A.PT.93.JAU1              ---D----V-----V-G-.M-K----V-----R----EQ------H---------------K------I-A--N---------S--------TEK-------DC--S-I-G---S-------------K-----NY-K---S------F---V--Q.QND-P-RKYT-  168
A.PT.x.ALI                ---G-S--V-----V-G-.M-K----V-----R----E-------H----------------GR----I-A--N---------S--------TEK-------DC--S-I-G---S-------------K-----NY-Q---S------F---V--Q.QNGKP-RNST-  168
A.SN.85.ROD               ---D----V-----V-G-.M-K----V-----------E-------H---------------KGN----I-A--N---------S-S------TEK-------DC--V-I-------------------K-----NY----RA------F---V--Q.QND-P-RDST-  168
A.SN.86.ST_JSP4_27        ---G----------V-G-.M-----------R-G--E-------H---------------KGE----I-A--N---------S-S--L---TEK-------DC--S-I-----S-------------K-----NY-Q----------F---V--Q.QNG-P-RD-T-  168
B.CI.88.UC1               --G--N--V-------G-.--K----V----HR--E--Q-S----H---------------K-EAY-------N----R-F--S----L---KRS-Y-----DV--Q---GS--S----N---------KI--Y-NY-S--EG------F---R-IQEGKDG----SA-  169
B.CI.x.20_56              -----N--V-------G-.--------V----HR--E--Q-S----H---------------K--AY--I----N----R-L--S--------RES-Y-----DV--Q---GS--S----N---------KI--Y-NY-S--DR------F---R--QEGKNG----SA-  169
B.CI.x.EHO                -----N----------C-.-------------R-----Q-S----H---------------K--A--------N----R-F--S----L---ERS-Y-----DV--R---GS--SS---N---------KI--H-NY-S--TG------F---R--QEGKDG----ST-  169
B.GH.86.D205_ALT          -----D--V-------G-.-------------R-GE--Q-S----H----------I----NK-AW-------N----R-F--S----L---ER--Y-----DV--Q---GS--S--S-N---------KI--Y-NY-S--EG------F---R--QEGKNG----SA-  169
B.JP.01.IMCJ_KR020_1      -----N----------G-.--K----V----HR--E-EQ-T----H---------------K--AY--I----N----R-F--A----L---K-S-Y-----DV--H---GS--F----N---------KI--Y-NY-S--RG------F---R--QEGGNGP---SAA  169
G.CI.92.Abt96             ----------------G-.--K------F--F-------AV----H---------------TKEA---I----N----------Q----L---TRK-Y-----ET--Q---GS--D---------------I----NY-T---R------F---Q--QKG.HG-K--SX-  168
AB.CM.03.03CM_510_03      -----D--V----K--G-.-A-------------GE--Q-S----H---------------DKXAX-------N----R-L-------L---ER--Y-----DV--Q---GS--S----N---------KI--Y-NY-S--EG------F---R-IQEGKNG---K-T-  169
H2_01_AB.CI.90.7312A      --G--N--V----K--G-.--K----V----HR-GE--Q-S----H---------------KG-AC-------N----R-F--S----L---EES-Y-----DV--Q---G---S----N---------KI--Y-NY-S--EG------F---R--QEGKNG----SA-  169
H2_01_AB.JP.04.NMC307_20  -----N--V-------G-.--K----V----HR-RE--Q-S----H---------------K-RAW-------N----R-F--S----L---EES-Y-----DV--Q---G---S----N---------KI--Y-NY-S--ER------F---R--QEGKNG-K--GA-  169
H2_01_AB.JP.07.NMC716_01  -----N--V-------D-.--K----V-H--HR-RE--Q-S----H---------------NK-AW--I----N----R-F--S----L---EE--Y-----DV--Q---G---S----N---------KI--Y-NY-S--ER------F---R--QEGKNG-K--GA-  169
H2_01_AB.JP.08.NMC842_10  -----D--V-------G-.--K----V----HR-GE--Q-S-I--H---------------RK-AW--I----N----E-F--S----L---KGS-Y-----DV--Q---G---S----N---------KI--Y-NY-S--ER------F---R--QEGKNG-K--SA-  169
U.CI.07.07IC_TNP3         --------V-----V-G-.M------V----HR--E-S-------H---------------EGE----I-A--N-----------------TRG--S----D---Q---G---S-Y------------------N--S---G------F----I-..Q---P-RK-T-  167
U.FR.96.12034             -----N--V-----V-R-.---------H-#-N--E---------H---------------KK-------A--N---------S----L---TRG-------DC--Q---GS--S--------------------N--S--RQ------F---RAL..Q-G--KR----  166
MAC.US.x.17EC1            -----------------.--------------------------------------------------------------------------------------------------------------------------------------------..----------  167
MAC.US.x.251_1A11         -----------------.---------------------------------------------------------------V--------------------------------------------------------------------------..----------  167
MAC.US.x.251_32H_PJ5      ---G-------------.-------------------Y-------------------------------R-----------R-------D-------------------V---------K-----R-----------R--..-Y---------  167
MAC.US.x.251_BK28         -----------------.---------------------------------------------------R-----------R-------D-----------------------------K-----R----------------..----------  167
MAC.US.x.MM142_IVMXX      --------V--------.---------------------A-----H------------------------N----R-------------D-------E-------------R------------H---------R--..-H---------  167
MNE.US.82.MNE_8           -----------------.--------------------------H------------------K-Q-------N----R-----H--------R--------DC------------------------------T---------R--..-Y----R----  167
MNE.US.x.MNE027           -----------------.--------------------------H------------------E-Q-------N----R--------------R--------D-------------------------------N---------R--..-Y----R----  167
SMM.CI.79.SIVsmCI2        -----S-VT-------R-.M--------H--FN--E---------N---------------EKE---------N---------S---------R--------DV--Q-I-----S-Y-Q-----------I----KY-S---R------F---RILHKQNGKCKR----  169
SMM.LR.89.SIVsmLIB1       --------V----K--G-.---------H---N--E-N-A-----H----------------GEAY-------N-------------L----R--------D---V---G-----SE--------------------K----------F-------..EH----R----  167
SMM.SL.92.SIVsmSL92A      -------------K--H-.---R---V--------E--K--V----H---------------EKEA-------N---------C----L---TR---------DL----I-GA--S-Y-Q-----------I----N--S--Q-K-S---F---RAL..Q.DG-KRTD-S  166
SMM.SL.92.SL92B           ---G---V---V---SR-IV-----C--FH----RE-E-A-----H------Y-A-------E--------V--N---------S---------EK-------DV--Q-A---------A-HA--Q------V--Y-GYAV--HSS-Q---L------L.QND-PK-K---  169
SMM.US.04.G078            -Q-----VV-------G-.---------H---N--E---------H----------------GET---I----N------------------TRG-------D---V---G---T--SE-------------------K---N-------------..-Y----R----  167
SMM.US.04.G932            ----------------G-.---------H---T--E---------Y------------------D-Y-------S----R---G----------R--------DC--T----------S-------------------K---N----------R--..-H----R-D--  167
SMM.US.04.M919            -----N--V-------G-.-------------------A-----H---------------RDDT---------N-------------------R--------DF--T---G-----SE----------K---R-K--K---N----------T--..-H------D--  167
SMM.US.04.M922            -----N--V-------GK.----------------R----A-----H---------------RDNT---------N-------------------R--------D---T---G-----SE----------K--------K---N----------T--..-H-----KD--  167
SMM.US.04.M923            -----N--V-------G-.--K------H---N------A-----H---------------KDET---------N-------------------R--------D---T---G-----SE----------K-----K--K---N--S-------T--..-H----R--S-  167
SMM.US.04.M926            --------V-------G-.------I--H---N------------H----------------GE----I----N----T--------------R--------D---T-----------G-------------------K---H------F---R--..-HG--------  167
SMM.US.04.M934            --------V-------G-.--K----------N------------H----------------GE----I----N-------------------R--------D---T-----------G-------------------K---Q------F---R--..-H---------  167
SMM.US.04.M935            -----N--V-------G-.-------------------A-----H---------------RDDT---------S-------------------R--------D---T---G-----SE----------K--------K---N----------T--..-H-----KD--  167
SMM.US.04.M940            -Q-----VV-------G-.---------H---N-RE---------H----------------GET---I----N------------------TR--------D---V---G---T--SE-------------------K---N----------E--..-Y----R-D--  167
SMM.US.04.M946            --------V-------G-.--K----------N-----R------H----------------GE----I----N-------------------R--------D---T-----------G-------------------K---Q------F---R--..-H----R----  167
SMM.US.04.M947            -----K--V-------G-.--K----------N------A-----H---------------KGEA---------N-------------------R--------D---T---G-----SE----------R-----K--K---N----------T--..-H--P--KD--  167
SMM.US.04.M949            -Q-----VM-------G-.---------H---N-RE---------H----------------GET---I----N------------------TR--------D---V---G---T--SE-------------------K--RH----------R--..-Y-----K---  167
SMM.US.04.M950            --------V-------G-.---------H---N------A-----H-----------------E----I----N--------------#----R--------D---T----------SG-------------------K---Q------F---R--..-H----R----  166
SMM.US.04.M951            -Q-----VV-------G-.--K----------N-RE---------H----------------GEA---I----N------------------AR--------D---V-------T--SE-------------------K---N----------R--..------R----  167
SMM.US.04.M952            --------V-------G-.---------H---N------------H-----------------E----I----N-------------------R--------D---T----------SG-------------------K---Q------F---R--..-H----R----  167
SMM.US.05.D215            -Q------V-------G-.--K------H--FN------A-----H---------------EKE---------N------------------TR--------DC--V---N------S-------------------K--TN--------------..EH-K--R----  167
SMM.US.06.FTq             -Q---N--V-------G-.---------H---N--E---A-----H---------------EGE---------N------------------TRR--S----D---T---G------------------------S----K---N----------RI-..-H----R---A  167
SMM.US.11.SIVsmE660_FL10  -----N--V---------.---------H---N-----MA-----H---------------RDET---------N-A--------H-------R--------D---T-----------SE----------K-----K--K---N----------T--..-H------D--  167
SMM.US.86.CFU212          --------V-------G-.--K------H---N-----QA-----H----------------GE----------N-------------------R--------D---T---GS--S----------------------K-----RN----------S--..-HG-----D--  167
SMM.US.x.F236_H4          -----N--V---------.---------H---N-----MA-----H---------------RDET---------N-A-----------------R--------D---T-----------SE----K-----K-----K--K---N----------T--..-H------D--  167
SMM.US.x.H9               -----N--V-------G-.--K------H---N------A-----H---------------RDEA---------N----X--------------R-----A--D---T---G-----SE----------K-----K--K---N----------T--..-H-X-----X--  167
SMM.US.x.PBJA             -----N--V-------G-.--K------H---N------A-----H---------------KDEA---------N-------------------R--------D---T---G-----SE----------K-----K--K---N----------T--..-H---------  167
SMM.US.x.PGM53            -----N--V----K--G-.--K---------FR------A-----H---------------RDE----------N-A------------------RK-------D---T---G-----SE----------K-----K--K---N----------T--..-H----R-D--  167
SMM.US.x.SME543           -----N--V---------.---------H---N-----MA-----H---------------RDKT---------N----------H--------R--------DC--T-----------SE---Q------K-----K--K---N----------T--..-H----R-D--  167
SMM.US.x.pE660.CG7G       -----N--V---------.---------H---N-----MA-----H---------------RDET---------N-A--------H--------R--------D---T-----------SE----------K-----K--K---N----------T--..-H------D--  167
STM.US.89.STM_37_16       --------V-------G-.---------H---N--E-S-A-----H----------------GEA---------N---------E---------TR---S----DC--Q---G-------------------K-------TK---N-------------..EH----R--TA  167
```

Vif end

MAC.US.x.239	WKQWRRDNRRGLRMAKQNSRGDKQRGGKPPTKGANFPGLAKVLGILA.....*	214
A.CI.88.UC2	R------HW----V-R-DY-SLE-------APR-H---V----E---.....	216
A.DE.x.BEN	R------HW----V-REDH-SL--G-SE-SAPR-H---V----E---.....	216
A.DE.x.PEI2_KR_KRCG	R-R--SNYW--F-L-RKDG--H----SE--AS--Y---V----E---.....	216
A.GH.x.GH1	R------HW----V-R-DY-SL----SE-SAPR-H---V----E---.....	216
A.GM.87.D194	R------HW----V-R-DY-SL--G-SE-SAPR-H---V-------.....	216
A.GM.x.ISY_SBL_6669_85	R-----NY-----L-R-DG-SH----SE--AQ--Y---V----E---.....	216
A.GM.x.MCN13	R------Y----QL-R-DG-SH----SES-APR-Y---V-E--E---.....	216
A.GM.x.MCR35	R------Y----QL-R-DD-SH----SES-APR-Y---V-E--E---.....	216
A.GW.86.FG_clone_NIHZ	R--R-G-Y-------R-D---Y----SES-PTR-H-----E--E---.....	216
A.GW.87.CAM2CG	R------Y-----L-R-D--SY----SES-AP--Y---V----E---.....	216
A.GW.x.MDS	R--R---Y-A---L-R-D--SH-----ES-APR-Y---V----E---.....	216
A.IN.07.NNVA	R--R---Y-----L-REDG-SH----SES-AQ--H---V-------.....	216
A.IN.95.CRIK_147	R--R---Y-----L-R-DG-SH-----ES-AQ-TH-----E------.....	216
A.JP.08.NMC786_clone_41	R------HW----V-R-DY-SLE---SG-SAPR-H---V----E---.....	216
A.PT.93.JAU1	R------YW----V-R-D-G-H----S---APR-Y---V----D---.....	216
A.PT.x.ALI	R------Y-----V-R-D---L-----ES-AP--H---V----E---.....	216
A.SN.85.ROD	R--R---Y-----L---D--SH---SSES--PRTY---V-E--E---.....	216
A.SN.86.ST_JSP4_27	R-----NY-----V-R-DG-SH----SE--APR-Y---V----E---.....	216
B.CI.88.UC1	R--R--N---SI-L-RK-NNRAQ-GSSQ-LAPRTH-----E------.....	217
B.CI.x.20_56	R--R-G-S---I---RK-TSRTQ-GSSQSFAPRTH-----------.....	217
B.CI.x.EHO	R--R--NS---I---RD-I-TSQ-SSSQSLAQ-TY-----E------.....	217
B.GH.86.D205_ALT	R--R--NS--SI-L-RK-NNRAQ-GS-Q-FAPRTY-----E------.....	217
B.JP.01.IMCJ_KR020_1	R--R--N---SI---RK-NNRAQ-SSCQSSAQ-TY-----E------RGTGHDN-	224
G.CI.92.Abt96	R--R-G-----I---RK--XR-Q-DSSQSF-Q--Y-----------.....	216
AB.CM.03.03CM_510_03	RE-R--N---SI-X-RK-NNRAQ-GSSQ-LAPRTH-----------.....	217
H2_01_AB.CI.90.7312A	R--R---S---I---RK-NSRTQ-GSSQ-FAPRTY-----------.....	217
H2_01_AB.JP.04.NMC307_20	R--R-------I-L-RK-NNRAQ-GSSQ-FAP-TH-----E------.....	217
H2_01_AB.JP.07.NMC716_01	R--R-------I-L-RK-NNRAQ-GS-Q-F-P-TH-----E------.....	217
H2_01_AB.JP.08.NMC842_10	R--R-------I-L-RK-NNRAQ-GSSQSFAP-TH-----------.....	217
U.CI.07.07IC_TNP3	--LR--NSG-S----RV--HSH-P--S-SSAQ-TY---V-------.....	215
U.FR.96.12034	R-FR--NC---FQL-RK-CERYQ-GSSATSSQ-TY--------A---.....	214
MAC.US.x.17EC1	----------.....	215
MAC.US.x.251_1A11	----------.....	215
MAC.US.x.251_32H_PJ5	-------S--S----------.....	215
MAC.US.x.251_BK28	----------S-------D----------.....	215
MAC.US.x.MM142_IVMXX	----------S--V----------E----------.....	215
MNE.US.82.MNE_8	----------S----------S------VD----------.....	215
MNE.US.x.MNE027	----------S----------S------D----------.....	215
SMM.CI.79.SIVsmCI2	RN--G-N----F-V-I-H-GSS--GSS-TSAQ--D----------.....	217
SMM.LR.89.SIVsmLIB1	------NY-----V-RK--GRN--T-SE-S---ID----------.....	215
SMM.SL.92.SIVsmSL92A	RE--G---------L-RK---RSQ-NSSE-FA---H----------.....	214
SMM.SL.92.SL92B	R-----N-------VT-Y-G-PESNSSTT---RV-----E------GR.....	218
SMM.US.04.G078	------------V-A--GSR---G-C--S---TD----------V.....	215
SMM.US.04.G932	----------SF--V--DG-RY----SESS----D----------.....	215
SMM.US.04.M919	------N---S----R-----Y--GSSESLAE--D----------.....	215
SMM.US.04.M922	------NT-----L-R----R---G-SESLAE-V----------.....	215
SMM.US.04.M923	------N--------RK----N--GSSESFAE-T----------.....	215
SMM.US.04.M926	------Y---N--V-TKD--RS--GSS-SS---ID---V--I----V.....	215
SMM.US.04.M934	----------S----TK---RN--GSSES----VD---V-------V.....	215
SMM.US.04.M935	-E----NS-----L-R----RN--GSSESLAE----------.....	215
SMM.US.04.M940	R----GN----F-V-TK--GRN--G-C--S---T----V-------.....	215
SMM.US.04.M946	----------SF---TK---RN--GSSES----VD---V-------V.....	215
SMM.US.04.M947	-------S-----L-RK---RH--GSSESFAE-V----------.....	215
SMM.US.04.M949	-------S-GS--L-RK-G-RNE-G-R-SS---T----V-------.....	215
SMM.US.04.M950	----------V-TK---RH--GSSESS---ID---V-------V.....	214
SMM.US.04.M951	------NS-GS----T--G--H--G-RE-S---TD---V-------V.....	215
SMM.US.04.M952	-----SS---SP-V-A----RG--GS-E-S---VD---V-------V.....	215
SMM.US.05.D215	R-------WGSF---RK--GRH--SSS-SS--R-D----------.....	215
SMM.US.06.FTq	------N--------RED--RY--GSS------TD----------.....	215
SMM.US.11.SIVsmE660_FL10	----G-N----------RN--GSS-S-AE----------.....	215
SMM.US.86.CFU212	R-----N---S--V-E--GGRT--GSS-S-A-R----------.....	215
SMM.US.x.F236_H4	------N--K----------RN--GSS-S-AE----------.....	215
SMM.US.x.H9	------NX---X-L-T--X-RN--GSSESFAX---------X----.....	215
SMM.US.x.PBJA	------N------L-R----RN--GSSESFAE-T----------.....	215
SMM.US.x.PGM53	------N----I--------RN--GSSESLAE----------.....	215
SMM.US.x.SME543	-----GN----------RN--GSS-S-AE----------.....	215
SMM.US.x.pE660.CG7G	----G-N----------RN--GSS-S-AE----------.....	215
STM.US.89.STM_37_16	R-----G--GSI-V-T--G--H-P--S--S-E-TD----------.....	215

```
                              Vpx start                                                                                              Vpx end
MAC.US.x.239                  MSDPRERIPPGNSGEETIGEAFEWLNRTVEEINREAVNHLPRELIFQVWQRSWEYWHDEQGMSPSYVKYRYLCLIQKALFMHCKKGCRCLGEGHGAGGWRP.GPPPPPPPGLA*  112
A.CI.88.UC2                   -T-----V-----------------D--I-AL--------------------R-------L-S--T-------M---M-I-L----T---G---P-----.----------V-  113
A.DE.x.BEN                    -T-----V-----------------E--I-AL--------------------R---------A--T-------M---I-T-F-R--T-W--DM-RE-LEDQ----------V-  114
A.DE.x.PEI2_KR_KRCG           -A---KTV----G--V------A--E-M--A---------------------R----DL---Q--T-----R-M-Y-M-I-V----T---G---P-----.-------.--V-  112
A.GH.x.GH1                    -T-----V-----------------D--I-AL--------------------R----D-------T-------M---V-I-F-R--T---G---P----S.----------V-  113
A.GM.87.D194                  -A-----V-----------------D--I-AL--------------------A---------T--T------IM---VYI-F----T---R---P-----.----------V-  113
A.GM.x.ISY_SBL_6669_85        -TN---T-----------E---D--D----A---------------------R---------R--T-------M---V---F----T-R-----P----S.----------V-  113
A.GM.x.MCNI3                  -T----TV----------E---N--D----A---------------------R---------Q--T-----------M-T---R--T---G---P-----.----------V-  113
A.GM.x.MCR35                  -T----TV----------E---N--D--I-A---------------------R---------Q--T-----------M-T---R--T---G---P-----.----------V-  113
A.GW.86.FG_clone_NIHZ         -T----TV----------E---A--D----A--------H------------R---------T--T---------M-MY--A-RDGT---G#--PE---S#----------V-  112
A.GW.87.CAM2CG                -T----TV----------E---A--D----A---------------------R---------Q--T-------M---V-I-F-R--T---G---P-----.----------V-  113
A.GW.x.MDS                    -A----TV----------EQ--A--D----A---------------------R---------T--T-------M---MYI-GR---T-R-----P-----.----------V-  113
A.IN.07.NNVA                  -T----TV--------------A--E----A---------------------G---------A--T-------M---M-I-ARR--T---G---G----S.--L------WV-  113
A.IN.95.CRIK_147              -T----TV--------------A--D----A---------------------G---------A--T-------M---I-I-VRR--T---G---P----S.----------V-  113
A.JP.08.NMC786_clone_41       -A-----V-----------------D--I-AL-----D--------------R------------T-------M---V-I-FMR--T-------P-----.----------V-  113
A.PT.x.ALI                    -AN---TV-----------------D----AL--------------------R---------Q--T-------M---MYT-FM---T---G---P----S.----------V-  113
A.SN.85.ROD                   -T----TV--------------A-------A---------------------R---------E--T------I----VY--VR---T---R---P-----.----------V-  113
A.SN.86.ST_JSP4_27            -AG---T------------------D----A---------------------R---------I--T-------M---M-I-S-R--T---G---P----S.----------V-  113
A.SN.x.A2057                  -A----TV----------E------D----AL--------------------R---------R--TI----------MY----R--T---G---P-----.----------V-  113
A.SN.x.A640                   -A----T---------------D--D----A---------------------R---E-----V--T-------M---MYT-F-R--T---G---P-----.----------V-  113
B.CI.88.UC1                   -.-----V-----D--------D--E--IT-L--V----------------C-A--RE-----S--T-----L-M---M-V-YT------Q----P----S.-------------  112
B.CI.x.20_56                  -.-----V-----D---V----A--E--LA-L-K---------------K--A--R------I--T-----L-M---MYI-IA---T-RR----P----Q.-------------  112
B.CI.x.EHO                    -.-----V---------V-------ET-L-HL--V--------------K--A--RE-----I--T-------M---M-I-FA---G--R----P----S.-------------  112
B.GH.86.D205_ALT              -.-----V-----D---V----A--E--IT-L--V-------------------A--RE-----I--T-----L-M---M-V-YT------Q----P----S.-------------  112
B.JP.01.IMCJ_KR020_1          -.-----V-----------------E--IT-L--V-------------------A---E----TI--T-------M----Y--LAR--A--R----P-----.-------------  112
G.CI.92.Abt96                 -.-----V--E--D-----------E---T----I--------------R-------------A--T-----L-M---M-V-------T--QK---P----Q.-------------  112
AB.CM.03.03CM_510_03          -.-----------D--------X--E--IT-L--V----------------C-A--R------S--T-----L-M---M-V-YT---T-RQ----P----S.-------------  112
H2_01_AB.CI.90.7312A          -.-----V-----D---V----A--E--IV-L----------------------A--R-D----I--T-----L-----M-V-FA---T--RG---P----R.-------------  112
H2_01_AB.JP.04.NMC307_20      -.-----V-----D--------D--E--IT-L----------------------A--RE-----S--T-----L-M---M-V-FS---T-RR----P----Q.-------------  112
H2_01_AB.JP.07.NMC716_01      -.-----V-----D--------D--E--IT-L----------------------A--RE-----S--T-----L-M---M-V-FS---T-RR----P----Q.-------------  112
H2_01_AB.JP.08.NMC842_10      -.-----V-----D--------A--E--IT-L----------------------A--RE-----T--A-----L-M---M-V-FS---T-RR----P----Q.-------------  112
U.CI.07.07IC_TNP3             --------------D--VE------ES--TA--Q--------------------G--------T---T------------Y----R--Q--R----P-----.-------------  113
U.FR.96.12034                 -G-----------E---V----S--E---RD---A--Q----------------Q--R------E--T-------------------Q-------P----S.----------I-  113
MAC.US.x.17EC1                ---------------------------------------------------------------------------------------------------.-------------  113
MAC.US.x.251_1A11             ---------------------------------------------------------------------------------------------------.-------------  113
MAC.US.x.251_32H_PJ5          -----------------V--------------------------------------------Q------------------------------------.-------------  113
MAC.US.x.251_BK28             --------------------------------------------------------------Q-----------M------------------------.-------------  113
MAC.US.x.MM142_IVMXX          --------------------------------------------------------------Q--T---------------------------------.-------------  113
MNE.US.82.MNE_8               --------------------------------------------------------------Q------------------------------------.-------------  113
MNE.US.x.MNE027               ---------------------------------------------------------------------------------------------------.-------------  113
SMM.CI.79.SIVsmCI2            -GSA-------T-------------YS---AV--G--K-----------R----F--------QA-TQ----L-------C-V----T--R----S----A.-------------  113
SMM.LR.89.SIVsmLIB1           -------------------------E--------Q--------------R----F--------A--T----------M-L------------.S----S.-------------  112
SMM.SL.92.SIVsmSL92A          -.----Q-------D-------D--E-----V--T--------------R-------------AA-T-----------------------G.--P-----.---------F--  111
SMM.SL.92.SL92B               -T-----------------------HN---AL-QT--Q-----------R-C----V----Y----A----VQ-M---M-Q-FR---T-R----SQ----T.-------------  113
SMM.US.04.G078                -------------------------H---A-------------------R------Y----R-E--T-----------W---M------R-----------.-------------  113
SMM.US.04.G932                ------------------------------D--------------------------------S-------------M---------------EI-----.-------------  113
SMM.US.04.M919                -------------------------D----D---A--------------R--------KI---V--T-----------------R-----E-E-------S.-------------  113
SMM.US.04.M922                -----------------L----D--D--------A--------------R---------M---E--T----------M-----R-------E-------S.-------------  113
SMM.US.04.M923                -------------------------E--------A--------------R---------I---V--T-------------------------GEY--------.-------------  113
SMM.US.04.M926                -A-------------D----T----R-----V-----------------R------YE-----Q---Q-------------------------------S.-------------  113
SMM.US.04.M934                -------------------------Q---A-------------------R------YE-----Q---Q-------------------------------S.-------------  113
SMM.US.04.M935                -----------------V----D--D--------A--------------R---------M---E--T----------M-----R-------E-------S.-------------  113
SMM.US.04.M940                -------------------------Q-----------------------R-------------VA-T-----------W---M------------------.---------R---  113
SMM.US.04.M946                -------------------------Q---A-------------------R------YE----PR---Q----------------R-----------------.-------------  113
SMM.US.04.M947                -----------------V----D--E----D---A--------------R---------I---V--T-------------------------GEY------S.-------------  113
SMM.US.04.M949                -----------------V----D--E-----M-----------------R-------------E--T-----------W----------R------------.-------------  113
SMM.US.04.M950                -------------------------Q----D------------------R------YE-----Q-------------------------------------.-------------  113
SMM.US.04.M951                -----------------V-------H----D------------------R------Y----R-E--T-----------W---M-------------------.-------------  113
SMM.US.04.M952                ----------------AA---VL---H-----V-----------------R------YE-----Q---E-----------I----------------------.-------------  113
SMM.US.05.D215                --N----------------------E----D---V--------------R----------------T-----------YI----------------D----.-------------  113
SMM.US.06.FTq                 -------------------------E----D---A--------------R-----------R--A--T-----------V---------R--------T--.-------------  113
SMM.US.11.SIVsmE660_FL10      ------------------E-----------G---A--------------R---------M---E--T-----------V--------------------T.-------------  113
SMM.US.86.CFU212              -A----------------------------L--V---------------R-C--F--------A--T-------M---M----R---N--R---E-----Q.-------------  113
SMM.US.x.F236_H4              --------------------K-------------A--------------R---------M---E--T-----------V------------E-------T.-------------  113
SMM.US.x.PBJA                 ----------------------D--D--------A--------------R---------M---V--T----------M-------------GE--------.-------------  113
SMM.US.x.PGM53                ----------------------------------A--------------R------R--M---E--T-----------V---R-------E-------S.-------------  113
SMM.US.x.SME543               ------------------E-----------G---A----------------------M---E--T-----------------------------------T.-------------  113
SMM.x.x.pE660.CG7G            ------------------E-----------G---A--------------R---------M---E--T-----------V------------E-------T.-------------  113
STM.US.89.STM_37_16           ----------------A-E------H----D------------------R------------PG-----------------S-R-----------------.-------------  113
```

	Vpr start	Vpr end	
MAC.US.x.239	MEER....PPENEGPQREPWDEWVVEVLEELKEEALKHFDPRLLTALGNHIYNRHGDTLEGAGELIRILQRALFMHFRGGCIHSRIGQPGGGNPLSAIPPSRSML*		101
A.CI.88.UC2	-M-GA.....#-MIV*I----*R-------I--D-Y------N-----R---K------S--L-V-GSR----*TRRRT-CP-T-TP-G-H-		83
A.DE.x.BEN	-T-APTEF---DGT-R-DLGSD--I-T-R-I-----R-------I---YY-H----------R---KT------V---A--NR----*TRRRT-CP-A-TP-A-H-		105
A.DE.x.PEI2_KR_KRCG	-T-APAEF---D-T-P-G-G----IGI-R--R-----------T---Y-CA-------S-R---NV------V---A--KI-----TR-ET-F----TP-G-Q-		106
A.GH.x.GH1	-T-APTEF---DGT-R--LGGD--IRI-G-I-----------I----Y-HS-----P---R----------V-L-A--NR---S-TRRRT-FP-A-TP-G-Y-		106
A.GM.87.D194	-T-APTEF---DGT-R--LGST--I-T-K-I----------C--I----Y-----------R----V------V-I-A--DR--K--TRRRA-CP-A-TP-G-H-		106
A.GM.x.ISY_SBL_6669_85	-T-APAEF---DGT-P---G----I-I-R-I----------YY--T---------R----V------T---A--G------R-R------TP-N-Q-		106
A.GM.x.MCN13	-T-APTEF---DGT-P---G----II-I-RKI-K----------Y-HT---------R---NV----------A--RL-----T--RT-FP-TSTP-T-Q-		106
A.GM.x.MCR35	-T-APTEF---DGT-P---G----II-I-RKI-K----------Y-HT---------R---NV----------A--RL-----T--RT-FP-TSTP-T-Q-		106
A.GW.86.FG_clone_NIHZ	-T-APTEL---DRT-P---G-A--I-I-R-IE----R----------RY--T---------R----------A---A--G------TR---------TP-G-H-		106
A.GW.87.CAM2CG	-T-APTEL---DGT-P---G----I-I-RDI----------G---A------R-R----V------T---A--N------TR---------TP-R-*-		105
A.GW.x.MDS	-T-APTEL---DGT-P---G----I-I-R-IT--------RS------R---T---------R----------T---A--G------TR---------TP---Q-		106
A.IN.07.NNVA	-T-APTEL---DRT-P---G----I-I-R-I-----R----------RY--T-Y--P---R----V------T---A--G------R---------TP---R-		106
A.IN.95.CRIK_147	-T-APTEL---DGT-P---G----I-I-R-I-----R----------IY--D-Y------R----V------T---A--G------R------T--TP---R-		106
A.JP.08.NMC786_clone_41	-T-APTEF---DGT----LGGD-IIRI----IR-------G--P-I----Y--A--------RK-VD----V--L--KA--SRPKVN-TR--TS-PVA-TP-N-Q-		106
A.PT.x.ALI	-T-APTEF--AGM--HQGAR----I---R-I-----R-----M-I---GY--T-------R-R---NA----------A--GR--V--TR-R-------TP-N-Q-		106
A.SN.85.ROD	-A-APTEL--VDGT-L---G---II-I-R-I----------I---KY--T---------R---KV------T---A--G------TR---------TP-N-Q-		106
A.SN.86.ST_JSP4_27	-T-APTES---DRT-P---G----I-T-R-I-*----------IT---Y--A---------RG--------LL---A--GR------R-R-------TP-G-R-		105
A.SN.x.A640	-A-APTEF---DGT-P---G----IQI-R-IG----------KY--A---------R----------T---A--G------TR--D------TP-D-Y-		106
A.US.93.7924A	-T-APTEL----GT-P---G----I-I-R-IQ---V----R----N--RY--T---------R----------T---A--G------TR---------TP-G-Q-		106
B.CI.88.UC1	-A-AAPET-----S------E---ED-M--I-Q---R----------F--S------A------K-------L---A--Q----------*--R-		105
B.CI.x.20_56	-A-AAPET-----D------E---ED----I-Q---R----------F--S------A------K-------L------Q----------*G-Q-		105
B.CI.x.EHO	-A-AVPEI---DKN------EQ---D----I-Q----------F------N---------KL------L------Q----------*G-Q-		105
B.GH.86.D205_ALT	-A-AAPEI-----N------E--IG-I---I-Q----------F--S------A------K-------L---A--Q------S-------T---P*G-R-		105
B.JP.01.IMCJ_KR020_1	-T-AAPET-----D------E---ED------Q---R----------F--G-----IA----------L------Q----------V----*--QK		105
G.CI.92.Abt96	-A-...EI---D-A----------I----------Y--D----------K-------L----D-RC----GNA------T-----GVF-		103
AB.CM.03.03CM_510_03	-A-AXXEI-----S----Q-E-RXG-I---I-Q----------X----F--S-Y----A-----VK------L---A--R----------X---P*G-Q-		105
H2_01_AB.CI.90.7312A	-A-AASET-----D------E---ED----I-Q---R----------F--S------A------K-------L------Q----------T---EG-Q-		106
H2_01_AB.JP.04.NMC307_20	-A-AAPET---D-N---G--E---EDI-----Q----------F--S------A--R---K-------L------Q----------PGG-*-		105
H2_01_AB.JP.07.NMC716_01	-A-AAPET-----N---G--E---EDI-----Q----------F--S------A------K-------L------Q----------PGG-*-		105
H2_01_AB.JP.08.NMC842_10	-A-AAPET---D-S---GS-E---EDI-----Q----------F--S-Y----A------K-------L------Q----------PGG-*-		105
U.CI.07.07IC_TNP3	-A-.....---D----------Q--------S-------Y--D-Y----------K-------V------A-----GQR-----A-V----NV--		102
U.FR.96.12034	-A-I....---DGA-----G----R-I--------VR----------Y----------K-----I-L------R-----SSR-----TTV---*GV--		101
MAC.US.x.17EC1	----------.----------.----------.----------.----------.----------.----------.----------.----------.----------		102
MAC.US.x.251_1A11	----------.-------*----G----------G----------.----------.----------.----------.----------R----------		101
MAC.US.x.251_32H_PJ5	----------.----------.----------.----------.----------.----------.----------N----------T-----GV--		102
MAC.US.x.251_BK28	----------.----------.----------.----------.----------.----------.----------N----------T----*GV--		101
MAC.US.x.MM142_IVMXX	----------.----------K----------.----------.----------.----------I---S--S----------T----------		102
MNE.US.82.MNE_8	----------D----------.----------.----------.----------K----------T------S------T-----R---		102
MNE.US.x.MNE027	----------D----------.----------Y--D----------K----------N------S------T---------		102
SMM.CI.79.SIVsmCI2	-A-H....---D---P-----A---D----I----------Y--D----------K-------I---A--L-----HS------TV----GV--		102
SMM.LR.89.SIVsmLIB1	-A-.....---D---P----------V----------N----Y--D-Y----------K----------R-----S--R---AT-----GV--		102
SMM.SL.92.SIVsmSL92A	-.SLSP....---D-A----------T-----I-T----------Y--D-Y----------K------L------TN----CRR-----AT-----NV--		103
SMM.SL.92.SL92B	---H....A--D-TNP-------IRD----------YV-DTY---I-----I-K---K---L---H--T-------R-P---GS--SA-DV--		101
SMM.US.04.G078	-A-.....---D----------T-----I-Q----------Y--D----------K----------N--V------S------AT-----EV--		102
SMM.US.04.G932	-T-.....---D----------I---------H--------Y--D-Y----------K-------T------V----------T-----RV--		102
SMM.US.04.M919	-T-.....---D-A----------I----------Y--D----------L------T-----S------T-----GV--		102
SMM.US.04.M922	-T-.....---D-A----------V----------Y--D----------L------V---V--S------A------GV--		102
SMM.US.04.M923	-T-.....---D-A----------I-----N----------Y--D----------I------R-----S------T-----GV--		102
SMM.US.04.M926	-A-.....---D-A----------L-----V----------Y--D-Y----------K------I------T-----S------T-----GV--		102
SMM.US.04.M934	-A-.....---D-A----------I-----I---S----------Y--D-Y----------K------I------T-----S------TV----GV--		102
SMM.US.04.M935	-T-.....---D-A----------V----------Y--D----------L------V------S------A------GV--		102
SMM.US.04.M940	-A-.....---D------S-----T-----I-Q---N----------Y--D----------K----------N--L-----AR-----AT-----EV--		102
SMM.US.04.M946	-A-.....---D-A----------I-----I---S----------Y--D-Y----------K------I------T-----S------T---A-GV--		102
SMM.US.04.M947	-A-.....---D-A----------V----------Y--D----------I------S-----S------T-----GV--		102
SMM.US.04.M949	-A-.....---D----------T-----V-Q---T----------Y--D----------K------L---A--A------AR-----A---GV--		102
SMM.US.04.M950	-T-.....---D-A----------I-----I---S----------Y--D-Y----------K----------T------S------T-----GV--		102
SMM.US.04.M951	-A-.....---D----------T-----V-Q----------Y--D----------KA--L------A------AT-----RV--		102
SMM.US.04.M952	-A-.....---D-A----------I---------S----------I--Y--D-Y----------K------I------T-----S------T-----GV--		102
SMM.US.05.D215	-AGV....---D----------G-----I----------N----Y--D----------K----------Q----------AT-----GV--		102
SMM.US.06.FTq	-A-.....---D----------T-----Y--D----------K------L------T-----TNR-----TV-T--NV--		102
SMM.US.11.SIVsmE660_FL10	-A-.....---D-A----------Y--D----------I---S--A------SR-----T-----D---		102
SMM.US.86.CFU212	-A-A....---D-A----------M----------KR---L--------Y--D----------K------I------A----H-R-----AT---T-DV--		102
SMM.US.x.F236_H4	-A-.....---D-A----------Y--D----------I---S--A------SR----------.....V--		90
SMM.US.x.PBJA	-T-.....---D-A----------I-----N----------Y--D----------I------R------S------T-----GV--		102
SMM.US.x.PGM53	-T-.....---D-A----------K----V----------Y--D----------I------A------SR-----T-----GV--		102
SMM.US.x.SME543	-A-.....---D-A----------I----------Y--D----------I---S--A------SR-----T-----A---		102
SMM.x.x.pE660.CG7G	-A-.....---D-A----------Y--D----------I---S--A------SR-----T-----D---		102
STM.US.89.STM_37_16	-TH-....---D----------I-Q---R-------S----Y--D----------K----------R----------AT---T-GV--		102

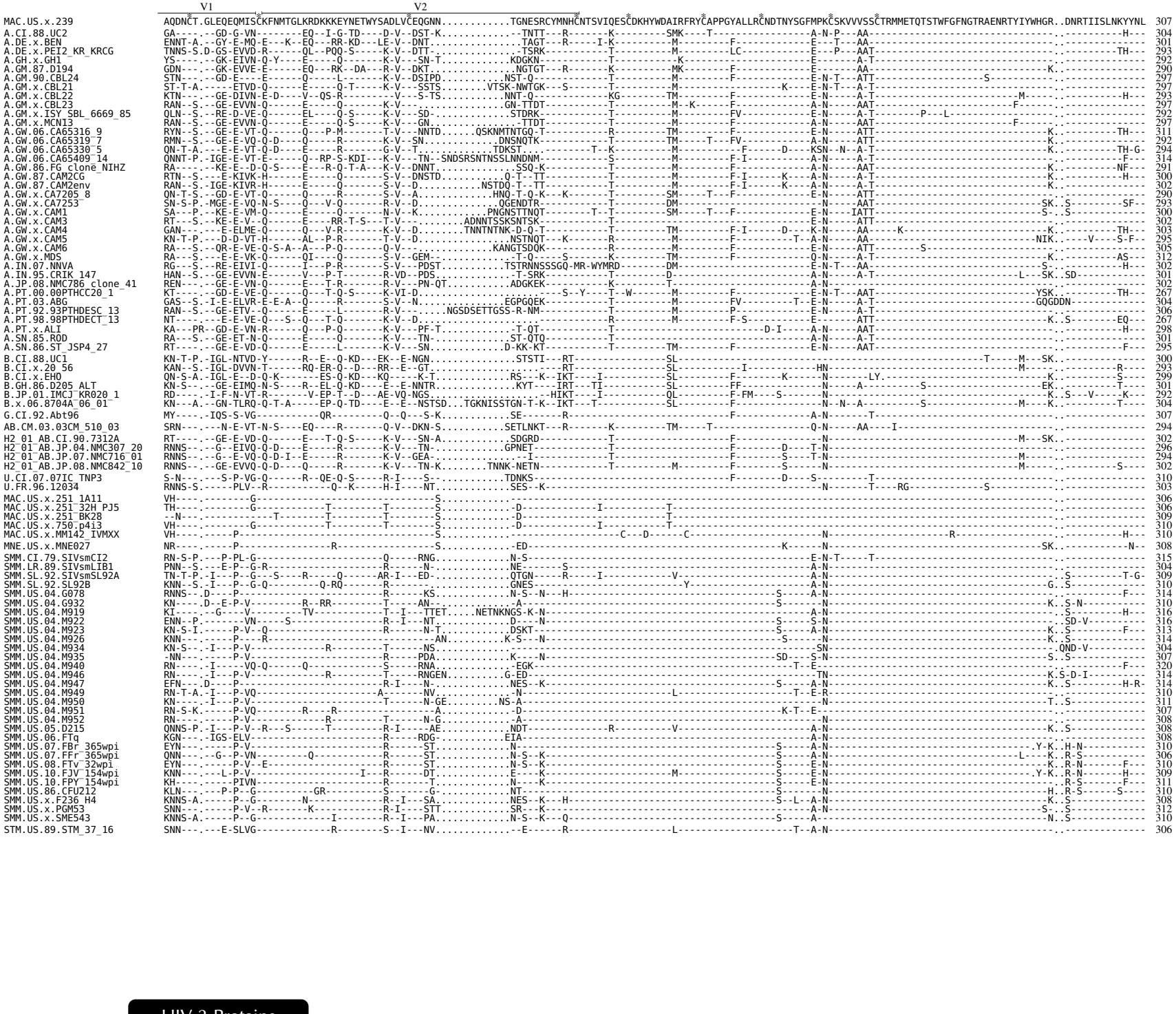
Tat start exon 1 end exon 2 start Tat end

```
MAC.US.x.239          METPLREQENSLESSNERSSCISEADASTPESANLGEEILSQLYRPLEACYNTCYCKKCCYHCQFCFLKKGLGICYEQSR.KRRRTPKKAKANTSSASNKPISNRTRHCQPEKAKKETVEKAVATAPGLGR.....*  130
A.CI.88.UC2           -----KAP-S--K-Y--P-P-T--WEVAAQ-L-KQ---L-A--H------T-P------SF---L---------W-VRKG.G------RT-THPP-TPD-S--IQ-GDSR-T-KQ---P-TP---TS-------------  131
A.DE.x.BEN            -----KAP-S--KPY--P---T--R-VTAQ-L-KQ---L-A--H----P-T-K----R-SF---L--S------S--RKG.R-----R-T-TPSP--PD-S--T--GDS--T-EQ-K-S-AT-V-TC---Q.----------  131
A.DE.x.PEI2_KR_KRCG   ----SKAP-S--M-C--P---T--Q-VKSQ-L-KQ--RL-----Q-----N-P-----------L---------W-ARKS.R-----R-T-THS----D-S--T--GNS-T--KQTK-P-T-LE--R---Q.----------  131
A.GH.x.GH1            ---H-KAP-S----Y--P---T--QGVTAQ-L-KQ---L----H------T-S----Q-SF---L---------W-ARKS.R-----R-T-THS----D-S--T--GDS--T-EQ-K-T-TTMV-TCS-----------  131
A.GM.87.D194          -----K-P-S----Y--P---T--R-VTAQ-R-KQ---L-A--H------T-S----Q-S----L---------W-ARQG.R-----R-T-THPPP--D-S--T--GDS--T-KQ-K-P-TT-VS-C---H.---------  131
A.GM.x.ISY_SBL_6669_85  -----KAP-S--G-Y--P--RT--Q-VA-Q-L--Q-----------T-N-K-F--G--F---L---N-------DRKG.R-----S---T--HS-P--D-S--T--GNS-T--KQ-K-LGTTLEAD-----------  131
A.GM.x.MCN13          -----KAP-S--M-Y--P---T--R-VGSQ-L-KQ---L----H----P-N-K----G--F---L---N-------DRKG.R-------T--HS----D-S--T--GNS----KQ-K-L--TTLE--R-----------  131
A.GM.x.MCR35          -----KAP-S--M-Y--P---T--R-VGSQ-L-KQ---L----H----P-N-K----G--F---L---N-------DRKG.R-------T--HS----D-S--T--GNS----KQ-K-L--TTLE--CS----------  131
A.GW.86.FG_clone_NIHZ -----KAP-S----C--P--RT--Q-VA-Q-L-RQ--------------T-S--------D--L---Q-----W-DRKG.R-------T--HP----D-S--T---NS----KQ-K-L-AT-E-DL------------  131
A.GW.87.CAM2CG        -----K-P-S--G-C--P--RT-GQ--T-Q-L-K-----------Q---E-D-S----R------L----------DRKG.R---------HS----D-S--T--NS--A-KQ-K-L-AT-E-D-------------  131
A.GW.x.MDS            --I--KAP------Y--P--HT--Q-VA-Q-L-RQ--------------N-N----R------I---N-----W--RKG.R---------THP----D-S--T--GNS-T--KQ-K-L-AT-E-DL-P-----------  131
A.IN.07.NNVA          -----KVP-S----C--PF-HT--Q-VA-Q--V-QE------FQ-----D-S-------F--EL---Q-------DRKG.R----R--T--HP--TPD-S--T---NS-T--EQ-K-L-TT-D-DL-P-----------  131
A.IN.95.CRIK_147      -----KAP-S----C--PFLHT--Q-VA-QG---Q--------------AS-----L------L---Q-------RKG.R-------T-THPTLT-D-S--T--WNS----E*-K-L-TT-V-DLDP-----------  130
A.JP.08.NMC786_clone_41 ---R-KA--S--I-YS-S---T-KQ--ADQRLIRQ---L-Y--H----T-N-K----Q------L---N-----W-DRKG.R-------T--HP----D-S--T--GNS-T--EQ-K-L-TTLE--C----------  131
A.PT.x.ALI            -----K-PGS--MPY--P---T--Q-VAVQ-L-KQ-------------T-N------E------L---N-----W-DRKG.R---S---I--HS----D-S--T---NS---EKQ-K-L-TTLG-DC-P--SHIYIS----  137
A.SN.85.ROD           -----KAP-S--K-C--PF-RT--Q-VA-Q-L-RQ-------------T-N-S----R------M---N---------RKG.R-------T-THP-PTPD-S--T--GDS--T-KQ-K---AT-E-DT-P-----------  131
A.SN.86.ST_JSP4_27    -----KAP-G--G-Y--P---T--Q--AAQGLVSP-D---Y---Q-----D-K----------M---N-----W--RKG.R-------T--HS----D-S--T--GNS----KQ-K-L-T-LE-IG-P-----------  131
B.CI.88.UC1           --I--Q---S--K--S-P--ST--PVVN-QG-D-Q--------------D-K----------L---------W-DH---.---.SS-R--VTA----DESL-AN-GDS--T-KQ-TK--TKGL-DL-P-----------  129
B.CI.x.20_56          --I-SQ------RF------ST-GEG---QGLD-Q--------------N---------F-----------W-DH---.---.SS-R--VTA---PDESL-TS-GDS--T-KQ-TK--EKGI-DL-P-----------  129
B.CI.x.EHO            --I--K---S--N--SGH--ST--GV-N-QGLD-R-----------K--S--------S----L-----------R---.---.SS-R--TT----P-ESL-A--GDS--T-KQ-KE--TTR--DL-P--SNTSTSRFAN-  139
B.GH.86.D205_ALT      --I--Q---S--K----P--ST--PVVN-QGLD-Q----------K--D-----------L----------DR--R-.SA-R--TTAP--PD-SL-A--GDS--T-KQ-KE--TTGT-D--P-----------  129
B.JP.01.IMCJ_KR020_1  -GI--Q-------F-S----ST--EG-N-RGLD-Q--------------R-K----------L---------DH---.---.SS-R--VTAPT---ESL-T-A-DG--A-KQ-KE--TTRT-D-----SDTSTS----  135
G.CI.92.Abt96         ----S-------K-CR-L---TF-ETVDA-GLEGTQA---Y----------S-K----------L---N----V----P-.R--.----T---SF----ES--A---NX--K-K*--K--TK---DL------------  129
AB.CM.03.03CM_510_03  --I--QG--S--K--S-P--ST--PVVD-QG-DSQ----------K--N-K----X------L---R----W-DH---.....-R--TTA-P-PD-S--T---DS--A-KQ-KE--TTGT-DL-P-----------  127
H2_01_AB.CI.90.7312A  --I--QG-----RF-S----ST---VVN-QGLD-Q-------H---R--N-------------L----------DR--.---.SS-RN-TA-----D-SV-T---NS--T-KQ-KE--TTG--DL-P--SNTSTS----  135
H2_01_AB.JP.04.NMC307_20 -----Q--G--SKF-K-H--ST---VVN-QGLD-Q----------G--N-R----A------L---N-------R-----.SS----TA-----D-SV-T-VGNS--T-KQ-KE--ATR--DL----SNTSTS----  135
H2_01_AB.JP.07.NMC716_01 -----Q--G--SKF-K-H--ST---VVN-QGLD-Q----------G--N-K----A------L---N-------R-----.SS----TA-----D-SV-T-VGNS--T-KQ-KE--ATR--DL----SSTSTS----  135
H2_01_AB.JP.08.NMC842_10 -----Q--G--SKF-K-H--ST---VVN-QGLD-Q----------G--N-K----A------L---N-------DR----.---.SS-R--TA-----D-S--TS-GDS--T-KQ-KE--ATR--D-----SSTSTS----  135
U.CI.07.07IC_TNP3     -----K---SLSK-CK-P-L-T--E-VRI---GDKE------------T---K----------L--------------.R-..S-----V--FP----S-PT---NS--K-KQ-----TTL-ST----K.----------  129
U.FR.96.12034         --I--K---S--K-CK-Q--ST--G-VG-QG--AA-------------S-K----------L------------P-.R-..S---V--YS-AP---SL--G--NS--KEK*-KAL-T----DL------------  128
MAC.US.x.17EC1        ------------------------------------------------------------------------------------------------------------------------------------------------  131
MAC.US.x.251_1A11     ----------SGH---T---A-----L-----------------------------------------------------------------------------------------------------------------  131
MAC.US.x.251_32H_PJ5  -------------------------T-----------------------------------------------------------------------------------------------------------------  131
MAC.US.x.251_BK28     --------------------L----T-----------------------------------------------------------L-P-----------------------------------------------------  131
MAC.US.x.MM142_IVMXX  --------------------Y----A-AI-------------------------------------S--K-H.R----------------E--P--I-L---K--------A-----------------------------  131
MNE.US.82.MNE_8       -----------K---G---T---A-P-L-----E-----------E-----------------------------------R-----------------D-------K-EQ-----A-----------------------  131
MNE.US.x.MNE027       -----------K---G----T---A-T-L-----E-----------------------------------------R-------V-------------D--K----K-EQ-----T----V-----------------  131
SMM.CI.79.SIVsmCI2    -----K--G-LSR-C--H-LST--Q-VF---LGTV-------------E-S-R----------L--------W--R--.R---A------Y-F----QS--T-D-NG--T-EQ-K---TT---DL---K.----------  130
SMM.LR.89.SIVsmLIB1   ----SK--GS--R-C--H---T--EAVG----L-SQE----W-------E-F-K----RS-------L--V------T-AR..-P-.-------H-PFT--QSVP-----S-SK-E*-KE-ATT---DL----------  127
SMM.SL.92.SIVsmSL92A  -----K----L-K-CR-PC-ST--EGVQ-Q-LDAEE-------------T-C-K-F-------L-----------PE-TR---AK--I-TDSFPS--*S-PT-S-N---K-E*TKK---T--VPD----K.----------  130
SMM.SL.92.SL92B       TGIQ*KG-GK*-RYYKRHC-YTLGM--PIQGLDSQ-DQ-PWD------T---S-F------F---L---R-----T-AKP-.R--.VK------PFDT--QS--SG--N---K-EQ--K--TE-D-DC----------  127
SMM.US.04.G078        ---------S--K-C--P---T--M-V----L-KAE----W-------K-C---------F-------R-----F--S---R-.-----------V--D*SL-E-A-N-P-KEE---K--T----DL----------  130
SMM.US.04.G932        ---------S--K--S----LT---AV-------E-----------E-S-----------------------Q-.--------I---F------L-T---NG--TEKE-K---AT---DL----------  131
SMM.US.04.M919        -----K----------G---YT--E-VL-----KVE----C---------F-----------L-----------Q-.R--------T--F----ESP-T-D-NR*-K-E---K--TE---DL----------  130
SMM.US.04.M922        -----K---------S----YT--E-VF-----KVE----W---------S-----------L-----------H-.R--------V----F----ESL-T-A-NR--K-E------TE--ADL----------  131
SMM.US.04.M923        -----K---S-----R-H--S---V--D-----S-E-----------------K---------H-----------Q-.R-.----T----F-----SL-R-A-DY--K-E------TE---DL----------  129
SMM.US.04.M926        ---------S--R-------YT--EAVL-----KVEG---C-----------------------L-----------PH-.R-------V----F----ESL-T-V-NS--TEK---K---TE---DL----------  131
SMM.US.04.M934        ---------S--R----H--YT--EAVL-----KVE----C---------S-A----------L-----------SH-.R-------V-I--F--P--SL-T-G-NR--KEE---K---TE---DL----------  131
SMM.US.04.M935        --I-WK--------------YT--E-VF-----KVE----W---------S-----------L-----------H-.R-------V----F-----SL-T-A-NR--K-E------TE--ADL----------  131
SMM.US.04.M940        ---------SL-K-CR-P---T--T-----G--KPE----W-------R-C---------F-------R-------S--.R-------------V--DQSL-E-A-NRP-K-EE--K--TT---D----------  131
SMM.US.04.M946        ---------S--R----H--YT--EAVL-----KVE----C---------S-A----------L-----------PH-.R-------V-I--F-----SL-T-G-N---KEE--KK--TE---DL----------  131
SMM.US.04.M947        -----K---S-----FR-H--S---V--A-------E----F-----------K----------H-----------Q-.R-.----T----F--P--SL-R-A-N-K-K-E------AE---DL----------  129
SMM.US.04.M949        --I--K---S--K-CK-L---T---P-VP--G--KP-----W-------E-C-----------L---R-------STK.R----Q-------F--D*S--V---NR-SKEKE-KK--T---D--R-----------  130
SMM.US.04.M950        ---------S--R----H---T--EAVL-----KVE----C---------F-----------L-----------PH-.R-------V-I--F-----SQ-T-D-N-----E------TE--ADL----------  131
SMM.US.04.M951        --I-------S--K-CK-H---T-KP----G--KP-----W-------E-C---------F-------R-------ST-.R-------------V--DQSL-E-A-NR---EK--KK--T---DL-R-----------  131
SMM.US.04.M952        ---------S--R----H--YT--EAVL-----KVE----C---L-----F-----------L-----------PH-.R-------V-T--F-T---SL-T-D-N---K-E------TK---DL-P-----------  131
SMM.US.05.D215        ----SK------R-F--L---T-G---N---L---E----W----------R---------Y-----------TAG.R---AK--ITPYSL-SR--S--TSS-DS--KEE*--K-AT----DL----------  130
SMM.US.06.FTq         ---------S--K-YK-H--ST--EA-PIQ---QIE----F---Q---T-C-----------L--------------.R-..----V-T-SF--P--S---T-D-NS--K-E*-K-L-TTL--DL----------  129
SMM.US.11.SIVsmE660_FL10 -----K---------CR-H--S---V-VP-------E----Y-------T---K----R-----H-----------H-.R-..----T---PLP----SL-T-A-NR--K-E---K--TE---DL----------  129
SMM.US.86.CFU212      ---RS--P-S--RY-S----ST--V-VP-Q--VTQE----W-------M-C-S----R------L---R----V--DRP-.R-..----V-T--LFT-D-S--T---NS----E*-KE-VST--SDL----------  128
SMM.US.x.F236_H4      -----K---------CR-H--S---V-VP-----------.----------K----R-----H-----------H-.R-..----T---PFP----SL-T-A-NR--K-E------TE---ADL----------  117
SMM.US.x.H9           -----K---S-----R-H--S---V--D-----S-E----------V-----K----------H-----------Q-.R-..----T----F---D-SL-X-A-N---K-E------XE---DL--X.----------  129
SMM.US.x.PBJA         -----K---S-----R-H--S---V--D-----S-E----------------K----R-----H-----------Q-.R-..----T--------D-SL-R-A-N---K-E------AE---DL----------  129
SMM.US.x.PGM53        -----K------G-------YT--E-VL-----KVE----C---------S-----------L-----------Q-.R-----------F--P-ESL-T-A-NS--K-E------TE---DL----------  131
SMM.US.x.SME543       -----K--------YR-H--S---V-VP-------E----------P---K----R-----H-----------H-.R-..----T-T-PLP----SL-T---NR--K-E---K--TE---ADL----------  129
SMM.x.x.pE660.CG7G    -----K--------CR-H--S---V-VP-------E----Y-------T---K----R-----H-----------H-.R-..----T---PLP----SL-T-A-NR--K-E---K--TE--ADL----------  129
STM.US.89.STM_37_16   -----K---S--R--S-P---T---V-A--GL--QE----W-------E-C-K-F---------L--VT-----T--R--.R-..VK----TYPI-----SL-T-A-NS--K-EQ-KE--TE-EST----K.----------  129
```

```
                              Rev start              exon 1 end exon 2 start
MAC.US.x.239                  MSNH.EREEELRKRLRLIHLLHQTNPYPTGPGTANQRRQRKRRWRRRWQQLLALADRIYSFPDPPTDTPLDLAIQQLQNLAIESIPDPPTNTPEALCDPTEDSRSPQD............ 107
A.CI.88.UC2                   -CEK.AD----QRE----R---------YR----G---N-R--R--Q-LR------KLHTA----A-SS--W---H--G-T-REL-----DL--S....DSNQGLAET............ 103
A.DE.x.BEN                    --ER.AD--G-QGK---LR---------Q-----S---N-R--R--Q-LR-V---NKLCAV------S---R---H--R-T-QEL-----DL--S....NSNQGLAET............ 103
A.DE.x.PEI2_KR_KRCG           -NGR.AD--G-QRKQ---R---------Q-L---R---N-R--RKQH-R--V---NS--T-----A-S---R---R--G-T-QEL------L--SSESTNNNQGLAETYNSLPAIWVRVDPRSAPGPCKDYERDSCERVERLVGGNGTDRQGNTCSSKKDQAGGRTCPPV 169
A.GH.x.GH1                    -HEK.ADG---QE-----R---------H-----S---N-R--Q----LR-V---KL-T-------S---R---D--R-T-HEL-----DL--S....NSNQGLAET............ 103
A.GM.87.D194                  -RDR.AD--G-QEK----R----I----H-Q---S---N-R--R--Q-FR-V---TKLHTI------S---R------G-T-QEL-----DL--S....NSNQGLAET............ 103
A.GM.x.ISY_SBL_6669_85        -TER.AD--GV-RK----R---------Q-L---R---N-R---E---K-I---------T-----A-P---QT-----G-T-QTL-----T..........Q-LAETQGSLPAVWVRVDPRSVPGPREGYKRDSYERGEELVGGSGTNRKGDTRSSTKDQAGSRNCPPV 158
A.GM.x.MCN13                  -TGR.AD-----RK-------------Q-L---S---N-R---KQ--R--V----K--T-----A-S--GRT--H--E-T-QDL-----SF--SSEGTDSNQ-LAKN............ 107
A.GM.x.MCR35                  -TGR.AD-----RK-------------Q-L---S---N-R---KQ--R--V----K--T-----A-S---RT--H--E-T-QDL-----SF--SSEGTDSNQ-LAKN............ 107
A.GW.86.FG_clone_NIHZ         -TER.AD--G-QRK----R---------Q-----S---N-R---KQ--R-I-----S--T-----A-S---R---H--G-T-QDL------L--SPESTNSNQ-LAEA............ 107
A.GW.87.CAM2CG                -TER.AD--G-QRK----R---------Q-----S---N-R----Q--R-I---------T-----A-P---RT--H--E-T-QDL-----HP--S.......Q-LAEA............ 100
A.GW.x.MDS                    -NER.AD--G-QRK----R---------Q-----R---N-R----Q--R-I-----S--T-----A-S---QTV----G-T-QDL-----HL--S.......Q-PAAT............ 100
A.IN.07.NNVA                  -TER.AG--GPERK----R---------Q-----R---N-R----Q--I-I-----S--T-----A-S-V-QT-----G-T-QEL-----HP--S.......Q-LAEA............ 100
A.IN.95.CRIK_147              -NER.AD--G-QRK----R---------Q-----S---N-R----Q--L-I-T---S--T-----ANS-V-QT-----G-T-QEL-----HP--S.......Q-LAET............ 100
A.JP.08.NMC786_clone_41       -TER.ADG---LRK----R---------Q-L---R---N-R----Q--R--V----K---------A-S---Q--RD--E-T-QDL-----............Q-LAES............ 96
A.PT.x.ALI                    -TER.AG--D-QRK----R---------Q---------N-R----Q--G-IV------FT-----ASS---R-V-H--G-T-QDL-----DL--SSESADNNQGLAET............ 107
A.SN.85.ROD                   -NER.AD--G-QRK----R---------Q-----S---N-R---KQ--R-I-----S--T-----A-S---QT--H--G-T-QEL-----HL--S.......Q-LAET............ 100
A.SN.86.ST_JSP4_27            -NER.AE-----RK----R---------Q-----S---N-R---KQ--R--V----K--T-----A-S--EQT--H--G-T-QEL------L--SSESIDSSQ-LAEI............ 107
B.CI.88.UC1                   -TTR.-KD..-Q-G---L----------QT----S---N-R---K--GL-I-------R-LS-S--EE-----V-R--E-TV-DL-N---S--T-...QAFTCIP-VWDQLVPRSNPSSNEGCERDSCEHRKSPMESSQKDSGSNHRDPQEDQTRT.............. 150
B.CI.x.20_56                  -TTR.-KD..-Q-G---L----------QA----S---N-R---K--GL-I-------R--S-S--EE-----V-R--E-T--DL-N---S--TT...QAFACIP-VWDQLAPRSNPSSNEGCGRDSCERGKSPMGSSQKNGGGNHRGPQENQTRT.............. 150
B.CI.x.EHO                    -NAR.--D..-Q-G---L----------Q-----S---N-R---KQ-GL-I-------HPL--S--EG-------R----I-KDL-N---S--T-...QASTCIP-IWDQLVPRSNPSSSQGCGRDSCERGEDLVGSPQESGRRDHCNTQEDQTRG.............. 150
B.GH.86.D205_ALT              -TAR.-GD..-Q-E---L----------Q-----S---N-R---K--GL-I-------H--STA-AEE--N----R----TV-DL-N--P-LNQSPTTQAPGCVP-VWDQLVPRSAPSGSKGYGRNSCECRRDLMGGSQESGESNHRDPQENQTRT.............. 153
B.JP.01.IMCJ_KR020_1          -TTR.-KD..-Q-G---LR---------Q-----S---N-R------GL-I-----Q-HPL--S--EE--N-------K-TV-DL-T---SI-T-...Q-STCIP-IWDRLVPRSNPSSDEGCGRDSYKHRKGPMGGSQKNSEGNRRDPQEDQTRTRTRPLVRDAVLQEH 164
G.CI.92.Abt96                 ---L.-E-.----------F--------Q---------N-R-K-KQ--L-I-----------X--ARE---------S-T-QDX--S--AV-KN-......QDP-SN............ 100
AB.CM.03.03CM_510_03          -TTR.---....-E---L----------H---------N-R---K--GL-I-------R--SAA-AEE--N----R----TV-DL-N------TT...QAPVCIP-VWDRLVPRSAPSSSEGYERDSCECGRDLMGSSQGGGESDPRDPQKNQTRT.............. 148
H2_01_AB.CI.90.7312A          -TAR.-KD..-Q-GI--L----------Q-----S---N-R---K--GL-I-------HPL-NS-AEE------RR----TV-DL-N---SS-TT...QAPVCVP-IWDRLVPRSSPSSSGGYGRDSCEHREDLMGGSQEDGEGNHRSPQKNQTGA.............. 150
H2_01_AB.JP.04.NMC307_20      -NAR.-KD..-Q-----L----------Q-S---S---N-R---K--GL-I-------HPL-NS-AEE-----VRR-------DL-N---SS-TT...QAPARTP-VWDQLAPRSSPSRREGCGRDSYERQQDHMGGSQEDGEGNHRNPQKNQTRT.............. 150
H2_01_AB.JP.07.NMC716_01      -NAR.-KD..-Q-----L----------Q-S---S---N-R---K--GL-I------VHPL-NS-AEE------RR-------DL-N---SS-TT...QAPARTP-VWDQLAPRSSPSRCEGCGRDSCERQQDLMGGSQKDGEGNRRNPQENQTRT.............. 150
H2_01_AB.JP.08.NMC842_10      -TAR.-KD..-Q-G---L----------RA--------N-R---K--GL-I------VHPL-NS-AEE------RR-------DL-N----S-TT...QAPARTP-IWDQLAPRSSPSRCESCGRDSYERHQDHMGGSQKDGEGNHRNPQKNQTRT.............. 150
U.CI.07.07IC_TNP3             -N-P.-E-..V-R------F--------R---------N-R----Q--L------N--S------A-S--------F-S---AEL-E--SGL--S...TDSNQ-TSEA............ 102
U.FR.96.12034                 ---P.-EG..VQR-------------------------N-R----Q--L-I-----K--T-------S-V--------G---QDL-E--AVV--H.......Q-TSAP............ 98
MAC.US.x.17EC1                -N--.--------------------------------------------------------------------------------------------------------............ 107
MAC.US.x.251_1A11             --S-.-------------------------------------R--------------------------------------------------K-------............ 107
MAC.US.x.251_32H_PJ5          --S-.-------------------------------------R--------------------------------------------G--N------............ 107
MAC.US.x.251_BK28             --S-.-------------------S-----------------R--------------------------------------------------KG-----............ 107
MAC.US.x.MM142_IVMXX          -RS-.TG-----R-----------------S-S------KR----Q-----------------------------------------I---------N-----A............ 107
MNE.US.82.MNE_8               --S-.AE-----R-------------------------N-R----Q---------------------N-----V------------------I--I-H----NP-----............ 107
MNE.US.x.MNE027               --S-.AE-----R-------------------S------N-R----Q----F------V---------------------------------I--I-H----SP-----............ 107
SMM.CI.79.SIVsmCI2            --DQ..EG-----------F--------Q-T-------N-R----Q--L-I----N----------ANSS----V----G-S--DL-E--R-L-TS-SSSQ-QD...-A............ 103
SMM.LR.89.SIVsmLIB1           -HED....R--QR-------------------------N-R----Q--L-I----------------A-----V-V----S---QDL-E--ATV--T-RSN-TSQ--N............ 102
SMM.SL.92.SIVsmSL92A          --QRGQE------------F--------Q---------N-R---KQ--Y-I----N---------A--SF--------S-T--DL-E-----I--T-KPDKQ-P............ 103
SMM.SL.92.SL92B               -Q-P...-----R-------I-------A-----S---N-R---K---I-IV-------------A--DI------FDS-S-Q-L-E---TV--T-R-QSA-............ 99
SMM.US.04.G078                --LQ.GE-----R---------------KE----H---R-R----Q--L-I--------------VSS-----V----S----DL-E----V--T-R-S--NP-VH-A............ 107
SMM.US.04.G932                --SN.---G---R------F------C-Q---------R-R----Q----I--------------A-SAF----R---------DL-----SV----QSAA-NP-DH-N............ 107
SMM.US.04.M919                --SS.AG-----R------F------H-Q-T-------R-R----Q----I--------------A-S-I--------D-SL-N--N---SV----H-I--SP--H-N............ 107
SMM.US.04.M922                ---I.AG-----R------F--------Q---------R-R---KQ----I--------------V-S----------D-SL-H-----AS---T-L-S--S---R--............ 107
SMM.US.04.M923                --SN...-----R------F--------D----T----R-R---KQ----I--------------VN-----------G----EL-N--AS---P-K-AA-SP*-H-S............ 104
SMM.US.04.M926                --HI.AG-----R------F--------Q-S-------R-RK---Q----I--------------A-------V----S-SL-NL----ASV--T-H-I--SP--H--............ 107
SMM.US.04.M934                -NHI.AG-----R------F--------Q-A-------R-R----Q----I--------------I-S-----V----D-SL-N------GV--T-R-A--SP--H--............ 107
SMM.US.04.M935                ---I.AG-----R------F------C-Q---------R-R----Q----I-------H------V-S----------D-SL-H-----AS---T-L-S--S---R--............ 107
SMM.US.04.M940                --LQ.GEG----R---------------KE----H---R-R----Q--L-I--------T--A----S-----V----S----DL-E--AGV--T-R-SADNQ--H-A............ 107
SMM.US.04.M946                -NHI.AG-----R------F--------Q-A-------R-R----Q----I------------------S-----V----D-SL-N------SA--T-RGTS-SP--H--............ 107
SMM.US.04.M947                --SN...D----R------F------S-E---------R-R----Q----I-----------N--A-------V----R----EL-----SA--P-K-T--SP*-H-N............ 104
SMM.US.04.M949                --LQ.KE-----R---------------SE--------R-R----Q--L-I--V--------T--A-S-----V----S----DL-E--ASV--T-RGS-DT---H-A............ 107
SMM.US.04.M950                --HI.AG-----R------F------N-Q-T-------R-R----Q----I-------H---N--A-S----------D-SL-N-----ASV--T-R-T--SP--H--............ 107
SMM.US.04.M951                --LQ.GE-----R---------------KE--------R-R----Q--L-I--V---VS---A---SS---------------DL-E---SV--T-R-S--NP--H-T............ 107
SMM.US.04.M952                -NHI.AG-----R------F--------Q-T-------R-R----Q----I--------------I-S----------D-SL-N------SV--T--ST--SP--H--............ 107
SMM.US.05.D215                -RQQ.-E-------------F--A-----QA-----R--N-R----Q--L-I------VS-LST--VA-----------S---TDL----ASV--D-RSSPDSP-DH--............ 107
SMM.US.06.FTq                 ---P..G-----R------F-----L--Q-T-------N-R----Q--L-I--------------A-S-I---VR---------QL-E--ASA-GD-QSVA-SP--H-T............ 106
SMM.US.11.SIVsmE660_FL10      --ST...------------F--------Q---------R-R----Q----I--------------------------G----DL-----S---P-N-VAKSP............ 100
SMM.US.86.CFU212              -TDQ...------------F--------Q---------N-R---Y---H-I--------------VHS-----V----------DL-E---SA--T-QCAASNKGDH-N............ 105
SMM.US.x.F236_H4              --ST...-----R------F--------Q---------R-R----Q----I--------------V------------G----EL-----SA--P-N-VAKSP*-H-N............ 104
SMM.US.x.H9                   --SN...-----R------F-------XX---------R-R----Q----I--------------A------------G----XL-N--ASA--P-K-AA-SP*-H-S............ 104
SMM.US.x.PBJA                 --SN...-----R---------------D---------R-R----Q----I--------------V------------R----EL-N--ASA--P-K-AA-SP............ 100
SMM.US.x.PGM53                --SS.AG-----R------F--------Q---------R-R----Q----I--------------A-S-----V----D-SL-N------SV---.....NP--R--............ 104
SMM.US.x.SME543               --ST...------------F--------Q---------R-R----Q----I-----------N---------------G----DL-----SA--T-K-AAKS-*-H-N............ 104
SMM.x.x.pE660.CG7G            --ST...-----R------F--------Q---------R-R----Q----I-----------N--A------------G----DL-----SA--P-N-VAKSP............ 100
STM.US.89.STM_37_16           --DQ...-----------QF--------Q---------N-R----Q--N------N---------AS------V----G-S-QDL----P-L-KD-Q-TA-N............ 99
```

	Rev end	
MAC.US.x.239	.....*	107
A.CI.88.UC2	.....-	104
A.DE.x.BEN	.....-	104
A.DE.x.PEI2_KR_KRCG	RGSGINRETL-	180
A.GH.x.GH1	.....-	104
A.GM.87.D194	.....-	104
A.GM.x.ISY_SBL_6669_85	RDRDISKETL-	169
A.GM.x.MCN13	.....-	108
A.GM.x.MCR35	.....-	108
A.GW.86.FG_clone_NIHZ	.....-	108
A.GW.87.CAM2CG	.....-	101
A.GW.x.MDS	.....-	101
A.IN.07.NNVA	.....-	101
A.IN.95.CRIK_147	.....-	101
A.JP.08.NMC786_clone_41	.....-	97
A.PT.x.ALI	.....-	108
A.SN.85.ROD	.....-	101
A.SN.86.ST_JSP4_27	.....-	108
B.CI.88.UC1	.....-	151
B.CI.x.20_56	.....-	151
B.CI.x.EHO	.....-	151
B.GH.86.D205_ALT	.....-	154
B.JP.01.IMCJ_KR020_1	KGEGNRV...-	172
G.CI.92.Abt96	.....-	101
AB.CM.03.03CM_510_03	.....-	149
H2_01_AB.CI.90.7312A	.....-	151
H2_01_AB.JP.04.NMC307_20	.....-	151
H2_01_AB.JP.07.NMC716_01	.....-	151
H2_01_AB.JP.08.NMC842_10	.....-	151
U.CI.07.07IC_TNP3	.....-	103
U.FR.96.12034	.....-	99
MAC.US.x.17EC1	.....-	108
MAC.US.x.251_1A11	.....-	108
MAC.US.x.251_32H_PJ5	.....-	108
MAC.US.x.251_BK28	.....-	108
MAC.US.x.MM142_IVMXX	.....-	108
MNE.US.82.MNE_8	.....-	108
MNE.US.x.MNE027	.....-	108
SMM.CI.79.SIVsmCI2	.....-	104
SMM.LR.89.SIVsmLIB1	.....-	103
SMM.SL.92.SIVsmSL92A	.....-	104
SMM.SL.92.SL92B	.....-	100
SMM.US.04.G078	.....-	108
SMM.US.04.G932	.....-	108
SMM.US.04.M919	.....-	108
SMM.US.04.M922	.....-	108
SMM.US.04.M923	.....-	105
SMM.US.04.M926	.....-	108
SMM.US.04.M934	.....-	108
SMM.US.04.M935	.....-	108
SMM.US.04.M940	.....-	108
SMM.US.04.M946	.....-	108
SMM.US.04.M947	.....-	105
SMM.US.04.M949	.....-	108
SMM.US.04.M950	.....-	108
SMM.US.04.M951	.....-	108
SMM.US.04.M952	.....-	108
SMM.US.05.D215	.....-	108
SMM.US.06.FTq	.....-	107
SMM.US.11.SIVsmE660_FL10	.....-	101
SMM.US.86.CFU212	.....-	106
SMM.US.x.F236_H4	.....-	105
SMM.US.x.H9	.....-	105
SMM.US.x.PBJA	.....-	101
SMM.US.x.PGM53	.....-	105
SMM.US.x.SME543	.....-	105
SMM.x.x.pE660.CG7G	.....-	101
STM.US.89.STM_37_16	.....-	100

```
                               signal peptide end gp120 start
                         Env start                                                                                                                                      V1
MAC.US.x.239                 MGCLGNQLLIAILLLSVYGIYCT...LYVTVFYGVPAWRNATIPLFCATKNRDTWGTTQCLPDNGDYSEVALNVTESFDAW.NNTVTEQAIEDVWQLFETSIKPCVKLSPLCITMRCNKSETDRWGLTKSITTTASTT..............STTASAKVDMVNETSSCI  152
A.CI.88.UC2                  -EPGR----AV---T-ACL---K...Q---------V----S---------------I------D--Q-IP-----A----.D-----------R-----------T---VA-N--PVTGNNTNA-AKP-AARP-..................-NPSYLTII--S-T-V  149
A.DE.x.BEN                   -EPGR---FVV---T-ACLV--S...Q-------I---K--S---------------I------D--Q-II-----A----.---------V----H-----------T---VA-N-SRVQGNTTTPNPRTSSSTTS................RPPTSAASII----N--  150
A.DE.x.PEI2_KR_KRCG          -.DSR---IV----T-ACL---A...Q-------I---K--S-------R-------I------D--Q-IP-----A----.---------V----N-----V-----T---VQ-E--STS-ESSNS-SEGS-.........................VPEIL---T---  140
A.GH.x.GH1                   -.-GKSL-CV-S--A-A-LV---...Q---------V----S---------------I--K---D--Q-IT-----A----.D-------V----S-----------T---VA-S--STTNNTTTTGSTTG..........................MSEI----P.S  137
A.GM.87.D194                 -EPGR----V----T-ACL---K...Q-------I------S---------------I------D--Q-IT-----A----.D-----------R-----------T---VA-N--ITSGTTATPSPPN...........................ITIID-N-T--  138
A.GM.90.CBL24                ....#----VTT--A-ACL-S---...K-------I------S---------------I------D--Q-I------A----.D-------V----S-----------T---VA-K--NTA-TQR-S-TTS----T--GAM....................EI----P-V  141
A.GM.x.CBL21                 ....#-L--V-S--A-ICL-----...Q-----------------S---------------I------D--Q-I------A----.D-------V----S-----------T---V--N-TTMKNISTTARPTTGISNT.......................ESILK-N----  138
A.GM.x.CBL22                 ....#I---V----T-ACLV----...Q-----------------S---------------I------D--Q-LT-----A----.D-----------H-----------T---VA-N-T-TRRNMTIAQSIT--AKAVV.........................I-D-DP--  137
A.GM.x.CBL23                 -MGGR----V----A-TCL-----...N-------I------S---------------I------D--Q-IT-----A----.D-----------N-----------T---VA----NTDARNTTTPT......................TA-PRTIKP-TEIS-N----  145
A.GM.x.ISY_SBL_6669_85       -.SGKI---V-F--T-ACL-----...K---------V-K--S---------------I------D--Q-IP-----A----.D-I-----V----N-----------T---V--N--A-TESAVAT-SPSG...........................P--I-D-DP--  138
A.GM.x.MCNI3                 -MGGR----V----T-TCL-----...N-------I------S---------------I------D--Q-IT-----A----.D-----------N-----------T---VA----NTDARNTTTPTTASPRT.....................IKP-TEIS-N----  145
A.GW.06.CA65316_9            ....#---------T-A-L-H---...K-------I------S-----------------------D--Q-I------A----.D-----------N-----------T---VA-T-SSTSARSTTTPTTTI--RTSNTNT...........TLNTTTSKNLLEDN-T--  150
A.GW.06.CA65319_7            ....#S---VTL--A-ACL-----...Q-------I------S---------------I------D--Q-I------A----.--------------------------T---VA-N-SDTS-STITN-NTTIDNTT........................EIEI-D--P-V  138
A.GW.06.CA65330_5            -.GGR----VT---A-A-L---A...Q---------V-K--S---------------I------D--H-IS-----A----.D-----------N--D---L-----T---VA-N-SN-TGTGNTT-SGAR.....................N-NTTTPLI--K--A--  144
A.GW.06.CA65409_14           -.DGR----F----A-TCL-----...Q-------I------S---------------I------D--Q-IT-----A----.D--I--------N-----------T---VQ-A--NT-S-NT-NNS.................TKVATT-KSTNDSKEL-N--P--  147
A.GW.86.FG_clone_NIHZ        -KGSK-------V-A-A-L-H-K...QF------I------S---------------I------D--Q-IT-----A----.---------V----N-----------T---VA-N-TRNM-TWT-R-DTQN..........................ITII-D.T-HA  138
A.GW.87.CAM2CG               -ERGR---------A-ACL---R..QQ-----------K--S---------------I------D--Q-IP-----A----.D--I--------N-----------T---VA-K--I-TS-TTMIRTTTPS-........................--EAPISDN-P--  144
A.GW.87.CAM2env              ---GR---------A-ACL---K..QQ-----------K--S---------------I------D--Q-IP-----A----.D--I--------N-----------T---VA-K--I-TS-TTMI..................RTTTPS--KEAPISDKFP-P--  146
A.GW.x.CA7205_8              -.GSKA---A----T-A-L-----...Q-------I---K--S---------------I------D--Q-I------A----.---------V----S-----------T---VE-S-TNTS-GSTTTPLNT..........................TLI---NP--  136
A.GW.x.CA7253                -TSEKT--------A-TCLL--K...Q-----------------S---------------I------D--Q-I------A----.-----------D--KR-----------T---VA-N-TNVTSTANTTIT.......................I-STNMI--DS-P-A  140
A.GW.x.CAM1                  -AYKR---------T-ACL---K.RKQ-------I------S---------------I------D--Q-IP-----A----.--------------N-----------T---VA-K-DGNI-STGNT-T-TRAR.......................ITSEEIK-N----  145
A.GW.x.CAM3                  -TRK----------A-A-L---R..QQ-------I-----------------------I------D--Q-IS-----A----.--------------S-----------T---VA-K--NTQSSNTATE.......................-PAK-MEINEI----P-M  143
A.GW.x.CAM4                  -A-GKS--------T-TCL---A...Q-------I------S---------------I------D--Q-IP-----A----.---I--------N-----------T---VA-K-KRMTGNTTTA.....................TTA---LLNDTTLL--S-T--  144
A.GW.x.CAM5                  -T-GR---------T-A-LVH-K...Q---------V----S---------------I------D--Q-I------A---N-----D--V---N-----------T----K-E-K-LGNATTLSN-T........................--S-TANLAIDDNTP--  144
A.GW.x.CAM6                  --RKR-------I-T-ACI---K...Q-------I------S---------------I------D--Q-IP-----A----.D-------V----N-----------T---VA-N-SR-RPSNSTGNVTT-IPS...............TTAA-TTPGMTVI----P-M  151
A.GW.x.MDS                   -TRKMH--------T-ACL---K.PQQ-------I------S---------------I------D--E-I------A----.D-------V----S-----------T---VA-N-SRLSDSASTRNTTTNA--A--TI...........RA--IAPRNTTIS-N-P--  157
A.IN.07.NNVA                 -TRERA---VV---T-TCL---K..QQ-------I---K--S-----------------------D--Q-I------A----.D-------V----N-----T-----T---VA-E-KNR--GNATT-APT--PE.......................INETTEIS-NT--V  146
A.IN.95.CRIK_147             -AHE-T-----F--T-ACL---K...Q-------I---K--S-------R-------I------D--Q-IV---A-A----.D-----------N-----T-----T---V-----T-T-TTTTAPT-TSAGST........................-TP-PI----NT--M  147
A.JP.08.NMC786_clone_41      -.-GKIL-IV-S--A-ACLV--K.KQQ---------V----S---------------I------N--Q-I------A----.D-----------S-----------T---VA-S--RTSNPTTPAGTTKS.....................FPTPGENKTI--NDT--  146
A.PT.00.00PTHCC20_1          -.GDR----V----A-ACL-H-K...E-------M---K--S-------R-------I------D--Q-IP---S-A--G-.DS----------N-----------T---VA---T..........................................I-------  121
A.PT.03.ABG                  -AHKM---------T-ACL---K...Q--A----I------S---------------I------D--Q-IT-----A----.--------------N-----------T---VA-N-TRNF-ASRNS-ATRN--TPS-P................KLGENLELK-N-T--  150
A.PT.92.93PTHDESC_13         -A-GRK----T--IA-A-LV--SSSKQ-------I---K--S---------------I------D--Q-IT-----A----.--------------N-----------T---VA-T--STRSNNTESNT........................NTT-EAPEI-DN----  144
A.PT.98.98PTHDECT_13         -.GDR----V----A-TCL---N...Q-------I---K--S----------G----I------D--Q-IT-----A----.---------V----N-----------T---VA---M..........................................I--N----  121
A.PT.x.ALI                   -MSSR----VT---A-ACLV--K...Q-----------K--S---------------I------D--Q-I------A----.D-------V----R-----------T----A-K-SNIS-ESTTTSP-PGS.........................TLKPLI--SDP--  142
A.SN.85.ROD                  -M...---------A-ACLV----...Q---------T-K---------R-------I------D--Q-IT-----A----.--------------H-----------T---VA-K-SST-SSTGNN-T-KS-STT-.....................--PTDQEQEIS-DTP-A  147
A.SN.86.ST_JSP4_27           -.-GR---FV-S--A-ACL---V...Q---------V----S---------------I------D--Q-I------A----.---------V----S-----------T---VA----STTAKNTTS-PTT-.......................TTANTTIG-N----  141
B.CI.88.UC1                  -AHTS-H-F-LL--I----FLGH.KKN-------I-------V------T-------V---------T-ISV-I--A----.---------VD---S-----------T---VA----NTG-NTTTKPITTPI-T.....................TKPSENLL-D--P--  148
B.CI.x.20_56                 -AHINSH---SL--I----CM-K...Q-------I-------V------A-------V-----S---T-ISV-I--A----.---------VD---N-----L-----T---VA-N-STNN-RTNNT-A-T--GN.......................-TTPIV---AIP-V  145
B.CI.x.EHO                   -AHVN-Y--VTL--I-I--YMGK...NF------I---K--S-------R-------V------D--T-IQ--I--A----.D----D--TK---S-----------T---V--K---TWSSASKE-TTSSA.....................SLR-STQTLL--D-K--  146
B.GH.86.D205_ALT             -AYFSSR-P--L--IGIS-FV-K...Q-------I-------V--I---T-------V---------T-IR--I--A----.D----Q--VD---R-----------T---VA-N-S-T--NPGNASSTT--KPT.......................-T-RGLKTI---DP--  148
B.JP.01.IMCJ_KR020_1         -AHINKY-FA-L--I-I--FMGE...DF------I---K---V---------I---V------D--D-I---I--A----.D--------VA---N-----------T---V--N-SNLN-TANTT-ANNI-A.......................-KNIS-L-GND---  144
B.x.06.8704A_06_01           -AYSSSR----L--I-I---V-R...Q-I-----I---K---V--L---T-------V------N--T-IK--I--A----.------Q--VD---R-----------T---VA-W-F-NG-NGTNQERTNI-.........................STII---DP--  140
G.CI.92.Abt96                -AY--------T----T--F----...K-------I------SV--------------------D----L------A----.D-------V----N-----T-----T----A-----T--NK----GTT-V-PA-V..............P--KMVTAEL--S--Q-L  152
AB.CM.03.03CM_510_03         -AYSSSC-P-TL----IC-FT-K...Q-I-----I---X--S---------------I------D--Q-IT-----A----.X-----------R-----------T---VA-N-TTT-ATPQNT-TRN...........................TTIIEDNDP--  138
H2_01_AB.CI.90.7312A         -.-GK-L-FV-S--A-A-L-----...K---------V----S---------------I------D--Q-I------A----.---------V----S-----------T---VA-S--STTATTTPPSTTNN--TT..................EPTTGGPEI---FP-M  147
H2_01_AB.JP.04.NMC307_20     -.-GK-L-FV-S--A-ACL--S-..KK---------V----S---------------I------D--Q-I------A----.D-------V----S-----------T---VA-N-TNTTSNTNTP-P-NI..........................EGPTLI----P--  141
H2_01_AB.JP.07.NMC716_01     -.-GK-L-FV-S--A-ACL--SA..KQ---------V----S---------------I------D--Q-I------A----.D-------V----N-----------T---VA-N-TNTTSNTNTP-SPNI..........................EGPTLI----P--  141
H2_01_AB.JP.08.NMC842_10     -.WGK-L-F--S--A-ACL--S-..KK---------V----S---------------I------D--Q-I------A----.---------V----N-----------T---VA-N-TNIASNTNTSASPP-.......................SPTTEGLI----P--  143
U.CI.07.07IC_TNP3            -AY--------L-C-NACV-----...Q-------I-------V--------------------D----L---I--A----.D-----------N-----------T----A-----T---K----GTVA--T-P-N............TTANITGRSEEI-G-DP--  154
U.FR.96.12034                -A---------L-C-NACV-----...Q-------I--------------N---------------------L------A----.---------V----S-----------T----A---S-T--N-----GTAA--T-P.................AMQESSKVI---EP--  149
MAC.US.x.251_1A11            -----------------------...-----------------------------------------------L-----------.E--------------------------T-------------------S---TAAP...............KAM-E-IN-I-------  151
MAC.US.x.251_32H_PJ5         -----------------------...Q-------------------E-----------------------L-----------.E-----------------------------------------K------L----P-A...............P--AS-I---------  151
MAC.US.x.251_BK28            -----------------------...Q-------------------------------------------L-----------.E------------------------------------------------S--ITTAAP............T-APV-E-I---------  154
MAC.US.x.750_p4i3            ----R------------------...Q-------------------------------------------L-----------.E-----------------------------------------K------S-I-TTAPT...........TPN--STKSI---------  155
MAC.US.x.MM142_IVMXX         ----------------------I...Q-----------------------------------D----L-----------.E-----------------------------------------K------S-------TT...........TTAKSVETR-I-----P-V  155
MNE.US.x.MNE027              -------------F--A-----I...Q-------------------V-R---------------D----L---I-------.E-----------H-----------T------K-------K------S----AP-T.............K--TTKEIEV---N-T-V  153
SMM.CI.79.SIVsmCI2           -E---------V--R--C--W---...Q-------I------SV---------------------Q-IRM---------.D-----------N-----------T----A-----T---K----GNT-AAPT--TTTTA.....VTATAV-TT-GG-II-D-TE--  161
SMM.LR.89.SIVsmLIB1          -A------------V-A-------...Q---------------------------------D----ISI----A----.D-----------N-----------T----A-K---N--N-----RTS---P--..................--ETNI--I-----P-M  150
SMM.SL.92.SIVsmSL92A         ----------------ASE-F---...Q-------------------Q--------------D----L-I---------.D-----------N-----------T----A-Q-DRN--R-----GGTA--TAA-TK...........A---LTVP-KEL-D--F-M  155
SMM.SL.92.SL92B              -A-P-LH---D--F---L-TW-A...Q---I---I------------Q------V----------L------A----.D-----------N-----------T----A-K---N--------RAA---S-P-TTS.........PLTAASPSGEEI--D-M--T  157
SMM.US.04.G078               --------------V-A-------...Q----------------------------V------D----L-I---A----.D-----------N-----------T----A-K---TD--K----GTT---QK--TTGP......ASKAKSATPVA-EV-TDS-P-V  160
SMM.US.04.G932               --------------I-A-----I...Q---I----------V---------------------D----L-----Q----.--------------N-----------T----A------D-------GVS--KNP-IVA...........PTSRTA-LKPEL---DP--  155
SMM.US.04.M919               ----------------A-------...Q-----------K---------------------D----L-I---R----.--------------N-----------T------K---T---Q----GITP-LST--T............Q---TTNTTEVI--S---L  154
SMM.US.04.M922               ----------------A-------...Q-----------K---------------------D----L-I---R----.D-----------N-----------T------K---T---K-----NV--ARNA-ATTPA....TTKTPA-VITVTPNV---NDT-V  162
SMM.US.04.M923               -----------L----AL----V...Q-----------K---V------------------------L-I---A----.--------------N-----------T----A-----N--------RTTAS-TIQKSTT.........STTP-IQAVAEKV--DSTP--  158
SMM.US.04.M926               ----------------A-------...Q-----------K---------Q------------D----L-I---R----.--------------N-----------T------K---T--------RHG--ATTP-TAA.........KST--EKVTNVE--D----V  157
SMM.US.04.M934               --------------I-A-------...Q-----------K---------------------D----L-I---R----.--------------N-----------T----A-----T---K----GTPSP-TAP..................-MQ-T-AEI-------  149
SMM.US.04.M935               ---------V------A-------...Q-----------K---------------------D----L-I---R----.--------------N-----------T------K---T--------GNT---TT-SS.............A-ATAVTPKI---SDP-V  153
SMM.US.04.M940               ---P-SR----L--V-A-------...Q-----------------------------------D----L-I---A----.D---------------------------T----A-K--RTQ-NK----RDTE-GNT-ATTPTTTPSPTTRNTTGNTTSRNEEIRANDT--  166
SMM.US.04.M946               --------------I-A-------...Q-----------K---------------------D----L-I---R----.D-----------N-----------T----A-----T---K----RGPS--TPSSTV...........KTTAKTRVP--EL------V  156
SMM.US.04.M947               --W--------L----AL----V...Q-------------V-----------------------D----L-I---A----.D-----------N-----------T----A------D-------GTR--MKP--TEAS......TGTA-P-TPITAKVL--SDP--  160
SMM.US.04.M949               -----------L--V-A-------...Q---------------------------V----D----L-I---A----.D-----------H-----------T----A----GT--NK----GTAI--PTP-TS..........TTKNRTTTQDRDLEDN-T--  156
SMM.US.04.M950               ----------------A-------...Q-------I-------V------------------D----L-I---R----.--------------N-----------T----A-----TD--K----GTPS--TTAST.............T-A-TPAAVEL---D---  153
SMM.US.04.M951               -----------L--V-A-------...Q-----------------------------I------D----L-I---A----.D---------------------------T-----------T---K----ATT-PATTIA..............TSSLTTTPETLTGNET--  152
SMM.US.04.M952               --Y------T----I-A--T--A...Q-------I---K----------------------D----L-I---R----.D-----------N-----------T----A-----T-----R--GTPS--TAK-P.............T--PT-VAVEI------V  153
SMM.US.05.D215               --------------V-AC--W---...Q-------------------S-------------------MPI----A----.D-----------N---S-------T----A------D-------GVQA--PTAST................---TTKAIET-TDN-A-V  153
SMM.US.06.FTq                -----------S----ACS---K...K-------I-------V-----------V------D----L-------A----.E-------V----N--K-------T----A----RT--N-----GNSD--TV-P...............-AKTTTEKGEI-D-DP-F  152
SMM.US.07.FBr_365wpi         ----------------A-------...Q-------I------SV--------------------D----L-----------.D-----------N-----------T----A-----T---K----GNA---TP-PTS.........SIAT--KPETAKVI---N---  157
SMM.US.07.FFr_365wpi         ----------------A-------...H-----------------V--------------------D----L-----------.D-----------N-----------T-----------T---K----RAA-S-TT..............PLA--P--T-KVLAKNDT--  152
SMM.US.08.FTv_32wpi          ----------------A-------...Q-------I------SV---V----------------N----L-----------.D-----------N-----------T----A-----T---K----GNA-A-TP-PT..........TSST--KPETALAI--NN---  156
SMM.US.10.FJV_154wpi         ----------------A-------...Q-------I------SV--------------------D----L-----------.D-----------N-----------T----A-----T---K----GNA---TP-PT..........SSTT--KPETAKVI--S----  156
SMM.US.10.FPY_154wpi         ----------------A-----I...Q-------I------SV--------------------D----L-----------.D-----------N-----------T----A---D-T---K----GNAIA-TP--TI.........SSKT--KPEIAKVI--S----  157
SMM.US.86.CFU212             --------------V-A-----A...Q-------------------E--------------D----L-I---A----.--------------N-----------T----A-----A---K----GTAAPPTTK-LK...........TTS-TKPP-LE--D-EP-V  155
SMM.US.x.F236_H4             -----------L--V--LE-C-V...Q-----------K---------------------D----L-I---A----.D-----------N-----------T----A-----T-----------GNAG--TTAIT............TTA-P-VAENVI--SNP--  154
SMM.US.x.PGM53               ----------------A-------...Q-----------K---------------------D----L-I---R----.--------------N-----------T-----------T---------NEIP-TT--TTT.........KSPKAEAITAKVI--SDP--  157
SMM.US.x.SME543              -----------L--V--LE-C-V...Q-----------K---------R------------D----L-V-I--A----.--------------N-----------T----A-----T-----------GRAE--TTAKST..........TST--TTVTPKVI--GD---  156
STM.US.89.STM_37_16          -A-P-----------ACLT----...Q------------------------------------------L-I----A----.D-----------N-----------T----A-----N---K----GKTV--VTP-...............AAA-AT-PEL-------V  152
```



HIV-2 Proteins

```
                                       V3 tip
                                     V3                                                                                                 V4
MAC.US.x.239                TMKCRRPGNKTVLPVTIMSGLVFHS..QP...INDRPKQAWCWFGGKWKDAIKEVKQTIVKHPRYT...G.TNNTDKINLT.APG.GGDPEVTFMWTNCRGEFLYCKMNWFLNWVEDRNTANQKP.....KEQHKRNYVPCHIRQIINTWHKVGKNVYLPPREGDLTCNS  461
A.CI.88.UC2                 --H-K-------V-I-L---HR---..-A..V--KK-R------K-N--G-MQ-----LAG----K...-.--D-S---FV.K--V-S-----Y---------F--N-T-------N-TSQK............Q---A---------------QY--------E-----  453
A.DE.x.BEN                  --R-K---------I-L--------.-..--T--R----R---R-RE-MQ-----L-Q----K...-.I-D-G---F-.K--A-S----A-------------N-T--------K-QT.............R---..---K--------------------E-A-E-  446
A.DE.x.PEI2_KR_KRCG         --H-K-----S---I-LR--R-----.R-..I--E-----------D--K-MQ-----L------R...-.--D-Q---F-.Q--K-S-A--VY------------N-TR----I-N-AHP.............Q---A------------R--Q-I-------E-V---  441
A.GH.x.GH1                  SIH-K-------V-I-L--------.-..--T--R------K---RE-MQ-----LI-----K...-.--D-KN--F-.K--R-S----AY------------N-T-------N-PNQT............QH--A-------------------------Q-----  440
A.GM.87.D194                --H-K-------V-I-L---RR---.R-..VY-KK-G------Q-N-IE-MR-----LA-----G...-.--D-G---F-.K--I-S-----Y------------N-T-------NKTNQT............HG--A---------------T---------E-----  439
A.GM.90.CBL24               -IH-K-------V-I-L--------.-..--N--R------K----E-MQ-----LA-----K...-.--D-S---FI.G--K-S----AY---------F--N-T-------N-TNTT............WH----------------------------E-S---  445
A.GM.x.CBL21                -IH-K-------V-I-L---HR---.R-..V--RK-M------E---RE-MQ-----L------R...-.--K-EN-TFA.G--E-S----AY------------N-T-------NTTNQT............QH----------------------------L-S---  446
A.GM.x.CBL22                -IH-K-------V-I-L---YK---.R-..VL-TK-R------Q-R--E-MQ-----LA-----K...-.----ES-TF-.---R-S-----Y------------N-T-------NKTGQT............Q--------K--------------------E-S---  442
A.GM.x.CBL23                -IH-K-------V-I-L----R---.-..V--K--R------K-E--G-MQ---E-LA-----K...-.--E-KN--F-.---K-S----AY------------N-T-----I-NKT...............H-----------------------------E-----  443
A.GM.x.ISY_SBL_6669_85      -IL----E----V-I-L---RR---..-K..I--KK-R----R-K-E-RE-MQ-----L------K...-.--D-N---F-.--EKDS----AY------------N-T-------NKTGQ.............QH-------E--------------------E-S-E-  440
A.GM.x.MCNI3                -IH-K-------V-I-L----R---.-..V--K--R------K-E--G-MQ---E-LA-----K...-.--E-KN--F-.---K-S----AY------------N-T-----I-NKTHR..............-----------------------------E-----  443
A.GW.06.CA65316_9           -IY-K-------V-I-L--------.-..--T--R------K-E--G-MQ---E-LK-----K...E.N-D-KN-TF-.---K-S-A--AY------------N-T-------N-TNQT............RH-------K--------------------E-----  459
A.GW.06.CA65319_7           --H-K-------V-I-L----I---.-..--K--R------K-N--E-MQ-----LA-----R...-.--Y-GN-TFA.---K-S----AY------------N-T-------NKSHP.............RY--A----K--------------------E---D-  439
A.GW.06.CA65330_5           --Y-K-------V-I-L----I---.-..--K--R------E-D--G-MQ-----LAN--K-K...-.K-ITD-HFE.---A-S----LY------------N-T-----L-N-TEQV............WH---------VV--------Y--------E-----  442
A.GW.06.CA65409_14          --H-K-------E-I-LL-------..-A..--T--R------E-E-RK-ME---E-LA-----K...-.--A-KN-TFA.---I-S----AY------------N-T-------NKTSKV............Q---A----K--------------------E-----  462
A.GW.86.FG_clone_NIHZ       --H-K---------I-F---FK---.-..V--KK-R------E-Q--E-MQ---E-LA-----K...-NRSR-EN-KFK.---R-S-----Y--------S---N-T-------N-TGQK............Q---A--R---------R----L-------E-----  441
A.GW.87.CAM2CG              S-Y---------V-I-L---QR---.R-..I--K--R------K-N-TE-MQ-----LAE----K...-.-K-ITD-TFK.--ER-S-----Y--S------F--N-T-------NKPNTT............----A-------------------------E-----  449
A.GW.87.CAM2env             S-Y---------V-I-L---QR---.R-..I--K--R------K-N-TE-MQ-----LAE----K...-.-K-ITD-TFK.--ER-S-----Y--S------F--N-T-------NKPNTT............----A-------------------------E-----  451
A.GW.x.CA7205_8             --H-K-------A-I-L--------.-..--R--R------K-E--K-MQ-----LA-----K...-.-KD-S--RF-.---R-S-----Y--A-----F--N-T-L---I-NGSQP.............Q---A----K----------RH--------E-----  437
A.GW.x.CA7253               SLH-K-------V-I-L--------.-..--T--R------E-E--E-MQ--RA-A-------...-.--D-G-------.---R-S----SY------------N-T-------N-TGI.............Q---A-------------------------E-V---  440
A.GW.x.CAM1                 -VH---------V-I-L--------.-..--R--R------K-N-TE-MQ-----L------R...-.--D-GN--F-.---K-S----AY---------F--N-T-------NKTGE.............-H-------K--------------------E-----  447
A.GW.x.CAM3                 SLH-K-------V-I-L--------.-..--T--R------K----E-MQ---E-LA-----K...-.--D-K---FA.---R-S----AY------------N-T-------N-TENRT...........WH----------------------------E-----  451
A.GW.x.CAM4                 SVH-K-------V-I-L--------.-..--M--R------K-E--G-M------LLE----K...-.AT-INN-R-A.E--K-S----AY------------N-T-------N-TGGP............QH----------------------------E-V---  451
A.GW.x.CAM5                 --H-K-------V-I-L--------.-..--K--R------K-N-RE-ME-----V------R...-.EPNKTV-FI.---K-S----AY---------F--N-T-------NKTGQE............QH--A----K----A----E-R--------E-I---  442
A.GW.x.CAM6                 -LH-K-------V-I-L--------.-..--R--------K-Q-RE-ME-----LKE----K...K.F-V-NN-TFE.---R-S----AY------------N-T-------N-TNQT............Q--------K----------R---------E-----  453
A.GW.x.MDS                  SLH-K-------V-I-L--------.-..--R--R------K-N-RR-MQ---E--I-----K...-.-S-ITN-TF-.---K-S----AY------------N-T-------N-TDKP............WH-------K----------R---------E-V---  460
A.IN.07.NNVA                SIY-K-------V-I-H-P------.-..--T--R------K-R--E-MQ------AE----R...-.--I-E--KF-.---K-S----AY------------N-T-----I-N-TNET............-H-------K--------A-RH--------E---D-  450
A.IN.95.CRIK_147            SIH-K-------T-I-L---YK---.R-..V--T--------K-R---MQ--K-VAE---.A...V.-KDPKN-TFA.---K-S----EY------------D-T---D-I-N-P-RK...........AW--------------------H--------E-----  450
A.JP.08.NMC786_clone_41     SIH-K-------M-I-LL--AR---.-..V--TK--------K-E--K-LE---E-LKE---R...-.P-D-KQ-RFI.K--E-S-A--KY------------N-D------GESKGK.............V--------S---R------I---------E-V---  450
A.PT.00.00PTHCC20_1         --H-K-----S-I-I-LG--MR--G..-RITS--KK--*-----E-N-TE-MR------M-------...-.V-D-EN-T--.E--K-S-A--SY------------N-T----LI-N-KGP.............----A---------------Q--------EE----  416
A.PT.03.ABG                 --H-K-------V---L--------.-..--T--R------K-Q--E-MA-----LI-----M...-.--D-T---F-.T--R-S----AY-R-------F--N-T-----I-NKTQP............Q--------K-------------------ES----  451
A.PT.92.93PTHDESC_13        -LHRK-------T-I-L--------.-..--T--R------R----E-MQ---E-LA-----K...-.-T--T---F-.---K-S----EY------------N-K-----I-NKTASG--..........WH----------V-----------------E-----  456
A.PT.98.98PTHDECT_13        --R---------V-I-L----I---.-..--K---------Q-N-TA-M------LA-------...-.N---EN-KFR.E--K-S----AY---------F--N-T-----I-N-TGQ.............QH--A----K-V--A------H--------E-----  414
A.PT.x.ALI                  --H-K-------V-I-L----I---.-..--K--R------K-E-RK-MQ---E-L------K...-.--D-NQ--F-.K--R-S-A--VY----------H-N-T-------NKTGQE............QH--A----K----I---A-----------E-----  446
A.SN.85.ROD                 SLH-K-----I-KQIML---H----HY--.-..--K--R------K-----MQ---E-LA----R...-.--D-RN-SFA.---K-S----AY------------N-T-----I-NKTHR...............--A----K----------R---------E-S---  448
A.SN.86.ST_JSP4_27          -VH-K-------V-I-L--------.-..--R--R------K-E--E-M----L-LA-----K...-.--D-E--RFI.---ERS----AY------------N-T-------N-TNQT............QH-------K--------------------Q-----  443
B.CI.88.UC1                 --H---------I-I------N---.-...L-T--R------K-N-IE--R---E--I-----K...-.----ER-R-V.G-SA-S----RH---------F--N-T-------N-TGTT............QK---T---K--V--------Y--------T-S---  448
B.CI.x.20_56                --H-----------I----RR---.R-..V--E--R------E-N-TE-MR---E-VM-------...-.IK-IT----V.G-SA-S--ARY---------F--N-T-------GK-GTK.............-----------V--------Y--------L-S---  441
B.CI.x.EHO                  --H-K-----M-V-IRTV--IL---.-..--K---------K-N-TE--Q---E--KN----S...-.-T-ISQ-R-A.EHARSS----RY------------N-TF------N-TGL.............----AS------V-----I-R---------E-S---  446
B.GH.86.D205_ALT            SIH-K-------V-IRTV---L---.-..--K--R------K-N-TE------R--I-----K...-GAK-ITSVK-VSEH-K-S---T-Y------------N-T-------NKTNTT............R---A-----------------I-------E-S---  451
B.JP.01.IMCJ_KR020_1        -IH---------V--------I---.-..--Q--R--------N-TQ------E--A---------...-.--D-A--K-A.EHSQ-S----RY------------N-T-------N-TGI.............-K---------V------------------H-----  439
B.x.06.8704A_06_01          SIH-K-------V-I-V----I---.-..--K--R------K-N-TE------E--I----HK...-GAK-IAN-S-VS-H-K-S-----Y------------N-T-----I-NKTNTT............Q---------------------I-------E-S---  454
G.CI.92.Abt96               -IX-----------I-----------.-..--E--R-----------E-MQ---E-V----S-K...-.--D-KX-TF-.T--E-S----K-----------------T-----I----MTLL--.....Q-RQR-----------------------------------  462
AB.CM.03.03CM_510_03        --H-K-------V-I-L-P------.-..--R--R------R-N-TE-M------L------K...-.--D-T---F-.K--E-S----AY------------N-T-------N-TGT.............QH--A----K--------------------E-----  441
H2_01_AB.CI.90.7312A        -IH-K-------V-I-L--------.-..--K--R------K-E-RE-MQ-----LI-----K...-.--D-RN-TF-.K--T-S----AY------------N-T-------N-TGQT............QH--A----K--------------------Q-----  450
H2_01_AB.JP.04.NMC307_20    --H-K-------V-I-L--------.-..--T--R------K----E-MQ-----LI-----N...-.NR-IS----K.K--E-S----AY------------N-T-------N-TNQT............Q---A----K--------------------E-----  444
H2_01_AB.JP.07.NMC716_01    --H-K-------V-I-L--------.-..--N--R------K---RE-MQ-----LI--S---...-.-K-ISLV-F-.K--E-S----AY------------N-T-------NKTNQT............Q---A----K--------------------E-----  442
H2_01_AB.JP.08.NMC842_10    --H-K-------V-I-L--------.-..--N--R------K---RE-MQ-----LI-----N...-.NK-IS---F-.K--T-S----AY------------N-T-------N-TGQ.............QH--A----K--------------------E-----  449
U.CI.07.07IC_TNP3           -VR----------------------.-..--E----------S--E--Q---E-L----------...-.I-D-K-----.Q--G-.---A----------------------------GENDT-DQNI..R-R------------V----NR------------------  467
U.FR.96.12034               --Y--------------------..-A..--NN---------N--G--Q---E--AN---H...-.K-ISQ---A.E-A.---S--K---------------T-------N-SMEG-TS.....R-R------------V--------R---------------  457
MAC.US.x.251_1A11           -I---------------Q-----.-..-------------R-----------------------...-.-----------.---.------------------------------------------.....---------------------------------I-  460
MAC.US.x.251_32H_PJ5        -------------------------.-..---------------N------M------------...-.-----------.--R.------------------------------DVT--R-......--R-R-----------------------------------  460
MAC.US.x.251_BK28           -------------------------.-..------------------------------------...-.-----------.---.------------------------------DVTT-R-......--R-R-----------------------------------  463
MAC.US.x.750_p4i3           -------------------------.-...V-E--N--------N--------------------...-.I----------.--R.------------------------------D.....VTTQRP--R-R-----------------------------------  464
MAC.US.x.MM142_IVMXX        -------------------A-----.-...V-E-------R---N--E-----------------...-.-----------.--R.------------------------------SLTT---......--R-------------------------------------  464
MNE.US.x.MNE027             -------------------------.-..--E-------R-E-N--E-----------------...-.-----------.--R.-----------------------------K-LTGTTQK...PQ------------------------------------  464
SMM.CI.79.SIVsmCI2          --R----------------------.-..--E-------------E-L----E-L------K...-G--D-R---F-.E--RKS----A--------------T-------E--MTE-N-.....E-KKRH--A-----------------------------  471
SMM.LR.89.SIVsmLIB1         -IR----------------------.-..--E----------S-RK--Q---E-L-------...-.--D-N-----.---.-------------------------------D-NHARWTSQTS--RK-K---------V-----R------------------  463
SMM.SL.92.SIVsmSL92A        ELH----------------------.-..--E---------K-E--K--R---E---Q-----...-.--R-E--Y--.---.------------------------------LTLPRWKNQKPE-KA----------V-----R-------------Q---  468
SMM.SL.92.SL92B             -I-----------------------.-..--E--R---------RE-MQ--K-----------...-.--D-R-----.---.------------------------------SSPRWTTQTK-----------------R------------------  469
SMM.US.04.G078              -IR----------------------.-..--E----------S--K--Q--E-----------...-.--D-E-----.--R.-------------------------------EKGSIWEQQNK---K-------------------------------  473
SMM.US.04.G932              -------------------------.-..--E-------R---D-RG-M----E--I------...-.--E-KN-T--.P--.-----------------------W-----NFGNTDPRWKQQNSR-K-----------V-----R------------N-E---  469
SMM.US.04.M919              --R----------------------.-..--E---------K-N-TE------E----------...-.--D-KQ----.---.-----------------------------D---ETGSRWMQQTQ--RK--------------------------------  475
SMM.US.04.M922              -IR----------I-----------.-..--E----------S--Q--Q---E--------N...-.--DIK-----.---.-------------------------------DIKGSRWKQQNV--RK-----------V--------H-------------  475
SMM.US.04.M923              --R----------------------.-..--E-------R---D--G--R---D-L-------...-.--E-K-----.--R.------------------------------IDITKDRWTRQNP--R---------------------------------  472
SMM.US.04.M926              -IR----------------------.-..--E--R----R---S--K--Q---E---------...-.-----K-----.---.-------------------------------ISHPRWTQQKKR-RE-----------V-----R----------------  473
SMM.US.04.M934              --R----------------------.-..--E--R--------S--G------E-L-N---N.ES-.-KDISQV---.---.-------------------------------ETG-RWTDQKS--RK---------------------------------  465
SMM.US.04.M935              -IR----------------------.-..--E--R----------Q--Q---E--------K...G--D-K-----.---.--------------------------------SAPRWKQQKK--RE-----------V-----R----------------  467
SMM.US.04.M940              -IR----------------------.-..--E--R--------E--G------E-LI----NNKT-..-K-EE-Y--.---.--------------------------------SHPRWTQQKK-------------------------------S-N-E---  482
SMM.US.04.M946              --R----------------------.-..--E--R--------S--G------E-L-N---Q...-.-K-ISQV---.---.-------------------------------ETSLRWLKQNP--RK--------------------------------  473
SMM.US.04.M947              -------------------------.-..--E--R----R-E--RE--Q---E-L--------...-.---------Y--.S--.------------------------------TKGYRWSSQN..K-DK------------V-----R---------------T-  471
SMM.US.04.M949              -IR----------------------.-..--E-------L-R-D--K--Q---E-L-------...-.--D-KR----.---.--------------------------------SS--K.....R--L-K--------V-------------S-N-E---  464
SMM.US.04.M950              -IA----------------------.-..--E--R----R------E------E--------K...-.REIKN----.---.-------------------------------QTHSRWKQQKP--RQ-K---------------------------------  470
SMM.US.04.M951              --R----------------------.-..--E--R------R-D--G--Q---E-L-R---D...-.--E-KN----.---.-------------------------------LTHL-KHGAKL--L-------------------------S-N-E-T-  466
SMM.US.04.M952              -I-----------------------.-..--E--R--------S-RK--E---E-L------H...-.--D-R--Y--.---.------------------------------YKNSSRWKQQKAR-KE-------------------------------T-  467
SMM.US.05.D215              -IR----------I---------A.-..--E--R------------Q------E-L-------...-.--D-S--K--.K--.----------------------------E--ATEMNARNKNNR-KK-H--------------R-----------------  467
SMM.US.06.FTq               -I---------M-------------.-..-----R----------S-MQ-----L-------...-.--D-KN--F-.E--KTS----A--------------T----------MSS-NK.....A--KR---------------------E-----  463
SMM.US.07.FBr_365wpi        --R-----------GT---FR--L..-Q...L-R--R----R------E--Q---E----------...-.----NE-Y--.---.--------------------------------SHPRWTEQTPG-R--------------------IH-------------  469
SMM.US.07.FFr_365wpi        --R--------I--I--R--R-Y--..-L..IR-EK------R-E-S--E--R---E----------...-.----KR----.--A.-----------------------------K--SHYRWRNQNPRDRY-------------------I--------------  466
SMM.US.08.FTv_32wpi         --R-----------GT---FR--L..-Q...--KK-R----R-E----E--Q---E----------...-.--D-E--Y--.---.--------------------------------SQSRWKNQTPG-R--------------------IH-------------  469
SMM.US.10.FJV_154wpi        --R--------I---GT---FR--L..-Q...--R--R----R-E-N--G------EA---------...-.--D-N--Y--.---.-------------------------------ASHPRWAGQTPG-K--------------------IH-------------  468
SMM.US.10.FPY_154wpi        --R--------I---GT---FR--L..-Q..I--KK-R----R-E-G--E--Q---E----------...-.--D----Y--.---..R---------------------------N---SHPRWTEQTPG-KY-------------------I--------------  470
SMM.US.86.CFU212            -IR----------------------.-..--E--R-----------S------E-L-------...-.--D-S-----.---.-------------------------------RTQMEARNNND--RR-----------V-----R----------------  469
SMM.US.x.F236_H4            --R----E-----------------.-..--E---------E-S--K--Q---E-L-------...-.--D-R-----.--A.------------------------------DQKGGRWKQQNR---Q-K---------------------------------  467
SMM.US.x.PGM53              -IR-----------------I----.-..--E----------S-RE--Q---E----------...-.--E-K-----.---.--------------------------------SGSLWTQQTKADEK----------------R-----------------  471
SMM.US.x.SME543             -------------------------.-..--E-------R---N-SE--Q---E-L-------...-.--D-R-----.--A.------------------------------DQNSNRWKQQKKP--Q------------------------------------  469
STM.US.89.STM_37_16         --S----------------------.-..--E-----------E-RG------E-L-------...-.--D-A--RIV.---.--------------------------I-N-S-SEMRDWNKNK---Q-----------V---------------Q-------  465
```

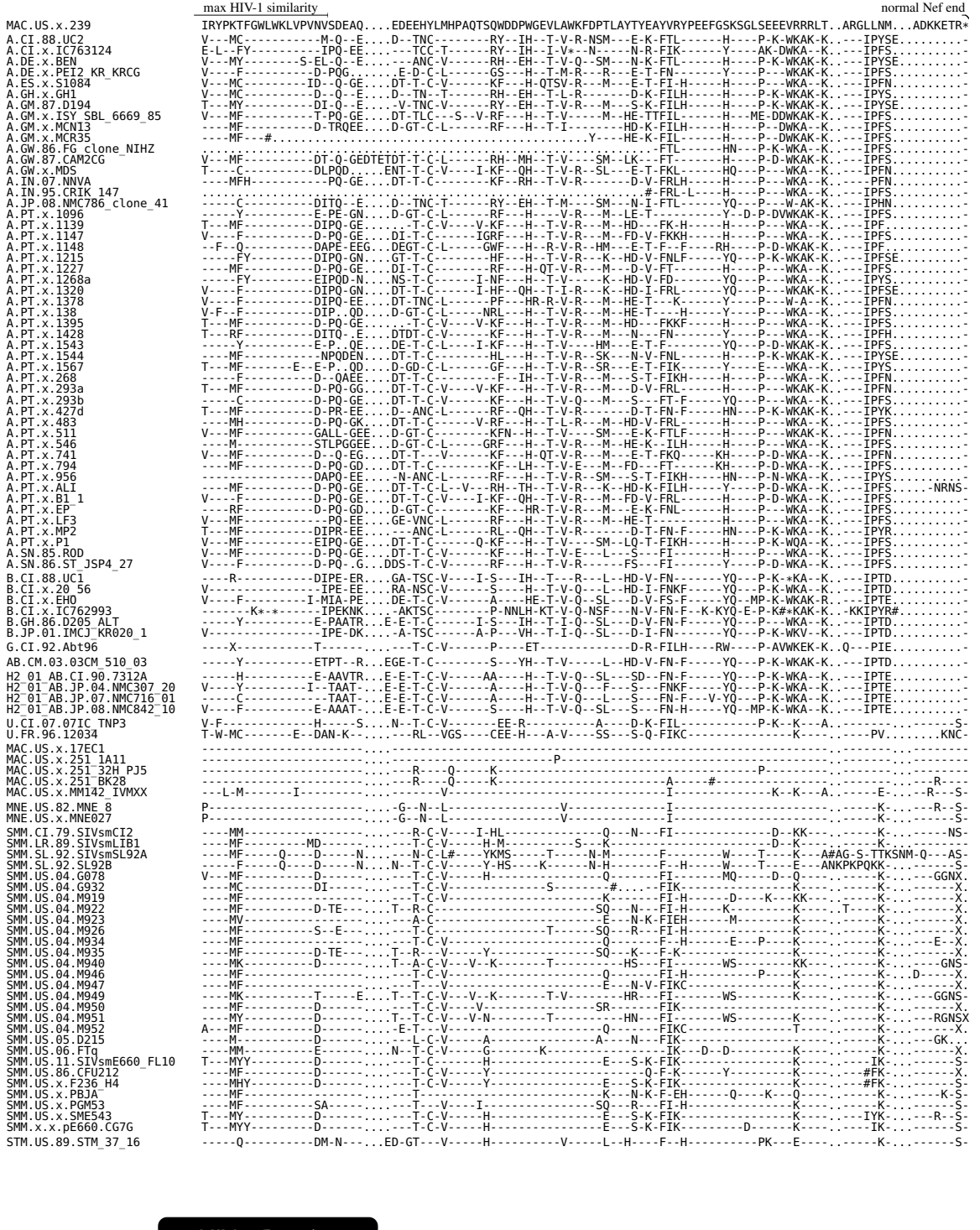
```
                               V5                                                gp120 end gp41 start
MAC.US.x.239                TVTSLIANIDW...IDGNQTNITMSAEVAELYRLELGDYKLVEITPIGLAPTDVKRYTTGG.TSRNKRGVFVLGFLGFLATAGSAMGAASLTLTAQSRTLLAGIVQQQQQLLDVVKRQQELLRLTVWGTKNLQTRVTAIEKYLKDQAQLNAWGCAFRQVCHTTVPW....  623
A.CI.88.UC2                 ----I-----T...---------F-----------------I------F---SE---SSA..PA----------L------------------S-------------------I-------------------A----------------S---T-----------....  613
A.DE.x.BEN                  ----I-----I..DKNRTH----F-----------------I------F----QR--SST..PV------------------------R----S--------------------------M------------A----------H-----S---------------....  608
A.DE.x.PEI2_KR_KRCG         ----I-----M....FD---S--F------------------------F---SE---SSA..PQ----------V------------------S-HPGLYW-------------------M----------------------R---R--S---------Y---L-....  601
A.GH.x.GH1                  ----I-----V...NS-------F-----------------I-V----F---RE---SSA..PV-----------------------------S--------------------------M------------A----------------S---------------....  600
A.GM.87.D194                ----I-----S...---------F-----------------I-V---PF---KE---SSA..PV-----------------------G-----S--------------------------M------------A----------------S---------------....  599
A.GM.90.CBL24               ----I-----V...EN-------F------------------------F---SEQ--LSS..PK---------------T---A---------S--------------------------M------------A-------------R--S---------------....  605
A.GM.x.CBL21                ----I-----T...Y-----D--F------------------------F---S----SSA..PG------P--------T---A---------S--------------------------M------------A----------------S---------------....  606
A.GM.x.CBL22                ----IL----V...N--------F------------------------F---TE---SST..PM------M----------------------S--------------------------M------------A----------------S---------------....  602
A.GM.x.CBL23                ----I-----AN....--N-D--F------------------------F---AE---SST..PM-----------------------------S--------------------------M------------A----------------S---------------....  603
A.GM.x.ISY_SBL_6669_85      ----I-----V...DGD-R----F-------------------V----F---AE---SSA..PG-H----L--------T---A---------S------FR------------------M------------A---------A---R--S---------------....  601
A.GM.x.MCNI3                ----I-----A....N--N----F------------------------F---AE---SST..PM-----------------------------S--------------------------M------------A----------------S---------------....  603
A.GW.06.CA65316_9           ----I-----V..DENNT-----F------------------------F---SER--SST..PV------I------------A---T-----S--------------------------M-----------------------------S---------------....  621
A.GW.06.CA65319_7           K---I-----V....K--E----F------------------------F---S----SSA..HG-HT------------T---A---------S--------------------------M------------A-------------R--S---------------....  599
A.GW.06.CA65330_5           ---------NEDK..-K-K----F----G------------I------F---EE---SST..PV-------M-------T---V---T-----S---------------------------------------A----------------S---------------....  604
A.GW.06.CA65409_14          --A-I-----VMG....-R----F------------------------F---------SA..HQ-P-------------TM------------S---------------------------------------A------R---------S-----------S---....  622
A.GW.86.FG_clone_NIHZ       ----I-----A.....-D-----F---A--------------------F---S----SSA..HQ-HT--------------------------S--------------------------M------------A----------------S-----------S---....  600
A.GW.87.CAM2CG              ----I-----E...R-NQT----F--D---------------------F---SQ---SPA..HG-P---A---------T---V---T-----S------------------------------------I--A----------------S---------------....  610
A.GW.87.CAM2env             ----I-----ER...-NQT----F--D---------------------F---SQ---SPA..HG-P---A---------T---V---T-----S---------------------------------------A----------------S---------------....  612
A.GW.x.CA7205_8             ----I-----IR....E------F------------------------F---KER--SST..QG---------------T---A---------S--------------------------M------------A-------------L--S---S-----------....  597
A.GW.x.CA7253               -------D---I...-K-G----F----S-------------------F---TE---SPN..HG-P---------------------------S--------------------------M------------A---------G---R--S---------------....  601
A.GW.x.CAM1                 ----I-----TDM..N-TE----F------------------------F---TE---SSA..HG-Q-------------T---V---T-----S---------------------------------------A-------------L--S---------------....  609
A.GW.x.CAM3                 ----I-----VE....A------F------------------------F---SE---SSA..HK-HT--------------------------S--------------------------M-----------------------------S---------------....  611
A.GW.x.CAM4                 ----------MR....NDS----F-----------------I-V----F----Q---SPA..PT-S--------------A--T---T-----S---------------------E----M------------A----------------S---V-----------....  611
A.GW.x.CAM5                 ----I----ETNM..HKT-----F-----D--K---------------F---SE---SSA..PG-----------------R-----------S--------------------------M------------A-------------R--S---------------....  604
A.GW.x.CAM6                 ----I----NV....-NQTS---F----G-------------------F---EQ---SSD..HG-Q-------------T---A---------S--------------------------M------------A----------------S---------------....  613
A.GW.x.MDS                  ----I-----D...KD-------F--D---------------------F---PE---SSA..HG-H-------------T---A---------S--------------------------M------------A----------------S---------------....  620
A.IN.07.NNVA                ----I----EE.......---V-F--G--D------------------F---PE---SFA..SR-QT---------------D----------S--P-----------------------M------------A----------A----DS---------------....  608
A.IN.95.CRIK_147            ----I-----VN.E--K.R----F---------V---------V----F---SE---SS-#.-G-HT---L--E-------------------S--P-----------------------M------------A----------------S---------------....  612
A.JP.08.NMC786_clone_41     ----I-----M....GK-S----F-------------------V----F---SQ---SST..PG------L--------T---A---------S----------M---------------M------------A----------------S---------------....  610
A.PT.00.00PTHCC20_1         -----------........----F------------------------F---SQ---SSA..PG-G---------------------------S--------------------------M-----------------------G--R--S---------------....  572
A.PT.03.ABG                 ----I----EVRNASQE-----AF-------------------V----F---SER--SSA..HR-QT------------T---T--------PS--------------------------M------------A----------------S---------------....  615
A.PT.92.93PTHDESC_13        ----I----EMTK....------F------------------------F---AE---SSA..HR-QT--------------------------S--------------------------M------------A----------------S---------------....  616
A.PT.98.98PTHDECT_13        ----S------........----F------------------------F---PQ---SSA..PG-G---------------------------S--------------------------M------------A----------------S---------------....  570
A.PT.x.ALI                  ----------T...---------F------------------------F---SER--SST..PR-----------------------T-A---S--------------------------M------------A-------------R--S---------------....  606
A.SN.85.ROD                 ----I------...QNN------F------------------------F---KE---SSA..HG-HT-------------------------VS---------------------------------------A---------Q---R--S---------------....  609
A.SN.86.ST_JSP4_27          ----I-----G...GE-------F-----------------I-V----F---P----SSA..PV---------------T---A---------S--------------------------M------------A----------------S---------------....  603
B.CI.88.UC1                 S---------VYYDGNDTK---------G-------------------F---EI---SST..-P------M------L--M-------T----S-----------------------------------------------------L--S---------------....  612
B.CI.x.20_56                ----I----E-...--S-E---------G-------------------F---NI---SSA..-P--R---M----------------------S-------------------A-------------------S-------------L--S---------------....  602
B.CI.x.EHO                  -----------...--K-L----V----S---K---------------F---SI---SSV..-P------L----------------------S-----------------V---------------------A----------------S---------------....  607
B.GH.86.D205_ALT            ---------NS....-NST---SV----S-------------------F-----R--SSV..KP------M---------M-------T----S----------------PV---------------------A----------------S---------------....  611
B.JP.01.IMCJ_KR020_1        ----I------...TNN-R---------S-------------------F---S----SSA..-P------P----------------------S---------------------------------------A-------------L--S---------------....  600
B.x.06.8704A_06_01          ---------ES....GN-T---SV----S-------------------F--------SSV..KP------M---------M------------S---------------------------------------A---------R------S---------------....  614
G.CI.92.Abt96               ---------EE..NRSN-H---IF------------------------F---N----SSV..-PK---------------M------T-----S--------------------------M--------------------------R--S---------------DALG  628
AB.CM.03.03CM_510_03        ----I-----T....K-------F-----------------I------F---TEG--SST..PG-----------------------------S---------------------------------------A----------------S---------------....  601
H2_01_AB.CI.90.7312A        ----------V..DVGN-R----F-----------------I-V----F---SE---SST..PG-H-------------T---A---------S--------------------------M------------A----------------S---------------....  612
H2_01_AB.JP.04.NMC307_20    ----I-----GP..DNIS-----F-----------------I-V----F---GE---SSA..PG-H-------------T---V---------S--------------------------M------------A----------------S---------------....  606
H2_01_AB.JP.07.NMC716_01    ----I------R..DNSS-A---F-------------------V----F---GE---SS...PG-H--DML--------T---V---------S--------------------------M------------A----------------S---------------....  603
H2_01_AB.JP.08.NMC842_10    ----I------...L-T------F-----------------I-V----F---EE---SSA..PE-H-------------T---V---T-----S--------------------------M------------A----------------S---------------....  610
U.CI.07.07IC_TNP3           -----------...T-T-R-----------------------------F---S----SSA..-P-------------------T---T-----S-----------------------------------------------------S------------------....  628
U.FR.96.12034               ----M-----R...-NK-----S---------------------R---F---P----SSV..-P-----------------------------S--------------------------M--------------------------S------------------....  618
MAC.US.x.251_1A11           -----------...T-------------------------------------N-R------.-----------------------------------------------L--------------------------------------------------------....  622
MAC.US.x.251_32H_PJ5        -----------...T----------------------------------------------.--------------------------------------------------------------------------------------------------------....  622
MAC.US.x.251_BK28           -----------...T-----S----------------------------------------.--------------------------------------------------------------------------------------------------------....  625
MAC.US.x.750.p4i3           -----------...T----------------------------------------------.--------------------------------------------------------------------------------------------------------....  626
MAC.US.x.MM142_IVMXX        ---------N-...T-----S-------------------------------N--------.------------------------------V-------------------------------------------S-----------------------------....  626
MNE.US.x.MNE027             -----------...--------------------------------------N--------.-P------------------------------------------------------------------------------------------------------....  626
SMM.CI.79.SIVsmCI2          ---------N-...N--G-----F------------------------F---N----SSAT.PA-----------------------------S---------------------T--------A----------------------R------------------....  633
SMM.LR.89.SIVsmLIB1         -----------...-NN-E--------I------------------------N----S-M....-P-------------------------------------------------------------------A----------------S-------------E-....  622
SMM.SL.92.SIVsmSL92A        -------D---...-NE-E---------------------------------N------S..----------------------------------------------R-----------------------------------------S---------------....  629
SMM.SL.92.SL92B             -----------...--N-E-----------------------------M---H------S..--K-------------------------------------------------------------------------------------S-------------L-....  630
SMM.US.04.G078              -----------...--N-------------------------------M---S-------..-------------------------------S--------------------------------------------------------S---------------....  634
SMM.US.04.G932              -------E-E-...-NK---------D--A------------------M---S------A..----------------------------------------------------------------------------------------S---------------....  630
SMM.US.04.M919              -------E-K-...NSN-----------------------------------N-R----TS.A------------------------------S--------------------------------------------------------S---------------....  637
SMM.US.04.M922              -------E-ES...RD------------------------------------N-R---NA-.-------------------------------S--------------------------------------------------------S-------------Q-....  636
SMM.US.04.M923              -------E---...-ND-E----------Q------------------F---G-R----T-.A------------------------------S--------------------------------------------------------S---------------....  634
SMM.US.04.M926              ----I--D---...--K-E-----------------------------M---N------TS.A---------------------------------------------------------------------------------------S-------------Q-....  635
SMM.US.04.M934              ---------NY..DKEK-T-------D-------------------------S------AS.A---------------------------------------------------------------------------------------S-------------Q-....  628
SMM.US.04.M935              -------E---...-HD-E---------------------------------N-R---NT-.AP-----------------------------S--------------------------------------------------------S---------------....  629
SMM.US.04.M940              -----------...-YK-E---------G-------------------M---S------S..----------------------------------------------------------------------------------------S---------------....  643
SMM.US.04.M946              ---------EYD.EHKNKT---------------------------------S------A-.A---------------------------------------------------------------------------------------S-------------Q-....  637
SMM.US.04.M947              N------E---...V-N-V---------------------------------S------T-.A------------------------------S-----------------------------------------------------R--S---------------....  633
SMM.US.04.M949              ----I------...-NE-----------G-------------------M---S-----S-..----------------------------------------------------------------------------------------S---------------....  625
SMM.US.04.M950              -------D--Y..VSSK-E---------------------------------S------TS.A------------------------------------------------------------------------------------R--S-------------Q-....  633
SMM.US.04.M951              N---I----EM...-NKD-----L----G-------------------M---S----NS-TS----------------------------------------------------------------------------------------S---------------....  629
SMM.US.04.M952              N---------Y...V-KDE---------------------------------S------TS.A----------------------------------------------------E----------------------------------S-------------Q-....  629
SMM.US.05.D215              ----I--E---...LNN-E-------D-------------------------S------Q..---QR--------------------------S--------------------------------------------------------S---------------....  628
SMM.US.06.FTq               ---------Y-...----R----F------------------------F--------SSVT.-P-----------------------------S-------------------------------------------------Q---K--S---------------....  625
SMM.US.07.FBr_365wpi        S----L-TM--...NSD-E-----G-D-------------------------T-R----T-.A------------------------------S-----------------------------------------------------R--S---------------....  631
SMM.US.07.FFr_365wpi        -----M-----...NSNDE-----G-D-------------------------N-R----T-.AP-----------------------------S--------------------------------------------------------S---------------....  628
SMM.US.08.FTv_32wpi         -----L---E-...RNNSE-----G-D-------------------------S-R----T-.A------------------------------S--------------------------------------------------------S---------------....  631
SMM.US.10.FJV_154wpi        -----L-TM--...NND-E-----G-D-------------------------S-R----T-.A------------------------------S--------------------------------------------------------S---------------....  630
SMM.US.10.FPY_154wpi        -----L--M--...-SNDE-----G-D-------------------------S-R----T-.A------------------------------S--------------------------------------------------------S---------------....  632
SMM.US.86.CFU212            -------E---...Y-K-------------------------------I---N------TT.---A------------------------------------------------------------------------------------S-------------E-....  631
SMM.US.x.F236_H4            -------E---...-NS-E----------------------I----------S-R----T-.A---------------------------V--S--------------------------------------------------------S---------------....  629
SMM.US.x.PGM53              -------D---...--------------------------------------S-R-----..A---------------------------V-RS------------------------------------------------R-------S---------------....  632
SMM.US.x.SME543             -------E---...-NN-E-----------------------------------R----T-.A------------------------------S------------------------H-------------------------------S---------------....  631
STM.US.89.STM_37_16         ----I------...TNN-E----A----------------------------N------S..---T-----------------------------------T------------------------------------------------S---------------....  626
```

```
MAC.US.x.239                  PNASLTPKWNNETWQEWERKVDFLEENITALLEEAQIQQEKNMYELQKLNSWDVFGNWFDLASWIKYIQY.GVYIVVGVILLRIVIYIVQMLAKLRQGYRPVFSSPPSYFQQTHIQQ..DPALPTREGKERDGGEGGGNSSWPWQIEYIHFLIRQLIRLLTWLFSNCRTL  790
A.CI.88.UC2                   V-D----R---M------KQ-RY--A--SQS------------------------------T--------.-------I-A---A--V--L-SRF-K----------G-L--I--HT..-RGQ-AN-ET-G-A-DAS-YDF---P-N--QL--HL-T----G-Y-I--D-  780
A.DE.x.BEN                    V-D--S-D-K-M------KQ-RY--A--SQS----------------------IL------T--V-----.--H----I-A---A--V--L-SRF-K----------G-L--I--HK..-RGQ-AN--T-E-V-GDS-YDL---P-N-VQ---HL-T---IG-YNI--D-  775
A.DE.x.PEI2_KR_KRCG           E-N-IV-D---M------QQTRD--A--SRS--Q---------------------------T--------.---VII-I-A------V--L-SR--K----------G-I--I--HK..-WEQ-D--ETDE-A-NSI-D-----P-A--------------G-Y-V-KD-  768
A.GH.x.GH1                    V-D--S-D---M------KQ-RY--A--SQS--Q---------------------------T--------.---------V---A-----L-SR--K----------G-L--I--HT..-RGQ-AN--T-E-DRDDD-YDL*--P-N------HL-T----G-YKI--D-  766
A.GM.87.D194                  V-D----D---M------KR-HY--A--SQS--Q---------------------------T--------.-------I-G---A-----L-SR--K----------G-L--I--HT..-RGQ-AN-ET-E-A-DDS-FGL---PLN--Q---HL-T----G-YNS--G-  766
A.GM.90.CBL24                 V-D--Q-D---M---Q--HQ-RY--A--SD---Q------------------------------------.---VII-I-A---IV-V----SR--K----------G-L--I-THR..HWEQ-D-GET-E-V-DNV-T-L---P---------L-----IG-YNI--D-  772
A.GM.x.CBL21                  V-DT-Q-D---M---K--QQIRY--A--STS--Q---------------------------T---R----.-------I-A---I--V----SRF-K-------P--A-L--V--HK..HRGQ-PS-ET-E-V-DSV-D----*P-A-------L-----IG-YNI----  772
A.GM.x.CBL22                  V-D----R-D-M------QQIRY--A--SGQ--Q---------------------------T--------.--F--I--VA------V----SR--K----------G-V--I--HK..-WEQ-D---T-E-V-DNA-S----*P-R---------V------Y-S--D-  768
A.GM.x.CBL23                  V-DT---E---M------G-IRD--A--SQQ--Q---------------------------T--------.----II-IVV-----------SR--K----------G-L--I--HK..-WEQ-A--ET-E-V-NNV-D-L*--P-R------H-------G-YNI--N-  769
A.GM.x.ISY_SBL_6669_85        V-DT---E---M------H-IR---A--SES--Q---------------------------T--------.--M----IVA------V----SR--K----------G-I--I--HK..-WEQ-D--ET-E-V-NDV-SR----P---------L------R-YNS--D-  768
A.GM.x.MCNI3                  V-DT---E---M------G-IRD--A--SQQ--Q---------------------------T--------.----II-IVV-----------SR--K----------G-L--I--HK..-WEQ-A--ET-E-V-NNV-D-----P-R------H------AG-YNI--N-  770
A.GW.06.CA65316_9             V-DT-E-N---M------KQ-R---A--SD---Q---------------------------T--------.----I--IVVF----------SR--K----------G-L--I-TPR..NWEQ-D-GEI-EGV-DNV-D-LH--P-A-------L--Q---G-Y-I--D-  788
A.GW.06.CA65319_7             V-DT-----D-M------QQ-RY-----SQS--Q--------T---------N-LS-----T---G----.----L--IVI-----------SR--K-----L----G-I--I--HK..-QEQ-A--ET-E-V-IN--DR----P-A-L----HL------R-YNI--N-  766
A.GW.06.CA65330_5             V-DT-Q-----M---D--QQ-RH--A--SD---Q------------------G---V----T--VR----.----I--IVAF-----V----SRF-K----------G-I--I--HK..-REQ-A--ET-E---DNA-D-L---P---------------IG-YNI--D-  771
A.GW.06.CA65409_14            V-D----D-K-M------E-TRY--A--SQS--Q-------------------------------R----.-------IVV------V----SR--K----------G-V--I--HK..-LEQ-AN-ET-E-V-SN--DRY---P-A-T-----L------R-YNI-SS-  789
A.GW.86.FG_clone_NIHZ         V-DT---D---M------Q--RY--A--SQS--Q---------------------T--L-FT--VR----.---V---IVA-----------SR--K----------G-I--I--HK..-QEQ-A--ET-E-V-SN--DR----P-A-------L------G-YNI--D-  767
A.GW.87.CAM2CG                A-E----D---M------Q--RY--A--SQS-----L--------------N---T-----T---S----.-------I-V------V----SR--K--------S-G-I--I--HK..--EQ-AS-ET-E-V-GN--DR---------------------G-YNI--N-  777
A.GW.87.CAM2env               A-E----D---M------Q--RY--A--SQS-----L--------------N---T-----T---S----.-------I-V------V----SR--K----------G-I--I--HK..--EQ-AS-ET-E-V-GN--DR-------..................I--N-  761
A.GW.x.CA7205_8               V-DT-I-N---M------NQ-R---A--SD---Q--------T------------------T--------.--F-II-I-AF-V--------SR--K----------G-I--I--HK..-WEQ-G--ETGE-A--SV-DNWL--P-R-------L------R-HNI--D-  764
A.GW.x.CA7253                 V-DT-M-D-K-M---K--EQIRY--A--SDS--Q--------T------------T--L--T--------.-------I-A------V----SR--K----------G-I--I--HK..GQEQ-----T-E-V-DN--DR----P-A-L-----L-----IT-YNS--D-  768
A.GW.x.CAM1                   V-D----D-A-M------Q--RY--A--SQS--Q-------------R-------------T--------.-------I-V-----------SR--K----------G-I--I--HK..-LEQLA--ET-E-V-SNV-DRF---P-A-------L------R-YNS--D-  776
A.GW.x.CAM3                   V-D----D---M------Q--HY--A--SEM--Q-------------------I-------T---R----.-------I-V---A--VI---SRF-K----------G-I--I--HK..-QEQ-D--ET-E-V-IN--DR----P-A-T-----L------R-YNS--D-  778
A.GW.x.CAM4                   V-D----N---M------LRTRY-DA--SQ---Q---------------------S-----T---R----.-------IVV----------ISR--K----------G-L--I--HT..-KEQ-A-GET-E-V-DNV-D-L---P-A-L-----L------R-Y-S--D-  778
A.GW.x.CAM5                   T-D----D---M------H-IRD--A--SDS--Q---------------------S-----T---R----.--F-II--VA------V----SR--K----------G-I--I--HK..-WEQ-ANGET-G---NNA-DN------R--------------G-YNS--N-  771
A.GW.x.CAM6                   V-D----D---M------Q--RY--A--SQS--Q---------------------------T--L--V--.--C--A-I-V------V----SR--K----------G-I--I--HK..-LEQ-AN--T-E-V-NN--YR----P-A-------L------R-Y-S--D-  780
A.GW.x.MDS                    V-S--E-D-E-M------Q--RY--A--SQK---------Q------------I-------T--------.-------I-V------V----SR--K----------G-I--I--HK..-REQ-D--ET-E-V-GN--DR----P-A-------L------R-YNS--D-  787
A.IN.07.NNVA                  E-DT-V-N-K-M-CK---Q--RH--A--SQS--Q-------------------I-------T--------.-------I-A--V---V--L-CSF-K----------G-I--I--HK..-QEQ-D---T-E-V-DN--YR----P-A------LL------R-YNS-KD-  775
A.IN.95.CRIK_147              V-D----D-K-M--*---Q--RY--A--SQS--Q------S--------------------T--------.--C----I-V------V----SR--K----------G-I--I--HK..-LEQ-A---I-E-V-DN--YR-*--P-A------HL------R-YNS-KD-  777
A.JP.08.NMC786_clone_41       V-D--S-D-D-M---Q--HQIS---A--SEK--Q---L-----------------------T---R----.--F--IA--A------V----SRF-R----------G----I--HK..-WEQ-D---T-E-A-DNV-D-L---P-R--------------R-YEI--S-  777
A.PT.00.00PTHCC20_1           V-D-IQ-N---M---Q--LQ-RD--V--SD---Q---------------------S-----T--VS--R-CR---IA--VA-------L---SR--K----------G-I--I-.SK..-REQ-P--EI-E-A-SNV-D-----P-A-------L------G-YNI--D-  739
A.PT.03.ABG                   V-D----D---M------Q--RY--V--SQS--Q-----------------N---S-----T--V-----.-------IVV-----------SR--K----------G-V--I--HT..-QEQ-A-GET-E-V-DN--Y-----P-A-------L------R-YNS--D-  782
A.PT.92.93PTHDESC_13          V-D--K-D---M----*--Q-RY--A--SQS--Q---L-----------------------T---R----.----I---VA--V---V----SR--K----------G-T--I--HK..--GQ-A--ET-E-V-DN--DRY---P-T-L-----L-----I--H-S-SN-  782
A.PT.98.98PTHDECT_13          V-D-IQ-D-K-M------LR-RD--V--SD---Q---------------------S-----T--VS--R-.----IA-IVA-------L---SR--K----------G-I--I--HK..-LEQ---GET-E-V-SNV-DNF---P-A-------L-----IG-YNI--D-  737
A.PT.x.ALI                    V-N--K-D-D-M------QQ-RY--A--SEQ--R--------T------------T--L--TA-V-----.----I--IVA------V----SR--K----------G-I--I--HK..-QEQ---GET-E-V-DNV-DRL---P-A-L----HL-A---IG-Y-I--D-  773
A.SN.85.ROD                   V-D--A-D-D-M------KQ-RY--A--SKS--Q-------------------I-------T--V-----.--L-I-A--A------V----SR--K----------G-I--I--HK..-RGQ-AN-ET-E---SN--DRY---P-A--------------R-Y-I--D-  776
A.SN.86.ST_JSP4_27            V-DT---D---M------QRIRN--A--SES--Q---------------------------T--------.-------I-V------V----SR--K----------A----I--HK..-REQ-A--ET-E-V-NSV-DNW---P-R-------------NR-YNI--D-  770
B.CI.88.UC1                   --ET---D-E-M---Q--KR-N--DA--------------R-------------------FT--MA--RL.-L-V-A-L-V-------M----R--K------------T--IP-RK..HRGQ-AN-ET-DE--NE-AYR--------A-------RN--I--YNG--N-  779
B.CI.x.20_56                  --E----N-TDM---Q--E--HY-DA--Q----------------------H---S----FT--MA--RL.-L-V---L-V------VI----R--K------------T--IP-HK..HRGQ-AN-ET-DE--RE-DYR--------A-------RN--I--YNG--N-  769
B.CI.x.EHO                    V-E--K-D---M---Q---Q-R--DA--K----------------------Q-I-S----FT--MA--RL.-L---I-IVV---A---I----R--K------------T--IP-RK..-RGQ-AN-ET-EG--NNE-YR-----------P----RD--I--Y-G----  774
B.GH.86.D205_ALT              --ET---N---M---Q--KQ-H--DA----------------------I------------T------HL.-L---A-LVV----V-------R--K------------T--IP-RK..-RGQ-AN-ET-EG--ND-DYR------------L---RN--I--YNG----  778
B.JP.01.IMCJ_KR020_1          --KNF--N-D-M---Q--NQ-R--D----K---V------E---K------Q---S----FT---A---I.-L-VI--LVV-------L----R--K------------V--IP-HK..GQGR-AS-ET-EG--DDEDYR------R---------RN--I--YNS--S-  767
B.x.06.8704A_06_01            --ET-K-D---M---Q---Q-NS--A-----------------------------------T------HI.-L---A-L-V---AV-------R--K------------T--IP-RK..GRGQ-AN-ET-EG--DD-DYR------------L---RN--I--YNG----  781
G.CI.92.Abt96                 A-KT-E-Q---M------KQIN---D---R-------------------------------T--V--VYL.-L-V-A---V------V--I-GR--------X------V--I--RK..-QEX--K-EI-GES-NK--YR------------X---GN-----Y----A-  795
AB.CM.03.03CM_510_03          V-E----N-D-M---Q--E--RD--A--SQI--Q---------------------S-----T---R----.-----A--VV---A-------SR--K---------SG-V--I--HT..-QGQ--S-ET-EG--ND-DYR--------V---L---RN--I--YNG----  768
H2_01_AB.CI.90.7312A          V-D----D-D-M---Q--KQIRD--A--SES--Q------------------------------V-----.-------IVA--VI--V---IGR--R----------G----IR-HK..-QEQ-AN-ET-EG--ND--YR----------------RN--I--YDG----  779
H2_01_AB.JP.04.NMC307_20      V-D--A-D-D-M---Q--KQIRY--A--SES--Q---------------------------T--------.--L----IVA---I--V--L-GR--R----------G-L--IR-HK..GREQ-AN-ET-EG--SDE-YR----------------RN--I--YEG----  773
H2_01_AB.JP.07.NMC716_01      V-D--A-D-D-M---Q--KQIRY--A--SIS--Q---------------------------T--------.--L----IVA---I--V--L-GR--R----------G-L--IR-HK..GREQ-AN-ET-EG--SDE-YR----------------RN--I--YEG----  770
H2_01_AB.JP.08.NMC842_10      G-D------D-M---Q--TQIRY--A--SES--Q---------------------------T--------.--L----IVA--LI--V--L-GR--R----------G-L--I--HE..HRGQ--N-ET-EG--SDE-YR----------------RN--I--YEG----  777
U.CI.07.07IC_TNP3             V-N--Q-D---M------V--KD------V--------------Q-------------L--T--V--VYL.-F-V-A-I-G--VI--V--L-SR--K----------L-V--I-THE..-QEQ--K-ET-G---DNV-FN--------L--------------Y--F-AW  795
U.FR.96.12034                 I-DT---N-D-M------E--NY------Q---A----------------N--I-------T--V--VYL.-L-V-A-I-I------V--L-GS--K--------H---VE-IP--R..-QEQ--KGEI-EGA-DS--YR------R------H---------Y-S--D-  785
MAC.US.x.251_1A11             ----------------------------------------------------------------------.----------------------------------------*-----..-Q---------G---------------------------------------  788
MAC.US.x.251_32H_PJ5          -------D---D----------------------------------------------------------.----------------------------------------------..-----------G---------------------------------------  789
MAC.US.x.251_BK28             -------D---D----------------------------------------------------------.-I-V------------------------------------*--T--..-----------G---------------------------------------  791
MAC.US.x.750_p4i3             -------N---D----------------------------------------------------------.---V---------------------------------------T--..-----------G---------------------------------------  793
MAC.US.x.MM142_IVMXX          -------D-----------------A-----------------------------------T--------.-I--I-----------------R-----------------*--T--..------K---KG---GS----------------------------------  792
MNE.US.x.MNE027               -------N---------------------------------------------------------R----.--------------------------------------------R-..-Q----K--T-G---DS----------------------------------  793
SMM.CI.79.SIVsmCI2            E-NT-Q-D---M------I-IRD--A--SES-----------------------------FT--V---YI.-F-V---I-G---IV-LI-L-G---K------------V--I--HK..GQER--N--T-E---DN--YR-----------------------YN---D-  800
SMM.LR.89.SIVsmLIB1           --KT-Q-N-E-M-------------A---Q-------------------------------T--V-----.--MV-L---G---M--VI---LR---------------H--IRT--..-Q-Q-IK--I-EG--DN--YR-------------H---------YN---AW  789
SMM.SL.92.SIVsmSL92A          --D----D---M-------------A---LM---------------------------------V----L.-----I-IV----I--VI---SR-SK------------H--I-THK..VQE---K--IDEE--NS--TR-----------------H-----YNS--A-  796
SMM.SL.92.SL92B               --D--V-D---M------K--E---A---QM---RL-------------------------T--VR----.--FL-I-IV-------V----SR---------------H--I---R..-QE--AK--T-GE--N---YRL-----------------I----YN-LTR-  797
SMM.US.04.G078                --D----N---M----------Y--A---EM---------------------S-----------------.---L---I-F------V--------K----------T-H--IP-RK..SQE---KG-E-GE--DS--YR----------------V-------N---AW  801
SMM.US.04.G932                --DT-I-D---M------K--N---A--SLM---------------------N------------R----.-----I-I------------FSR-GK------------H---PVHK..-QER--NGEG-E---GN---R-----------------------YDS----  797
SMM.US.04.M919                -----V-S---M---K---Q-----A---Q----------------------NI-------T--------.--L--L--VG------V---V------------P--A----IP-HK..GQEPL-K--E-GE--DR---R------------------------N----Y  804
SMM.US.04.M922                ---T-L-N-E-M---------N---A---Q-----------------------I-------T--------.--L--L---G-K----V-----R-------------A-V--IP-HK..GQEP--K--E-G---NR--SR------------------------NS---C  803
SMM.US.04.M923                --DT---D---M------KQ-N---A---Q-------------------------------T---R----.--L--L---G------L----TR-------------A-V--IP--T..GQG---K--E-G---NR---R------------------------NS--DW  801
SMM.US.04.M926                --D--I-D---M-------------A----M----------------------------------R----.---V---I--------V-----R-------------T-I--IP-HK..GQEQ--NGEE-EN--DR---R------------------------NS--AC  802
SMM.US.04.M934                --G--V-D---M-------------A----M------------------------------T--------.-I-V-----M------V-----R--K-------P--T-I--IP-HK..GQEP--KG-E-E---DR---R-----------------------------C  795
SMM.US.04.M935                --ET-V-N-S-M-------Q-----A---Q-----------------------I-------T---R----.--L--L--VG------V-----R-------------A-V--IPVHK..GQEP--K--E-G---DR--SR-------------------------S---C  796
SMM.US.04.M940                --DT-I-D---M-------------A---LM-----V---------------S--------T--------.--FL---I--------V-----R--K----------A-H--IP-RK..SQEP--K--G-GE--DN--YR------------L-----------N---AW  810
SMM.US.04.M946                --D--V-D---M-------------A--------------------------------------------.---V-----M------V-----R--K-------P--T-I--IP-HK..GQE---KG-E-EE--DR---R------------------------N----C  804
SMM.US.04.M947                --ET---N---M-------Q-----A---Q-----------------------I-------T---R----.--L--L---G------V-----R-------------V-V--IPL-K..GQE-Q-K--E-G---GR---R--------------------I---N---DW  800
SMM.US.04.M949                --D--I-D---M---Q---Q-----A----------------L------------------T--------.---L---I--------V-----R--K------P---A-H--I--R-..NQEP-IKGEG-EE--DS--YR------------L-----------N--KAW  792
SMM.US.04.M950                --S--M-E---M-------Q-----A----M------------------------------T---R----.-----I---M------V-----R--K-------P--T-V--IP-HK..GQE------E-G---DR--SR------------------------N----C  800
SMM.US.04.M951                --D----D---M-------------A---L------V------------------------T---R----.---L---I--------V-----R--K----------A-H--IP-RK..SQEP---GEE-EE--DS--YR--------L---L---V-------NS---W  796
SMM.US.04.M952                --D--V-D---M-------Q-----A--------------------------------------------.---V-----M------V-----R--K-------P--T-V--IP-HK..GQE---K--E-G---DK---R-------------H----------N----C  796
SMM.US.05.D215                --DT---N---M-------Q-----A---QM------------------------------T---R----.--I--L---G---I--V--L-GQF-K------------H--I--HK..LQGQ--EG-I-GE--DS--YR--------L---L---L------YNS--AW  795
SMM.US.06.FTq                 V-D--Q-D---M------KQ-AY--A---QQ------L------D----------------T--V--VYF.-F-V-A---V------VI---G---K----------T-R--IY-HK..GQEQ--K--I-E-V-DNV-YR-----------------------YD----W  792
SMM.US.07.FBr_365wpi          ---T-Q-N-S-M------KQ-N---A---Q-------------------------------T--V-----.--F--L---G------V-----R-------------A-V--IP-RK..GQEQ--E--E-G---DR--SR-----------------------------Y  798
SMM.US.07.FFr_365wpi          --T-VM-N---M---Q--KQ-N---A---Q-------------------------------T--------.--F--LA--G---I--V-----R-------------A-V--I--SK..GQEQ-IK--E-G---DR---R-----------------------------Y  795
SMM.US.08.FTv_32wpi           ---T-L-N---M------KQ-N---A---Q----------------------------------V-----.--F--L---G------V-----R-------------A-V--IP-RK..GQEQ--K--E-GN--DR--SR-----------------------------Y  798
SMM.US.10.FJV_154wpi          ---T-L-N---M------KQ-N---A---Q-------------------------------T--V-----.--F--L---G------V-----R-------------A-V--IP-RKGQEQEQ-IE-EE-G---DR--SR-----------------#-----------Y  798
SMM.US.10.FPY_154wpi          ---T-L-N---M------KQ-N---A---Q---D------------------------------V-----.--L--L---G------V-----R-------------A-V--IP-HK..GQEQ--E--E-E---DR--SR-----------------------------Y  799
SMM.US.86.CFU212              -A-NV--D---M----------Y--A---L-------------------------------T--V-----.-I-----I--------V-----R--K----------N-R--I--HK..-QEQ-----I-EG--ID--IR---------------F-------Y-S----  798
SMM.US.x.F236_H4              --ET-V-N---M-------Q-----A---Q-----------------------I-------T---R----.--L--L---G------V-----R-------------A-V--IP-HK..GQEP--K--E-G---DR--SR-------------------------S--DW  796
SMM.US.x.PGM53                -----V-N---M-------Q--D--A---QA----------------------I-------T--------.--L--L--VG------V---I-R-------------V-V--IP-HK..GQEQ--K--E-G---DR---R------------------------N----Y  799
SMM.US.x.SME543               --D--V-N-D-M------G------A---Q-----------------------I-------T---R----.--L--L--VG------V-----R----------P--A-V--IP-HK..-QEP--K--E-GE--DR--SR-------------------------S--DW  798
STM.US.89.STM_37_16           --D--V-D---M-------------A---Q------V------------------------T--VR----.---L-I-LVM--VA---M-L--R--K-----------CR--IP-HK..GQEQ--K--T-EG--DR--IN--------T-------V------YN-F-AC  793
```

Env end
gp41 end

```
MAC.US.x.239              LSRVYQILQPILQRLSATLQRIREVLRTELTYLQYGWSYFHEAVQAVWRSATETLAGAWGDLWETLRRGGRWILAIPRRIRQGLELTLL*  879
A.CI.88.UC2               --ANSPTRRL-S-N-....TA--DW--LKAA-----CEWIQ--F--IA-T-R-------RG-CKAVQ-I--G---V-------A-IA---  866
A.DE.x.BEN                --KNSPTRRL-S-S-....TA--DW--LKAAQ----CEWIQ--F--FA-TTR--------W---AA--I--G---V-------A--A---  861
A.DE.x.PEI2_KR_KRCG       ---SFPT--L-F-S-QRA-TT--DW--LTIA-----CEWIQ-VL-VLA-TTR----S--R---GAMG-I--G---V-------A--A---  858
A.GH.x.GH1                --TNSPTHRL-S-N-....TA--DW--LKAA-----GEWIQ--F--FAKTTR---S---G-CAAVQ-V--G---V-------A-IA---  852
A.GM.87.D194              --KNSPTRRL-S-S-....TA--DW--LKAA-----CEWIQ--FR-FA-T-R--I----RG-C-AAQ-I--G---V-------A-IA---  852
A.GM.90.CBL24             ---SSL-------S-QRA-TA--DW--L-AA-----CEWIQ--L--LT-ATR------RN---GA-Q-I--G---V-------A--A---  862
A.GM.x.CBL21              I-KSF-T-----S-G-QRA-TA--DW--PGAA-----CEWIQ--L--FA-ATR---TSV-RNFCG-MGQI--G----------A--A---  862
A.GM.x.CBL22              ---SFL-----------A--DW--LKAA-----CEWIQ--FR-FAKFTR------R---GG-G-I--E---V-------A-IA---  847
A.GM.x.CBL23              ---ISLT-R-VF-S-QRA-TA--DW---DAA-----CEWIQG-F--FA-ATR-----T-R---GA-Q-I--G---V-------A-IA---  859
A.GM.x.ISY_SBL_6669_85    ---L-L---L----------DW--LKAA-----CEWIQ--F--LA-VTR---TS-GRS--GA-G-I--G---V-------A-IA---  847
A.GM.x.MCN13              ---ISLT-R-VF-S-QRA-TA--DW---DAA-----CEWIQG-F--FA-ATR-----T-R---GA-Q-I--G---V-------A-IA---  860
A.GW.06.CA65316_9         ---IFP-----F-G-QRA-AAT-DW--LNTA--R--CEWIQ--F--FA-ITR---TSI-R-F-R--G-I--GV----------A--A---  878
A.GW.06.CA65319_7         ---SFL----V--N-.......DW--LKIAF----CQWIQ--F--AA-ATR---TS-CRS--G--GQI--G----------A-IA---  849
A.GW.06.CA65330_5         ---TSL-----SRS-QRA-TA--DW--LKIA-----SEWIQ--F--FA-A-R---TSM-RN--GA--WI--EVF-V-------A--A---  861
A.GW.06.CA65409_14        ---IF-T--LT--N.....L-DW--LNTAS----CEWIQ--F--FA-ATR------CR---GL-G-I--G----------A-IA---  872
A.GW.86.FG_clone_NIHZ     ---ISP-----F-S-QRA-TA--DW--LKAA-----CEWIQ--F--LA-TTR------GR---RA-Q-I--G---V-------A--A---  857
A.GW.87.CAM2CG            ---IF-T------N-.......DW--PKVAF----CEWIQ--F--AA-A-R------CR-V-GM-Q-I--G---V-------A--A---  860
A.GW.87.CAM2env           ---IF-T------N.....L-DW--PKVAF----CEWIQ--F--AA-A-R------CR----V-Q-I--G--VV-------A-IA---  844
A.GW.x.CA7205_8           ---GFLT--LLP-S-QRA-TAT-DW--LKTAH----FEWIQ--FN-FAGT-R---T---R---GA-QWV--R----------A-IA---  854
A.GW.x.CA7253             ---IF-T-----RN.....L-DW--IKTAL----CEWIQ--F--AA-TTG------CRG--R---G-I--G-F-V-------A-IA---  851
A.GW.x.CAM1               ---IFLT----F-N.....L-DW--ASTAF----CEWIQ--F--FG-A-K---TS-CRS--R---G-I--G---V-------A--A---  859
A.GW.x.CAM3               ---SFLT----F-N.....L-DW--LRTAF----RQWIQ--F--FA-ATR---TS-CRG--R--DNF--G--S--------A-IA---  861
A.GW.x.CAM4               ---SFPT--------RA-TA--DW--LKAA-----CEWIQ--F--FA-T-R------RN---AMG-L--G---V-------A--A---  861
A.GW.x.CAM5               --KSFLT--L-----------DW--LNTAF----CEWIQ--F--FA-ATR---TS-GRS--G--G-I--G---V-------A--A---  850
A.GW.x.CAM6               ---SFLT------N.....L-DW--LRTAF----CEWIQ--F--FA-F-R----S-CRS--G--G-I--G----------A-IA---  863
A.GW.x.MDS                ---IFLT----F-N-.....DQ--LRIAF----CEWIQ--F--FA-VTR---TS-CRG--G--G-I--G---V-------A-IA---  870
A.IN.07.NNVA              ---SFLT------N-.....DW--LRTAF----CEWIQ--F--AA-TTG----S-CRS--G----I--G---V-------A--A---  858
A.IN.95.CRIK_147          ---SFLT------N-.....DW--LKTAF----CEWIQ--F--FA-T-GK---S-GRN-CG-VE-I--G----------A-VA---  860
A.JP.08.NMC786_clone_41   ---IF-T--L-----------DW--VKTAH----CEWIQ--F--FA-T-R--------EF-TA---F--G---V-------A-IA---  856
A.PT.00.00PTHCC20_1       ---SSL-------S-QRA-TA--DW--LKAAL----CEWIQ--FG-FA-ATR---TS-GR---GA-Q-I--G---Y-------A-IA---  829
A.PT.03.ABG               ---SFLT------S.....L-DW--LRTAF----CEWIQ--F--FA-A-R---T--CRS--G--G-I--G---V-------A--A---  865
A.PT.92.93PTHDESC_13      I--TFLT--SAF-N.....L-DW--LKIAF----CQWIQ--F--AA-ATR------GRS--GA-G-I--E---V----G--A-IA---  865
A.PT.98.98PTHDECT_13      ---SSL-----F-N-QRA-TA--DW--LKAAF---RCE*IQ--FR-FT-ATR---TS-GR---GA--QI-EGM----------A-IA---  826
A.PT.x.ALI                ---ISP-----FRS-QRA-TT--DW--LKAA-----CEWIQ--FR-FA-I-R---TNT-R---GAVQWV--R---V-------A-IA---  863
A.SN.85.ROD               ---SFLT--L-Y-N-.....DW--LRTAF----CEWIQ--F--AA-ATR------CRG-RV-E-I--G---V-------A-IA---  859
A.SN.86.ST_JSP4_27        ---SF-T--L-S-S-RRA-TAV-DW--FNTA-----GEWIQ--FR-FA-ATG---TN--RGF-G---GQI--G---V-------A-IA---  860
B.CI.88.UC1               -LKTS-----A-...........QP--LS-A-----I-W-Q--I--AT-A-R----NTGRA--KA---TAEA-I-----------A---  858
B.CI.x.20_56              -L-TS-----A-...........QP--LS-A-----I-WLQ--I--AT-A-G----N-GRA---A---TAGA-I-V---------A---  848
B.CI.x.EHO                --KTF----LV-..........QP--LPPA--R--I-W-Q--I--AA-A-G----S-ARTS-GV---AAGE-I---------A--A---  853
B.GH.86.D205_ALT          -LKTF---HQ-STN-.......QP--LPVA-----I-W-Q--LR-AA-ATG----S-GET---A---AA-A-I-----------A---  860
B.JP.01.IMCJ_KR020_1      -LKIF-P--LVF..........QQ--LP-A-----IGW-Q--I--AT-A-G---TNTGRA---A---TA-A-V-----------A---  846
B.x.06.8704A_06_01        -LKTF-----------------P--LP-A-----IGW-Q--LR-AA-ATG----S-GETF--A---MA-A-I-----------A---  860
G.CI.92.Abt96             ---IX-T---LF--I-R---A---H--L-AA-FS--FRWLQ--CT-AT-A-Q---TST-RA--K--G-V--G----------------  885
AB.CM.03.03CM_510_03      -LKTS------.....QP--LPFA-----IGW-Q--LR-AA-ATR----S-GET---A--GAA-A-------------A---  847
H2_01_AB.CI.90.7312A      -LKTF-T---A-...........QP--LLFA-----IGW-Q-----AAGATG----STGRT---A---TA-G-I-V---------A---  858
H2_01_AB.JP.04.NMC307_20  --KTF----LA-R..........P--LP-AH----I-WLQ-----AA-A-G---TS-SRTI--A---TA-G-I-----------A---  852
H2_01_AB.JP.07.NMC716_01  --KTF----QA-R..........P--LP-AH----I-WLQ-----AA-A-G----S-SRT---A---TA-G-V-----------A---  849
H2_01_AB.JP.08.NMC842_10  --KTF-T--LT--.........P--LP-AH----I-WLQ-----AA-A-G---TS-IRTI--A---TA-G-I-----------A---  856
U.CI.07.07IC_TNP3         --QSF-S--AAC-...RV-TA---H--L-GA-----CQW-Q--L--G-KAXR----S--R------G-V--G----------F--A---  882
U.FR.96.12034             ---IC-S---LF-SI.........R-HL-IA------Q--K--F--FGKA-R---SRTGRE-----G-V---LR--------F--A---  868
MAC.US.x.251_1A11         ------------------R---------------N-------VD----------*-A-----------------------------  877
MAC.US.x.251_32H_PJ5      ---A------------A-----------------Q----VG-----------------------------------------------  879
MAC.US.x.251_BK28         ---A--------------R-V-----------------------G-----------R-----------------------------  881
MAC.US.x.750.p4i3         ---A------------V-R-------------------Q-----G-----------------------------------------  883
MAC.US.x.MM142_IVMXX      ---A-------F-------R-------L-----------Q-----AQ----------E---A-Q---------------------  882
MNE.US.x.MNE027           ---A-------F--F-T----V-----------------Q----VA-----------------G-V----------E------  883
SMM.CI.79.SIVsmCI2        VL-IF-S--G-F-LV-....AA-NRI-PRIA------Q--Q--C-EAA-A-R----S-GR------G-I--GL--A---------AF--  887
SMM.LR.89.SIVsmLIB1       -F-IC-N-RQLF----EAI-PV---I-R-F--------W-ED-L--A--A-R----S--R---Q--G-I--GLW-------------  879
SMM.SL.92.SIVsmSL92A      -L-IF-S---TY---LSLIS---DI--RAVA-IR---Q--I-SL-EAAIA-RQ--VSTGRTI-Q--GHI--G-A-----------N-  886
SMM.SL.92.SL92B           A--A--N--QLC----EIS-P---LV-R-AG-IR---N--I--C-EA----Q-AIV----LI----G-V--G-A----------M-N-  887
SMM.US.04.G078            ----I--S---TF----E-A-----FI-LRV--------W-K-----A-K--R----S--R-I----G-V--GL---------F------  891
SMM.US.04.G932            -L-IC-T---VF---YKA------II----A------RW-K-----A----R------GR------EGV--E-------------  887
SMM.US.04.M919            -L-I------AF---FT----VQ--I---IS-----CLWLQ-----AS-A-R---TST-R------G-V---V----------S---  894
SMM.US.04.M922            --SI------V-----LIA--VQ--V---IS-----CRW-Q-----A--A-R----S--------G-V-------------------  893
SMM.US.04.M923            -L-SC-----V--S--R----V---I-V-I-------R--Q--A--W-KF-R----S--R------G-V--G--------------  891
SMM.US.04.M926            -L-I------VF---FT----VQ--I---IA-----CRW-Q-----AT-A-R----S--R--R---G-V--G----------F--  892
SMM.US.04.M934            -L-I------VF---F-----VQ--I--SIA-----CQWLQ--AH-AY-A-R----S--R------G-I--G--------------  885
SMM.US.04.M935            -LSI------V----FSIV--VQ--V---IS-----CRW-Q-----A--A-R---TS---------G-V---V-------------  886
SMM.US.04.M940            -L-IF-S---VF----E-V-T--DLI-L-V-------W-K-----A----G------GDY----S---V--G-G-----------  900
SMM.US.04.M946            -L-I------VF---F-A---VQ--I---RIA-----CQWLQ-----AC-A-R----S--R--R---G-V--G-G----------  894
SMM.US.04.M947            -L-S------A--S-LK----V---I-I-I-------R--Q--A--W-KF-R----S--------G-V--GV----------F--  890
SMM.US.04.M949            ---IC-S---VF----EAV-IL---I-L-V-------W-K-----A----R----S-GDY---S---V--G-G----------F--  882
SMM.US.04.M950            -L-I------VF---F-I---VQ--I---IA-----CRW-Q-----AY-A-R---TS--R----I-G-V--------------------  890
SMM.US.04.M951            -L-IC-S---VF----E-V------I-L-V-------W-K-----A----G------GDY---N---L--G-G--------------  886
SMM.US.04.M952            -L-I------VF---F-A---VQ--I---IA-----CRWLQ-----AC-A-R---TS--R------G-V-G--------------  886
SMM.US.05.D215            --QIC-T---VF--I-EV--TV--II---T--------W-Q--I--FFKF-R----S--R------G-V--G-A---------F---M--  885
SMM.US.06.FTq             -LNS--S---V-RGI-RA---V---I-L----------W-Q--A--AFKA-R----S--R----I---V-G-M--------------  882
SMM.US.07.FBr_365wpi      -L-I------VF--FFS-A---Q--I---IA-----CRW-Q-----A--A-R---ES--RN-----G-V--G----------------  888
SMM.US.07.FFr_365wpi      -L-L------VF---FSVI---Q--I---VA-----CRW-Q--A--A--A-R----S--R------G-V--G----------------  885
SMM.US.08.FTv_32wpi       -L-I------VF--FFS-A---Q--I---IS-----CRW-Q---R-A--A-R---ES--R----I-G-V--G----------------  888
SMM.US.10.FJV_154wpi      -L-I--L---TF--FFS-A---Q--I---IA-----CRWLQ--A--A--A-R-A-ES--R------G-V--G----------------  888
SMM.US.10.FPY_154wpi      -L-I------VF---FS-A---Q--I---IA-----CRWLQ--------A-R---ES--R------G-V--G----------I---  889
SMM.US.86.CFU212          -L-I--N---V----SV--A---TI-----IG--CLW-K--L--AG-A-GQ---S--R------G-V--G--N-------------  888
SMM.US.x.F236_H4          -L-S------V--S--T----V---I-I-IA------R--Q-----W-KL-R----S---I----G-V--G----------------  886
SMM.US.x.PGM53            -L-I------VF---.---------V---IS-----CRW-Q-----A--ARG----S--R-S----G-V---V------------  886
SMM.US.x.SME543           -L-I------V-----R----V---I-I-I----------Q--A--W-KF-R----S--R-I----G-V--G-------V------A---  888
STM.US.89.STM_37_16       ---I--T-H-TF--I-RI-------V-LGAA-----CIWIQ--A--A--A-G----S-GR------G-V--R-G----------------  883
```

Draft of 30 September 2015, not for public release
hiv2prot_main.tex r8876–10546


max HIV-1 similarity
normal Nef end
MAC.US.x.239 IRYPKTFGWLWKLVPVNVSDEAQ....EDEEHYLMHPAQTSQWDDPWGEVLAWKFDPTLAYTYEAYVRYPEEFGSKSGLSEEEVRRRLT..ARGLLNM...ADKKETR* 262
A.CI.88.UC2 V---MC----------M-Q--E....D--TNC--------RY--IH--T-V-R-NSM---E-K-FTL------H----P-K-WKAK-K.----IPYSE.....- 258
A.CI.x.IC763124 E-L--FY----------IPQ-EE....---TCC-T------RY--IH--I-V*--N-----N-R-FIK------Y----AK-DWKA--K.----IPFS.....- 254
A.DE.x.BEN V---MY--------S-EL-Q--E....---ANC-V------RH--EH--T-V-Q--SM---N-K-FTL------H----P-K-WKAK-K.----IPYSE.....- 258
A.DE.x.PEI2_KR_KRCG V----F----------D-PQG.....E-D-C-L------GS---H--T-M-R---R---E-T-FN------Y----P---WKAK-K.----IPFS.....- 254
A.ES.x.S1084 V---MC---------ID--Q-GE....DT-T-C-V------KF---H-QTSV-R---M---E-T-FI-H-----H----P---WKA--K.----IPFN.....- 242
A.GH.x.GH1 V---MC----------D--Q--E....D--TN--T------RH--EH--T-L-R------D-K-FLH-----H----P-K-WKAK-K.----IPYS.....- 256
A.GM.87.D194 T---MY----------DI-Q--E....-V-TNC-V------RY--EH--T-V-R---M---S-K-FILH----H----P-K-WKAK-K.----IPYSE.....- 258
A.GM.x.ISY_SBL_6669_85 V---MF----------T-PQ-GE....DT-TLC---S--V-RF---H--T-V-----M--HE-TTFIL-----H---ME-DDWKAK-K.----IPFS.....- 257
A.GM.x.MCN13 ----MF----------D-TRQEE....D-GT-C-L------RF---H--T-I--------HD-K-FILH----H----P--DWKA--K.----IPFS.....- 257
A.GM.x.MCR35 ----MF---#.....Y----HE-K-FIL------H----P--DWKAK-K.----IPFS..... 214
A.GW.86.FG_clone_NIHZ .....FTL------HN---P-K-WKA--K.----IPFS..... 181
A.GW.87.CAM2CG V---MF----------DT-Q-GEDTETDT-T-C-L------RH--MH--T-V----SM---LK---FT------H----P-D-WKAK-K.----IPFS.....- 261
A.GW.x.MDS T----C----------DLPQD.....ENT-T-C-V----I-KF--QH--T-V-R--SL---E-T-FKL------HQ---P---WKA--K.----IPFN.....- 256
A.IN.07.NNVA ----MFH----------PQ-GE....DT-T-C--------KF---RH--T-V-R-------D-V-FRLH-----H----P---WKA--K.-----PFN.....- 256
A.IN.95.CRIK_147 .....#-FRL-L----H----P---WKA--K.----IPFS..... 172
A.JP.08.NMC786_clone_41 -----C----------DITQ--E....D--TNC-T------RY--EH--T-M----SM---N-I-FTL------YQ---P---W-AK-K.----IPHN.....- 257
A.PT.x.1096 -----Y----------E-PE-GN....D-GT-C-L------RF---H---V-R---M--LE-T----------Y--D-P-DVWKAK-K.----IPFS.....- 257
A.PT.x.1139 T---MF----------DIPQ-GE.....T-C-V----V-KF---H-T-V-R---M--HD---FK-H-----H----P---WKA--K.----IPF.....- 254
A.PT.x.1147 V---F----------D-PQ-GE....DI-T-C------IGRF---H--T-V-R---M--FD-V-FKKH----H----P---WKA--K.----IPFS.....- 257
A.PT.x.1148 --F--Q----------DAPE-EG....DEGT-C-L----GWF---H--R-V-R---HM---E-T-F--F----RH---P-D-WKAK-K.----IPF.....- 256
A.PT.x.1215 -----FY---------DIPQ-GN....GT-T-C--------HF---H--T-V-R---K--HD-V-FNLF----YQ---P-K-WKAK-K.----IPFSE.....- 258
A.PT.x.1227 ----MF----------D-PQ-GE....DI-T-C--------RF---H-QT-V-R---M---D-V-FT------H----P---WKA--K.----IPFS.....- 257
A.PT.x.1268a -----FY---------EIPQD-N....NS-T-C------I-NF---H--T-V-----K--HD-V-FD------YQ---P---WKA--K.----IPYS.....- 257
A.PT.x.1320 V----F----------DIPQ-GN....DT-T-C------I-HF--QH--T-I-R---K--HD-I-FRL------YQ---P---WKAK-K.----IPFSE.....- 258
A.PT.x.1378 V----F----------DIPQ-EE....DT-TNC-L------PF---HR-R-V-R---M--HE-T---K------Y----P---W-A--K.----IPFN.....- 257
A.PT.x.138 V-F--F----------DIP..QD....D-GT-C-L-----NRL---H--T-V-R---M--HE-T----H-----Y----P---WKA--K.----IPFS.....- 255
A.PT.x.1395 T---MF----------D-PQ-GE.....T-C-V----V-KF---H--T-V-R---M--HD---FKKF----H----P---WKA--K.----IPFS.....- 255
A.PT.x.1428 T---RF----------DITQ-.E....DTDT-C-V------KF---H--T-V-R---M---N---FN------Y----P---WKA--K.----IPFH.....- 256
A.PT.x.1543 -----Y----------E-P..QE....DE-T-C-L----I-KF---H--T-V----HM---E-T-F--------YQ---P-D-WKAK-K.----IPFS.....- 255
A.PT.x.1544 ----MF----------NPQDEN....DT-T-C--------HL---H--T-V-R--SK---N-V-FNL------H----P-K-WKAK-K.----IPYSE.....- 258
A.PT.x.1567 T---MF-------E--E-P..QD....D-GD-C-L------GF---H--T-V-R--SR---E-T-FIK------Y----E---WKA--K.----IPYS.....- 255
A.PT.x.268 -----F----------D--QAEE....DT-T-C---------F--IH--T-V-R---M---S-T-FIKH----H----P---WKA--K.----IPFN.....- 257
A.PT.x.293a T---MF----------D-PQ-GG....DT-T-C-V----V-KF---H--T-V-R---M---D-V-FRL------H----P---WKAK-K.----IPFN.....- 257
A.PT.x.293b -----C----------D-PQ-GE....DT-T-C-V------KF---H--T-V-Q---M---S---FT-F-----YQ---P---WKA--K.----IPFS.....- 256
A.PT.x.427d T---MF----------D-PR-EE....D--ANC-L------RF--QH--T-V-R-----D-T-FN-F-----HN---P-K-WKAK-K.----IPYK.....- 256
A.PT.x.483 ----MH----------D-PQ-GK....DT-T-C-------V-RF---H--T-L-R---M--HD-V-FRL------H----P---WKA--K.----IPFS.....- 257
A.PT.x.511 V---MF----------GALL-GEE....D-GT-C--------KFN--H--T-V----SM---E-K-FTLF----H----P---WKAK-K.----IPFN.....- 258
A.PT.x.546 ----M-----------STLPGGEE....D-GT-C-L----GRF---H--T-V-R---M--HE-K--ILH-----H----P---WKA--K.----IPFS.....- 258
A.PT.x.741 V---MF----------D--Q-EG....DT-T---V------KF---H-QT-V-R---M---E-T-FKQ----KH----P-D-WKA--K.----IPFN.....- 257
A.PT.x.794 ----MF----------D-PQ-GD....DT-T-C--------KF--LH--T-V-E---M--FD---FT------KH----P-D-WKA--K.----IPFS.....- 257
A.PT.x.956 ----------DAPQ-EE....-N-ANC-L------RF---H--T-V-R--SM---S-T-FIKH----HN---P-N-WKA--K.----IPYS.....- 257
A.PT.x.ALI ----MF----------D-PQ-GE....DT-T-C-L--V---RH--TH--T-V-R---K--HD-K-FILH----Y----P-D-WKA--K.----IPFS.....-NRNS- 263
A.PT.x.B1_1 V----F----------D-PQ-GE....DT-T-C-V----I-KF--QH--T-V-R---M--FD-V-FRL------H----P-D-WKA--K.----IPFS.....- 257
A.PT.x.EP ----RF----------D-PQ-GD....D-GT-C--------KF---HR-T-V-R---M---E-K-FNL------H----P---WKA--K.----IPFS.....- 257
A.PT.x.LF3 V---MF----------PQ-EE....GE-VNC-L------RF---H--T-V-R---M--HE-T----------H----P---WKA--K.----IPFS.....- 257
A.PT.x.MP2 T---MF----------DIPR-EE....---ANC-L------RL--QH--T-V-R-------D-T-FN-F-----HN---P-K-WKA-K.----IPYR.....- 257
A.PT.x.P1 V---MF----------EIPQ-GE....DT-T-C------Q-KF---H--T-V----SM--LQ-T-FIKH----H----P-K-WQA--K.----IPFS.....- 257
A.SN.85.ROD V---MF----------D-PQ-GE....DT-T-C-V------KF---H--T-V-E---L---S---FI------H----P---WKA--K.----IPFS.....- 257
A.SN.86.ST_JSP4_27 V----F----------D-PQ-.G...DDS-T-C-V------RF---H--T-V-R-----FS---FI-------Y----P-D-WKA--K.----IPFS..... 256
B.CI.88.UC1 ----R----------DIPE-ER....GA-TSC-V----I-S---IH--T---R---L--HD-V-FN------YQ---P-K-*KA--K.----IPTD..... 238
B.CI.x.20_56 V----------IPE-EE....RA-NSC-V------S----H--T-V-Q---L--HD-I-FNKF----YQ---P-K-WKA--K.----IPTD..... 238
B.CI.x.EHO V----F----------I-MIA-PE....DE-T-C-V------A----HE-T-V-Q--SL---D-V-FS-F-----YQ--MP-K-WKAK-R.----IPTE..... 239
B.CI.x.IC762993 -----K*-*------IPEKNE....-AKTSC-------P-NNLH-KT-V-Q-NSF---D-V-FN-F--K-KYQ-E-P-K#*KAK-R--KKIPYP#..... 234
B.GH.86.D205_ALT -------Y----------E-PAATR..E-E-T-C-----I-S---IH--T-I-Q--SL---D-V-FN-F-----YQ---P---WKA--K.----IPTD..... 241
B.JP.01.IMCJ_KR020_1 V----------IPE-DK....A-TSC------A-P---VH--T-I-Q--SL---D-I-FN------YQ---P-K-WKV--K.----IPTD..... 239
G.CI.92.Abt96 ----X----------T------...---T-C-V------P----ET------------D-R-FILH----RW----P-AVWKEK-K..Q---PIE..... 257
AB.CM.03.03CM_510_03 -----Y----------ETPT--R...EGE-T-C--------S---YH--T-V-----L--HD-V-FN-F-----YQ---P-K-WKAK-K.----IPTD..... 239
H2_01_AB.CI.90.7312A -----H----------E-AAVTR...E-E-T-C-V-----AA----H--T-V-Q--SL---SD---FN-F-----YQ---P-K-WKA--K.----IPTE..... 240
H2_01_AB.JP.04.NMC307_20 V----Y---------I--TAAT-...E-E-T-C-V------A----H--T-V-Q---F---S---FNKF-----YQ---P-K-WKA--K.----IPTE..... 240
H2_01_AB.JP.07.NMC716_01 -----C----------TAAT-...E-E-T-C-V------A----H--T-V-Q---L---S---FN-F---V-YQ---P-K-WKA--K.----IPTE..... 240
H2_01_AB.JP.08.NMC842_10 V----F----------E-AAAT-...E-E-T-C-V------S----H--T-V-Q--SL---S---FN-H-----YQ--MP-K-WKA--K.----IPTE..... 240
U.CI.07.07IC_TNP3 V-F-----------H-----S....N--T-C-V------EE-R----------A---D-K-FIL----------P-K--K---A.--------------S- 264
U.FR.96.12034 T-W-MC-------E--DAN-K---------RL--VGS----CEE-H---A-V----SS---S-Q-FIKC----------K----------PV.......KNC- 261
MAC.US.x.17EC1 ---------- 264
MAC.US.x.251_1A11 ----------P---------- 264
MAC.US.x.251_32H_PJ5 ----------R----Q----K----------P---------- 264
MAC.US.x.251_BK28 ----------R----Q----K----------A----#----------R--- 263
MAC.US.x.MM142_IVMXX ---L-M-------I----------V----------I------K--K---A.------E-----R---S- 263
MNE.US.82.MNE_8 P----------G--N--L----------V----------I----------K-------R--S- 264
MNE.US.x.MNE027 P----------G--N--L----------V----------I----------K----------S- 264
SMM.CI.79.SIVsmCI2 ----MM----------R-C-V----I-HL----------Q---N---FI----------D--KK----------K---------NS- 263
SMM.LR.89.SIVsmLIB1 ----MF---------MD----------T-C-V----H-M----------S--K----------D--K----------K---------- 262
SMM.SL.92.SIVsmSL92A ----MF-----Q----D-----N.....-N-C-L#----YKMS-----T-----N-M------F--------W---T---K---A#AG-S-TTKSNM-Q---AS- 258
SMM.SL.92.SL92B -----F-----Q----D-----N....N--T-C-V----Y-HS----K------N-H------F--H----W---T---E---ANKPKPQKK----------S- 267
SMM.US.04.G078 V---MF----------D----------T-C-V----H----------Q-------FI------MQ-----D--Q----------K-------GGNX. 263
SMM.US.04.G932 ----MC----------DI----------T-C-V----------S--------#.....-FIK----------K----------K---------X. 258
SMM.US.04.M919 ----MF----------T-C-V----------K-------FI-H------D---K---KK----------K---------X. 263
SMM.US.04.M922 ----MF----------D-TE----....T--R-C----------SQ---N---FI-H-----K---------K-----.T---K---------X. 256
SMM.US.04.M923 ----MV----------A-C----------E---N-K-FIEH------M--------K----------K---------X. 263
SMM.US.04.M926 ----MF----------S--E----....-T-C----------T-----SQ---R---FI-H----------K----------K---------X. 263
SMM.US.04.M934 ----MF----------T-C-V----------Q-------F--H------E---P---K----------K-------E--X. 263
SMM.US.04.M935 ----MF----------D-TE----....T--R-C-V-----Y----------SQ---K---F-K----------K----------K---------X. 263
SMM.US.04.M940 ----MK----------D-------....T--A-C-V----V--K-------T-------HS---FI-H------WS--------K----------K----.D----GNS- 266
SMM.US.04.M946 ----MF----------T-C-V----------Q---------FI-H---------P---K----------K---.D----X. 263
SMM.US.04.M947 ----MF----------T---V----------E---N-V-FIKC----------K----------K---------X. 263
SMM.US.04.M949 ----MK----------T-----E....T--T-C-V---V--K-------T-V--------HR---FI-------WS--------K----------K------GGNS- 266
SMM.US.04.M950 ----MF----------D-------....-T-C-V----V----------SR-------FIK----------K----------K---------X. 263
SMM.US.04.M951 ----MY----------D-------....T--T-C-V---V-N--------T----------HN---FI-------WS--------K----------K-------RGNSX 266
SMM.US.04.M952 A---MF----------D-------....-E-T---V----------Q-------FIKC----------T----------K---------X. 263
SMM.US.05.D215 ----M-----------D-------....-L-C-V-----A----------A----N---FIK----------K----------K-------GK... 260
SMM.US.06.FTq ----MM----------E-------....N--T-C-V-----G-------K----------IK---D--D---------K----------K---------X. 263
SMM.US.11.SIVsmE660_FL10 T---MYY---------D-------....T-C-------H----------E---S-K-FIK----------K----------IK---------S- 264
SMM.US.86.CFU212 ----MF----------D-------....-T-C-V----Y----------Q-F-K------Y---------K----------#FK---------X. 262
SMM.US.x.F236_H4 ----MHY---------D-------....T-C-V-----Y----------E---S-K-FIK----------K----------#FK---------S- 263
SMM.US.x.PBJA ----MF----------T----------K---N-K-F-EH------Q----K---Q----------K-------K-S- 262
SMM.US.x.PGM53 ----MF----------SA-------....T---V-----I----------SQ---R---FI-H----------K----------K---------S- 264
SMM.US.x.SME543 T---MF----------D-------....T-C-V-----H----------E---S-K-FIK----------K----------IYK-------R--S- 264
SMM.x.x.pE660.CG7G T---MYY---------D-------....T-C-V-----H----------E---S-K-FIK---------D------K----------IK---------S- 264
STM.US.89.STM_37_16 -----Q----------DM-N----...ED-GT---V-----H----------V-----L--H----F--H----------PK---E----.------K---------S- 265

VII

PLV Proteins

Contents

VII-1	Introduction	**405**
VII-2	Sequences	**406**
VII-3	Alignments	**413**
VII-3.1	Gag	413
VII-3.2	Pol	417
VII-3.3	Vif	424
VII-3.4	Vpx	426
VII-3.5	Vpr	427
VII-3.6	Tat	428
VII-3.7	Rev	430
VII-3.8	Vpu	432
VII-3.9	Env	433
VII-3.10	Nef	440

VII-1 Introduction

The selection of Primate Lentivirus protein sequences for the following alignments was based on the sequences in the complete genome alignment as a starting alignment, and complete or nearly complete genes from other isolates were added if they increased the diversity of samples represented.

The annotation is mainly based on knowledge from HIV-1, and should therefore be taken with caution for HIV-2 and SIV sequences.

VII-2 Sequences

Sequences included in the PLV protein alignments.

Name	Accession	Proteins	Author	Reference
H1B.FR.83.HXB2	K03455	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Wong-Staal, F.	*Nature* **313**(6000):277-284 (1985)
H1O.CM.91.MVP5180	L20571	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Gurtler, L.G.	*J Virol* **68**(3):1581-1585 (1994)
H1N.CM.95.YBF30	AJ006022	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Simon, F.	*Nat Med* **4**(9):1032-1037 (1998)
H1P.CM.06.U14788	HQ179987	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Vallari, A.	*J Virol* **85**(3); 1403-7 (2011)
CPZ.CD.06.BF1167	JQ866001	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Li, Y.	*J Virol* **86**(19):10776-10791 (2012)
CPZ.TZ.00.TAN1	AF447763	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Santiago, M.L.	*J Virol* **77**(3):2233-2242 (2003)
GOR.CM.07.SIVgorCP2135con	FJ424863	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Takehisa, J.	*J Virol* **83**(4):1635-1648 (2009)
MAC.US.x.239	M33262	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kestler, H.	*Science* **248**(4959):1109-1112 (1990)
H2A.SN.86.ST_JSP4_27	M31113	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kumar, P.	*J Virol* **64**(2):890-901 (1990)
H2B.CI.x.EHO	U27200	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Rey-Cuille, M.A.	*Virology* **202**(1):471-476 (1994)
H2G.CI.92.Abt96	AF208027	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Brennan, C.A.	*ARHR* **13**(5):401-404 (1997)

Name	Accession	Proteins	Author	Reference
H2U.CI.07.07IC_TNP3	KC693505	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Ayouba, A.	*AIDS* **27**(15):2488-2491 (2013)
H2U.FR.96.12034	AY530889	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Damond, F.	*ARHR* **20**(6):666-672 (2004)
ASC.UG.10.RT03	KJ461716	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Lauck, M.	*Retrovirology* **11**(1):55 (2014)
ASC.UG.10.RT08	KJ461715	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Lauck, M.	*Retrovirology* **11**(1):55 (2014)
ASC.UG.10.RT11	KJ461714	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Lauck, M.	*Retrovirology* **11**(1):55 (2014)
COL.CM.x.CGU1	AF301156	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Courgnaud, V.	*J Virol* **75**(2):857-866 (2001)
COL.UG.10.BWC01	KF214240	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Lauck, M.	*Retrovirology* **10**(1):107 (2013)
COL.UG.10.BWC07	KF214241	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Lauck, M.	*Retrovirology* **10**(1):107 (2013)
DEB.CM.04.SIVdeb04CMPF3061	FJ919724	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Aghokeng, A.F.	*Infect Genet Evol* **10**(3); 386-96 (2010)
DEB.CM.99.CM40	AY523865	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
DEB.CM.99.CM5	AY523866	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
DEN.CD.x.CD1_CMO580407	AJ580407	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Dazza, M.C.	*J Virol* **79**(13):8560-8571 (2005)

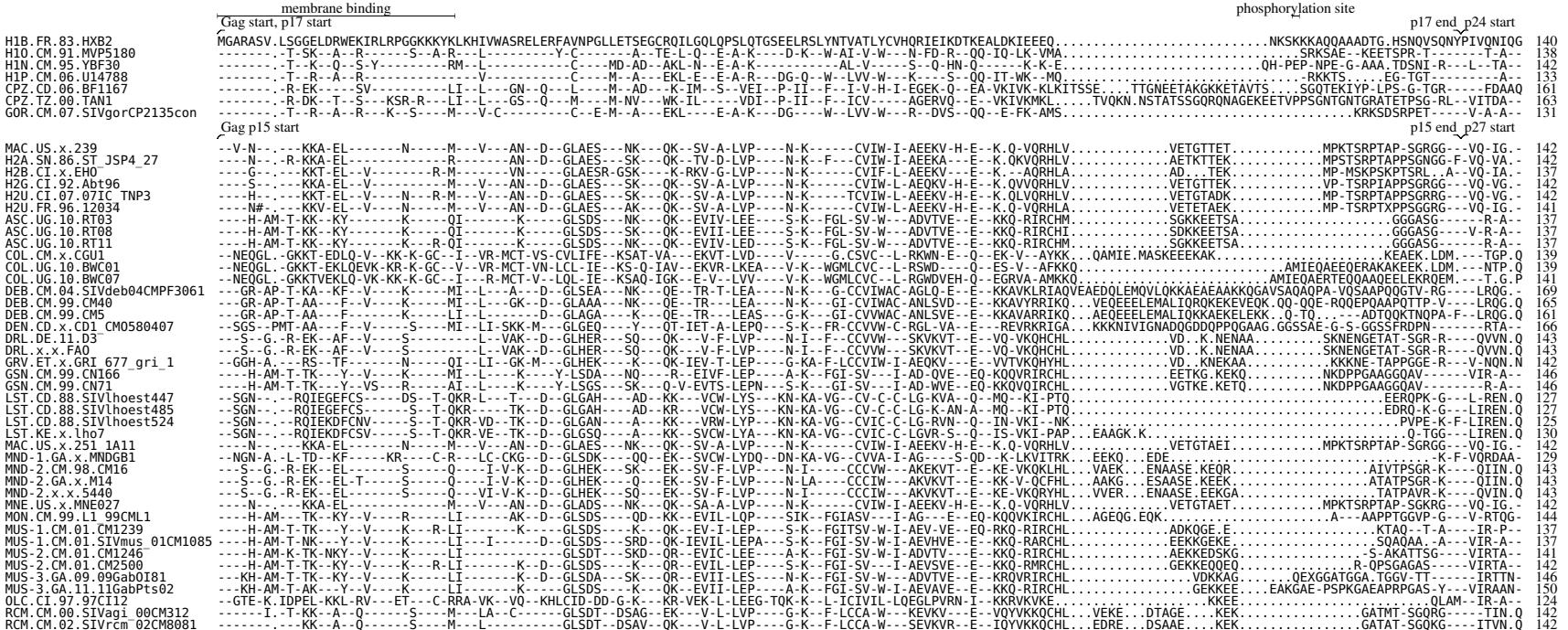
Name	Accession	Proteins	Author	Reference
DRL.DE.11.D3	KM378563	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Dietrich, U.	*ARHR* **31**(3); 353-7 (2015)
DRL.x.x.FAO	AY159321	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Hu, J.	*J Virol* **77**(8):4867-4880 (2003)
GRV.ET.x.GRI_677_gri_1	M66437	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Fomsgaard, A.	*Virology* **182**(1):397-402 (1991)
GSN.CM.99.CN166	AF468659	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Courgnaud, V.	*J Virol* **76**(16):8298-8309 (2002)
GSN.CM.99.CN71	AF468658	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Courgnaud, V.	*J Virol* **76**(16):8298-8309 (2002)
LST.CD.88.SIVlhoest447	AF188114	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.CD.88.SIVlhoest485	AF188115	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.CD.88.SIVlhoest524	AF188116	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.KE.x.lho7	AF075269	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Hirsch, V.M.	*J Virol* **73**(2):1036-1045 (1999)
MAC.US.x.251_1A11	M76764	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Marthas, M.L.	*J Med Primatol* **18**(3-4):311-9 (1989)
MND-1.GA.x.MNDGB1	M27470	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Tsujimoto, H.	*Nature* **341**(6242):539-541 (1989)
MND-2.CM.98.CM16	AF367411	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Takehisa, J.	*ARHR* **17**(12):1143-1154 (2001)

Name	Accession	Proteins	Author	Reference
MND-2.GA.x.M14	AF328295	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Souquiere, S.	*J Virol* **75**(15):7086-7096 (2001)
MND-2.x.x.5440	AY159322	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Hu, J.	*J Virol* **77**(8):4867-4880 (2003)
MNE.US.x.MNE027	U79412	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Kimata, J.T.	*J Virol* **72**(1):245-256 (1998)
MON.CM.99.L1_99CML1	AY340701	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Courgnaud, V.	*J Virol* **77**(23):12523-12534 (2003)
MUS-1.CM.01.CM1239	EF070330	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
MUS-1.CM.01.SIVmus_01CM1085	AY340700	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Courgnaud, V.	*J Virol* **77**(23):12523-12534 (2003)
MUS-2.CM.01.CM1246	EF070329	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
MUS-2.CM.01.CM2500	EF070331	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
MUS-3.GA.09.09GabOI81	KF304707	Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef	Liegeois, F.	*Viruses* **6**(7):2880-2898 (2014)
MUS-3.GA.11.11GabPts02	KF304708	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	*Viruses* **6**(7):2880-2898 (2014)
OLC.CI.97.97CI12	FM165200	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
RCM.CM.00.SIVagi_00CM312	HM803690	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Ahuka-Mundeke, S.	*J Gen Virol* **2010** Dec;91(Pt 12):2959-64 doi: 101099/vir0025767-0 Epub 2010 Aug 25

Name	Accession	Proteins	Author	Reference
RCM.CM.02.SIVrcm_02CM8081[redacted]	HM803689	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Ahuka-Mundeke, S.	*J Gen Virol* **2010** Dec;91(Pt 12):2959-64 doi: 101099/vir0025767-0 Epub 2010 Aug 25
RCM.GA.x.SIVRCMGAB1	AF382829	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
RCM.NG.x.NG411	AF349680	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	*J Virol* **75**(24):12014-12027 (2001)
SAB.SN.x.SAB1	U04005	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Jin, M.J.	*EMBO J* **13**(12):2935-2947 (1994)
SMM.CI.79.SIVsmCI2	JX860430	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.LR.89.SIVsmLIB1	JX860431	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.SL.92.SIVsmSL92A	JX860432	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.SL.92.SIVsmSL92B	JX860433	Gag, Pol, Vif, Vpx, Vpr	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.SL.92.SL92B	AF334679	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Chen, Z.	*J Virol* **70**(6):3617-3627 (1996)
SMM.US.04.G078	JX860415	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.05.D215	JX860413	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)

Name	Accession	Proteins	Author	Reference
SMM.US.06.FTq	JX860414	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.86.CFU212	JX860407	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Fischer, W.	*J Virol* **86**(24):13217-13231 (2012)
SMM.US.x.H9	M80194	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Courgnaud, V.	*J Virol* **66**(1):414-419 (1992)
SMM.US.x.PGM53	AF077017	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Novembre, F.J.	*J Virol* **72**(11):8841-8851 (1998)
SMM.x.x.pE660.CG7G	JX648292	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Lopker	Unpublished
STM.US.89.STM_37_16	M83293	Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef	Novembre, F.J.	*Virology* **186**(2):783-787 (1992)
SUN.GA.98.L14	AF131870	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Beer, B.E.	*J Virol* **73**(9):7734-7744 (1999)
SYK.KE.x.KE51	AY523867	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
SYK.KE.x.SYK173_COMGNM	L06042	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Hirsch, V.M.	*J Virol* **67**(3):1517-1528 (1993)
TAL.CM.00.266	AY655744	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	*Virology* **349**(1):55-65 (2006)
TAL.CM.01.8023	AM182197	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	*Virology* **349**(1):55-65 (2006)

Name	Accession	Proteins	Author	Reference
TAN.UG.x.SIVagmTAN1	U58991	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Soares, M.A.	*Virology* **228**(2):394-399 (1997)
VER.DE.x.AGM3	M30931	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Baier, M.	*Virology* **176**(1):216-221 (1990)
VER.KE.x.9063	L40990	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Hirsch, V.M.	*J Virol* **69**(2):955-967 (1995)
VER.KE.x.AGM155	M29975	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Johnson, P.R.	*J Virol* **64**(3):1086-1092 (1990)
VER.KE.x.TYO1_patent	DJ048201	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Omori, T.	Patent: WO 2007049749-A 13 03-MAY-2007; DNAVEC CORPORATION
WRC.CI.97.97CI14	AM745105	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
WRC.CI.98.98CI04	AM713177	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
WRC.GM.05.Pbt_05GM_X02	AM937062	Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef	Locatelli, S.	*Virology* **376**(1):90-100 (2008)
MON.NG.x.NG1	AJ549283	Pol, Vif, Vpr, Vpu, Env, Nef	Barlow, K.L.	*J Virol* **77**(12):6879-6888 (2003)



Draft of 30 September 2015, not for public release
plvprot_main.tex r8876–10546

```
                                                                                                                          CyPA binding                                                                       major homology region
H1B.FR.83.HXB2             QMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVH.AGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTN.....NPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTL  304
H1O.CM.91.MVP5180          -----------------A------N--I----M------V-Y-I-----AI----G-L-V---V----------T--PA.M--LP---I---T----------Q----I-T-R....GANS----D--RK--V-----M-K----V-----------------------  303
H1N.CM.95.YBF30            -----PLT----------I-------------M--------S----------------------V------D---T---P.V--LP---L-D--------------A--VA---A.....---V---D--R---V----R-------V---E-K-----------------  306
H1P.CM.06.U14788           --T-MPL----------A------N--I----M------I-D-I-----A-----G-L-V---V----------T---P.V--LP---L---T-G--------K----T---R.....-N-V---D--RK--V-----V-K--C-V-----K-----------------  297
CPZ.CD.06.BF1167           -PR--P----------T----N-K----------A---I-H-------AI-D--G---V---V------D---L--T-.---LV--TL---T----------IV---Q-V-T..PQNQGGV---D-----------L-K----V----------------------I  328
CPZ.TZ.00.TAN1             VAR--P---------R-I---G-N---------------Y---S---A--E----------V----------A--A-.---QQA-ML---T-A-------------VL---T..PQAQGGV---D-----------L------V-----K-----------------I  330
GOR.CM.07.SIVgorCP2135con  -----PL----------A------N--I----M-----SI-Y-I-----AI----G-L-V---V----------L--AV.V--LPV--L---S-G--------Q---VA-I-R....N-N-V---D--RK--V-----V-K--C-V-----K-----------------  296

MAC.US.x.239               NY--LPL----------LI---K-GA--V-G-Q-----C--Y-I-Q---C--D------IIRDI------D--LQ--QP..A-.QQ--L---S---------SVD---Q--YR....QQN-----N--R---Q---Q-C----N--N---VK-------QS-------S-  305
H2A.SN.86.ST_JSP4_27       NYT-VPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--D------IIR-I------D--AQ--IP..--LPA--L-------------VE---Q--FR....PQN-V---S--R---QI--Q-C----N--N----K-------QS-------S-  306
H2B.CI.x.EHO               NYS-LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--QQ--SP..--MPA--L-------------VE---Q--YR....PQN-V---N--R---Q---Q-C----N--N----K-------QS-------S-  301
H2G.CI.92.Abt96            NY--LPL----------L----K-GA--V-G-Q-----CL-Y-I-Q---C--E------IIR-I------X--LQ--QQ..--PPA--X---T----------IE---Q-THR....QQN-----N--R---Q---Q-C----N--N---VK-------QS-------S-  306
H2U.CI.07.07IC_TNP3        NY--LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--IQN-QP..--LPA--L-------------VE---Q--YR....QQN-----N--R---Q---Q-C----N--N---VK-----S-QS-------S-  306
H2U.FR.96.12034            NYA-LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I--D---D--TQ--NQ..--PPA--L-------------VE---Q--YR....PQN-----N--R---Q---Q-C----N--N----K-------QS-------S-  305
ASC.UG.10.RT03             -HQ---L---I-KT-ITTI---K--A--VAL-Q-----CI-Y-I-QLM-AI-E--G-V-LI-DVV-----A--QL--QP.-Q-QPNAGL-Y-S------VS-SI---LE---R....AVD-VN-----RK--VM--QRC-K--N-VN----------S-K--------C-  302
ASC.UG.10.RT08             -HQ---L---I-KT-ITTI---K--A--VAL-Q-----CI-Y-I-QLM-AI-E--G-V-LI-DVV----HT--QL--QP.-Q-QPNAGL-Y-S------VS-SI---LE---R....AVDQVN-----RK--VM--QRC-K--N-VN----------S-K--------C-  302
ASC.UG.10.RT11             -HQ---L---I-KT-ITTI---K--A--VAL-Q-----CI-Y-I-QLM-AI-E--G-V-LI-DVV-----N--QL--QP.VQ-QPNAGL-Y-S------VS-SI---LE---R....TVDQVN-----RK--VM--QRC-K--N-VN----------S-K--------C-  302
COL.CM.x.CGU1              GP-.HQPLSPRTLGAWVKCV-GGIA-SLA---L-Y-T--IAY-M-L---ILDT--GFP-V--DE--KK-E-Y-LL---Q..Q-QQQ-AL-Q-TA---T-N--SVA--VA-GE..........-IAN--RG--VQS-E-VIQIAR-S-V------S--D-KS------SA-  296
COL.UG.10.BWC01            GPA.HQPLSPRTL-AWVKCV-GGIA-SLA---L-Y-T--IAY-M-L---ILDT--GFL-V--DE--KK-E-Y-LQN--P..Q-QQQ--L-Q-TA---T-N--SVA--VA-GE..........--AN--RG--VQS-E-VIQIAR---V------S--D-KS------TA-  296
COL.UG.10.BWC07            -GPMHSPLSPRTL-AWVKCV-QGITASLA---L-Y-T--IAY-M-L---ILDT--GFL-V---EL-KK-E-Y-LT---P..Q-QQQ-AL-Q-TA---M----SVA---A-GE..........--AN--RG--VQS-E-VIQIAR-S-V------A--D-KS-----FSA-  299
DEB.CM.04.SIVdeb04CMPF3061 -Y--MPL---IVKT-INA----K----IV-L-QV-A--C--Y-I-GL--AI-DL-G---II-DV---------LQ--QQ..Q-.PQ--L---S-A-----N--VE---A---RAQGVGGG--D--Q--R--V----QRC-K--N--N---VK-------K----------  336
DEB.CM.99.CM40             -F--MPL---IVKT-INA----K----IV-L-QV-A--C--Y-I-GL--AI-DL-G---II-DV---------LQ--QQ..Q-.PQ--L---S-A-----N--VE---A---RPAGQGQG--D--Q--R--V----QRC-K--N--N---VK-------K----------  332
DEB.CM.99.CM5              -Y--MPL---IVKT-INA----K----IV-L-QV-A--C--Y-I-GL--AI-DL-G---II-DV---------LQ--QQ..Q-.PQ--L---S-A-----N--VE---A---RPAGQGAG--D--Q--R--V----QRC-K--N--N---VK-------K----------  328
DEN.CD.x.CD1_CMO580407     -YQ---L-T-V-KT-I-E--DRK----IV-L--V-A--LI-Y-I-QL--AI-E--G---V--DV-----HD--LA--ANLQQ-LP-AAQ-D-----------SIED--T---R....QQG--N--D-----V-M--Q-V-K--Q-V-----K-----S-K---------A-  332
DRL.DE.11.D3               -A----------------I---K--A--V---L-----CISY-I-Q---AI-D--G-L-IV-DI------D---R--QV..--LPQ-VL-N-S---------SIE---E-T-R....AQDS-N--A--RQ-VV---QRC-T--N-VN---VK-------K---------A-  307
DRL.x.x.FAO                -A----------------I---K--A--V---L-----CISY-I-Q---AI-D--G-L-IV-DI------D---R--QV..--LPQ-VL-N-S---------SIE---E-T-R....AQDS-N--A--RQ-VV---QRC-T--N-VN---VK-------K---------A-  307
GRV.ET.x.GRI_677_gri_1     AW---PL----------C----RWGA--V---Q-----CLSY-V-Q---VI-D--G-L-I---V----------T-RPP.---LPA--L-D-T---------SI----E-TF-....A--R-D--AQ-RK-V----Q-V-Q--N-QKV----------Q---------A-  307
GSN.CM.99.CN166            -YQ---LNS-I-KT--SI----K-A--TVAL-Q-----CI-Y-M-Q---AI-DY-G-V-II-DV---Q-----LL---P.-A-QPVAGL-D-S-A----V---PN---E-I-R....Q-Q-VN-SD--RK------QRC-K--N-VN----K-------KE-----F-C-  311
GSN.CM.99.CN71             -FQ---LNS-V-KT--SI----K-A--TVAL-Q--T--CI-Y-M-Q---A--DY-G-V-II-DV---Q-----LL--QP.-A-QPVAGL-D-S-A----V---PN---E-I-R....Q-N-VN-AD--RK------QRC-K--N-VN--G-K-------KE-----F-C-  311
LST.CD.88.SIVlhoest447     RW--TPL----IQT---I--DRGWK--TVA-----T-K-L-D---V---AI-D--G---II-DH-V--G-----K--QQ.QPAQPG-GL-T-NAT-V--V---VE--LA-T-.....ADT-VD--K---E-V-QAME-V--IHQ-V-VI--K-------KE-A---F-A-  291
LST.CD.88.SIVlhoest485     RW--TPL----IQT---I--DRGWK--TVA-----T-K-L-D---V---AI-D--V---II-DH-V--G-----T--QQ.QPAQPG-GL-T-NAT-V--V---VE--LA-T-.....ADT-VD--K---E-V-QAME-V--IHQ-V-VM--K-------KE-A---F-A-  291
LST.CD.88.SIVlhoest524     RW--TPL----IQT---I--DRGWK--TVA-----T-K-L-E---V---AI-D--G---II-DH-A--G-----Q--QQ.QPAQPG-GL-T-NAT-V--I---VE--QA-T-.....ADT--D--K---E-V-QAME-V--IHQ-V-VM--K-------K--A---F-A-  289
LST.KE.x.lho7              RW--TPL----IQT---I--DRGWK--TVA-----T-K-L-D---V---AI-D--G---II-DH-V--G-----Q--QQ.QPAQPG-GL-T-NAT-V--V---VE--LA-T-.....ADT-VD--K---E-V-QAME-V--IHQ-V-VM--K-------KE-A---F-A-  294
MAC.US.x.251_1A11          NY--LPL----------LI---K-GA--V-G-Q-----C--Y-I-Q---C--D------IIRDI------D--LQ--QP...APQQ--L---S---------SVD---Q--YR....QQN-----N--R---Q---Q-C----N--N---VK-------QS-------S-  305
MND-1.GA.x.MNDGB1          -YQYTP----IIQT---T----KWK-----L----T---ISH---I---A--D--G---V--DV---Q-----LT--QQ.QPAQPG-GL-T-S-----------VE--LA--NM....QQNA-N--T---S-----M-RL-KSHC-I--T-V------A-K-------NVM  294
MND-2.CM.98.CM16           TP---G----------CI---K----IV---I-----CL-Y---G---AI-D--G-L-IV-DV------D-NLR--QV..--MPQ-VL-N-T---------SIE---E-T-R....EQDAVN--G---Q--V---Q-C-S--N-VN----K-------K---------A-  307
MND-2.GA.x.M14             TP---G----------CI---K----IV---I-----CL-Y---G---AI-E--G-L-IV-DV------D--LR--QV..--LPQ-VL-N-T---------SIE---E-T-R....QQEQVN--A---Q--V---Q-C-S--N-VN----K-------K---------A-  307
MND-2.x.x.5440             TP---G----------CI---K----IV---I-----CI-Y---G---AI-D--G-L-IV-DV------D--LR--PV..--MPQ-VL-N-T-N-------SIE---E-T-R....QQDQVN--G---Q--V---Q-C-S--N-VN----K-------K---------A-  307
MNE.US.x.MNE027            NYT-LPL----------LI---K-GA--V-G-Q-----CI-Y-I-Q---C--E------IIR-I------D--LQ--QQ...APQQ--L---S---------VD---Q--YR....QQN-----N--R---Q---Q-C-K--N--N-----------QS-------S-  305
MON.CM.99.L1_99CML1        GFQ---VE--L-KT--Q-I---K-A---VAL-Q-----MI-Y-I-QL--AI-D--G-L-II-DV------N--L---QP.PQ-QPNAGLGD-T-A----VS--P-Q--E-I-R....A-N-VQ-SD--RK-V----QRC-K--N-VN----K-------K---------C-  309
MUS-1.CM.01.CM1239         -FQ--SLNA-I-KT--SI----K-AA--VA--Q--A--TI-Y-M-QL--AI-D--G-I-II-DV---Q-----LL--QP.QQ-QPNAGL-S-S------V---PN---E-I-R....Q-D--N-AD---K---M--QRC-K--N-V-----K-------K---------C-  302
MUS-1.CM.01.SIVmus_01CM1085 -FQ---LNA-I-KT--SI----K-AA--VA--Q--A---I-Y-I-QL--AI-E--G-I-II-DV---Q-----LLN-QP.QQ-QPNAGL-N-T------VS--P----E-I-R....A-N--N-A----K-V-M--QRC-K--N-VN----K-------K------F-C-  302
MUS-2.CM.01.CM1246         -YQ--SL---L-KT--ST----K-A---VAL-Q-----CI-Y-V-Q---AI-D--G-I-LI-DVV--Q--DR-VL--QP.TQ-QPNAGL-Y-S-A-V--VS--PA---E---R....QQD-VN--N--RK------QRC-K--N-VN----K-------K------F-C-  306
MUS-2.CM.01.CM2500         TYQ--SL---I-KT--STI---K-A---VAL-Q-----CI-Y-V-Q---A--D--G-V-LI-DVV--Q-----LL--QP.-Q-QPNAGL-Y-S-A---VS--PA---E---R....QQD-VN--D--RK------QRC-K--N-VN----K---R---K------F-C-  307
MUS-3.GA.09.09GabOI81      -FQ---LE--L-KS--ST----K-A---VAL-Q--T--LI-Y-M-Q---AI-E--G-L-II-DV---Q--D--LL--QP.-Q-QPNAGL-S-S------VS--PA---E-I-R....Q-E--N-SD--RK------QRC-K--N-VN-----------KE--------C-  311
MUS-3.GA.11.11GabPts02     -YQ---L-T-V-KS--NTI---K-A--TVAL-Q-----LIAY-M-Q---AI-E--G-L-II-DV---Q--D--ML--QP.-Q-QPNAGL-S-S------VS--PA---E---R....Q-D--N-S---RK---V--QRC-K--N-VN------------K---------C-  315
OLC.CI.97.97CI12           IFE-VPLAT---QS---AL-DRG-KS-IV-L---FA--LL-EEI--L--V--DY-G---V---A-----E-Y--Q--QP.I--LPV--IPP--A---V----N--Q-LA-VNA....G...-D--AL----V-I--Q-L-K----VTV------A----S------F-A-  286
RCM.CM.00.SIVagi_00CM312   -PE--P---------------K-GA--V---------CI-Y-V-Q---AI-E--G-L-IV--V--D---D--LR--NP..--LPA--L---T---------NIA---A-T-R....A-N--A--N--RN--V---Q-C-K--N-VN----K-------K---------A-  306
RCM.CM.02.SIVrcm_02CM8081  -PE--P---------------K-GA--V---------CI-Y-I-Q---AI-E--G-L-IV--V------D--AR--NL..--LPA--L---T---------SVA---A-S-R....A-N--I--N--RN------Q-C----N-VNV---K-------K---------A-  306
RCM.GA.x.SIVRCMGAB1        -PE-NP---------------K--A--A---------CI-Y-I-Q---AI-E--G-L-IV--V------D--AR---P..---PA--L---T---------SIA---A-T-R....A-N--N--NL-RN---V--Q-W-K--N-VN----K-----S-K---------A-  306
RCM.NG.x.NG411             -PE--P---------------K-GA--V---------CI-Y-V-Q---AI-E--G-L-IV--V--D---D--LR--NP..--LPA--L---T---------NIA---A-T-R....A-N--A--N--RN--V---Q-C-K--N-VN----K-------K---------A-  306
SAB.SN.x.SAB1              -W---PL-----------I---K--A--V------A---I-Y-I-Q---A--E--G-L-IV-DV------D--LR--PP.QQ-P-Q-VL-D-Q----------I----E-T-R....AQNAVN--N---G------Q-C-K--N-VN----K-------K---------A-  318
SMM.CI.79.SIVsmCI2         NYT-LPL----------LI---K-GA--V-G-Q-----C--Y-I-Q---C--E------II--I------D--LQ--NV..--LPA--L---S----------VD---Q--YR....QQN-V---N--R---Q---Q-C----N--N---VK-------QS-------S-  306
SMM.LR.89.SIVsmLIB1        NY--LPL-------*--L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D*-LQ--QP..S--PA--L--------------VD---Q--YR....QQN-----N--R---Q-R-Q-C----N--N---VK-------QS-------S-  304
SMM.SL.92.SIVsmSL92A       NY--LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--LQ--QP..--LPA--L--------------VE---Q--YR....AAN-----N--R---Q---Q-C----N-VN----K-------QS-------S-  306
SMM.SL.92.SIVsmSL92B       NY--TPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-M------D--LQ--RGQQPAQPA-GL---S---------NPN---E--YR....AQN-V---D--R---Q---Q-C----N--N---VK-------QS-------S-  306
SMM.SL.92.SL92B            NY--TPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--LQ--RGQQPAQPA-GL---S----------PS---E--YR....AQN-V---D--R---Q---Q-C----N--N---VK-------QS-------S-  306
SMM.US.04.G078             NYT-LPL----------L----R-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--LQ--QP..--LPA--L--------------VD---Q--YR....QQN-----N--R---Q---Q-C----N-MN---VK-------QS-------C-  306
SMM.US.05.D215             NYI--PL----------L----K-GA--V-G-Q-----C--Y-V-Q---C--E------IIR-I------D--LQ--QP..--LPA--L---T----------VD---Q--YR....QQN-V---N--R---Q---Q-C----N--N----K----K--QS-------S-  306
SMM.US.06.FTq              NY--LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--D------IIR-I------D--LQ--QP.M---PA--L-D------------VE---E--YR....QQN-----N--R---Q---Q-C----N--N---VK-------QA-------S-  307
SMM.US.86.CFU212           NY--LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--LQ--QP..---PA--L-D------------VE---Q--YR....QQN-----N--R---Q---Q-C----N--N---VK-------QS-------S-  306
SMM.US.x.H9                NYT-LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--LQ--QP..---PA--L--------------XD---Q--YR....QQN-----N--R---Q---Q-C----N--N---VK-----S-QS-------S-  306
SMM.US.x.PGM53             NYT-LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--LQ--QP..---PA--L-D------------VE---Q--YR....QQN-----N--R---Q---Q-C----N--N---VK-------QS-------S-  306
SMM.x.x.pE660.CG7G         NY--LPL----------L----K-GA--V-G-Q-----C--Y-I-Q---C--E------IIR-I------D--LQ--QP..--LPV--L--------------VD---Q--YR....QQN-----N--R---Q---Q-C----N--N---VK-------QS-------S-  306
STM.US.89.STM_37_16        NY--LPL----------L----K-GA--V-G-Q-----C--Y-I-QL--C--E------II--I------D--MQ--QP.P--LPA--L---S---------SPE---Q--YR....QQN-----N--R---Q---Q-C----N-VN---VK-----T-QS-------S-  307
SUN.GA.98.L14              NYI--PL----VQT---I----GWK--TVA--A--MT--I-E-I-V---AI-E--G----I-DH-V--S-----M--QQ.QPAQPG-GL-T-N-T----V---VE---Q-IYG..LNGATRVA-QD---G-V-E-MERV--LHQT--VVE------------T---F-A-  294
SYK.KE.x.KE51              -WI-VGVNT---KT--EA-NS-K-DASLV-L-QI-T--FI-Y---D---AI-D--G---VI-DV----GS---LQ--QP.QQQQPVAGL-D-SA---------I----E-I-R....Q-N--Q--Q--RQ------Q-C-QV-N-VN----------T-K-------HC-  327
SYK.KE.x.SYK173_COMGNM     -WI-TGVPV---KT--EA-NS-K-DASIV-L-Q--T--FI-Y---G---A--D--G---VI-DI----G-A--LE--QP.QQ-APQAGL-D-SA--------SIA---E-I-R....Q-N-VQ-----R-------Q-C-QV-N-V-----K-----I-K-------HC-  330
TAL.CM.00.266              -YQ---L---L-KT-IATI---K-A--TVAL----A--CI-Y-I-Q---A--E--G---II-DI---Q-S---QM--Q..I--LPA--L---A----------PA--VE-T-R....AQN-VD--A-----V-M--QRC-K--N-VN---VK---R---K-------RC-  309
TAL.CM.01.8023             -YQ---L---L-KT-IATI---K-A--TVAL----A--CI-Y-I-Q---A--E--G---II-DI---Q-----QM--Q..I--LPA--L---A----------PA--VE-T-R....QQN-VD--S-----V-I--QRC-K--N-VN---VK---R---KE------RC-  313
TAN.UG.x.SIVagmTAN1        -W--TPL----------T----R-GA-IV---Q-----CLSY-I-Q---VI-D--G---II--V--D---Q--IT--PP.---LPA--L-D----------SVA---E-TF-....A--RVD--R--RG-V----Q-C-K--N-I-V------A----K-------QA-  313
VER.DE.x.AGM3              AWI-VPL----------A----K-GA-IV---Q-----C--Y-I-Q---VL-D--G-L-IV--I------Q--IA--PP.---LPA--L-D------------V---LE-IYT....A--RVD--A--R-------Q-C-K--N-V-V---------A-K---------AI  317
VER.KE.x.9063              AW--VPL----------A----K-GA-IV---Q-----C--Y-I-Q---VL-D--G-L-IV--I------Q--VT--PP.---LP---L-D------------V---LE-IYT....A--RVD--A--R-------Q-C-K--N-V-V-----------K---------AI  313
VER.KE.x.AGM155            AW--VPL----------A----K-GA-IV---Q-----C--Y-I-Q---VL-D--G-L-IV--I------Q--VT--PP.---LPA--L-D-G----------V---LE-IYT....A--RVD--A--R-------Q-C-K--N-V-V-----------K---------AI  313
VER.KE.x.TYO1_patent       AW--VPL----------A----K-GA-IV---Q-----C--Y-I-Q---VL-D--G-L-IV--I------Q--VT--LP.---LPA--L-D-----------SV---LE-IYT....A--RVD--A--R-------Q-C-K--N-V-V-----------K---------AI  313
WRC.CI.97.97CI14           GPQWDPL----IQT---HI-DRG-GA-SV-F-A--TDR-LSH-I--L--NI-D------II-DIV---IQ----T--QA.----------Q-T------V-ASVEDELA-LAQ.....---VD--S-----VLR--Q-V-QI-V-V-----K-----D-KA-------V-  287
WRC.CI.98.98CI04           GPQWDPL----IQT---HI-DRG-GA-CV-F----TDR-LAH-I--L--NI-D------II--IV---IQD---T--QP.----------Q-T------V-ASVEDELA-LAQ.....----D--AV----VLR--Q-V-QI-V-V-----K-----D-KT-------V-  291
WRC.GM.05.Pbt_05GM_X02     GPQWEPL----IQT---HI-DRG-GA-CV-F----TDR-LTH-I--L--NI-D------IV--IV---IT----T--PP.V--LP--A--N-T------V---ME-ELA-LAQ.....Q---DI-A-------R--Q-V-QI-V-------K-----D-KS-----F-V-  289
```

```
                                                                                                                  p24 end p2 start        p2 end p7 start                      Zn motif                                                  Zn motif
H1B.FR.83.HXB2             RAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNSAT.....IMMQRGNFRNQR.............KIVKCFNCGKEGHTARNCRAPRKK..........................GCWKCGKEGHQMKDCT..  427
H1O.CM.91.MVP5180          -----T---------------S-----Q-------E-------V--------T---KI-----ASAQQDLKGGYTAVF----QNP-RK.............GPI----------I-K-------R..........................------Q--------KN.  432
H1N.CM.95.YBF30            -----T--------------------QL------G-----------------A-------------QQPT-S....VFA-----KGI-.............-PI----------L----K---RG..........................------Q--------KN.  431
H1P.CM.06.U14788           ----------S---D-----------Q-------G-------N--------T----------AAANQASQELK.GGYTTVFMQSG--.............-P--------V--I-K--K---RR..........................------Q--------KS.  425
CPZ.CD.06.BF1167           -------P-------------------------G-------L---------A----------Q-TQQNVNMVQG..RAAPQGK-GA..............GNI-------I------------R-..........................------Q----L----..  453
CPZ.TZ.00.TAN1             -------P--T---------------H------QG------L---------S---KI-----ASA-AGG.....VN-L-.-GK-PPLK.............KGQLQ------V--------------..........................---R--Q---------TR  456
GOR.CM.07.SIVgorCP2135con  -------D---------I--------Q-------G-------N--------T----------AAAHSDLK.GGYTAVFMQKGGX--GI..........KKGP--------V--I-K------R-..........................------Q--------RS.  427
                                                                                                                  p27 end p2 start        p2 end p8 start
MAC.US.x.239               ----TDAA------Q---I--------LV--G--VNP-----L----------Q---LM---LKEALAPVPIP..FAAA-QRGP................R-PI--W-------S--Q-----RQ..........................-------MD-V-AK-P..  428
H2A.SN.86.ST_JSP4_27       ----TDPA------Q------------LV--G--INP-----L----------Q---LM---LKEAMAP-PIP..FAAA-Q...................RRTI--W-------S--Q-----RQ..........................-------A--I-AK-P..  426
H2B.CI.x.EHO               ----TDPA------Q---I--------LV--G--MNP-----L-----I----Q---LM---LKEALTPS-NP..FAAA-PRAG...............KRT-T-W----A-----Q-K---RQ..........................------QQ--I-SK-P..  424
H2G.CI.92.Abt96            ----TDPA------Q---I--------LV--G--MNP-----L-----X----Q---LM---LKEAL-PTALP..FAAA-QKTGG.............KRSTI--W--------V-Q-----RQ..........................-------P--I-AK-P..  431
H2U.CI.07.07IC_TNP3        ----TDPA------Q---I--------LV--G--MNP-----L----------Q---LM---LKEALAPGPLP..FAAV-QRTGG.............KKSMI--W-------S--Q-----RQ..........................-------P--I-AK-P..  431
H2U.FR.96.12034            ----TDPA------Q---I--------LV--G--MNP-----L-----I----Q---LM---LKEALQPMPIP..FAAV-QRGGP.............RRP-R-W----------Q-K---RQ..........................-----E-P--N-AS-P..  429
ASC.UG.10.RT03             -----D-S------T---------E--A----M.-G-S----L---------HY-TKL-----ASAF.QK-GNYNMVQGP--PKGAR-GPM........KNPNPR-----QF--MSKQ-PH-K-I........................K-F---R---MARQ-RAS  435
ASC.UG.10.RT08             -----D-S------S---I-----E--A----M.-G-S----L---------HY-TKL-----ASAF.QR-GNYNMAQGPQ-PKGAR-GPM........KNPNPR-----RF--MS-Q-PQ-K-M........................K-F---R---MARQ-RAS  435
ASC.UG.10.RT11             -----D-S------T---------E--A----M.-G-S----L---------HY-TKL-----ASAF.QK-GNYNMVQGP--PKGAR-GPM........KNPNPR-----QF--MS-Q-PQ---I........................K-F---R---MARQ-RAS  435
COL.CM.x.CGU1              ---P-AG-I-A--ANN--I-H------R---G-.QKPS--D-LA--------D---K------Q-FQQER.....TN-IE...................V-TA-----QGI--L--M-PKRPIGGAGR......GRGRGRGGFRGAPRRPVR-FT-NQ---MQR--P..  432
COL.UG.10.BWC01            ---P-GS-I-Q--ANN--I-H---E--R---G-.QKPS--D-LA--------DY--K------Q-FRQE.......T-NMIE.................V-TA---Q--EI--LQKQ-PKRGAGGFRGGGMRGRGRGGRGAWGRGRNPGPLK-FG--QL--IQR--Q..  438
COL.UG.10.BWC07            ---P-GG-I-Q--ANN--I-H---E--R---G-.QKPS--D-LS--------DY--K------QGMQKIGLVE.........................VSTAR-----QV--L----PK--QFPTRG...RGSSRGRGIMRGGMRRGGQAK-FG--QL--IQAT-P..  436
DEB.CM.04.SIVdeb04CMPF3061 -----D-A------T--MI-------R--------NP-----LH--------QQ---LM----VSAMREGGG...LT-V-QR.GPGGP..........RPGNRIR-Y---QI--VQKD--K--TV........................R-F--QQT--MA-N-P..  463
DEB.CM.99.CM40             -----D-A------T--MI-------RI---G--QNP-----LH--------QQ---LM----ASALKE-GS...LG-V-QRRGPPGS............RRQIR-----QI--LQKD-KR---T........................K-F---Q---IA-N-G..  458
DEB.CM.99.CM5              -----D-A------T--MI-------RI---G--QNP-----LH--------QQ---LM----ASALKE-GS...LG-V-QRGGQGGG...........PRRQLR-----QI--VQ-D-KK---V........................K-F-------IA-N-G..  455
DEN.CD.x.CD1_CMO580407     -----D-S--L---TS------------Q--E----NP-----LQ---------Q---PT---LASAFQKQGLCHGSRKGAQPPVER.............RGLR-----QI--L-KD--K-KRVT...................QPGA-F----L---A-N-R..  462
DRL.DE.11.D3               ---HTDAA------Q---I--------V---G--TNPS----LL---------------M----REQQTA...........QNPP-GPP.....RGPRGPPPRNPR-P---QY---L-Q-TK----..........................--FR--ALD-MLRN-...  435
DRL.x.x.FAO                ---HTDAA------Q---I--------V---G--MNPS----LL---------------M----REQQAA...........QNPP-GPP.....RGPRGPPP-NPR-P---QF---L-Q-TK----..........................--FR--ALD-MLRN-...  435
GRV.ET.x.GRI_677_gri_1     -----P-D------Q---I--------L---G--MNP-----LI--------Q---KLMV-M--NGQ.........N-V-V-PQKKGP..........RGPL-------F--MQ-E-K---QI........................K-F----I--MA---K..  427
GSN.CM.99.CN166            -----D-A------QS--------E--L----M.-G-N----L---------T--S-L-----ATALKGTSS...YN-V-K-PPGR-.............G-TP--Y---QF--L--D-PK-KER........................K-F---RA--FS-P-R..  435
GSN.CM.99.CN71             -----D-A------QL--------E--L----M.-G-S----L---------T--S-L-----ATALKG.SS...YN-V---PPGR-.............G-TP--Y----F--I--D-PK-KER........................K-F----A--LARQ-KT.  435
LST.CD.88.SIVlhoest447     ---GG-H---E--K-KM---------RMVI----EG-------R--------A--G-I-----ATAMQ-QMR...QN-V-VTPP--AQ.GRFVRTGGGGPR-PLT------P-----M--Q--QE..........................---N--SKE-RFAQ-PKP  430
LST.CD.88.SIVlhoest485     ---GG-H---E--K-KM---------RMVI----EG-S--D--K--------A--GKI-----ATAMQ-QMR...QN-V-VTPP--EQ.GRFVRTGGGGPR-PLT------P-----M--Q--QE..........................---N--SKE-RFAQ-PKP  430
LST.CD.88.SIVlhoest524     ---GG-H---E--K-KM----------MVI----EG-------K--------A--RKI-----ATAMQGQMK...HL-V-QTPP--AQ.GRFVRTGGGGPRRPLT------P--S-KM--Q--QE..........................---N--AKD-R-AQ-PKP  428
LST.KE.x.lho7              ---GG-H---E--K-KM---------RLVI----EG-S-----K--------A--GKI-----ASAIQQQVR...QN-V-VTPL--AQ.GRFVRTGGGGPR-PLT------P-----M--Q--QE..........................---N--SKE-RFAQ-P..  431
MAC.US.x.251_1A11          ----TDAA------Q---I--------LV--G--VNP-----L----------Q---LM---LKEALAPVPIP..FAAA-QRGP................R-PI--W-------S--Q-----RQ..........................-------MD-V-AK-P..  428
MND-1.GA.x.MNDGB1          ------G---M--QQH--IE----E--Q--RS--KG------LE---------Q----LM--M-RT-VGQ..S...QNFV-QRGPQR.......GPVRQPTGRKPI----N----V--FFK---R-..........................---N--AMD--KAQ-P..  423
MND-2.CM.98.CM16           ---RTDPQ--T---Q---I--------S---G--MNPS----PL--------KY--QMM----KEAQS.......AV---NSGGPPRG.......PPRQPPRNIR-P----F--GL-D-IS----..........................--F---DL--I-RN-P..  434
MND-2.GA.x.M14             ---RTDPQ--T---Q---I--------S---G--MNPS----LL--------KY--QMM----QKMQS.......EV---NSGGPPRG.......PPRQPPRNPR-P---RY--VL-D--L----..........................--F---DT--M-RN-P..  434
MND-2.x.x.5440             ---RTDPQ--T---Q---I--------AT--G--MNP-----LL--------KY--QMM----Q-AQAA.......V---NSGGPPRG.......PPRQPPRNPR-P----F--VL-D------R..........................--F---DP--L-RN-...  433
MNE.US.x.MNE027            ----TDPA------Q---I--------LV--G--MNP-----L----------Q--KLM---LKEALAPGPLP..FAAA-QKGP................R-PI--W-------S--Q-----RQ..........................------QM--V-AK-P..  428
MON.CM.99.L1_99CML1        ----TD-A------Q------------L----M.-G-S--D-LQ--------M--S-I-----AGAIAN....MPMN-V-ARGPPQ.............RKGQP-------F--M-K--K--QRR........................K-YN--QP--LA---PQP  434
MUS-1.CM.01.CM1239         ----SD-A------S---------E--L---SM.-G------LQ--------A--S-LM---ASAL.RQNS.QL.N-V-G...ARGK.....GSQGGPRGNPR-Y---QF--M--S-PK-KTR........................K-F---R---LA-Q-R..  430
MUS-1.CM.01.SIVmus_01CM1085 ----SD-A------S---------E--L---SM.-G------LQ--------T--S-LM---LASAF.KQAG.TL.N-V-GAKGARG-.....GNGSQGNRGNPR-Y---QY--V--D-PK--N-........................K-F---R---LARQ-R..  434
MUS-2.CM.01.CM1246         -----D-A------QS--------E--LV---M.-G---K--LQ----IE--M--S-LM---TNAL.RQNTINTIN-V--..QSPR......GVMGKKRENSTR-Y---QF--L--D-PK-KST........................R-F-------LARQ-R..  437
MUS-2.CM.01.CM2500         -----D-A------QS--------E--LV---M.-G------LQ--------I--SKLM---ASAL.RQNTIETIN-V-G..QGPR......GGAGGGPRRTPR-----QF--I-KD-PK--VR........................K-F---R---LA-Q-R..  438
MUS-3.GA.09.09GabOI81      ----TD-A------T---------E--A----M.-G------LQ--------AS---L-----VDAF-KTGATA...-V--AAPIPR..........GPNNRQPR-----QF--L-K--PK--VV........................K-F---R---MA-Q-N..  438
MUS-3.GA.11.11GabPts02     ----TD-A------S---------E-------M.-G------LQ--------AQ---L-----VSAF-KSGASTAMMQRGPPPG--P............QKRAPR-----QF--L--S-PK-KVV........................K-F---R---MA-Q--..  443
OLC.CI.97.97CI12           ----G-----T----K--I--------GV-----.VKPS----L----------Y--------LRP-Q...........-V-QRPAP..........GGGPQQGRLR-Y---RF---KN--SQKGP.......................T-Y----P--IA-N-R..  405
RCM.CM.00.SIVagi_00CM312   -----DPA------QS--I---------V--G--MNP-----L---------QY---------Q-ANSA.........--A-MRGGQR.....SGPPRGQGGKTFL--Y----L--IS-Y-K---R-..........................------E-D-IL---P..  435
RCM.CM.02.SIVrcm_02CM8081  ----TDPA------Q---I---------V--G--MNPS----LL--------Q----------QHANSM.......M-A-MRGQGR.....SGPPRRGGERPVP--Y----T--IS-F-K---R-..........................------D-D-L-----..  435
RCM.GA.x.SIVRCMGAB1        -----DPA------QS--I--------MV--G--MNPS----L---------Q----------QMMQSN........--A-QSAN.......RGPPRRSGGNPNLR-Y----P--IS-Y-K---R-..........................------SPD-LL----..  433
RCM.NG.x.NG411             -----DPA------QS--I---------V--G--MNP-----L-----A---Q----------QM-Q.......SN--A-Q-GP-KGP......PKLGGGPRFL--Y----T-----Y-KTS---..........................---R--E---L----P..  434
SAB.SN.x.SAB1              ----TDPA------QS--I---------V--G--MNP-----L-----I--AQ----LM----TAAFQQQ-V..GN-FV-Q-ARPRGP....LGGRGRPLNPNI--Y----P--L--F-K---RQ..........................------SPD------Q..  453
SMM.CI.79.SIVsmCI2         ----TDAA------Q---I--------LV--G--MNP-----L-----I----Q---LM---LREAFQPGPLP..FAAA-QRGGGVR..............RPI--W-------S--Q-K---RQ..........................-------S--V-AK-P..  431
SMM.LR.89.SIVsmLIB1        ----TDPA------Q---I--------LV--G--MNP-----L----------Q---LM---LKEAFQP-SLP..FAAV-QRGGTK...............RT-R-W-R-----S--Q-K---RQ..........................-------T--I-AK-P..  428
SMM.SL.92.SIVsmSL92A       ----TDPA------Q---I--------LV--G--MNP-----L----------Q---LM---LKEALGPTQLP..FAAA-QKG..P...............SK-R-W-------S--Q-K---RQ..........................-------M--I-AK-P..  428
SMM.SL.92.SIVsmSL92B       ----TDPA------Q---I--------LV--G--MNP-----L----------Q---LM----KDALTGPLV.....AA-FRGAAKGQ..........GNKP-IR------T--S--Q-----R-..........................------E---IQAN-P..  432
SMM.SL.92.SL92B            ----TDPA------Q---I--------LV--G--MNP-----L----------Q---LM----KDALTGSLV.....AA-FRGAAKGQ..........GNKP-IR------T--S--Q-----R-..........................------E--RIQAN-P..  432
SMM.US.04.G078             ----TDPA------Q---I--------LV--G--MNP-----L----------Q---LM---LKEALAPGPLP..FAAA-QRGGPR.................-T---W----------Q-K---RQ..........................-------T--V-AK-P..  430
SMM.US.05.D215             ----TDPA------Q---I--------LV--G--MNP-----L----------Q---LM---LKDALQPGHLP..FAAA-QKGGPR................-MI--W-------S--Q-----RQ..........................-------M--T-AN-P..  430
SMM.US.06.FTq              ----TDPA------Q---I--------V---G--MNP------S---------Q---LM---LKSALAPGVLP..FAAA-QKGGLR..............HT---W-------S-KQ-----RQ..........................-------M--I-AK-P..  431
SMM.US.86.CFU212           ----TDPA------Q---I--------LV--G--MNP-----L----------Q---LM---LKEALTPGPLP..FAAA-LRGGPK..............RT---W----D-----Q-K---RQ..........................-------M--V-AK--..  430
SMM.US.x.H9                ----TDPAX-----Q---I--------LV--G--INP-----L----------Q-X-LM---LKDALTQGPLP..FAAV-QKGQ................X--IX-X--X----S--QF----RQ..........................-------A--V-AK-P..  429
SMM.US.x.PGM53             ----TDPA------Q---I--------LV--G--MNP-----L----------Q---LM---LKEALTPGQLP..FAAV-QKGP................R-TI--W-------S-KQ-----RQ..........................-------T--V-AK-P..  429
SMM.x.x.pE660.CG7G         ----TDPA------Q---I--------LV--G--MNP-----L-----I----Q---LM---LKEALRPDQLP..FAAV-Q.KGQR..............RTI--W-------S--Q-----RQ..........................---G---T--V-AK-P..  429
STM.US.89.STM_37_16        -----DPS------R---I--------LV--G--MNP-----L----------Q---LM---LKE-FQPDPLP..FAAA-QQG.................RRT---W----A----KQ-KG--RQ..........................------QQ----AK-P..  429
SUN.GA.98.L14              ---SG-E---E--K-KM----------LVI----ETPS-----R--------S--GKI-----ASAFRQQVG..RQA-V-QNLPPRNSQGRFVRIGGGGPR-PMT------P--L-NQ--E-K-G...................PPGS------M--KQAQ-P..  436
SYK.KE.x.KE51              -----DPA----L-QS--I-----E-RS----MVK.P-----LQ--------QY--KL-----VMTQH-L......G-I-GPRQGSNP..........RRGPTR-----QL--LQKD-PR-K-L........................K-FN--GT--IARQ-RQP  452
SYK.KE.x.SYK173_COMGNM     -----DPS--G-L-Q---I-----E-RQ----MVK.P-----LQ--------L---KLM----VMAQQ-V......N-V-GPSKG..............RSMI--Y---QI--MQKD-KK-L-A........................K-FN---T--LARA-RQP  451
TAL.CM.00.266              ----TD-A------Q---------E--A----M--G------LQ--------Q--S-LM---AEALK.QNTQQVIA-V-Q-QGGP....RGRGGPRRNPPGQIR-Y----F--I-K--P--PRQKV...................PPGT-Y----P--IA-Q-S..  450
TAL.CM.01.8023             ----TD-A------Q---------E--A----M--Q------LQ--------Q--S-LM---AEALK.QNTQQVIA-V-.-QGGP....KGRGGPRRTPPGQIR-Y----F--I-K--P--Q-PRG...................PPGS-F---QM--RAAQ-R..  453
TAN.UG.x.SIVagmTAN1        ----TP-D-----------I--------LV--G--IHP-----L--------------KLMV---Q-MQG.......VN-V-GAPRGGRG..........RGRGPPR--K--QI--IQKD-PRAGPN........................K-L----P--LA---R..  437
VER.DE.x.AGM3              ------G---Q----S--I--------V---G--MHP-----L---------SY--K-M--M-QNMQSQ.......N---Q-GQ-GRP............RPP---Y----F--MQ-Q-PE---M........................R-L----P--LA---R..  439
VER.KE.x.9063              ------G---Q----S--I--------V---G--MHP-----L---------SY----M--M-QNMQTQ.......N-V-Q-AAGGV-..........RQRPPL--Y----F--MQ-Q-TE---I........................K-L----L--LA---R..  437
VER.KE.x.AGM155            ------G---Q----S--I--------V---G--MHP-----L---------SY--K-M--M-QNLQSQ.......N-V-Q-GG-GRP............RPPP--Y----F--MQ-Q-PE---I........................K-L----P--LA---R..  435
VER.KE.x.TYO1_patent       ------G---Q----S--I--------V---G--MHP-----L---------SY--K-M--M-QTMQ-Q.......N-V-Q-GPKR-.............RPPLR-Y----F--MQ-Q-PE---T........................K-L----L--LA---R..  434
WRC.CI.97.97CI14           ----GAP---Q---DK--I-------------G-.QQP-----L----------Y-SKIM-Q-LQEAL.H.SQ....A-V-QAGAKP.............KGPIR-----QI--M-KE-PK-K-PVPMR.................RGQGPT--G--EV--VQRN-P..  416
WRC.CI.98.98CI04           ----GAP---Q---DK--I--------A---G-.QQP-----L----------Y-SKIM-Q-LQEAL.-.SQ....A-V-QQPVKP.............KGPIR-----QI--M-KD-PK-K-PVPLR.................RGQGPT--G--EI--VQRN-P..  420
WRC.GM.05.Pbt_05GM_X02     ----GVP---Q---DK--I--------Q---G-.QSP-----L----------Y-SKIM-Q-LQEAL.-QSQ....A-V-QQAVK...............RGPLR-----QI--M-KD-KK-KRSPQQR..................GPPT--G---T--IQ-N-P..  416
```


PLV Proteins

```
                               p7 end p1 start      p1 end p6 start  Vpr binding                                                             Vpr binding
                                                                                                                                        p6 end, Gag end
H1B.FR.83.HXB2             .......ERQANFLGKIWPSYKGRP...GNFLQSRP...EPTAPPEESFRSGVETT................................TPPQKQE.PIDKE..LYPLTSLRSLFGNDPSSQ..*   500
H1O.CM.91.MVP5180          .......G--------Y--PGGT--....--YV-K...QV..SPSAPPMEEAVK-.....................QENQ.SQKGDQ........-E.........---FA--K----T-Q......-   499
H1N.CM.95.YBF30            ......EG--------S-SPF----.....---P-TTTRK.------L--YGFQE-...........................KSTQGKEMQEN--R.TENSLY.P-----------------..-   513
H1P.CM.06.U14788           .......G-----------PGGK--....--YV-KQV...Q-----M.................................EEEEMTQNK-XE-KEDE--L.----A--K----T-Q......-   493
CPZ.CD.06.BF1167           .....QSKSGV-----RT-WWGC--....---V-NQSQTQNQEQSQTQEKLVEKA...............EPSAPP.................M.EEVNHWNSKDGF--K-I--S-Q......-   529
CPZ.TZ.00.TAN1             N.....NSTGV-----RT-LWGC--....---V-NT-....EKGK....................................AQEQETAQTPVVPTAPPL-MTMKGGF--K-I--S-Q......-   525
GOR.CM.07.SIVgorCP2135con  .......G--------Y--PGGE--....--YV-KQV...Q-----M.......................................EEGEKEMRKPEKEEL...-------K----T-Q......-   490

                               p8 end p1 start      p1 end p6 start       PTAP motif                          PSAP in HIV-2 B, U                     p6 end
                                                                                                                                                  Gag end
MAC.US.x.239               .......D---G---LGPWG..KK-...R--PMAQ.......VHQGLMPTAPP................EDPAVDLLKNYMQLGKQQREK-RESREKPYKEVTED-LH-N----G-Q......-   511
H2A.SN.86.ST_JSP4_27       .......----G---LGPWG..KK-...R--PVAQ.......I-QGLTPTAPPID..PVEDLLEKYMQQGKRQREQRERPYKEVTEDFLQLE..KQETPCRETTED-LH-N----K-Q......-   522
H2B.CI.x.EHO               .......----G---FGPWG..KK-...R--PVQA.......PQ.GIVPSAPPMNP.AFGMTPQGAIPSAPPADPAEEMLKNYMQLGKKQKE..NRERPYKEVTED-LH-N----E-Q......-   520
H2G.CI.92.Abt96            .......----G---FGPWG..KK-...R--PMTQ.......V-QGLTPSAPP................MDPAVDLLKNYMQLGRKQKE....QRNKPYKEVTEX-LH-S----D-Q......-   510
H2U.CI.07.07IC_TNP3        .......----G---FGPWG..KK-...R--PMTQ.......I-QGLTPTAP................PADPAVDLLKNYMQMGRKQKE....SRERPYKEVTED-LH-N----D-Q......-   510
H2U.FR.96.12034            ........-TKVG---LGPWE..KK-...R--PMAQ.......V-QGLTPTAP.....PAEPAVDLLTPTAPPADPAVDLLKSYMQQGKKQKE..NRERPYKEVTED-LH-N------Q......-   522
ASC.UG.10.RT03             ....EGAQGG------AFLD..KHLGPRKGPRNF.........................PVVQPSAPPLTLQADTFQEENLGPAVGIQKIP..-G...EKTRESLYP--S----D-Q......-   517
ASC.UG.10.RT08             ....DSAQGG------AFLD..KHLGPRKGPRNF.........................PVVQPSAPPLTLQADTFQEENLGPAVGIQKTP..-DNNGEKTRESLYP--S----D-Q......-   520
ASC.UG.10.RT11             ....EGAQGG------AFLD..KHLGPRKGPRNF.........................PVVQPSAPPLTLQADTFQEENLGPAVGTQKTP..-GNNGEKTRESLYP--S----D-Q......-   520
COL.CM.x.CGU1              .......NK------ATRGV..ELQT.....AIFP-.......KMSKDLPPR.............................................R-KGESLYP--K----D-Q......-   483
COL.UG.10.BWC01            .......NK------AGRGI..DLQ...TAIFPPA.......MSKMQKQEGR..............................................-RDENTLYP-------E-Q-.....-   492
COL.UG.10.BWC07            .......NK-V----TPGGM..DLQ...TAVFPPK.......MSQRKEPSK-.............................................EMDREGIYP--K----E-LL-GQN-   494
DEB.CM.04.SIVdeb04CMPF3061 ...KAAPQSKN-----LPWG..R-AP..R-..........FIEQ.................QHQTWGIGKYIPEGGDEIFSSSRVPEQK-...-QQEETRRGKYPD-SD-K----E-Q......-   545
DEB.CM.99.CM40             ......QTPR-----NTPWA..S-RP..R---EGIGKY.I-ESTADQVYST............................TRVPVQGQENNPQR-ETTRRGKY.PD-MG-K----E-Q......-   536
DEB.CM.99.CM5              ......QVPR-----NAPWA.SR--...R---EGIGKY.I-ESTADQVYST............................TRVPVQGERNAPQEDTTTRRGKY.PD-MG-K----E-Q......-   533
DEN.CD.x.CD1_CMO580407     ..SNTNDQG-V----NGLGR..KN-P..R--V-YY-QMNQDLPWSPVRV.............QQQPQYPQNFQEGGNRQSCQEDKRQNKQS...HQEDEETTRESLYP--K----D-Q......-   559
DRL.DE.11.D3               .......PK-V----NTPWG..SGKP.RNFPQA-.......................LTPSAPPLPGYLQEDPAERMLQKYMEQGAQQKRQ-R.QQQQK-GPYKEAYN--S----S-QLQ....-   522
DRL.x.x.FAO                .......PK-V----NTPWG..SGKP.RNFPQA-.......................LTPSAPPLPGYLQEDPAERMLQKYMEQGAQQKRQ-R.QQQQKRGPYEEAYN--S----S-QLQ....-   522
GRV.ET.x.GRI_677_gri_1     .......NG------YGHWG..-AKP.RNFVQYRGDTV.GLEPTAPP................METAYDPAKKLLQQYAEKGQRLREER....-QTRKQKEKEVEDV--S----G-Q......-   514
GSN.CM.99.CN166            ...TGTPIKELY--DLGRGR..R-AP..R--PVTFLLP.TAP-E.--EPQQ..............................LNLSTPVGEP-SKGQKE-SRNTLYP--S----E-Q......-   513
GSN.CM.99.CN71             ....GTPAKEL---DLGRGR...-AP..R--PVTSLLP.SAP-E.-VEPQQ..............................LNLSPPVGEP-NKEQKA-KRNTLYP--S----E-Q......-   511
LST.CD.88.SIVlhoest447     .......KGKV----YGPWR..SG-P.----PLMGGN..AGV.V-SAPP................MERSPTQAERALETYRNLGQQLRR.....QQQIPQKCVDEPCLSFFF-PD-Q......-   513
LST.CD.88.SIVlhoest485     .......KGKV----YGPWR..NG-P.----PLMGGN..AGV.V-SAPP................MERNSTKAERALETYRNLGQQLRK.....QQQVPQKCVDEPCLSFFF-PD-Q......-   513
LST.CD.88.SIVlhoest524     .......KGKV----YGPWR..NG-S.----PIMGGA..AGV.I-SAPP...............MEKVPPTRAERAIETYRNLGQQLRK.....QQQVPQKCVDEPCLNFFF-PD-Q......-   512
LST.KE.x.lho7              .....KPKGKV----YGPW-.SKP-G..NYP-LGGA...AGR.I-SAPP..................MESAPTKAERALETYRTLGQQLKR...QQQQVPQKCVDEPCLNSLF.PD-Q......-   516
MAC.US.x.251_1A11          .......D---G---LGPWG..KK-...R--PMAQ.......VHQGLTPTAPP................EDPAVDLLKNYMQLGKQQREK-RESREKPYKEMTED-LH-N----G-Q......-   511
MND-1.GA.x.MNDGB1          ...KPAQQQRV----YGPWG..PSKP..--YP............................AQEVTPTAPPLEEKPLQKTLSTYQKLGRGL.....RQKMKE-KREEDFH--ST--QE-Q......-   503
MND-2.CM.98.CM16           ........KMV----NTPWG..SGKP..R--P.......AMPLT-T.AP............PMPGMEDPAERMLLDYMKKGQQQKAESK.....QEKKERGPYEAAYN--S----T-QLQ....-   518
MND-2.GA.x.M14             ........KMV----NAPWG..SGKP..R--P.......AMPLT-SAPP.............MPGMEDPAEKMLLDYMRKGQQQRAAKET....KQEK--GPYEAAYN--S----T-QLQ....-   519
MND-2.x.x.5440             .......PKMV----NTPWG..SGKP..R--P.......AMPLT-SAPP.............MPGLEDPAEKMLLDYMKKGQQQRAAAGA....QKEEK-GPYEAAYN--S----T-QLQ....-   519
MNE.US.x.MNE027            .......D---G---FGPWG..KK-...R--PMAQ.......MHQGLTPTAPP................EDPAVDLLKNYMQLGKQQRE....SKRKPYKEVAED-LH-N----E-Q......-   507
MON.CM.99.L1_99CML1        P....KQNKGV----NPFGP..KKGP..R--PLTSVQP.SAPTL-A.........................GTAEPVNLNQEIGS-...KTGTP--TRRDLYP--A----E-Q......-   511
MUS-1.CM.01.CM1239         ....SEGAKS-----DPLGL.QR-RP..R--P-TSLLP.TAPPEQ-GDLVNLSPPV......................................GPKDGASDRKSLYP--S---QD--LEK..-   506
MUS-1.CM.01.SIVmus_01CM1085 ....SDNTKAV----DPFGL.QK-RP..R--P-TSLLP.TAPPEQ-GELVNLSPAV.....................................GPGEG-QDRKSLYP--S---QD-Q-D-..-   510
MUS-2.CM.01.CM1246         ...TDTGKSAV-----DLGF.RS-KP..R--PVTSLMP.TAPPDTA-PVNFSPLVGPKE...................................-P-K-RKDRKTLYP--S----D-Q-K-..-   517
MUS-2.CM.01.CM2500         ...TNSDKGAV-----GFGF.HR.KP..R--PVTSLMP.SAPPDTA-IVNFSP-..........................VGPRKEN...QT-QAPPDRKNLYP--S----D---E-..-   519
MUS-3.GA.09.09GabOI81      .....TSTGRV----NPFGP.QK-GP..R--PVTQLVP..-----GTGEPVNFSP...........................KVGEPQKEKE-TRKTLYP--S----D-Q-E-..-   514
MUS-3.GA.11.11GabPts02     ......SQGKV----NPFGP.QK-GP..R--PVTQLVPTA-PDFGKPVNFAPKVG..........................-PTTSGEEGETTRNSLYP--S----D---E-..-   521
OLC.CI.97.97CI12           .VGGQQQQSKV-----WRGG.-AE-PLNMQQRTEMDIF.PSAPTL-...........................QILEKGTR......-KREM--EKKPLYP--A----E-Q......-   481
RCM.CM.00.SIVagi_00CM312   ........K-V----NAPWK..QGKP..R--PLTSVTPSA-GMDQYDP.................AEAMLKRYLEQGKKQKQQKRQE....-KSQQGRAYEEA-S--N------Q......-   518
RCM.CM.02.SIVrcm_02CM8081  ........K-V----SAPW-..RGKP..R--PLTSLTPSA-GVEQDNH.................ASQYDPAEAMLRKYLEQGRMQKQQQR-E-KAQKGKAYEEA-A--N----S-Q......-   524
RCM.GA.x.SIVRCMGAB1        ........K-I----RLPWG..QGKP..R--PLTSLTPSA-GMESNYDP.................AEEMLKNYLRRAGEQKRQQRQE....-SKKREGAYQEA----N----S-QLQ....-   519
RCM.NG.x.NG411             ........KEV----NAPWK..QGKP..R-LPLTSLRPT.....APG...........REEGWNPQYDPAEEMLRKYLALGRQHKQEQRKE...NKEKVGRAYEDA-S--N----S-L......-   522
SAB.SN.x.SAB1              ........K-V----FGPWG..RGKP.RNFP-T-IRPT.A-PMERDY-RPEENWYADRPPTRGPGPDDPATALLKQYAVQGKRQKQQWQ.....NHSPQQSPYEEAYS-------E-Q......-   555
SMM.CI.79.SIVsmCI2         ..........----G---FGPWG..KK-...R--PMTQ.......LQQGLTPTAPPMDPA........VAAPPMDPAVDMLRNYLKLGQQQRE...RKERPYKEVTED-LH-N----D-Q......-   519
SMM.LR.89.SIVsmLIB1        ..........----G---FGPWG..KK-...R--PMAQ.......MHQGLTPTAPP................ADPAVDLLKNYMQLGKKQRE...NRERPYKEVTED-LH-N----D-Q......-   507
SMM.SL.92.SIVsmSL92A       ..........----G---FGPWG..KK-...R--PMAQ.......M-QGLTPTAPP................ADPAVDLLKDYMKLGRKQKE...SRERPYKEVTED-LH-N----E-Q......-   507
SMM.SL.92.SIVsmSL92B       .......NQK-G---LGPWG..KK-...R--PMQT.........TSLTPSAPP................DPAAQMVK.EYLQKAQREKT...RRSRPYKEVTED-LH-N----E-Q......-   508
SMM.SL.92.SL92B            .......NQK-G---LGPWG..KK-...R--PMQT.........TSLTPSAPP................DPAARIVK.EYLEKAQREKT...RRSRPYKEVTED-LH-N----E-Q......-   508
SMM.US.04.G078             .......D---G---FGPWG..KK-...R--PMAQ.......V-QRLVPTAPP................EDPAVDLLKNYMQLGRKQRE...NRERPYKEVTED-LH-S----E-Q......-   509
SMM.US.05.D215             .......D---G---FGPWG..KK-...R--PMAQ.......VHQGLTPTAPP................ADPAVDLLKNYMQLGRKQRE...NRERPYKEVTED-LH-N----E-Q......-   509
SMM.US.06.FTq              ..........---VG---FGPWG..KK-...R--PMTQ.......M-QGLTPTAPP-DLA............RDPAEDLLKRYMQQGRKQRE...SRERPYKEVRED-LH-S----E-Q......-   514
SMM.US.86.CFU212           ..........----G---FGPWG..KK-...R--PMAQ.......VHQGLTPTAPP................ADPAVDLLKEYMKQGRKQRE...NRERPYKEVTED-LH-N----D-Q......-   509
SMM.US.x.H9                ..........----G---LGPWG..KK-...R--PMAQ.......M-QGLTPTAPP................EDPAVDLLKNYMKVGRRQRE...NRERPYKEVTED-LH-N----E-Q......-   508
SMM.US.x.PGM53             ..........----G---LGPWG..KK-...R--PMAQ.......M-QGLTPTAPP................EDPAVDLLRNYMKMGRKQRE...NRERPYKEVTED-LH-N----G-Q......-   508
SMM.x.x.pE660.CG7G         ..........----G---LGPWG..KK-...R--PMAQ.......M-QGLIPTAPP................EDPAVDLLKNYMKMGRKQRE...NRERPYKEVTED-LH-N----E-Q......-   508
STM.US.89.STM_37_16        ..........---VG---FGPWG..KK-...H--PMAQ.......I-QGLTPTAPP................EDPAADLLRSYMQLGKKQRE...SRKTPYKEVTED-LH-N----E-Q......-   508
SUN.GA.98.L14              ......QGK-V----YGPWN..RG-P..----PVMPSAP.P.................LEDLTLGNRMTPPQSKAERALETYRLLGQGLR....AQQKR--RGECQEPC-NM--PE--......-   522
SYK.KE.x.KE51              RKGQGNPPP-------G-GN..RK-P..A--PVM-ETP.TAPPAEDWPWQQQTWGN.YASPQQKHIAMATPLQPQSSPKTVKDLLVPGK-S....-KKEEMKEKGPLYP--Q----D-Q......-   562
SYK.KE.x.SYK173_COMGNM     KRNQGPPVA-------G-GV.SR--P..A--PVRSEPS.A-PLEDI........-DGPWLTWSAQMSQQAQAKAQNSPSKKPPTNREVLS-....KESSG--ETKSLYP--S----E-Q......-   555
TAL.CM.00.266              .......SA-------SFLG..KI-P..VRGPRNF-.........VVQPSAPPI................EPLTASWQEGTTTWTPPAKK.....PQETTTTRESLYP--A----D-Q......-   528
TAL.CM.01.8023             .......QP-------AFLG..KI-P..VKGPRNF-VV.Q-----I.......................EPLTASWQEGTTTWTPPTKD....AAAPKQQTRESLYP--A----D-Q......-   532
TAN.UG.x.SIVagmTAN1        .......SG------RMPTW..-TKP.RNFLE.QGGAV...PTA-PMPAHGFP...............TGSPAAGAYDPARKLLEQYAKKG..DQLRKQKEKELEDY--S----G-Q......-   524
VER.DE.x.AGM3              ........G-V----YGRWM..-AKP.RNFPAATLGVE.PTAP--PSP...................YDPAKKLLQQYADKGKQLREQRK...KP-AVNPDWTEGYSL.N----E-Q......-   522
VER.KE.x.9063              ........G-V----YGRWT..-TKP.RNFLAATHGVE.PSAP--PTP...................YDPAKKLLQQYAEKGKQLREQGK...RT-PTNPDWAEGYSL.N----E-Q......-   520
VER.KE.x.AGM155            ........G-V----YGRWM..-TKP.RNFPAATLGAE.PSAP--PNN................STPYDPAKKLLQQYAEKGKQMRNQNR...NP-ANNPDWNEGYSL.N----E-Q......-   521
VER.KE.x.TYO1_patent       ........G-V----YGRWM..-AKP.RNFPAATLGAE.PSAP--PSGTTP................YDPAKKLLQQYAEKGKQLREQKR...NP-AMNPDWTEGYSL.N----E-Q......-   520
WRC.CI.97.97CI14           .....KIR--V----NGRGQ..MPSNNYPVQKLRSLNN.FGPTTTMP-APPMR-DPNI.........QELEKYLLKGAQLREQQEHKK......TEGQKR-VGVPEWP--K---EP-Q......-   512
WRC.CI.98.98CI04           .....KSR--I----NGRGQ..MPSNNYPVQKVRSLNN.FGP-.TIP-APPMR-DPNI.........QELEKYLLKGAQLKEQAQAK.......KEEKKR-TG-PEWP--K---ET-Q......-   514
WRC.GM.05.Pbt_05GM_X02     .....ETR--V----NGRG...Q--....S-NYPVQ..........KMNAGVNTFGPVV.....PTAPVDPTMQELEQYLLRGAQMKRE......KEKEMQKEGKVEYP--K---EP-Q......-   503
```

Pol start
Gag-Pol TF start
Gag-Pol TF end_protease start

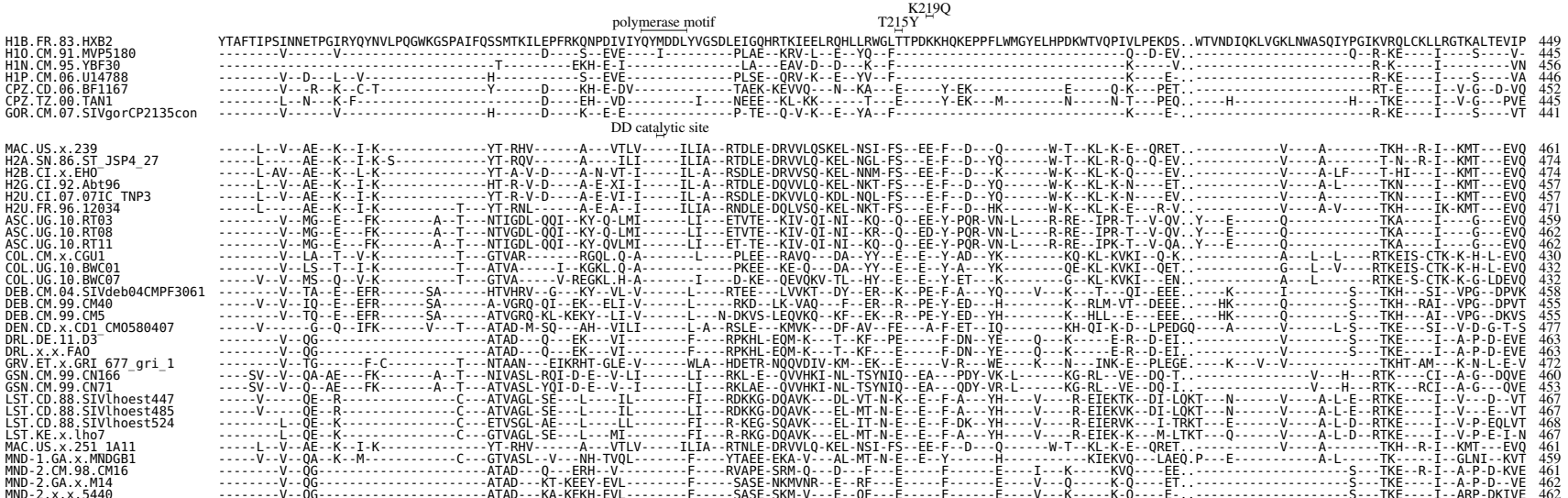
```
H1B.FR.83.HXB2        FFREDLAFLQGKA...REFSSEQT...RANSPTRRELQVWGRDN.....................................NSPSEAG.ADRQG..TVSFNFPQVTLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMSL....PGRWKPKMIGGIGGFIKVRQYDQILIEICGH  125
H1O.CM.91.MVP5180     ----V--SGGHE----QLCA-...TS..VPISPTNGGGSEG.....................TRES.ESEGGS.....-R.....A-PICL--IP--D--I--A-V--H-C------------NNIQ-----E---T------------KE-NNVTV-VQ-K  121
H1N.CM.95.YBF30       ----E-VS--RET....KLPPDNNKE...H--AT---W-S-G.....................................EEHTGEGDAG-P-E.--ELSV.PT------I------VI-V---KEVR-------------I--LQ-----E-K-------------------N-TVD-Q-R  132
H1P.CM.06.U14788      -------SGGQE-----QLCT-TS...TTD---Y.....................................GGGGDDSEQA-G-KGG-ERV...-P-SL--I---D--VIPV----HIC------------VDNLP-----E-----------------KEFEEVKX--E-R  122
CPZ.CD.06.BF1167      ----THPLVGVQT...K-CA-SESDSESGTESDTGETSGES.....................RTLSSS.....................N.GGGESLEQQGWL--EIH-----MMEVTVQ--VCT----------FADLK-----K-SCN-RT-------VP-QE-YNVP-T-A-K  128
CPZ.TZ.00.TAN1        ----THPLVGVQT.----LCA-HP....RERE.....................................GSGAGDSTDTSGANCPTT-DDDERRVL---N-----MM-V-VQ--VCQ---------S-FCNIK----K-Q-T--T-------VP-SE-YN-PVQ-GNK  121
GOR.CM.07.SIVgorCP2135con  ----I--SRGRE----QLCA-AS...AT----N.....................................GGGRERNEETREGGT...L-S----I---D--IIPV----H-C------------V-HL-----E-----R----------KE-EDVE---EHR  117

Pol p15 start
p15 end_protease start

MAC.US.x.239          ---PWSMG..KE-...PQ-PHGS.......-ASGADANCSP.....................RGPSCGSAKELHAVGQAAERKA-RKQREALQGGDRG-AA--FS--R--V--AH-E--PV-V--------SIVTGIE-....GPHYT--IV-------NTKE-KNVE--VL-K  137
H2A.SN.86.ST_JSP4_27  ---VGPMG..KE-...PQ-PCGP.......N-AGADTNSTPDR..PSRGPTREVHAARKEAERAERAIQRSDRGLPAAR..ETR-TMQRDDRGLAA--FS--K--V--AHVE--PV-V--------SIVAGVE----GSNYS--IV-------NTKE-KNVE-RVLNK  150
H2B.CI.x.EHO          ---VRPLG..KE-...SQ-PRPG.......T-.GDSAICAPDEP.SIRHDTSGCDSICTPCRSSRGDAKELHATREEAEG..EQRETLQGGDRG-AA--FS--R--V-KAT-E--SV-V--------SIVAGIE-....GSNYT--IV-------NTNE-KNVE--VV-K  150
H2G.CI.92.Abt96       ---VWTLG..KE-...SQLPHDP.........--SGSDSISTT.................DGPSSRPTEELHAXGEEAKG....AEKQTLQGGDRXXAA--LS--R--V--AY-ES-PV-V--------SIVAGIE-....GSNYS--IV-------NTKE-KNVE-KVV-K  133
H2U.CI.07.07IC_TNP3   ---VWPMG..KE-...SQLPHDP.......D-SGVDTNSA.................PSRPSSGSAEELHADGQETKG....EQGKTIQGGDGG-AA--FS--R--V--AH-E--PV-V--------SIVAGIE-....GVNYH--IV-------NTKE-TNVE--VL-K  133
H2U.FR.96.12034       ---AWPLG..K-....SQLPHGP.......--AGTDPNSP.....SSRTSSGSADPDSPSSGPSSGSAKELHATGQETKG..EQRETIQGGDGG-AA--FS--K--I--AY-E--PVKV--N-----SIVAGIK-....EPYYT--IV-E-E---NTKKFHNVN-KVL-K  147
ASC.UG.10.RT03        ----GFFR..QASGAQEGPQEL.....................................PSSTAISSTFDSAGRHFSGGELRPSCRDTKDP..RR...GEDKG--LSL-VIP--G--I-SCLVER-EVDL--------IIHDQDIEL..GDN-H--I---V--N-R-KA-HSVE--WQ-K  135
ASC.UG.10.RT08        ----GFFR..QASGAQEGPQEL.....................................PSSTAISSTFDSAGRHFSGGEPRPSCRDTEDP..RRQQWGKNKG-PLSL-VIP--G--I-SCLVE--EVDL--------IIHDQDIEL..GDN-H--I---V--N-R-KA-HNVD--WQ-K  138
ASC.UG.10.RT11        ----GFFR..QAPGAQEGPQEL.....................................PSSTAISSTSDSAGRHFSGGELRPSCRDTEDP..RREQWGEDKG-PLSL-VIP--G--I-SCLVE--EVDL--------IIHDQDIDL..GDN-H--I---V--N-R-KA-HSVS--WQ-K  138
COL.CM.x.CGU1         ---GNQGG..-ATD....-HLPP.....EDEQRPATQ.....................................E-EGGE-LSLLE-S-RR--I-IVEVE--KCQ----------FK-EDVQ..LT-A-TAVQ-Q----A-R-K--RDKILKMG-K  107
COL.UG.10.BWC01       ---GRQGD..RP-...DRNL-PS.......NEQDA-AGGT.....................................EERREH-VSLLE-S-RR--I-IAE-E--KCQV---------IK-SDID..IK-T-TAVN-Q----N-R-K--KDKKVL-G-K  109
COL.UG.10.BWC07       ---DTRGD..-PT...DSCLPPK.......DESEEGTLQG.....................................GDGQGRNLSLLEIP--R--F-WAELE--RFQV-V---------K-EEIQ..LT-T-TVVT------S-RAK--KNRVMK-G-K  109
DEB.CM.04.SIVdeb04CMPF3061 ----TSLG..EESP..Q-...........FHRA.....................ATPDMGHREVHTGGRGDLLLFQGTGTET..PAAGGDEEREIPR-IGLE-P--R--VT--E-E-RKVQV---------IIH-KDI..QLTSPFV--TV--L----N---C-PG-E-QCQ-K  134
DEB.CM.99.CM40        ----YSLG..LQET..Q-L-RRNWEV.HPGEHS-SGILD.....................................YQGTSTGAGEQPPE-GNNEEREV.PG--GAEIS--R--MT-VH-E-RKVQV-I-------IIS-KDI..DLEAP-T--TV--L----N-KC-PG-E-AMANK  133
DEB.CM.99.CM5         ----CSLG.LPET...Q--FGRDWEI.HPREHS-SG-FN.....................................NQGTCARGKEC-PG-HN-QEREV.PR--GAEIS--R--MTAVYVE-KRVQV-I-------IIS-KD...DLDAP-T--TV--L----N-KC-PG-E-KMADK  133
DEN.CD.x.CD1_CMO580407 ----WSWE..EESS..Q-LCTVLPSDESGFAMEPS-G.....................PAATTVPTKFSRGGKQTILSGRQEAEQAV...PSGG-GDNEGE-LS-VE-S--R--IGQ-WVQ--RV-V-I-------IIA-QDI.NLGTEG-I--TV-----Y-N-QA-PG-E-KLMDK  149
DRL.DE.11.D3          ----YSLG..QRET.QEFS-GL.....................ADPISSSIARIPPGGSGREDAPKVHGAGSSAEETT..TTTAEE-PLQGGLQL--FS-RL--TTIVE-E--KV-----------IQGLE----.T-K---QI------T-N----FNCE-K-A-K  139
DRL.x.x.FAO           ----YSLG..QRET.QEFS-GL.....................ADPISSSIARIPPGGSGREDAPKVHGAGSSAEETT..TTTAEE-PLRGGLQL--FS-RL--TTIVEVE--KV-----------IQGLE----.T-K---QI------T-N----FNCE-K-A-K  139
GRV.ET.x.GRI_677_gri_1 ---VWPLG..RSET.KKFCAIQRRHS.WSGTNSPP.....................NGNSLRSSKEAPPAVCREGTAPERGE....RTDKETEGERSGGC-LELP--R--MKRVI-E-TPVQ----------IIQ-KD-HFPPHKP-RS-VV-----G-H-KE-QGVQVQLEDK  145
GSN.CM.99.CN166       ---PWPGA..PPSP..Q-L--NFPAA.NSPGG.GGRTPA.....................................AQSLDT-RGTPEQGTEG-EQEH-LSL-VIS--G--MKEVS----VVSM--------IVQDSAI..ELDTP-T---V-----L-Q--EHKHVEVIFNEK  130
GSN.CM.99.CN71        ---PWPGA...TSP..Q-L--NFPAA.ISPGG.GGGTPA.....................................AQPVSAGGGAPEQGAEGRKEEH-LSL-VIS--G--MREVSM---VVSM--------IVQ-SAI..ELDTP-T---V-----L-Q-KEHRHVEVIFNEK  129
LST.CD.88.SIVlhoest447 ---LWTLE..EWSS...-KL-FDGGE..CRS.S-ISSP.....................NGEEPNTSRESLRNIQEPRTTAEE.....ATADPTEMCGRAMPEFLL-SR--IEN-LVN--PVS-------IFS-N-V..RIE-PYT-RVV-----Q-R-KE-KDVF---A-K  139
LST.CD.88.SIVlhoest485 ---LWTLE..EWSS...-KL-FDGGE..CRS.S-ISSP.....................NGKELNKSRESLRNIQEPRTTTEE.....ATAGSTEMCGRAMPEFLL-SR--IEN-LVNR-PVS-----R----IFS-N-V..KIE-PYT-RVV-----Q-R-KE-KDVF---A-K  139
LST.CD.88.SIVlhoest524 ---VWSLE..EWTF...-QL-NHGRG...CRS.DSFSAT.....................HGESASNKSRENGNRDIQEPRTTTEE.....ATTSSTEVCGRALPEFLL-SR--IEKVLVN--EVT----------IFS-N-V..RIE-PYT-RVV-----Q---KEFRDVF-Q-A-K  140
LST.KE.x.lho7         ---IWSLE..QTSR..QLS-TGRS...CRE.NSISSP.....................NGECSNKSRESTGDLQDFRTTAEK...AA-ASPTEMCGRALPEFTL.SR--IES-LVN--PVS----------IFS-N-V..RIE-PYT-RVV-----Q-R-KE-RDVF---A-K  139
MAC.US.x.251_1A11     ---PWSMG..KE-...PQ-PHGS.......-ASGADANCSP.....................RGPSCGSAKELHAVGQAAERKT-RKQREALQGDDRG-AA--FS--R--V--AH-E--PV-V--------SIVTGIE-....GPHYT--IV-------NTKE-KNVE--VL-K  137
MND-1.GA.x.MNDGB1     ---VWPLG..SLQT..G-L-.....................................GTRGDSNSSTIRGETSAENSEHLSEIRERA.....QAE-EG-EERGG-S--EYS-SR--IEEVSVD-VTIR---------IFN-RNI..KLK-N-Q--I------NLR-K---NVYV--R-K  132
MND-2.CM.98.CM16      ----YPLG..QWQT..Q-LP.......GDAIDPN.GP.............SNARDGRPSRENAVRLHEERATAESGKQ.....TGKE-E-SIRG-LQL--FS--N--TTVVE-E--KV-----------IKDLD----.K-N---QI------S-N-K-FFNCKVT-A-K  137
MND-2.GA.x.M14        ----CSLG..QWQA..Q-L-........CHATDPISTP.............DARDGGSSREDVIRLYEKGATTEGSKGD....QAGEG--SLRG-LQL--FS--N--TTVVE-E--KV-----------IKDVD----.T-N---QI------A-R-K--FNCK-TVA-K  138
MND-2.x.x.5440        ----YPLG..QWQA..Q-LP.......CHAIDPISTP.............DARVGRPGRDKAVRLYEERATTEGSSR-....PERGKE-SLRG-LQL--FS--N--TTVVE-E--KV-----------IKDLE----.T-N---RI------A-R----FNCKV--A-K  138
MNE.US.x.MNE027       ---LWPMG..KE-...PQ-PHGP.......NASGADTNCSP.....................RGPSCGSAKKLHAVGQTAER....KQKEALQGGGRG-AA--FS--R--V--AH-E--PA-V--------SIVAGIE-....GPHYT--IV-------NTKE-KNVE--VL-K  133
MON.CM.99.L1_99CML1   ----SFRS..QEGA..K-LP-HLCAA.QRTDSAS.....................................RDCGASQSQPRDWF----EDG-PK-DEAG-LSLLSIS--G--IRSVL----KVSM--------IIQDQHV..NLGEN-T--VV-----M-E-K--KA-QVIFEDR  130
MON.NG.x.NG1          -RESFWPR..E.GA..K-L--NLLAA.NSP-G-NPXGLQGELLT.....................................SSGTTRDP....SADSPE-DEEG-LSLLSIP--G--VR-VE-A--SVSM--------IVQ-QDI..NLGEN-T---V-----L-Q-K-FKGVTVSFQ-K  133
MUS-1.CM.01.CM1239    ---GSSWP.PE-EA..K-LPPDLAFT.NGPARAGGRPGQSLPAS.....................................SQGWSERQEE-LS--VIS-SG--IREVM-E--VVHM--------IVQDSNI..QIDKP-N--LV-----N-S--E-RGVQVSFNEK  126
MUS-1.CM.01.SIVmus_01CM1085 ---GS-RP.PE-EA..E-LPPDLSLA.NCSAR-GGGIGQSLSSS.....................................SRGGETGQEE-LS--VIP-SG--VRPVE-E--VVSM--------IIQDAHI..HLE-P-N--SV-----N-S-KE-RGVQVCFN-K  126
MUS-2.CM.01.CM1246    ----GSRL.PVSEA..E-LP-YIAHA.YGTPRYS-AGELLPTSRAQG.....................................RATQG-ERQKD-LS--VIS--G--IK-VD-N--PVSM--------IIQDSDI..KIS-ELT--VV--L--L-N-K--KGVIVKFNEK  129
MUS-2.CM.01.CM2500    ----RFRL.PQ.EA..E-LPCYIAHA.ISSPRYS-DS-LLP-.....................................GGSTEGE....PNPPSPTGQEE-LS--VIP--G--IR-VMVH--PVSM--------IIQ-SEI..EIDKE-T--VV--L--L-N-K--KAVKV-FNDK  130
MUS-3.GA.09.09GabOI81 ----SFWS.PE-GT..K-LP-HPVGA...DG---DWGASQLFA.....................................EGRGAAEGEGRNQEN-LS--VIP--G--IR-VE-E---RVSV--------II--SHI..VLK-E-T---V-----L-N--E-RNVA-RFQDK  127
MUS-3.GA.11.11GabPts02 ----SFRS.PEEGT..K--PCHPAGTHSPTGFWQASQLCPESR.....................................T-NHLGGRGRDDKEL-LSL-V-P--G--FR-VS-E--EVSV--------II--G-V..QIK---T---V-----M-S--E-ENVQ-KFYDK  130
OLC.CI.97.97CI12      ----MEGR.ESR-SPQHATEDRDGHL.SLCTYPG.....................................TDFGEGNK.....KERDE-GEET-LSL-SIS--R--MIELY-E--QVT----------IKGKDI..YME-NGI-RI------A-R-M--RCRIRIF.EK  125
RCM.CM.00.SIVagi_00CM312 ----CSLE..AGET..E---FNLRDTLCSRDGSV-P.....................CGSHVKEILGAGEKAEAAKEAG.....EKPAGESIRRGIVL--FS--K--VTWVT-E--KV-----------ISDIE----.GDN---RI------Y-N-K--YNCK---A-K  138
RCM.CM.02.SIVrcm_02CM8081 ---KCALE..SGET..E-L-VDFLDSLCSRGG-GQS.....................CLPVRSSRSHAKEIPGTGQNAETATE-GREGSEGESLRG-TCL--FS--K--VTMVN-E--KV-----------INGIE----.GEN----I------Y-N-K--YNCK---A-K  144
RCM.GA.x.SIVRCMGAB1   ----TPLG..SGEA..E-L-FDFLDSLCSRDGEQLRP.....................CRRDAKELSEEGRGTKETTEAG....EQRAEER-SISG-L-L--FA--K--TTIAQ-E--KV-V--------I-GIE----.LDA-T--I------Y-N-K--KNCE---A-K  139
RCM.NG.x.NG411        ----CSLE..AGET..E-P-LDFLETY.....CS-.....................ERGRMESPVRSSRRDAQEIPSIGETAQTGAEER...EQGESGESLRGCVVL--FS--K--VTWVE-E--KV--------G--IQNLD----.GEN----I------Y-N-K-FNHCK-T-A-K  143
SAB.SN.x.SAB1         ----VWPLG..QRET.QEFP-DLHQTN.SSPNG-GLQQAGGKLVCRQTSDQRTRARRSSNSPVKAVCCSGETAETAVA....KPLATTEPLRGGLQL---S--R--MK-VY-E--KVT----------S-IQGIE----.GDN---RI------C-N-KA-HNQEVK-EDK  157
SMM.CI.79.SIVsmCI2    ---IWPMG..KE-...SQLPHDP.......T-AGADTNGAPNGSS.........CGGAPNGSSCGHVEELSETRSATER.....EERKALQRGNGRLAA--FS--R--V--AY-E--PV-V--------SIVAGIE-....GLGYT--IV-------NT-E-TNVN--VL-K  142
SMM.LR.89.SIVsmLIB1   ----LWSLG..KE-...SK-PHGP.......NASGADTNCSP.....................SRPSCGSAEELHAIGQEAEG....EQRKTLQGGDRG-AA--FS--R--V--AH-E--PV-V--------SIVAGIE-....GPNYT--IV-------NTKE-ENVE-TVL-K  133
SMM.SL.92.SIVsmSL92A  ---IWSLG..KE-...SQLPHGP.......NASGPDTDSPP.....................SGPSSGSAEGLHEVGKETER....EQRETIQGGDGGLAA--FS--R--V--AH-E--PV-V--------SIVAGIE-....GLNY---IV-------NTKE-QKVK-QVL-K  133
SMM.SL.92.SIVsmSL92B  ---PWPMG..KE-...SQLPHAD.........NIIDTVSTP.....................RSSSPDGQ.GVSTESTEGKD....KEEQALQGGDGG-AA--FS-RR--V-RAL-E--PL-V--------SIVAGIE-....GPGYT--IV-------NTKE-RTVK--VL-K  130
SMM.SL.92.SL92B       ---PWPMG..KE-...SQLPHAD.........NIIDTVSTP.....................RSSSPDSQ.GVSRESTEGKD....KEEQALQGGDGG-AA--FS-RR--V-RAL-E--PL-V--------SIVAGIE-....GPGYT-RIV-------NTKE-RTVK--VL-K  130
SMM.US.04.G078        ---VWPMG..KE-...SQ-PHGP.......TSEAGANGSP.....................RGSGSGSAKELHAARREAEG....EQRETLQGGDGG-AA--FS--R--V--AY-E--PV-V--------SIVAGIE-....GPQYT--IV-------NTKE-QDVE-KVL-K  133
SMM.US.05.D215        ---IWPLG..KE-...PQ-PHGP.......-TSGADADCPS.....................GGSSGRPAEELHAARQEAEG....EQRETLQGGDRG-AA--FS--R--I--AY-E--PV-V--------SIVAGIE-....GPNYT--IV-------NTKE-TEVE-KVL-K  133
SMM.US.06.FTq         ---LWSMG..KE-...SQLPHDP.......DASGAYANGPS-RPS.....................KGSGGGSLEEVHAAGQEAEG....EQRETIQGGEGG-AA--FS--R--V--AH-E--PV-V--------SIVAGIE-....GLNYT--VV-------NTKE-KNVA--VL-K  137
SMM.US.86.CFU212      ---XWPMG..KE-...PQ-PHGP.......-TSGADANCSP.....................SRSSCRXVKGIHETGQETER....EQRETLQGGDGG-AA--FS--R--V--AQVE--PV-V--------SIVAGVE----GTNYT--IV-------NTKE-KNVE--VL-K  133
SMM.US.x.H9           ---AWPMG..KE-...PQ-PHGP.......DASGADTNCSX.....................RGSSCGSTEELHEGGQKAEG....EQRETLQGGNGG-AA--FS--R--V--AY-EE-PI-V--------SIVAGIE-....GPNYT--IV-------NTKE-KDVK-KVL-K  133
SMM.US.x.PGM53        ---AWPMG..KE-...PQ-PHGP.......DASGADANCSP.....................RGSSCGSTEELHEDGQKAEG....EQRETLQGGDGG-AA--FS--R--V-AAY-EE-PV-V--------SIVAGIE-....GPNYT--IV-------NTKE-KDVK-KVL-K  133
SMM.x.x.pE660.CG7G    ---AWPMG..KE-...PQLPHGP.......DASGADTNCSP.....................RGSSCGSTEELHEDGQKAEG....EQRETLQGGDRG-AA--FS--R--V--AY-EE-PV-V--------SIVAGIE-....GPNYT--VV-------NTKE-KDVK-KVL-K  133
STM.US.89.STM_37_16   ---LWPMG..KE-...PQLPHGP.......NTSGADANCPS.....................RRPSCGSAEKLHAAGQEAER....EQE-TLQGGDRG-AA--FS--R--V--AH-E--PV-V--------SIVAGIE-....GLQYT--VV-------NTKEFKNVN--VL-K  133
SUN.GA.98.L14         ---VWPLE..QG-A...--L--HAFCS.A.....................SRGSDIGEQNDTTPVKGRESSGDLQTVRTGPQ....ST-EEE-KGGM-GAMLEYA-SR---EKVF-N--PVT----------IVS-D-V..SIE-H---RVV-----Q-R-KE-ADVFV--RDK  141
SYK.KE.x.KE51         ----RVGE..QETP..SK--CDEGDT.NSPTSGGLAMAATDMGK.LCFTPTEAHSNGHPATATVLSKDSQGPLGPWETF...REER-DEGEGT-LSL-SIS--R--MME-NVE-TKVTM---------IIR-QDI..QLHQP-T--IV--L--N-R----N--RFQ-ERP  159
SYK.KE.x.SYK173_COMGNM ----RMGG.LQE-P..SQLP--ERAI.CSP-GGH.........RWAMAHMVSTDEPTSTSEGTELSLEETPHQQGSALA....EGEQWE-GDQE-LSLS-LS--R--MIEVDVE-D-VQM-V-------IIR-EDI..QLHQP-S--IV--L--N-T----RN-RFTVVKP  153
TAL.CM.00.266         ----VFFR..ENPT..GQRTE-FS........SGATIGTAN.....................RASHSIMAGGDYHVDPTSQE.....APRNNHNKGE-LSLFGQS--R--I-KVL-E--EV-V--------IIQVGEI..ELE-VPHQ-TV--LA-Y-Q-QS-SA-E-QWENK  134
TAL.CM.01.8023        ----GFFR..EDPP..GQGTE-LPSG.ATHRTAD.....................RASHSIMAGGDYHMDPSDQG....CRSTQATDEGE-LSLVSQS--R--I-KVL-E-REV-V--------IIQ-GDI..ELE-VPHQ-TV--LA-Y-Q-QS-TGVE--WEHK  135
TAN.UG.x.SIVagmTAN1   ----NANL..-NKA.AQF-R.ARRSR...SDGSPDACARIP.....................HRFASCGSIRPSQETAGAVCQKG..RSIEKAEGEGTGGL-LEFP--R--MSK-Y-Q--PVQV---------IIK-EDI..HLE---R--VV-----G-S-K---N--VN-D-K  143
VER.DE.x.AGM3         ---VWPVD..-SET.QKFSRRYSWGG.ANCA-STES.....................IRPCKEAPAAICRQGEAVEGTKE...KTTSSESRLDRGIF-.ELP--R--IK-VY-E-VPIR---------IIK-AD-..QLS-T----I------GLN-KE-SDREVRLEDK  140
VER.KE.x.9063         ---VWTVD..-DKT.EKLSRCHSWSG.TERA-STDT.....................IRPSKEAPAAVCREGKTVEGTRE...ENTTNKSRLGGRIF-.ELP--R--IK-VI-E-VPVKV---------IIK-TDI..QLT-P-R--V------GLN-KE-SNVEVK-EDK  140
VER.KE.x.AGM155       ---VWPVD..-DKT.KKFSR-HSWGG.TKCA-STEQ.....................LYTLRPSKEAPAAVCRERETNEKSEQ...KPPSE-SRLERGIF-.ELP--R--IR-CI---TAVK---------IIKDTD...QLR-S-R---V-----GLN-KE--NVEVQLEDK  143
VER.KE.x.TYO1_patent  ---VWTVD...-GKT.EKFSRRYSWSG.TECASSTERHHP.....................IRPSKEAPAAICRERETTEGAKE...ESTGNESGLDRGIF-.ELP--R--IK-VY-E-VPIK---------IIK-ND-..QLS-P-R--I------GLN-KE-NDREVK-EDK  143
WRC.CI.97.97CI14      ----WQRA..DALEQLSSSEA-ELEQ.FWANNNNAFCSSNE-GSKH.........SGAGKIFIEGSTIEGTTGTQE.....IDVNL----.KKN-RI--MIQCI-E-HSVS----------IDVNL----.NKN-Q-RI------TMN-KLHR-VK-R-L-K  147
WRC.CI.98.98CI04      ----WERA..DALKQLPSSKG-ELEQ.FWAG.NNTVCSSNEGGSKH.........TRIREIPPERSSIEGAGTGK.......KGREEERNGAARVA---IS-RN--MIQCI-E-NSVS----------IDIHLN....NKN-Q-RI------TMN-KL-RK-Q-R-L-K  146
WRC.GM.05.Pbt_05GM_X02 ----WPR...AE-...LQQL-SS.........KN-CRGEYLWASG.....ANSTSGSYHARTRAISSERSSDEEG......ERERNAEGGESRISL--IS-RT--MIQCI--RESVS----------VDVQL.....E-N-Q-RI------TMQ-KL-KDVR-K-L-K  138
```

```
                                          protease end p66, p51 RT start                        M41L             K70R
                                                                                                              D67N                                        D110 catalytic site
H1B.FR.83.HXB2               KAI.....GTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETV...PVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGL......KKKKSVTVLDVGDAYFSVPLDEDFRK  281
H1O.CM.91.MVP5180            EVQ....-----------L---I--GL-----------AP-...-----------------SR---E--TA--Q---Q-----R----------I----------------------------------G--......-QRQ------------C---P----  277
H1N.CM.95.YBF30              --V....----------------------------------...-----------------T---E--R-------------R----------I------------------------------------......-Q--------------C---K----  288
H1P.CM.06.U14788             QVY....---------------I--L-----S-------I-...-----A-----R------SK---E--RA--Q---Q----T----------I-------GS---------------------------G--......Q-R----I---------C---P----  278
CPZ.CD.06.BF1167             TVK.....TK--I-----------V-R-L---------TV-I-...K----E-------------SR---E--T---ENLRA----IEV--D------I------------------------------------......--R-VI-----------I---PE-Q-  284
CPZ.TZ.00.TAN1               EVR.....A-----E--I-------I-K-L-----------V-...K-Q--E-------------SK---E--T---KTL-------AV---------I-------TS---------------L-------------......R-RNM-----------I---P----  277
GOR.CM.07.SIVgorCP2135con    -VV....----------------I--K----------S-DV-...--S-------------SK---E--TA--Q---Q----------------------I-----------------------------------G--......---R------------C---P----  273
                                         protease end p51 RT start                                                                                          D catalytic site

MAC.US.x.239                 RIK.....--IMT-D--I--F------AL-MS-----AKV-P-...K-A----K----L-----SK---V--R---EK---D-QLEEAP-T------T-------KN---M-I------RV----T---------------......A-R-RI----I------I----E--Q  293
H2A.SN.86.ST_JSP4_27         RVR.....A-IMT-D--I--F---I--AL-MS--L-VAK--PI...KIM----K----LR-----K---E--K---EK--R--QLEEAP-T------T-------KN---M-I-------V----T-I-------------......A---RI------------I--H----Q  306
H2B.CI.x.EHO                 RVR.....A--MT-D--I--F---I-NSL-M-----VAR--P-...K-Q---EK----IR----SK---L--K---EK-----QLEEAP-T----S-T-------KN---M-I-------V--E-T---------------......AS--RI---------------P---Q  306
H2G.CI.92.Abt96              RIQ.....A--MT-D--I--F---I-VKL-MS-----AKV-P-...K-R----------IR----SK---Q--R---DK--Q--QLEEAP-T------T-T-----KN---M-I------RV----T---------------......AE-RRI----I------I---V---Q  289
H2U.CI.07.07IC_TNP3          -VR.....A-IMT-D--I--F---I-VKL-MS----VAKV----...K-T----------IR----SK---E--K---EK--Q--QLEEAP-T------T-----R-KN---M-I------RV----T---------------......AE-RRI-----------I----Q--P  289
H2U.FR.96.12034              TIR.....--IMT-D--I--F------EL-MS-----GKEAL-...IGG----K----LR----SK--LE--K---EK-----QLEEAP-T------T-------KN---M-I------RV----S---------------......AE--RI-----------I--C----Q  303
ASC.UG.10.RT03               RVT.....EE----D--I--M---F--KF-V---MVVGKFKPT...E-R-RE-K-----------SA---Q--K--VQD-VDK-QLE---------S---C-R---K----M-I---K--EH---LA---------------......P--AQ-SIV-IK---YAI--H---Q-  291
ASC.UG.10.RT08               RVT.....-E----D--I--M---F--KF-V---MVVGKFKPT...E-R--E-K---R-------SA---Q--K--VQD-VNK-QLE---------S---C-----K----M-I---K--EH---LA---------------......P--AQ-SIV-IK---YA---H---Q-  294
ASC.UG.10.RT11               RVT.....EE----D--I--M---F--KF-V---MVVGKFKPT...E-R--E-K-----------SA---Q--K--VQD-VDK-QLE-V-------S---C-R---K----M-I---K--EH---LA---------------......P--AQ-SIV-IK---YA---H---QR  294
COL.CM.x.CGU1                EYR.....-DI-I-H--I--V---M-KEGKLVVAQLSDR-PVT...K-T--E-------------SK---EG-QK--ERL-A----E-AELG------I-C-R---KNE----I-----------M--------G--......IE-THI----I------I---P-YQ-  263
COL.UG.10.BWC01              EY-.....-DI------I------I-KDGKIV-AQLSDR-PIT...K----E-----R------SK---EG-QK--ERL------E-AAHG------I-C-R---KNE----I-----------F---Q-----G--......VE--NI----I------I---P-YQQ  265
COL.UG.10.BWC07              EYK.....-DI-LAD--I--L----IGEGKIV-AQLSAN-PVT...K-S----L---R------SK---EG-QK--DRL-A----E-AALG------I-C-----KNE----I---------------I-------G--......HE--NI----I-----------PEYQ-  265
DEB.CM.04.SIVdeb04CMPF3061   V-V.....-E----E--I--L---F-APL-FSI-MLE-Q-QYT...------E-KS-------ASSQG-DRGSQGNSGKDGGGRPMAPAS-T----S---------KN---M-I-L------A--PV----T---------------......PQ-RQI----IA---Y-I---PS-AQ  290
DEB.CM.99.CM40               V-V.....SDL---D--I--L---Y-AKM-VSI---VLDYT....------E-A----I-----SR---E--KD-LKP-IEA-Q-VPAA-T----S---V------KS---M-I-L-A---A--EV----T-----SA-......PQMEQI----LA---Y-IA--PT-A-  287
DEB.CM.99.CM5                V-V.....SD----D--I--L---Y-AKM-VSI---VLDFT....------E-A---RI-----SR---E--K--IRP-VEA-Q-VPTE-T---SS---V------KN---M-I-L-A---A---V----T-----S--......PQMEQI----LA---Y-I---PQ-AR  287
DEN.CD.x.CD1_CM0580407       V-V.....D--------I--L---Y-SKL-V---MAVQQMRPQ.VT-------F---RI------T--L---QQ-V-Q-----RLE-AP-T---------V-P--K-QY-M-I---K--EA------------------......QQ-EHI----LK---Y-----P--AQ  307
DRL.DE.11.D3                 ITH.....AS--------------V-CKL-A-----V-QV-P-...K-T----------I-----SK---E--K---EQ-----Q--R-------------C-----G----------Q---A----F---------G--......-ACEQI-----------C---P----  295
DRL.x.x.FAO                  ITH.....AS--------------V-CKL-A-----V-QV-P-...K-T----------I-----SK---E--K---EQ-----Q--R-------------C-----G----------Q---A----F---------G--......-ACEQI-----------C---P----  295
GRV.ET.x.GRI_677_gri_1       IIT.....-SI-I-S--I------I-A-A-MK-VMGVLSSQIE.ETK-Q--E-K----L-----SR---E--T---KQ--E---L-R--G-----------------K-Q--M-----------A----F-----------------......Q---QI--I-I----Y-I--CKE---  303
GSN.CM.99.CN166              RIK.....A-----------L---C-SKL-V---MVQQK--PI...E-H--A-KE---I-----SR---E--TQ-TQ--LRL-QLE---------S---------KSQ--M-I---K--EA---LA-----------IPHPAGLEQ-EH--II-MK---Y-I--Y-----  292
GSN.CM.99.CN71               RIK.....A-----------L---C-SKL-V---MVQQKL-P-...E-H--S-KE---I-----SK---E--TQ-TQ--L-L-QLE---------S---------KSQ--M-I---K--EA---LA------------------......EQ-EH--II-MK---Y----Y-E---  285
LST.CD.88.SIVlhoest447       TTV....----L-----D-V---I--NV-AR-VLSQLSDKIP.ITK------C---R------SK---EG-QA--DRL---------ADAG------I-------KNE----I---Q--AL----H-----------I.......--C-Q---V-I-------I---P-Y-Q  297
LST.CD.88.SIVlhoest485       TTV....----L-----D-V---I--SV-AK-VLAQLSDKIP.ITK-R----C---R------SK---EG-QA--DRL---------ADAG------I-------KNE----I---Q--AL---LH-----------I.......--C-QI--V-I-------I---P-Y-Q  297
LST.CD.88.SIVlhoest524       TTV....----L-----D-V---I-NPL-AR-VLAQLSEKIP.ITK-H----C---R------SK---EG-QA--DRL-------PVD-G------I-------KNE----I---K--EL----H-L---------I.......--C-QI--V-I-------I---P-Y-Q  298
LST.KE.x.lho7                TTV....----L-----D-V---I--AM-AK-ILAQLSDKIP.ITK-S----C---R------SK---EG-QA--DRL-------PVD-G------I-------KNE----I---K--EL----H-L---------I.......--C-QI--V-I-------I---PNY--  297
MAC.US.x.251_1A11            RIR.....--IMT-D--I--F------AL-MS--L--AKV-P-...K-A----KV---L-----SK---V--R---EK---D-QLEEAP-T------T-------KN---M-I------RV----T---------------......A-R-RI----I------I----E--Q  293
MND-1.GA.x.MNDGB1            GTF....----I----ID-----IMEKL-GK-ILAQLSDKIP.ITK-------V---RI-----SK---VG-QK--DRL-E-----RVD-G-N----I-------KNE----I--------L----H-L---------I.......--C-RI----I------I---P-Y-P  290
MND-2.CM.98.CM16             TTH.....AS----------V---V-KKL--------V-KV---...K-T----T----I-----SK---L--Q---SQ-----Q----------------C-----G-S---------Q---V----F---------G--......-QCEQI----I------C---------  293
MND-2.GA.x.M14               STH.....AS--------------V-KKL---------K-----...K-T---------RI-----SK---L--QT--D---------R------------C-----G-S---------Q--RV----F---------G--......-QCEQI-I--I------C----S---  294
MND-2.x.x.5440               ITH.....AA--------------V-KKL---------K-----...K-T----T---RI-----SK---L--Q---NQ---------RV------------C-----G-S---------Q---V----F-----T--G--......-QCEQI----I------C---V----  294
MNE.US.x.MNE027              RIK.....--IMT-D--I--F------AL-MS-----AKV-P-...K-T----K----LR-----SK---V--R---EK---D-QLEEAP-T------T-------KN---M-I-------V----T---------------......A-R-RI-----------I----E--Q  289
MON.CM.99.L1_99CML1          EVW.....A--------I--L---V-AKM-V---MAAGGDLQPPSLE-T--A-KE---L-----SR---E--TQ-TQ--L-L-QLEPTE-N----S-C-------KS---M-I-------A----F---------------......E-MDH--IV-MK--FY-I--W-P---  289
MON.NG.x.NG1                 TVR.....A---L----I--L---V-SRM-I---MXQGIAQSP.SLT-T----K----L-----SK---E--TT-XR--IEL-QLEEAE-T----S-C---R---K----M-I-------A----F---------------......QNMANI-IV-MK--FY-I----N---  291
MUS-1.CM.01.CM1239           TIK.....A--------I--M---C-SKF-I---MIQEK--PI...K-A--E-AK--M-------Q------EG-VQQ-L-LDQ-EE---D----S-C---R---KS---M-I---Q--EA--E-T---------------......AEHEH--IV-IK--FY-----PA---  282
MUS-1.CM.01.SIVmus_01CM1085  TIR.....A--------I--M---C-SKF-I---MIQGKL-PI...K-T--E-AK--M-------Q----R--ED-TQQ-L-V-Q-EE----------S-C---R---KS---M-I---Q--EA--ELA---------------......IEHEH--IV-IK--FY-----P----  282
MUS-2.CM.01.CM1246           RIV.....A----S---I--L---C-SKLEI---MVMAGQQLK.PTK-S--ERKK--M-------SK---E--K--TQK-I-L-QLEEA--N----S-----R---K----M-I---K--EV--E-A---T-----S--......AQ-AH--IV-MQ--FY-----KE--P  287
MUS-2.CM.01.CM2500           RI-.....A--------I--L---C-SKL-I---MVGGQL-PT...K-G--A-KE--M------SK---E--E--TQR-L-L-QLEET--Q----S-A---R---KS---M-I---K--EA----S---T-----S--......S--AH--IV-MQ--FY-I--Y-E--P  286
MUS-3.GA.09.09GabOI81        YIV.....A-----------M---C--KL-F---CVQTSL-PL...K-R--E-KE-----------A---D--IN-TQ--L-L-QLEEA--D----S-C-------KS---M-I---K--ES----S---------------......SR-RH--IV-MK---Y-----P----  283
MUS-3.GA.11.11GabPts02       EVY.....AN--I-------L---V-S-M-V-----VTG-K-I...-----E-K-----------A------EQ-TQ--V-L-QLEE---------S-C---R---K----M-I---K--EN---LF---------S--......AEHEH-----MK---Y-I--------  286
OLC.CI.97.97CI12             EIQ.....EA----D----VV---V-AAV-IR-VCAQLSDKIP.ITK----E-HKP-QI-----SK---E--EK--DRL-E----Q-AEAS-QW---I-C-----KNEY---I---A--AA-------------------......-EREQI----I------I---PEY-E  283
RCM.CM.00.SIVagi_00CM312     VTH.....AH--------------V-HKL-AS------KA-II...K-E--E-Q---R------SK---E--T---NA--------R----------I-C-R--------------------------F---------G--......--C-QI----I------C--Y-P---  294
RCM.CM.02.SIVrcm_02CM8081    VTH.....AH--------------V-HKL-AS------KA--I...E------Q----------SK---E--T---NA-------------------I-C-R--------------------------F---------G--......--C-QI----I------C--Y-P---  300
RCM.GA.x.SIVRCMGAB1          RTH.....AH--------------V-KKL-A-------Q---I...K-E--S-Q---R------SK---E--T---NA-------------------I-C----------------------------F---------G--......-QCERI----I------CL-Y-P---  295
RCM.NG.x.NG411               TTH.....AH--------------V-SKL-AS------KA--I...K-E---Q---R------SK---E--T---NA--------R----------I-C----------------------------F---------G--......-QCRQI----I------C--Y-P---  299
SAB.SN.x.SAB1                TCK.....A-I---E---------V-A-L-V---LTQRE--PI...K-H----Q---RIR----SK---E--KA--EDL--Q-HLER--------------R---K-Q--I-M---Q---S----Q---------------......QQREQI----I------C---P--Q-  313
SMM.CI.79.SIVsmCI2           RVK.....--IMT-D--I--F---I--AL-MS--C-LAKV-P-...K-T----KE---L-----SK---E--R---EK-----QLEEAP-T------T-------KN---M-I-S-----V----N---------------......AE-RRI----I------I--W-E--Q  298
SMM.LR.89.SIVsmLIB1          RIR.....--IMT-D--I--F------EL-MS----VAKV-P-...K-------K----LR----SK---E--K---EK-----QLEEAP-T------T-------KN---M-I-------V----T-I-------------......A-RRRI------------I----E--Q  289
SMM.SL.92.SIVsmSL92A         EVR.....A-LMT-D--I--F---I--AL-IS--Y-VAKVQP-...K-T----K----QR----SK---E--K---EK-----QLEEAP-T------T-------KN---M-I-------V----T-I-------------......QQRRRI------------I---P---Q  289
SMM.SL.92.SIVsmSL92B         VIR.....E-LMT-D--I--F---I-AAL-VS--Y-VAKV-YT...K----E-----RL-----SK---Q--T---EK-----QLERAP-T------T------NKD---M-I-------M----T---------------......-EMERI----I-----------PE--Q  286
SMM.SL.92.SL92B              VIR.....E-LMT-D--I--F---I-AAL-VS--Y-VAKV-YT...K----E-----RL-----SK---Q--T---EK-----QLERAP-T------T------NKD---M-I-------M--E-T---------------......-EMERI----I-----------PE--Q  286
SMM.US.04.G078               -IK.....--IMT-D--I--F------AM-MS--L-VAKV-P-...K------KE---L-----SK---E--K---EK-----QLEEAP-T------T-------KS---M-I------RV----T-I-------------......A-RRRI------------I----E--Q  289
SMM.US.05.D215               -IK.....--IMT-D--I--F------TM-MS----VAKV-P-...K-T----K----LR----SK---E--K---EK---D-QLEEAP-T------T-------KN---M-I-------V----T---------------......A-RRRI------------I----N--Q  289
SMM.US.06.FTq                TIR.....---MT-D--I--F------RL-MS----VARV-P-...K-T----Q----LR----SK---E--K---EK-----QLEEAP-T------T-------KN---M-I------RV----T---------------......A-RRRI------------I---PE--Q  293
SMM.US.86.CFU212             RIR.....--LMT-D--I--F------AL-MS----MAQV-P-...K-T----K-----R----SK---Q--K---EQ---D-QLEEAP-T------T-------KN---M-I-------V----T---------------......A-R-RI------------I-------Q  289
SMM.US.x.H9                  VIK.....--IMT-D--I--F------AM-MS--L--AKV-PI...K-T----K----LR----SK---I--R---EK---D-QLEXAP-T------T----XX-GN---M-I-------V----T---------X--......A-RRRI------------I----E--Q  289
SMM.US.x.PGM53               VIK.....--IMT-D--I--F------AM-MS--L-VAKV-PI...K-T----K----LR----SK---V--R---EK---D-QLEEAP-T------T-------KN---M-I------RV----T---------------......A-RRRI------------I----E--Q  289
SMM.x.x.pE660.CG7G           VIK.....--IMT-D--I--F------AM-MS--L--AKV-PI...K-T----KE---LR----SK---I--R---EK---D-QLEEAP-T------T-------KN---M-I-------V----T---------------......A-RRRI------------I----E--Q  289
STM.US.89.STM_37_16          -IK.....--IMT-D--I--F------AL-MS----VAKV-P-...K-T----K----I-----SK---E--K---EK---D-QLEEAP-T------T-------KN---M-I-------V----T-I-------------......A-RRRI------------I----G--Q  289
SUN.GA.98.L14                T-T.....---------D-----I-SAL-GR-VLATLSEKIP.ITK----E-AA----------SR---EG-QK--DRL-A-----RAD-G------I-------KNE----I---V--GM----H-L-------------......--C-QI--V---------I---P-Y--  299
SYK.KE.x.KE51                DGRIKEVE-SL---------L---I-SKL-AK-VMIAQKL-PM...K-Q----KEL--L-------R--LE--RA-VSD-LEK-QLE-AS-T---------V-Q---KL---MPM---K--EA--------------G--......ST-R.M----LK---YT-----E--Q  319
SYK.KE.x.SYK173_COMGNM       SGKRKQVE--L---------L---I--KL-VK-VMVQTAL-P-...K-S---DKEL-RL-----SV--LE--KA-VED-L-A-QLE-AS-T---------V-R---KK---M-I---K--EA----F---------G--......-RQ-.L-II-LK---Y-----KE--P  313
TAL.CM.00.266                R-R.....D---I----I--L---F-AKFSV---MVS-D-PIT...K----E-KEP--I------R---EG-QI-IDD-V-A-QLEEV-------S-----R---K-Q--M-I---A--QA----A---Q-----S--......PQM-QI----MK---Y-I---P--SQ  290
TAL.CM.01.8023               R-R.....D---I----I--L---F-AKF-V---MVSTE-PVI...K----E-KEP--I------K---EG-QA-IDD-V-A-QLEE---------S-----R---K-Q--M-I---A--QA----A---Q-----S--......PQ--QI----MK---Y-I---P--AR  291
TAN.UG.x.SIVagmTAN1          ETR.....-N--I----I--V---I-APL-AR-IMGQLNEKIP.ITK-Q------S---L-----S-------T--TQAL--------V-G---------C-R---K-Q--M-I---A---A----F---------------......-R--QI----I----Y-I---PE--P  301
VER.DE.x.AGM3                ILR.....--I-I-S--I------I-APA-AK-VMGQLSEQIP.IT-----E-AR--FL-----SK------Q---DQL----------G--A------C-----KSQ--M---------A----F---------S-F......E-MTEI----I----Y-I---PE---  298
VER.KE.x.9063                ILR....----L-A--I--------APA-AR-VMGQLSEQIP.IT-----D-AR--HI-----SK------T---EQL----------G--A----I-C-R---K-Q--M---------V----F---------------......Q-M-QI----I----Y-I---PE---  298
VER.KE.x.AGM155              ILR....----I-A--I------F-A-A-AK-VMGQLSQTIP.IT--R--E-AR--RL-----SK---I--Q---KTL-E---L-RV-GD-A------C-R---KSQ--M---------A----F---------------......--M-QI-II------Y-I---PE---  301
VER.KE.x.TYO1_patent         ILR.....--I-L-A--I--------APA-AR-VMGQLSEKIP.VT-----E-AR--C-R----SK---E--Q---SQL-Q-----RV-G--A----I-C-----KSQ--M---------A----F---------------......R-MRQI---------Y-I---PN---  301
WRC.CI.97.97CI14             TIE.....ADIMT-K--I------V-NKM-MH-A-AQLSDKIP.ITK-S----CS--R------SKK--EG-QK--DRL-A-----RAD-G------I-------KNE----I---K--EV----F---------G--......R-N-NIS-V-I------I---PEY-Q  305
WRC.CI.98.98CI04             TVE.....ADIMT-K--I------V-EKM-MH-A-AQLSDKIP.ITK-S----CT--R------SK---EG-QK--DRL-A------AS-G------I-------KNE----I---K--EV----F---------G--......R-N-NIS-V-I------I---P-Y-Q  304
WRC.GM.05.Pbt_05GM_X02       EVR.....ANLMT-K--I------V-EKL-MH-A-AQLSDKIP.ITK-S----CQ--RI-----SK---EG-QK--DRL-A-----RAEAN------I---R---KNE----I---G--AV----F---------G--......H--RNI----I-----FI---PEY-Q  296
```


K219Q
T215Y
polymerase motif
DD catalytic site

Draft of 30 September 2015, not for public release
plvprot_main.tex r8876–10546

```
                                                                                                                                                                              p51 RT end p15 RNase H start
H1B.FR.83.HXB2              LTEEAELELAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEP.FKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTP.KFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTNRG 617
H1O.CM.91.MVP5180           -SK------E----K----------Q-D---WVS---H-E------V--DE.H----------QKAS----IR--A-V---VSQ-A------L-.--R--VTR-------A--------------S----I-------T-----------------N----------EQ- 613
H1N.CM.95.YBF30             F-----------------L-------G-E-V------------------L.H---------K--S-----I---V-V-R-VA-----------.--R--V---V--A---DH-------------------------T---S-------------------F--D-- 624
H1P.CM.06.U14788            F-R------E--K-----------Q-E-E--VD-----A------VF--E.H----------QKAT----IR--A-VI--VSQ--------L-.--R--VNRNV-----SD--------------S----I--------D--P-T---------------------D-- 614
CPZ.CD.06.BF1167            --A------E--KQ----T-Q-T----KEP-K-AV----K------V--KE.NQI---------NH-T---EL---AGV----G----I---RV-.V-Q--VT--V-QQ--SD---VS-------------I----N-LT--MPE---Y--------DS---------D-- 620
CPZ.TZ.00.TAN1              M-R------E--KQ----K-Q-A----KLP-Q-A-----------------E.G---------KSP-T---EIR--AGLI---GN---I---IV-.--L--VS----SQ--D---V--V-----I-----IR---N-LSD--PE-----------DS-K-R--------- 613
GOR.CM.07.SIVgorCP2135con   -SK------E---------------Q-E-E--VNV----E-------F-DE.Y----------QKAT----IR--A-VI--VSQ--------L-.--R---CRS--DAF--D----S-------S----I-----------P-------------S-Q-------D-- 609
                                                                                                                                                                             p51 RT end p15 RNase H start
MAC.US.x.239                W--M--A-YE--KI--SQEQE-C--QEG-P-E-TVI-SQDN--S-K-H--D..-I--V--F-KIKNT---G-RL-AHVI---GK-A-----QV-.--H--VE-DV--Q---D---V------D-IS-----R-VFN-V-D--E-E--Y-T--SC-KQS-E-----I-D-- 628
H2A.SN.86.ST_JSP4_27        W--L--A--E--KI--SQEQE-C--QEE-E-E-TV--DQDN----K-H-GG..-I--V----KVKNT---G-RL-AQV----GK-AL----RI-.--H--VERD---Q--DN---V----D-D-IS-----R-VFN-V-D--L------T--SC-KQSRE-----I-D-- 641
H2B.CI.x.EHO                W--L--A-FQ--KI--EQEQE-S--KEGVP-E-TV--NLAN----K-H-GD..-I--V----KVKNT---G-RL-AHV----GK-AL----EI-.M-H--VER---DQ---D---V------D--S----IR-A-N-V-D-LE-V--Y-T--SC-KAS-E-------D-- 641
H2G.CI.92.Abt96             W--L--A--Q--KI--XQEQE-A--KEXEP-E-TV--NLDN----KVH-GN..RI--V----KIKNT---G-RL-AHV----GK-AL-----L-.F-H--VERD---Q---D---V----D-D--S--H--R-V-N-V---LEQT--Y-T--SC-KTS-E-------D-- 624
H2U.CI.07.07IC_TNP3         W--L--A--Q--KI--EQEQE-T--REGEP-E-TV--NLDN----K-H-GN..-I--V--F-KIKNT---G-RL-AHVI---GK-A------V-.I-H--VER----Q---D---V------D--S-----R-A-N-V---LERV--Y-T--SC--NS-E-------D-- 624
H2U.FR.96.12034             W--M--A-F---KI--SQEQE-A--REEN-E-TVL-NQDN----K-H-GD..RT--V--F-KVKNT---G-RL-AHV----GK-A------I-.M-H--VER-V--Q---D---V--VLD-D--S-----R-A-N-V---LL.E--Y-T--SC--AS-Q-------D-- 637
ASC.UG.10.RT03              V-----T-----QA--DQEAK-G---AE-P-EVD-IQL-GT--G-TVR-DK..EV-----F-KQ-S--S-PF---VD-IM--GK--L-Y--RI-.Q-NV-VN--Q-DA--N-H--VS---DIKP-H----IQM----VQD--SE---W-I-------S---------D-- 626
ASC.UG.10.RT08              V-----T--V--QA--DQEAK-G---AE-P-EVD-TQL-ET--G-TVR-DK..EV-----F-KQ-S--S-PF---VD-IM--GK--L-Y--RI-.Q-NV-VN--Q-DA--S-H--VS---DIKP-H----IQM----VQD--P----W-I-------S---------D-- 629
ASC.UG.10.RT11              V-----T-----QA--DQEAK-G---AE-P-EVD-IQL-GT--G-TVR-DK..EV-----F-KQ-S--S-PF---AD-IM--GK--L-Y--RI-.Q-NV-VN--Q-DA--S-H--VS---DIKP-H----IQM----VQ---PE---W-I-------S---------D-- 629
COL.CM.x.CGU1               -SA---E--QD-KQ-I--E-Q----I-HE-IWVD-FRL-K---G-AVC-KH..GT-RR--HNTGKNQ-F-SMQE-ASVI---GR-A--T---V-.-M-V-AKR-D--Q--SD---SA----I---SSSYV---VWN-VL--LKE-P-YWT--GCS-KLGA----WINSK- 597
COL.UG.10.BWC01             -SA---E--E--KK-V--EQK----R-EE-IWV--MRL-K---G-AVQ--H..GT--R--INTGKS--F-SMQE-ATVI---GR-A------V-.RM-V-AKR-D--Q--SD---S-----I---S--YV---VWN-ML--VRS-P-YWT--GYS-KDNW----WINSS- 599
COL.UG.10.BWC07             -SGQ--E--EQ-------EQE----K-EEEIWVT-FRL-K---G-AVE-VH..GT--R--VNTGKS--F-SMQE-AMVI---GR-AL-----V-.LM-I-AKR-D--Q--SD---S-----I---S-S-V---VWN-VLS-LEKEV-YWT--GYS-KEGI----W-NSK- 599
DEB.CM.04.SIVdeb04CMPF3061  -----QA--E-A-Q----Q-S-T--KEQE----D-T-LSE---G-T-R-SK..GI--------KAK---F--FH-IAKLMM-VG-----T--RL-.T-R--V--QD-DG--H-H----------A-H-----R---S-VS---KE-D-Y---------S-E-----ISE-- 625
DEB.CM.99.CM40              -S---QA--R--K----QE-S----KEQEP----LS-L-K---G-V-R-PK..GI-----FS-DK---Y--FH--AK-MY--G-----F--RI-.Q-R--VV--E-DN--HNH---A---D--AIH--H------E-VS---PD-D-Y---------S----------EW- 622
DEB.CM.99.CM5               -S---QA--R--K-TP-QE-S----KEDEP--S--S-I-R---G-V-R-TK..GL-----FSKDK-T-Y--FH--AK-MY--G-----Y---V-.--R--VV--E-DN--HNH-------D--A-H--Q--R---E-VS---PD-D-Y---------S-M-----ISEW- 622
DEN.CD.x.CD1_CM0580407      ---G--A--K-AQ-----SIS-S--R-E-P-----S-L-EN--G-T-K--K..HL------HKDK-S-Y-PYQ--AKVMA--GR-AL----RL-.T-R--VS-VD-DA--S----VN---D--M-S----LRNF-N-VS---DK-P-Y-T------Q--Q-R----ATN- 644
DRL.DE.11.D3                W-R-----YE--KL----QM-----Q-E-P-K-K---L-N---S---E-DD.N-P-------KVKNS----MRM-AGL----AK-AL----RL-.T-Y--VER-V--Q--Q----V----D----S--H-I----N-LSD--P-E-VY---------VS-E-------S-- 631
DRL.x.x.FAO                 W-R-----YE--KL-I--QM-----Q-E-P-K-K---L-N---S---E-DD.N-P-------KVKNS----MRM-AGL----AK-AL----RL-.I-Y--VER-V--Q--Q----V----D---IS--H-I----N-LSD--P-E-VY---------VS-E-------S-- 631
GRV.ET.x.GRI_677_gri_1      W-----A-YKN-QG-VQ-TQE-T----L-E---TV----E------FT--G..AV--V-R--KQ-ET----LRT-AHL----CK-ALT---RL-.RVQ--VD-K--DM--QD---VS---------S--L------S-V----K-EDVY------SKV---------LSE-- 639
GSN.CM.99.CN166             -----QI----------QSSE-G---AE-P-VV--TSL-EQ--G-MFS--G..-M--SA-F-KT-N--I-SYQ-FAD-LT-VGR--Q-T---V-D--RI-VV--Q-DA--MN---------TI-AIH--H-LRQ--T-VS--LED-V-Y---------TS-----------T- 628
GSN.CM.99.CN71              -----QI----------QSSE-S---AE-P-VV--TSL-EQ--G-MFS-DS..-M-RS--F-KT-T--I-SYQ--AD-MT-VGR--Q-T---V-D--RI-VV--Q-D---MN---------TI-A-H--H-LRQ--T-VS--LEN-V-Y---------TS-----------T- 621
LST.CD.88.SIVlhoest447      W-DA-LE-YEQ-QQ----KLQ-A----T-E-VVRV-QSKK-II-F-WR-GN..SI-RA-R-Q-QKS----PLQK-A------GK-------HV-.-IQV-VTR-V-DQ--SDH--V-----L--IS--Q-EQE--S--A---E-VD-Y------EKVG-T-------QS- 634
LST.CD.88.SIVlhoest485      W-DA-LE-YEQ-QQ----KLQ-A----T-E-VVRV-QNKK-VI-F-WR-GN..SI-RA-R-Q-QKA----PLQK-A------GK-------HV-.-IQV-VTR-V-DQ--SDH--V-----L--IS--Q-EQE--S--A---E-VD-Y------EKVG-T-------QS- 634
LST.CD.88.SIVlhoest524      W----LE-YET-KT----K-Q-A------E-VVRV-QNKK-VI-F-WR-GN..AI-RA-R-Q-QKS--M-PLQ--A------GK-------FV-.-IQV-VTR----N--INF--V----DL--IS--Q-EQE--S--A---E-TD-Y------E-VG-T-----I-QS- 635
LST.KE.x.lho7               W----LE-YGQ-K-V---KMQ-A----E-E--VRV-QNKK-II-F-WR-GN..NI-RA-R-Q-QKA---PLQK-V--I---GK-------FV-.-IQV-VTR-V--H--SDH--V----DL--IS--Q-EQE--IW-A---I-VD-Y------EKVG-T-----I-QS- 634
MAC.US.x.251_1A11           W--M--A-YE--KI--SQEQE-C--QEG-P-E-TVI-NQDN--S-K-H--D..-I--V--F-KIKNT---G-RL-AHVI---GK-A-----QV-.--H--VE-DV--Q---D---V-------N-IS-----R-VFN-V-D--E-E--Y-T--SC-KQS-E-----I-D-- 628
MND-1.GA.x.MNDGB1           M----R--YEQ-K---A-EQE-S----N-E-YVRF--TTG-DISF-WK-GN..-V-RA---GKQKT--S--LMK-AG-T--VGR-------FV-.-MQI-TTR-I--D--H----C-----V--IS--M-ERE--S-SP--LE-V--Y---------DS-M-----I-D-- 626
MND-2.CM.98.CM16            W-R-----YE--KL-VQ-E------Q-E-P-M-KV--LT----S---E--D.N-P--V-----TKN----ELRV-AGL----AK-AL-----L-.--Y---ER-V-DQ--P--------------S--H-IG---N-LR--VP-EDVY---------NS-E-------A-N 629
MND-2.GA.x.M14              W-R-----YE--KL-VQ-E------Q-G-P-M-KV--LT----S---E--E.N-P--V-----TKH----ELRV-AGL----AK-C-----EL-.--Y--LER-V-DQ--HD-------------IS----IR---N-L----P-EDVY---------NS-E-----Y-A-- 630
MND-2.x.x.5440              W-R-----YE--KL-VQ-E------Q-E-P-M-KV--LT----S---E--D.N-P--V-----TKN----ELRV-AGL---VAK-CL-----L-.--Y--LER-V-DQ--HD---V------D-IS----IR---N-L-D--P-EDVY---------NS-E-----Y-S-- 630
MNE.US.x.MNE027             W--M--A-YE--KI--SQEQE-C--QEG-P-E-TVI-NQDN--S-K-H--D..-I--V--F-KIKNT---G-RL-AHVI---GK-A-----QV-.--H--VE-DV--Q---D---V------D-IS-----R-VFN-V-D--E-E--Y----SC-KQS-E-----I-D-- 624
MON.CM.99.L1_99CML1         -SDL-QA-----------EST-G----NEP-KV-LTSLAEN--G-RFF--K..SV-----F-KI-ST-S-TYQ--AD-LA-LGK-A--T--RL-.I-R--VV--Q-DA--ADN--IN-V-DI-A-YS-H-LRQ--T-VQD--E--P-Y---------TS-E--------F- 624
MON.NG.x.NG1                -SDL-QA---------SQQST-G----QL--KV--TSLXE---G-XYF--X..RV-----F-KIKN--S-TYQ--AD--T-LGK-AL-----L-.I-R--VN--Q-D---ADN--V-----I-P-HS-H-LRQ--T-VQX--EH-P-Y---------AS-E-------ST- 626
MUS-1.CM.01.CM1239          -----Q------Q----QA-Q-A-F--EQP-V---VSL-DS--G-NFL-NK..GI--S--F-KV-SV---SYQ--SD-IAR-GR--L-----P-E-VRI-VI--Q-DQ----H--VS---DI-A-H-TH-LRQ-FT-VP--LSE-P-Y------HKVS-----------T- 618
MUS-1.CM.01.SIVmus_01CM1085 -----QA-----Q---QQT-Q-A----EQP-V---VSL-DS--G-NFT-SK..GV-R--F-KV-SV---SYQ--AD-MA--GR--L-----P-D-VRI-VI--Q-DQ--IAH--VS---TI-A-H-TH-L-Q-FT-VP--LKE-P-Y---------VS-----------T- 618
MUS-2.CM.01.CM1246          -S-Q-Q------------QEAG-A----E-P-VL--VSL-EQ--G-TFT-DR..NM-R---F-KI-T--S-PYQ--A--LSRASK-AL-C-----D-CRI-VV--Q-DN--ADS--T----DI-A-H--Y-LRQ-FT-VP---ET-P-Y---------NS---R-----D-- 623
MUS-2.CM.01.CM2500          ---Q-Q----------M-QE-K-A----E-P--L--VSL-EQ--G-TFT--K..NM-----F-KV-T--I-TYQ--A--LA-AAR--L-C-----D-CRI-VV--Q-DH--A-S--T--V--I-A-H--H-LRQ-FT-VP--LEA-P-Y---------NS---R-----D-- 622
MUS-3.GA.09.09GabOI81       -S-L----M------M-QE-Q-D----E---VV--TSL-G---G-MFS-NS..-P-----F-KV-SV---PYQ---D-MT-LGR-A--Y---V-D--RI-VI--Q--K--ADH--V----DV-A-HS-H-LRQY-T-VQ---EK-P-Y---------TS-A-R------T- 619
MUS-3.GA.11.11GabPts02      ---A--A--------I-QE-Q-R----G-E-TV--TSL-EQ--G-TFS--N..-H-----F-KV-TT-S-SYQ--AD--TRLGK-AL-Y---I-DV-RI-VI--Q-DN--A----V--V-NV-AIH--H-LRQY-T-VQ---E--P-Y-T------NS-----------T- 622
OLC.CI.97.97CI12            W----KE--E---------EQE--M-KEEEP-QCT.LCYTK-TVA-R-W-KH..GT----RFQ-KMT----KFR-M-G-IE--GK---------L-.I-QV-TTRKD--A-AFDD--VH---DI---SNSENQRTFWE-VPA-M.EV---------Y--KS-E-------SSG- 616
RCM.CM.00.SIVagi_00CM312    F-K-----YE--KLL---KL-------E-P---KV--LEG---S--VE-GD.G-P-------KQKT----EIRM-AGL----CK-A------L-.--E--VER-V--Q--AD---VS---------S----IR---N-V-D--P-E-VY---------NS--------I-D-- 630
RCM.CM.02.SIVrcm_02CM8081   F-R-----YE--KLL---QM-------E-P---KV--LTG---S--VE-GD.G-P-------KQKT----EIRM-AGL----CK-AV-----L-.--E---ER-V--Q--AD---VS---------S----IR---N-V-D--P-E-VY---------NS-Q-----I-D-- 636
RCM.GA.x.SIVRCMGAB1         M-K---I-YE--KM----KL------EK-P-V-N---LEG---S---E--S.G-P-------KQKT----EIRM-AGL----AK-A-----RL-.T-R---ER-V-D*-RSQ---V----D----S----IR-G-N-V-D--P-E-VY---------NS-I---------- 630
RCM.NG.x.NG411              F-R-----YE--KAV---KI-----N-E-P---KV--LEG---S---E-GD.S-P-------KQKT----EIRM-AGL----CK-A------L-.--E---ER-V--Q--SD---VS---------S----IR---N-V-D--P-E-VY---------NS---------D-- 635
SAB.SN.x.SAB1               W--------E---Q--RQKQQ-Q----ALP-R-KVL-L-D---G-----PE.N-I--V----KIKT----ELRM-AGL----GK-------QI-.IME--VER-L--Q--SD---V-------M-S--Q-IR---K-V-D--P-EAVY---------NS-E-----L-D-- 649
SMM.CI.79.SIVsmCI2          W--L--A------I--EQEQE-AF-REEEN-E-TV--NQDN----K-H-GS..RV--V--F-KIKNT---G-RL-AHV----GK-AL-----I-.F-H--VER-V--Q---D---V----D-D--S-----R-VFN-V-D--P-E----T--SC--VS-E-------D-- 633
SMM.LR.89.SIVsmLIB1         W--M--A-YE-SKI--SQEQE-C--QED-P-E-TVV-NQDS--S-K-H--D..-I-RV--F-KVKNT---G-RL-AHV----GK-AL--*-QI-.--H--VER-I*-Q---D---V-----*D-IS-----R-VFNPV---LE-V------RSC---S-I-----I---- 620
SMM.SL.92.SIVsmSL92A        W--L--A-----KI--SQEQE-S--QEGEP-E-TAI-NQEN--S-K-H-GE..-I--V----KIKNT---G-RL-AQVI---GK-A------V-.--H--VER-L--Q---D---V--V-D-D--S----IRMAFN-V-D--P-E--Y----SC-KQS-A-------D-- 624
SMM.SL.92.SIVsmSL92B        W--L--A------I--NQEQE-R--REDEP-E-TVL-NQDN--S-K-H-GD..RI--V--F-KIKNT---GIRL-ANV----GK--L--------.F-H--VER-V-DQ---D---------D-IS----IR-VFN-V-D--EKE-VY-I--SC--NS-E-------D-- 621
SMM.SL.92.SL92B             W--L--A------I--NQEQE-R--REDEP-E-TVL-NQDN--S-K-H-GD..RI--V--F-KIKNT---GIRL-ANV----GK--L--------.F-H--VER-V-DQ---D---------D-IS----IR-VFN-V-D--EKE-VY-I--SC--NS-E-------D-- 621
SMM.US.04.G078              W--M--A-YE--KI--SQEQE-C--QED-P-E-T-V-SQDN--S-K-H-GE..-V--V----KIKNT---G-RL-AHV----GK-A-----QV-.--HM-VER----Q---D---V------D--S-----R-VFN-V-D--E---------SC-KQS-E-----I-D-- 624
SMM.US.05.D215              W--M--A-YE---I--SQEQE-C--QED-Q-E-TVV-NQDN--S-K-H--G..-T--V--F-KIKNT---G-RL-AHVI---GK-A-----QV-.--H--VER-V--Q---D---------D--S-----R-VFN-V-D--E-E-------SC--QS-E-----I-D-- 624
SMM.US.06.FTq               W--M--A-FE---I--NQEQE-C--QED-P-E-T-V-SQDN--S-K-H--N..-I--V----KIKNT---G-RL-AHV----GK-A------V-.--HM-VER-I--Q---D---V------D-IS-----R-VFN-V-D--E---------SC--Q--V-----I-D-- 628
SMM.US.86.CFU212            W--L--A-YE--KI--SQEQE-C--QEG-P-E-TVV-NQDN--S-K-H--D..-V--V--F-KIKNT---G-RL-AHV----GK-A-----RI-.--H--VER-I--Q---D---V----D-D--S-----R-VFN-V-D-LE-E----T--SC--QS-E-------DG- 624
SMM.US.x.H9                 W--M--A-YE--KI-PSQEPE-C--QEG-P-E-PVI-SQDN--S-K-H--D..-I--V--F-KIKNT---G-SL-AHV----GK-A-I---QV-.R-H--VEX-I--Q---D---V------D--S-----R-VFN-V----Q--------RSC--QSRE-------D-- 624
SMM.US.x.PGM53              W--M--A-YE--KI--SQEQE-C--QEG-P-E-TVI-SQDN--S-K-H--D..-I--V--F-KIKNT---G-RL-AHV-E--GK-A-----QV-.--H--VER-I--Q---D---V------D--S-----R-VFN-V----Q----------SC--QSRE-------D-- 624
SMM.x.x.pE660.CG7G          W--M--A-YE--KI--SQEQE-C--QEG-PIE-TVI-SQDN--S-K-H--D..-V--V--F-KVKNT---G-RL-AHV----GK-AL----EV-.--H--VER-I--Q---D---V----D-D--S-----R-VFN-V----Q----------SC--QSRE-------D-- 624
STM.US.89.STM_37_16         W--M--A-Y---KI--SQEQE-Q--RED-P-E-TVV-DQDN--S-K-H--D..-I--V--F-KIKNT---G-RL-AHVI---GK-A-----QI-.--H--VER-V--Q--AD---V------D--S-----R-VFN-V---LE-T-------SC--QS-E-----I-D-- 624
SUN.GA.98.L14               M----RE-YQS-Q-V-Q-S-S-S--E-D-E--CRV--VK--IL-F-WL-GK..QV-RV-R-QKRGA--E-PCQ--AA-L---GR-------FV-.-IQV---R-I-SQ--AD---C-----L---S--K-EQE--T-AT--VP-.D-Y------EKLE-R-----I-QG- 634
SYK.KE.x.KE51               ------A-M--------VEQT-S--Q-E-P-E-H-S-L--Q--G-I-K-G-QE-P-I---TGKAYAT-H--YQA-AQLMN--GIQALW----I-.E-H--VKR-E--K---DH-----L--VKC-H--I--RWY-N-VS--V.E------------NS-E-Q---I-D-- 656
SYK.KE.x.SYK173_COMGNM      ---A--A--E---Q----EQA-S----K-P-E-H-T-L-SQ--G-M-K--QKGPP-I---T-KTFA--S--YQS-AQLLN--GIQ-LWY---V-.T-H--VKR-E--K---D------V--VK-IS-----RWY-N-VP---PE-V----------DS-T-N----ASD- 651
TAL.CM.00.266               -S----A---------A-ET--T--QEGEP-E--LT-LAE---G-M-K-GK..RL-----F-KQ-T--S-PYQ--IG-M---GK----F--QV-.V-RI-VV--E--Q----H--CS---TIVPIH-----R---N-VQ---E--D-Y------H--S-E-------AT- 625
TAL.CM.01.8023              ------A------A--A-ET--S--KEGEP-E--LT-LAE---G-MVK-GK..RL-----F-KQ-T--S-PYQ--VG-M---GK-A--F--QV-.I-RV-VV--E--Q--S----CS-L-TIVPIH-----R---N-VS---E----Y------H--S-E-------TS- 626
TAN.UG.x.SIVagmTAN1         W--Q--V-FRQ-Q-----EQQ-A------S-R-T-T-LDD---G--FK--G..-V------TKSKNT---EFRV-AGM--RVCK-A-TF---L-.IMEV-VER-V--Q--SD---V--L--I--IS----I----N-V-D--PK-KVY-C---C--NS------FL--T- 636
VER.DE.x.AGM3               W-----A-Y---K----TEQE-T--K-GRPIR-AV--LEG---S--FK--G..QV--V---TKQKNT---EFRV-AGL---LCK--L----EL-.VLE---ER-V--Q--AD---VS---D----S---------T-T----PKEDVY-----C--NSRE-----I-QY- 633
VER.KE.x.9063               W-----A-Y---K----TEQE-T--A-G-P-R-AV--LED---S--FK--G..-I--V----KQKNT---ELRT-AGL----CK-AL----EL-.I-E--VER-V--Q--AD---VS--------SV--------T-T----P-EDVY-----C--NS-E-----I-QY- 633
VER.KE.x.AGM155             W-P---A-YE--K----TEQE-T--A-E-P-R-AV--L-D---S--FK--G..-I--V--F-KQKAT---ELRV-AGV----GK-AL----QL-.T-E--VERD---Q--AD---VS-----D--SV----T---T-T----P-EDVY-----C--QS-E-----I-QQ- 636
VER.KE.x.TYO1_patent        W-P---A-Y---A----TEQE-T--K-GIPIR-AV--LEG---S--FK--G..QV--V---TKQKNT---ELRT-AGL----CK-AL----IL-.VLE---ER-V--Q--AD---VS-----D--S----L----T-T----PKEDVY-----C--NS-E-----ISQY- 636
WRC.CI.97.97CI14            W----LE--EI-KQ----EQS-A--KGKDLIAEL-KTGPGQWAYTIREIGE.K-P--V---G-Q-HT-S-ELRSMAH-M---G--------RV-.--RI--E-DI-DS--SD---VS---DT-H-S--F-ARE--K-VL--LPEV--Y-I--S---DS-E-----.WTKS 639
WRC.CI.98.98CI04            W----LE--EI-KS----EQM-A--TGKELVAEL-KTGSGQWAYSIREAGE.K-P--V---G-Q-HT---ELRS-AH-M---G---------V-.---I--E-DI-DS--SDH--VS---DT-H-S--Y-ARE--KIVL----EV--Y-I--S---DS-E----F.WTNS 638
WRC.GM.05.Pbt_05GM_X02      W-M---A--EI--A----SQQ-S--GGGDLIAEV-KTGTGQWAYTIKEAGE.T-P--V--FT-I-HT---ELRT-AHCI---G--A------L-.--RI-VE-DI-DM--S----------V-H-S--F-KRE--T-VS---E-V-EM-I--S--K-S-E-----.WTKT 630
```

Draft of 30 September 2015, not for public release

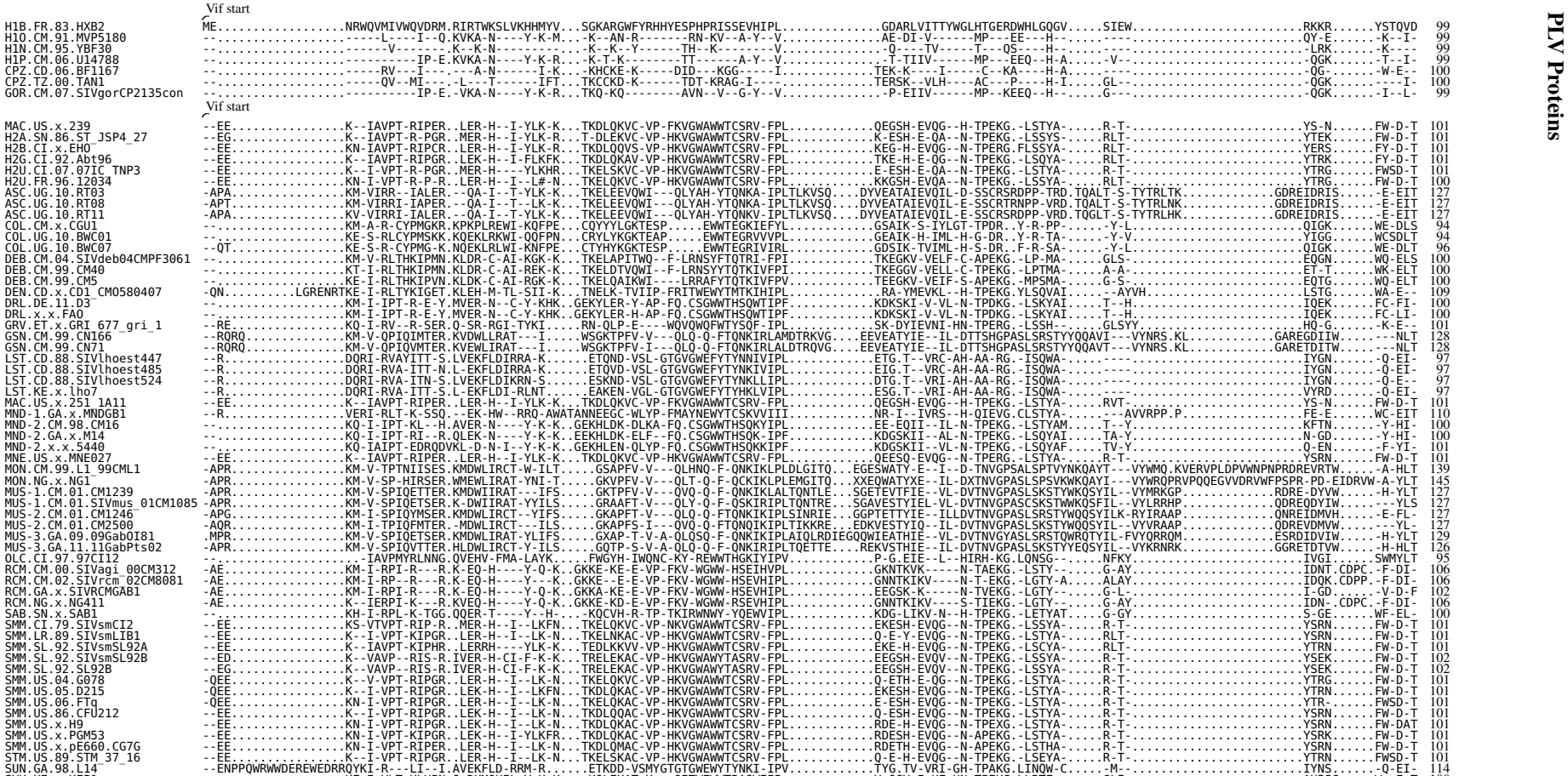
plvprot_main.tex r8876–10546


```
                                                                                                                p66 RT, p15 Rnase H end p31 Integrase start
H1B.FR.83.HXB2                RQKVVTLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIR..KVLFLDGIDKAQDEHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLK.GEAMHGQVDCSPGIWQLDCTHLE 784
H1O.CM.91.MVP5180             K-NIIK-EE-----A--M-VLI-----KEQ---------V----SS--T--D-PI-Q----E-T---R---T------------KI-----KD--..R----E---Q--ED----------L----G---I-----I---P--HI-.---T-----Y--E---M------ 780
H1N.CM.95.YBF30               -----SIA------A-----LM---E--RD----------M---HS---K------S----E-----R---S------------------S---..-I-----E---ED-DR-----K----------I---------------.--------N----V-------- 791
H1P.CM.06.U14788              ---IIK-EE-----A--E-VL---KE--EQA--------V----S-T----D-P--QK---EMT-------S------------NI-----KD--..R----E---Q--ED----------L--E-G---I-----INN-P--HV-.---------------------M- 781
CPZ.CD.06.BF1167              KSR-IE--E----QA-----L---K-AKEK---------V---L-GT----D-P--EE--QE-MN--A---N------------E------KN--..-----E---Q--ED--RF----KML-DEY----I-----I-Q-N--HR-.---I--------E--------I- 787
CPZ.TZ.00.TAN1                -YRSKD-EN----QA--W-VD---K---AQ---------VM-VL-GL----D-PI-E---QK-TQ-TAI---------------E------KN--..-I------NE--ED-D------K-L-DEY----------I-Q-P--HI-.---I-----Y--E---I------ 780
GOR.CM.07.SIVgorCP2135con     ---IRK-EG-----A--E-VL---EE--KN---------V----A-S----D-P--QK---EMTT------------------EI-----KD--..R----E---Q--ED----------L----G---------IN--P--HV-.---R----------------- 776
                                                                                                                  p15 RNase H end p31 Integrase start
MAC.US.x.239                  KD--KV-EQ----QA--E-FLM--T---PKA--IV----VM---TGC-TE---R-------EM---SEI-V-----------QEI-H---Q---..Q----EK-EP--E--D-----VKELVFK-G--RI--RQ--DT----HQ-.---I---ANSDL-T--M------ 795
H2A.SN.86.ST_JSP4_27          -D--RL-EQ----QA--E-FAM-VT---PKA--IV----VM--VAG--TE---KI------EM----AI-V------------QE--H---Q---..Q----EK-EP--E--------VKELSHK-G--KL--RQ--NT-TQ--Q-.---I----NAEL-T--M------ 808
H2B.CI.x.EHO                  KD--KP-EQ----QA--E-FA-------PQ---IV----VM--VA---TET--PI-RE---EM-----I-VG-------L---QE--H---Q---..QI---EK-EP--E------N-VKELVHK-GI-QL--RQ--N------Q-.---I----NSEL-T--M------ 808
H2G.CI.92.Abt96               KD--KV-EQ----QA--E-FAM------PK---IV----VM---AG--TET--P---K---EX----XL-VG-----------QE--X---Q---..Q----XK-EP--E----F---VKELTHK-GI-QL---Q--N-YHS--Q-.---I----NAEL-T--M------ 791
H2U.CI.07.07IC_TNP3           KD--KV-EQ----QA--E-FLM------SQA---V----VM---AG--TET--P---K---EM-M--AL-XG-----------LE--H---Q---..Q----EK-EP--E--D-F---VKELVFK-GI-RL---Q--DT----HQ-.---I----NTEL-T--M------ 791
H2U.FR.96.12034               KDR-KP-EQ----QA--E-FAM--E---PK---------IM---AG--TET--P--SK---E--R--A--VG-----------QE--H---Q---..Q----EK-EP--E--------VKELQHK-GI-QL---Q--NT-HV--Q-.---------NTEV-T--M------ 804
ASC.UG.10.RT03                KE---A-EH-----A----LL-----G-TKQ--------V----TGA-TETDHPILE---T--QS--AI--S-----------A------K---..R----EQ-PQ--E---R--N--QDLRDR-QI-ALI----LKA-P---G-.--P-----NMEV-L--M------ 793
ASC.UG.10.RT08                KE---A-EH-----A----LL-----G-TKQ--------V----TGA-TETNHPILE---T--QS--AI--S-----------A------K---..-----EQ-PQ--E---R--N--QDLRDR-RI-ALI----LKA-P---G-.--P-----NMEV-L--M------ 796
ASC.UG.10.RT11                KE---A-EH-----A----LL-----G-TKQ--------V--VVTGA-TETDHPILE---T--QS--AI--S-----------A------K---..R----EQ-PQ--E---R--N--QDLRDR-QI-ALI----LKA-P---G-.--P-----NMEV-L--M------ 796
COL.CM.x.CGU1                 EEE--N-HEGS--QA--TGVL---KHGPKRM-L----L------TG--YDQLPSSRGE--QAGMA--AIHV--C----------LI-QK-..-V-..Q-MWI-K-EA-EED-Q-F---VQYLKEQ-G--T------WER-SE--N-.-Q-V---L-Y-Y-L-------E- 762
COL.UG.10.BWC01               EEE-I--KEGS--QA--R-VLM--EHGPKRM-L----L--F--VTG--EE-N-P--E---QKGLE--AIHVS-C----------II--K-..-V-..Q-MWMEK-EE-ELD-Q-----VQY-QEQ-GI-QL-----WER-PQ--NR.-QP---SL-Y-Y-V--M----E- 764
COL.UG.10.BWC07               EEEIEV-HEGS--QA--KG-LM--KHGPKRM-L----L-----VSG--EETN-P--E---QEG----ALHV--C----------LI-QK-..-V-..-AMWIEK-EE-EED-S-F---VQHLRETYG--AI-----WDR-SQ--NR.--PV--SL-Y-Y----M----E- 764
DEB.CM.04.SIVdeb04CMPF3061    K-S-KA-EN-----A--T-VLM--E---EK---------V-N-LTEH-TTT-H---EK--Q--Q--QE---Q-------L-----I-----K---..-I---ERVEE--ED-D------KQLREE----TL---Q-INQ-P---IH.--PK----NADL----M------ 792
DEB.CM.99.CM40                K-S-KC-EN-----A--E--L---EEGPSKM-------------LEH-SET-HKI-EKV-QA-QG--QI---------------D-------K---..RI---ERVEE--ED-----G--KQLRDEYQ--TLI--Q-ITQ-P---TR.--PT-----A-M----M------ 789
DEB.CM.99.CM5                 K-G-KS-EN-----A--E--L---EEGPSKM-------------LLEH-TET-H-I-EK--QA-QE--Q---N-------------D-------R---..RI---ERVEE--ED-----G--KQLRDE-Q--TLI--Q-I-Q-P---I-.--P------A-M-V--I------ 789
DEN.CD.x.CD1_CMO580407        K------EE-----A--IGVLM-----PNK--------------AGT-TT--HQI-E---QAMQG--EI-IT-------------TEI-----K---..QI---EQ-PQ--E--D-----KEIRDKYG--TM-----INV-P---TH.--PIR----A-L-T--M------ 811
DRL.DE.11.D3                  KE--IA-EE-----A-----L---K---PK----------M---SSS-EI-DNPI-------M-S--A---N-------------E-------R---..Q-----NME---E--DL--N---SL-QE-G--GI-------Q-P---IH.--PI-----A---T--M------ 798
DRL.x.x.FAO                   KE--IAIEE-----A-----L---K---PK----------M---SSS-EI-DNPI-------M-S--A---N-------------EE------R---..Q-----NME---E--DL--N---SL-QE-G--GI-------Q-P---IH.--PI-----A---T--M------ 798
GRV.ET.x.GRI_677_gri_1        KSRIRE-EN----QA--T-VKM--E---EN---------VMN-LT-C-QE-N-P--E---QA-M--RQ---Q-------------TEI-----K---..QI---R-EE---D-A---N---S-VQE-G--NI-------A-P---IR.--PK-----A-IET--M------ 806
GSN.CM.99.CN166               KHRAIE-EE----QA--H-VL---KEGPPKM-L------VM--L-S--EV-T-A--E---QE-L---AI--S-------------TE------K---..Q---EN-EP-VED--------KYLRDQYKI-ALL-----NK-S---VH.--PK----NAEL-M--M------ 795
GSN.CM.99.CN71                KYRAIE-EG----QA----VL---KEGPPRM-L------VM--L-S--EV-T-P--E---QE-LG--A---S-------------TE------R---..Q---MEN-EP-VED--------KYLRDQYKI-ALL-----NK-S---IH.--PK----NAEL----M------ 788
LST.CD.88.SIVlhoest447        KE--KE-------QA--E-VL---R--KSR---------VIKVLSQR-TET-HPV-K----ECK--DQ---G-------------QE--H---K---QKQ----GK-EP--E-YS-F-N-AKDLEEK-HI--M---Q--N--N---T-.---IT----T------I----M- 803
LST.CD.88.SIVlhoest485        KE--KE-------QA--E-VL---K--KSR---------VMKVLSQR-TET-HPV-K----ECK--DQ---G-------------QE--H---K---QIQ----EK-EP--E--S-F-N-AKDLEEK-H---M---Q--N--N---T-.---IT----T------I----M- 803
LST.CD.88.SIVlhoest524        KE-IKE-Q-----QA--E-VL------KST---------VMKVLSQR-TET-HPI-KN--KE-Q--DA---G-------------QEI-Q--IK---TKQ-M--EK-EP-VE--G-F-N-AKDLEEKYGI--M---Q--NE-A---Q-.---IT----V-V----------- 804
LST.KE.x.lho7                 KE--KE-------QA--E-VLM-----NSK---------VMK-LSQR-TET-HPI-KD----CKQ-DQ---G-------------QE--H---K---QKQ-M--EK-EP-VE--S-F-N-AKDLEEK-----M---Q--ND-AN--K-.---IT----V-V----------- 803
MAC.US.x.251_1A11             KD--KV-EQ----QA--E-FLM--T---PKA--IV----VM---TGC-TE---R------KEM---SEI-V-----------QE--H---Q---..Q----EK-EP--E--D-----VKELVFK-G--RI--RQ--DT----HQ-.---I---NSDL-T--M------ 795
MND-1.GA.x.MNDGB1             F-R-EEYLN----Q---H-VK---E---SY---------VV--LASR-TETDHPI-KE---LMKG---I--S-L-------------I-----S---..------QN-EP--E--------EAQLREK-H--AL---Q--Q--S--CHH.--PIK--T-A-L-V--I------ 793
MND-2.CM.98.CM16              KSR-IA-EN-----A--E--KM------PK-------------M--LS-A----DNPI-RE---LM-H--A-----------------------R-V-..Q----E------E--D---N----L-Q--CI-NI-------Q-P---T-.--PI-----T---T--M----M- 796
MND-2.GA.x.M14                KS--IA-EN-----A--E--L---K---PRA-------------LS-T--R-DNPI-RE--N-M-A--A-----------------I-----K---..Q----E------E--D---N---SLSQE-SI--I-------Q-P---V-.--PI-----AA--T--M------ 797
MND-2.x.x.5440                KS---A-EE-----A--T--KM------PR-------------LS-A----DNPI-RE---LM-G--EI--S-------------I-----K---..Q----E------E--D---N----LSQ------I-------Q-P---T-.--PV----NAD--T--M------ 797
MNE.US.x.MNE027               KD--RV-EQ----QA--E-FLM--T---PKA--IV----VM---TGC-TE---R-------EM---TEI-V-----------QEI-H---Q---..Q----EK-EP--E--D-----VKELVFK-G--RL---Q--DT----HQ-.---I----NSDL-T--M------ 791
MON.CM.99.L1_99CML1           AR-A---EQ-----A--E-VL---R-GPPQM--I-----V---LASC-EI---PI-EA--QE-L--DQ-F-S-------------E--R--GQ---..Q---MENLEP-RED--------KYLRDTYHI-TLL-----NH-H---NH.--PKS---NAEV-V--M------ 791
MON.NG.x.NG1                  QR--IP-EQ-----A--E-VL----NGPASM--------VM--LASK-EV---PI-EK--Q--LG-KQ---S-------------E-------Q---..Q---MEN-EP-RED--------KYLRDTYGI-TLL-----NH-P--HTQ.--PK-----SDI----M----Y- 793
MUS-1.CM.01.CM1239            KE---S-EN-----A--E-VL---KEGPPSM--------V---VAS--QE-T-P--EE--Q--LT--A---S-------------D-------H---..Q---MEQ-EP-KED--------KYLRDKY-I-ALL-----NL-P---TH.--PKT---NAEL-V--M------ 785
MUS-1.CM.01.SIVmus_01CM1085   KE--IS-ES-----A--E-VL---KEGPPRM-L------V---VAS--QE-T-P--E---QE-LS--AI--S-------------E------NS---..Q---MEQ-EP-KED--------KYLRDKYQI-ALL-----NH-P---TH.--PKT---NAEI-V--M------ 785
MUS-2.CM.01.CM1246            QERAIN-EN----QA-----L---M-GPP-M-L------------T-A-EV---P--E---Q-MLS-NAIFISR-------------E--H---R---..Q---MEN--P-VED--------KYLRDKY-I-TIL-----NK-SA--TH.--PK----NADL-V--M------ 790
MUS-2.CM.01.CM2500            QE-AID-EN-----A-----LM--T-GPPKM------------T-A-ET---P--EK--Q--LT--AI--S-------------QE--H---K---..Q---MEN-EP-IED-D------KYLRDKY-I-TLL-----NR-SA--TH.--PK----NADI-V--M------ 789
MUS-3.GA.09.09GabOI81         ---AID-EN----QA-----L---K-GPPKM-L-----------VT-C-EA---P--E---Q--LT--EI-IT-------------TE--R---Q---..Q----ENVEP-KED-D------KYLRDVYQI-ALL-----NH-S---TQ.--PKT---NAEI----M------ 786
MUS-3.GA.11.11GabPts02        QR---E-QG-----A--E-VL---K--PSKL-L------V----LSE-ET-D-P--E---Q-MLQ-QAI-VS-----R-----ME--R---Q-V-..QI--MEN-EP-VED--------KYLRDTYKI-TLL-----NK-G---TQ.--PK----NAEL----M------ 789
OLC.CI.97.97CI12              KH-I-HIQESS---A-IR--IE--R--KS---------L--M--VLEK-EV--EP-IQEL-KE-E---RI-IG-------L---QE--Q---K---TKQ-M--AD-EP--ES-N---Q-AQSLQQE-QI-LI---R--DE-EE-KG-.-VPR--I-AVGVYR--I------ 785
RCM.CM.00.SIVagi_00CM312      KE--KE-E------A--E-VL---K---KR---------VF--LAGS--T---P--Q-------G--E---S-------------E------R---..Q-------E---E--D-----------E--RI-Q--------Q-P---V-.---V---T-A---T--M------ 797
RCM.CM.02.SIVrcm_02CM8081     -E--KE-E------A--E-VL---K---KK---------VF--LAGS--T-D-P--Q-------G--E---S-------------E------K---..Q-------E---E------N------E--QI-Q--------Q-P---V-.----------A--KT--M------ 803
RCM.GA.x.SIVRCMGAB1           KE--KE-EE-----A--E-VL---K---PK---------VY--LE----T-D-G--TE--N-M-G--A---S-------------E------K---..Q-------E---E------N----L-E--QI-QI-------Q-P---V-.---I-----A---T--M------ 797
RCM.NG.x.NG411                KE--KE-EN-----A--G-VL---K---KR-------RHVF--LAGS--T---P--Q-----M-G--E---S------------------Q---..Q-------G---E------------EE-QI-QI-------Q-P---V-.---V-----A---T--M------ 802
SAB.SN.x.SAB1                 D----A-EN-----A--E--L---R---SK---I------M---AGE-TE-DNNI-Q----E-----A--I--------V----EI-----Q---..Q-----R-EE--E--D---A---S-QQE-G--AI-------A-P---I-.--SV-----A---V--M------ 816
SMM.CI.79.SIVsmCI2            -E--KP-EQ----RA--E--AP--R-A-PKA-V------VM--VAGH-TETD-P---E---E-----AI-VG-------L---QE--H---Q---..Q----EQ-EP--E--------VKQLVHK-GI-QL---Q--NT-TQ--Q-.---I----NAEL-T--M------ 800
SMM.LR.89.SIVsmLIB1           -D--KA-EQ-----A--E-FLM-M----P-A--IV----VM--VAG--TE--NN-------EM---TAI-V-*----------RE--H---QR--..Q--#-EK-EP--E--------IKELVFKYGI-RL---Q--DT----HQ-.---I----NAEL-T--I------ 785
SMM.SL.92.SIVsmSL92A          KD--KV-EQ----QA--E-FAM--A---PKA--------V---VLG--TE---PI--E--NHMML--EL-V--------L---QE--H---Q---..Q----EN-EP--E--D-----IKELVYK-GI-#L---Q--NT-----Q-.---I----NADI-T--M------ 790
SMM.SL.92.SIVsmSL92B          KE--LP-EQ----QA----LL---K--PSK--V------V-N--TG--SE-D-DI-A------VQ--A--IG-------------NE--R---Q---..Q----ES-EP--ED-D-----VKEL-QKY-I-QL---Q--NA-N---Q-.---I---TNAEV-T--M------ 788
SMM.SL.92.SL92B               KE--LP-EQA---QA----LL---K--PSK--V------V-N--TG--SE-D-DI-A------VQ--A--IG-------------NE--R---Q---..Q----ES-EP--ED-D-----VKEL-QKY-I-QL---Q--NA-N---Q-.---I---TNAEV-T--M------ 788
SMM.US.04.G078                -N-IKS-EQ----QA--E-FLM-VT---T-A--IV----VM--VTG--TE---KI------EM---TAI-V--------L---QEI-H---Q---..Q----EK-EP--E--------VKELVFK-GI-RL---Q--DT-----Q-.---I----NSEL-V--M------ 791
SMM.US.05.D215                -DR-K--EQ----QA--E-FLM-VK---PKA---V----VM--VAG--TE--NSI------EM---TAI-V------------QE--H---Q---..Q----EK-EP--ED-------VKELVFKYGI-RL---Q--DT----HQ-.---I----NAEL-T--M------ 791
SMM.US.06.FTq                 KD--KL-EQ----QA--E-FLM-VT---P-A--IV----VM--VAG--TE---RV-----EM---TAI-V------------QEI-H---Q---..Q----EK-EP--E--------VKELVFK-RI-RL---Q-IDT----HQ-.---I----NAEL-T--M------ 795
SMM.US.86.CFU212              KDR-KV-EQ----QA--E-FLM--T---PKA--IV----VM--VAG--TE--NKI-S----EM---TAI-VT-------L---QE--H---Q---..Q----EK-EP--E--------VKELVFKYGI-RL---Q--DT----HQ-.---I----NADL-T--M------ 791
SMM.US.x.H9                   -G-TKL-EQ----QA--E-F----A---PKA--IV----VM--VAG--TE---R-------KM----AI-V-------X--E-QE--H---QE--..Q----EK-EP--E--------VKKLVFK-X--RL---Q--DT----H--.---I----NAXL-T--M------ 791
SMM.US.x.PGM53                -D-AKP-EQ----QA--E-F----A---PKA--IV----VM---AG--TE---R-------EM----AI-V------------QE--H---Q---..Q----EK-EP--E--------VKELVFK-GI-RL---Q--DT-H--HQ-.---I----NAEL-T--M------ 791
SMM.x.x.pE660.CG7G            -D-AKL-EQ----QA--E-F----A---PKA--IV----VM---AG--TE---R-------EM----AI-V------------QE--H---Q---..Q----EK-EP--E--------VKELVFK-S--RL---Q--DT----HQ-.---I----NAEL-T--M------ 791
STM.US.89.STM_37_16           KN--KA-EQ----QA--E-FAM--A---PKA---V----VM---TG--TE---K-------EM----AI-V------------QE--H---Q---..Q----EK-EP--E--------VKELVFK-GI-RL---Q--DT----HQ-.---I----NAEL-T--M------ 791
SUN.GA.98.L14                 -SR-KK-EN----QA--E--KM--E--RSS-----------RLLSKR-TETD----KE-V-LIRQ-DQ---G-----------QEI-Q---Q---KRQ-M-IEK-EP-VE--G-F-N-AASLQEM-DI-L----Q--NE-AQ--Q-.---IT----A-V----I----M- 803
SYK.KE.x.KE51                 K-TLKR-EN-----A--E-VLM------P---------------LMNC-TS--HPV-E--MQKAME---I-IT-----------QE--Q---K---..-I----N-P---E------T-VEYIRQE-H--RQ---A--EM-P---IR.--PK-----TDL-T--M------ 823
SYK.KE.x.SYK173_COMGNM        T-R-QY-EQ----QA--EGLLM-----KDK----V----SY--LMTC-TNT-HPI-E---QEA----AI-VT------------A------K---..-I---ER-PQ--ED--R----MEYLRQE-H--RQ---A-IQQ-P---NR.--PK-----VDIYN--M----E- 818
TAL.CM.00.266                 KEH-IK-EK-----A--E-VL---K---PK---------VY--LAGS-TETDNKIIED--QL-LS--A-----L------------------Q---..-I--VEQ-PE--E---R--N---DLKAR-K--TI---A-IEA-P---VQ.--PKT--NNAAV-T--M------ 792
TAL.CM.01.8023                KEH--G-SN-----A--E-VL---KH--PK---------VY--LVGN-TETDNKI-ED--Q--LG--A---T-L------------------Q---..RI--VEQ-PE--E---R--N--KDLKAR-K--TI---A-IEA-P---VQ.--PKT--TNAAV-T--M------ 793
TAN.UG.x.SIVagmTAN1           ESQ-IE-E-----QA----VL---KH-KSK---------VM--LT---T--D-V-------EM-N-DAI--T-----------QEI-----Q---..Q-----RMEE--ES-D---T--QFIRDA-GI-AL-------A-P---IR.--PI-----A-V-V--M------ 803
VER.DE.x.AGM3                 K-R-EK-EN----QA--M--KM--E---PN----------M--LT---T--D-P-IE---ALMVQ-HQI--Q----D-------EI-----Q-M-..-I---EK-EE--E---R--N---NL-DTYG--QI-------M-P---I-.--PV-----A---V--M------ 800
VER.KE.x.9063                 K-R-EK-EQ----QA--V--KM--E---EK----------M--LT---T--D-P--E---ALMV--RAI--Q-------------EI-----Q---..R---IGR-EE--E--DR-----KNL-DT-G--QI-------M-PNAKI-.--PI-----A---V--M----V- 800
VER.KE.x.AGM155               K-R-QQ-EN----QA--T--KM--E---PK----------M--LT---T--D-P--E---A-MVQ--AI--Q-------------EI-----K-V-..RI--IGR-EE--E--DR------NL-DT-G--QI-------M-P---V-.--PI-----A---V--M----I- 803
VER.KE.x.TYO1_patent          K-R-E--EN----QA--T--KM--E---PN----------M--LT---T--D-P--E---ALM-Q-QQI--Q-------------EI-----K---..R----EK-EE--E---R--N--KNL-DTYG--QI-------M-P---I-.--PV-----A---T--M------ 803
WRC.CI.97.97CI14              KE---SVQN----Q---I-LK---E-G-QKM--I-----VM--VKDH-TM---P--QE---L-KG--EI-ID-----------TEI-R-------S.-IM--EN-TP-IES-T---QG-K-LQEE-H--QLI--R-IEE-ST--G-G-AP---M--YEV-T--M------ 808
WRC.CI.98.98CI04              KE--IN-YN----QA--A-LK---E-G-SKI--I-----VM--VKDH-TI---A--QE--DL-KQ--EI-IE-----------TEI---------S.--M--EN-TP-IES-T---QG-K-LQEE-HI-QLI--R--EE-ST--G-G-APL--M--YEV-T--M------ 807
WRC.GM.05.Pbt_05GM_X02        -N-IQKI-N----Q---I-LRM--E--SDK---I-----V----KDH-T----P-MQE--DL-KG--EI-IT-----------TEI-A-------S.R-M--EN-TP-IES-T---QG-G-LEEE-H--TIL--R--EE-AT--G-G-HPI--M--YEI-T--M------ 799
```

```
H1B.FR.83.HXB2              GKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTGATVRAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQI.TKIQNFRVYYRDSRNPLWKGP  953
H1O.CM.91.MVP5180           --I-I-------DF---------------------A-----V------P---S-AMK-----T--QH-----------------A------S--Q------------------V------------T----LI--L-SQ---T------.L--N----------D-I----  949
H1N.CM.95.YBF30             --I----------L----------------I----------V----------S---K------N---------------A--------------I----------------------------------T-----I---------TN--T--.L-V-----------D-I----  960
H1P.CM.06.U14788            --I----------FM------V-----A---V----S-----V------P---S-A-K-----LN-TH----------------------------H------------------V------------T--D--I--L--Q---T------.L-----Q-------D-I----  950
CPZ.CD.06.BF1167            ---V---------F-----LSS---K---F-I----------K------A---S-A-K------N-Q--------------------Q--I-VD-I-E-----R---A---Y-------------TP-D--L--LT------------.F-L-----------D-V----  956
CPZ.TZ.00.TAN1              ----I--------F-------E---R-----I----------K------P---STA-K------Q-QH-------------------Q--Q--E-------Q-R---I---Y-----------E-T----LL--LT-N----Q-----.L-V---------A-D-I----  949
GOR.CM.07.SIVgorCP2135con   --I----------F---------------------A-----I------P---S---K-----LNVTH-------------------------Q------------------V------------T--D--I-ML--Q---T------.L-------------D-I----  945

MAC.US.x.239                --I-I--------F-------Q---RQ--L-----------ITHL-----A--ASQE-KMVA-----EHT--V-----------A--HH--NQ-DRI-E--NSVE-I-L---HCM----R----DMTPA--LINM-T-EQEIQFQ-SKN.S-FK-------EG-DQ-----  964
H2A.SN.86.ST_JSP4_27        --I-I--------F-------Q-S-RQ--L------S---ITHL-----A---SQE-KMVA--I--E-S--V-----------A--HH--NQ-SRI-E--NTVE-I-L---HCM----R----DMTPA--LINMVTAEQEIQF--AKN.S-L------F-EG-DQ-----  977
H2B.CI.x.EHO                ----I--------F-------Q---RQ--L------S---ITHL-----A---SQD-KM-A--I--E-T--V----E------A--HH--NQ-DRI----VSIE-V-L--THCM----R----DMTPA----NM-T-EQEIQF--TKN.L-F--------EG-DQ-----  977
H2G.CI.92.Abt96             ----I--------FV------Q---RQ--L----------ITHL-----A---SQE-KMVA--V--E-A--V--------------HH---Q-DKI----NSIE-I-L--HCM----R----DMTPA---INM-T-EQEIQFQ-TKN.S-FK-------EG-DQ-----  960
H2U.CI.07.07IC_TNP3         ----I--------F-------Q---RQ--L------T---ISHL-----A---SQK-KMVA--V--E-S--V-----------A--HH--NQ-DRI----NSME-I-L--THCM----R----DMTPA--LINM-T-EQEIQFQ-TKN.S-FK-------EG-DQ-----  960
H2U.FR.96.12034             ---VI--------F-------Q---RQ--L----------ITHL-----A---SQE-KMVA--V--EHT--V-----------A--HH--NQ-ERI-E--NTME-I-L--AHCM----R----DMTPA---INM-S-EQEIQF--SKN.S-FK-------EG-DQ-----  973
ASC.UG.10.RT03              --I-I---------T--KL--Q---K---I---Q-CA----TQ------P---SQELA--A---NVQHST-V-----------NK--Q--ET-SKI--EVQY-E---A--LL-L---KR-----MTPA--LINM-H--LELQQ..TKF.P-FS-------TGKS-E-L--  960
ASC.UG.10.RT08              --I-I---------T--KL--Q---K---I---Q-CA---ITQ------P---SQELA--A---N-QHST-V-----------NK-RQ--ET-SKI--EVQY-E---A--LL-L---KR-----MTPA--LINM-H--LELQQ..TKF.P-FS-------TGKS-E-L--  963
ASC.UG.10.RT11              --I-I---------T--KL--Q---K---I---Q-CA----TQ------P---SQELA--A---NVQHST-V-----------NK-RQ--ET--KI--EVQY-E---A--LL-L---KR-----MTPA---INM-H-ELELQQ..TK-.P-FS-------TGKS-E-L--  963
COL.CM.x.CGU1               ----------CTLFCW-TILKR---E--GRA-I---SQ-E-RQV-----P--VSQHFK--V--L--AHTT-H-----------QR--DV-RK-KKMK----T-ESK-A---YAL-------L--K-PW--Q-ERAIIELD-QN-T-LQNQ-FK--KA-WKEH.TGE-Q--  931
COL.UG.10.BWC01             ----------CTLFCW-TILKR-S-E--GRA-I---SQ-E--QV-----P---SQAF---V--L---HTT---------I--QR--MI-EG-KKMEE---T-ESK-A---YAL-------L--K-PW--Q-ERAIIELE-QNITQIQNK-F---K--WKEH.TGE-Q--  933
COL.UG.10.BWC07             -AI-------CTLFCQ-FMLKR--AE---RG-I---SM-E--QV-----T----QAFK--V--L--THTT-T---------I--QR--M--ET-RKMK----T-ESR-A---YAL-------L--K-PW--Q-ERAIIELD-QNITQIQNK-FK--K--WKEH.TGE-Q--  933
DEB.CM.04.SIVdeb04CMPF3061  ----I--------FT-TQ-LKE---KQ--L---Q-GA---IQQV-----P--ISQAFA-----L--EHTT-V-----------NK-RQ--DT-T---ED-QR-E---A--TH-L----R----DMTPS----NM-N-ELEI-Y--QK-.S-FSG-K----EGAD-S----  961
DEB.CM.99.CM40              N-I-I-----------TKIL-----K---L---Q-GA----TQ------P--ISQP-K-----L--EHTT-V------------K-RI--DT-K-I--D--R-E---A--TY-I----R----DM-PI--L-NM-T-E-E--T--QK-.S-LLG-K----EGAD-T----  958
DEB.CM.99.CM5               N---I--------FM-TK-LS----K---L---Q-GA---I-Q------P--ISQP-K-----L--DHTT-V------------K-RV--ET-K-I-ED--R-E---A--T--I----R----DMA-I--L-NM-T--LE-QT--QK-.S-LLGYK----EGAD-T----  958
DEN.CD.x.CD1_CM0580407      N---I---NA----L-TKIL-----K---L---QI-A---ITKL-----P--ISQAMQ------N--HTT-V---------I--NK--Q--EV-AL--ED--R-E--LA--TL-L----R--L-DMT-A--YINMLN-HLEIQH..TKQ.P-FS--K----EGA--Q-Q--  978
DRL.DE.11.D3                --I-I-------------------EK-------------SHL-----P---SEK-ATV----K-EHTT-V---------------NQ-------I-----K-E-------L------------E---A--------S-LL-NK--QN-.L----------EG-DQ-----  967
DRL.x.x.FAO                 --I-I--------H----------EK-------------SHL-----P---SEK-ATV----K-EHTT-V---------------NQ-------I-----K-E-------L------------E---A--------S-LL-SK--QN-.L----------EG-DQ-----  967
GRV.ET.x.GRI_677_gri_1      ----I--------F-------R---K---H-----LA-----HL-----P---SQN-A-V---GN-EHTT----------S-----RQ--E--S-I--DC-R-E------TH------------I-SA--L-NMLT-QLELNT--N--.Q--L--K---EG-D-V----  975
GSN.CM.99.CN166             -------------VW-RI--Q---RQ--LK--E--AT---THL-----P--VSKELE------D-QHTT-V-----------N---Q--ET-TKI-EEVTY-E---AQ-CY------------DMCPT----NM-H-ELE-QH-NT-N.S-F-K------QGA----Q--  964
GSN.CM.99.CN71              -------------VW-RI--Q---RQ--LK--E--AT---THL-----P--ISKELE------N-QHST-V-----------N---Q--ET-QKI--EVTY-E---AQ-CY------------DMCPT--L-NM-H-ELE-QH-NT-S.S-F-K------QGA--H-Q--  957
LST.CD.88.SIVlhoest447      -Q--IN-------FMV-----D---KT--H-----CS----QQV-----P--VSKE-Q-VI--L--EHTT-------------AK--A--Q--ER--ED-QQ-----L--TH-----QR--L--L-PA--FINM-NA-LE-QY---LN.S--LK-K----QG-D-Q----  972
LST.CD.88.SIVlhoest485      -Q--IN-I---C-FMV-----D---KT--N-----CS----QQV-----P--VSKE-Q-VT--L--EHST-------------AK--V--Q--ER--ED-QQ-----L---H-----QR--L--L-PA--FINM-NA-LE-QY---LN.S--LK-K----QG-D-Q----  972
LST.CD.88.SIVlhoest524      -QI-MN---M---FMV-----D---KT--N-----CS-----Q------P--VSKD-Q-VT--LS-EHTT-------------AK--V--Q--ER--ED-QQ-----L---H-----QR--L--L-PA--YINM-NA-LE-QYI--LN.S--LK-K----QG-D-Q-R--  973
LST.KE.x.lho7               -Q--IN-------FMV-----D---KT-SH-----CS-----Q------P--VSKE-Q-VT--I--EHTT-------------AK--V--S--ER--ED-QQ-----L---H-----QR--L--LTPA--FINM-NAELE-QY---LN.S--LK-K----QG-D-Q----  972
MAC.US.x.251_1A11           --I-I--------F-------Q---RQ--L-----------ITHL-----A--ASQE-KMVA-----EHT--V-----------A--HH--NQ-DRI-E--NSVE-I-L---HSM----R--T-DMTPA--LINM-T-EQEIQFQ-SKN.S-FK-------EG-DQ-----  964
MND-1.GA.x.MNDGB1           NQI-I--------FMK----T----KK--E------AQ--ISKL-----P---SQE-ETM---L--EHT--------------NK--Y--EL-EKI-EDCKE-----A--T------QR--L--MT------NM-N-ELEYQYQ-N--.S-NL--K--F-EG-DQ-----  962
MND-2.CM.98.CM16            ----IA------R-L------T---K---H-------------HL-----P---SEK-ATV----Q-EHTT-----------I-AK-HH--Q----------K-E-------L------------E--P-----------LL-TK--HN-.Q-----------EG-DQQ----  965
MND-2.GA.x.M14              ----I---------L----------K---H-------------HL-----P--VSEK-ATV----Q-EHTT-------------AK-HH--V--E------R-E-------L------------E--P--------T---L-TK--QN-.S-----------EG-EQ-----  966
MND-2.x.x.5440              --I-I---------L----------K---H-------------HL-----P--VSEK-ATV----Q-EHTT-V-----------AK-HH----E------R-E-------L------------E--P--------T---L-TK-RQN-.S-----Q----EG-DQQ----  966
MNE.US.x.MNE027             --I-I--------F-------Q---RQ--L------S---ITHL-----A---SQE-KMVA-----EHT--V-----------A--HH--NQ-DRI-E--NSME-I-L---HCM----R----DMTPA--LLNM-T-EQEIQFQ-SKN.S-FK-------EG-DQ-----  960
MON.CM.99.L1_99CML1         -SI-----------VW-KIL-R---KC-GIA--E--AM---TQ------P---SQEFE--A---N-HHTT-V-----------N--RQ--ET-KKI--EV-Y-P--LAQ-L--L----------DMAPVD-FINM-H-ELELQTSNN--.--FSK----F-TGAD-Q----  960
MON.NG.x.NG1                -SI-----------VW-KIL-R---K--GIA--E--AT--IHHV-----P---SQEFA--A---N-EHST-V-----------N---QI-TTL-KI--EVQF-S--LAQTLY-L----X----DI-PI--L-NM-H--LELQHSNN--.--FSK----F-TGAD-S----  962
MUS-1.CM.01.CM1239          --I-----------TW-KIL-R---RQ--LG--E--AL---TQ------A--ISEEFG------S-DHTT-V-----------NK--Q--ET-QKI-EEVTY-E---AQ----M---K-----DLA-A---INMLH-ELELQH--S-K.S-F--------TGSD-S----  954
MUS-1.CM.01.SIVmus_01CM1085 --I-----------MW-KIL-R---RQ--LG--E--SL--ATQ------A--IS-EFE------S-EHTT-V-----------NK--Q--EA-HKI-EEVTY-E---AQ----L---K---V-DL-NA---INMLH-ELELQH--Q-K.Y-F--------TGSD-S----  954
MUS-2.CM.01.CM1246          ------------SF-W-RI-SQ---RL--LE-SN--AT--ISQ--P---T---SKEFQ-VA---N-QHTT-V-----------NA--Q--ET-HK--EEVTY-E---AQ----L-N-------DMTPT--L-NMLY-ELEIQQ--NHN.--FSK------RGA--Y-LR-  959
MUS-2.CM.01.CM2500          --I-----------F-W-RI-SQ-N-RL--LE--N--AT--ITQ------T---SKEFQ-VA-----EHST-V-----------NA--Q--ET-NK--EEVQY-E---AQ----L-N-------DMTPT--L-NMLY-E-E-QQI-NTK...FSK------RGA--S-Q--  956
MUS-3.GA.09.09GabOI81       E-I-------------WTKIL-R---RN--LA--Q--SQ---SH------T---SEEFA--A-----EHTT-------------NA--Q--TQ-KKI--EVQY-E---AQ-TY-L----R--M-DL-P----INM-H-ELEIQY-NS-N.--F--------QGSD------  955
MUS-3.GA.11.11GabPts02      -------------VW-RLL-R---RD--LA--Q--AM---TH------P---S-QFA--T--GN-EHTT-V-----------NI--Q--NT-SKI-EEV-Y-E---AQ-A--L----------DL-PI---INMLH-ELELQHT-N--.S-F--------QGGD------  958
OLC.CI.97.97CI12            ----CL-A-TG--FLYG-ILR--D-EN-VK-V-R--SQ----I------T--VNQKMQ-V---L--EH-T---W------II-AK-RV--E-L-RI-E--T--E--LA--I------N---Q--GTVA-EL-NQ-Y--T.QI-YTQNK.F-KFSGYRV-YKNKEGQ-C--  953
RCM.CM.00.SIVagi_00CM312    ----I--------------------K----------A-----HL-----A---S-A-Q-V----Q-EHT--V---------------HQ--T----I-----K-E-------L-------------------I----SEL--NK--N--.S--------F-EG-DQQ----  966
RCM.CM.02.SIVrcm_02CM8081   ----I--------------L-----K---H------A-----HL-----A---SSA-Q-V----Q-EHT--V---------------HQ--T--T-I-----KIE-------L-------------------I----S-L--TK--N--.S--------F-EG-DQQ----  972
RCM.GA.x.SIVRCMGAB1         --I-I--------------------K---H------A----RKL-----A---S-A-Q-V----Q-EHA--V----------------Q--I--E---E---K-E-------LV------------------I------LA-NK--N--.S----------EG-DQ--R--  966
RCM.NG.x.NG411              --I-I--------------------K---H------A-----HL-----T---S-A-Q-D----Q-EHT--V---------------HQ--I----I-----K-E-------L-------------------I----S-L--TK--N--.S--------F-EG-DQQ----  971
SAB.SN.x.SAB1               --I-I--------F-------Q---KA--H------S---ITQL-----T---SQQ-A-I---GK-EHT--V----------------Q--E----I--D--R-E---I---H---------------A--LIN--H-ELE--T--QK-.S----------EG-D-V----  985
SMM.CI.79.SIVsmCI2          ----I--------F-------Q---RQ--L------S---ITHL-----A---SQE-KMVA--V--E-T--V--------------HH--NQ-DRI-E--NSIE-I-L--CHCM----R----DMTPA----NM-S-EQEIQF--TKN.S-FK-------EG-DQ-----  969
SMM.LR.89.SIVsmLIB1         --I-I--------F-------Q---RQ--L-----VS---ITHL-----A---SQE-KMVA---N-E-T-RV-----------A--HH--TQ-DKIKE--NSIE-I-L--THCM----R----DMTPA--L-NM-T-EQEIQFQ-SKN.S-FK-------EG-DQ-----  954
SMM.SL.92.SIVsmSL92A        --I-I--------F-------Q---RQ--L------A-----HL-----A---SQE-KMVA-----E-S--V-----------A--HH--NQ-DRI----VS-E-V-L--THCM----R----DMTPA--LINM-TAEQEIQYH-TKN.S-FKK------EG-DQ-----  959
SMM.SL.92.SIVsmSL92B        ----I-------SF-------R---RQ--L----I-S---I-HL-----A---SQE-KMVA--L-VE-S--V-----------A--LH---N-DKI-E---SVE-L-L--AHCM----R----DMTPA----NM-T-ELE-QY-TS-N.S-F--------EG-DQ-----  957
SMM.SL.92.SL92B             ----I--------F-------R---RQ--L----I-S---I-HL-----A---SQE-KMVA--L-VE-S--V-----------A-DLH---N-DKI-E---SVE-L-L--AHCM----R----DMTPA----NM-T-ELE-QY-NS-N.S-F--------EG-DQ-----  957
SMM.US.04.G078              --IVI--------F-------Q---RQ--L------S---ITHL-----A--ASQE-KMVA-----EHT--V-----------A--HH--NQ-ERI-E--NSVE-I-L---HCM----R----DMTPA--LINM-T-EQEIQFQ-SKN.S-FK-----F-EG-DQ-----  960
SMM.US.05.D215              ----I--------F-------Q---RQ--L------S---ITHL-----A--ASQE-KMVA-----E-T--V-----------A--HH--TH-DRI----NSIE-V-L--THCM----R----DMTPA--L-NM-T-EQEIQFQ-SKN.S-FK-------EG-DQ-----  960
SMM.US.06.FTq               ----I--------F-------Q---RQ--L------S---ITHL-----A---SQE-KMVA-----E-T--V-----------A--HH--TQ-DRI-E--NSVE-I-L---HCM----R----DMTPA--LINM-T-EQEIQFQ-SKN.S-FK-------EG-DQ-----  964
SMM.US.86.CFU212            --IVI--------F-------Q---RQ--L------S---ITHL-----A---SQE-KMVA--TN-EHT--V-----------A--HH--TQ-DRI-E--NSVE-I-L---HCM----R----DMTPA--LINM-T-EQEIQFQ-SKN.S-FK-------EG-DQ-----  960
SMM.US.x.H9                 --I-I-----G--F-------X---RQ--L------S---ITHL-----A---SQE-KMVA-----E-T-XV-------X---A--HH--TQ-DRI-E--NSIE-I-L--IHCI----R----DMTPA--L-NM-T-EQEIQFQ-SKN.S-FK-------EG-DQ-----  960
SMM.US.x.PGM53              --I-I--------F-------Q---RQ--L------S---ITHL-----A---SQE-KMVA-----E-T--V-----------A--HH--TQ-DRI-E--NSIE-I-L---HCM----R----DMTPA--L-NM-T-EQEIQFQ-SKN.S-FK-------EG-DQ-----  960
SMM.x.x.pE660.CG7G          --I-I--------F-------Q---RQ--L----------ITHL-----A---SQE-KMVA-----E-T--V-----------A--HH--TQ-DRI-E--NSIE-I-L---HCM----R----DMTPA--L-NM-T-EQEIQFQ-SKN.S-FK-------EG-DQ-----  960
STM.US.89.STM_37_16         --I-I--------F-------Q---RQ--L------S----THL-----A---SQE-KMVA-----E-T--V-----------A--HH--TQ-DRI----NTVE-V-L---HCM---KR--L-DMTPA--L-NM-T-EQEIQFQ-SKN.S-FK-------EG-DQ-----  960
SUN.GA.98.L14               E---I---------MV---L-N-Q-KT--TW----CAM----Q------P--ISKD-E-V---L--QHTT-------------AK--V--Q--SRI-ED-QE-----L--LH-----QR--L--MT-A--FINM-NA-LE-QY---IN.S--LK-K----EG-D-Q----  972
SYK.KE.x.KE51               ----C---NT----T-TKILKR------GL---QI-A---I-H------P--ISDAFA-------VEHTT----------I--NK-RQ--ET-N-I-EEV-R-E---A--T--L-----------TTPAD-YINMLY-EL-LQNTHT...Q-FS--K----HGT.SD-Q--  989
SYK.KE.x.SYK173_COMGNM      ----C---NT------TKILKR---D---L--MQI-S---I-Q------P--VSDKFK-----C--EHTT----------I---K-RY--EA-S-I--DVT--Q---A--T--L-----------I-P---YINMLY-EL-LQQNTTS..P-FS-------QGK.NE----  985
TAL.CM.00.266               -QI-C-----------TKIL-R---R---L---QV-S---ISHL-----P--VS-EMQ-MV--LK-EHSA-V---------A--NK--Q---T-T-I--EVQY-S---AQ-T--L-Y--R--L-DMCPA-A-INM-Y-EL--TQ--N--.QNFSD-K----KGA----Q--  961
TAL.CM.01.8023              -Q--C-----------TKIL-R---R---L---QV-S---IEHL-----P--VS-EMQ-TA--LK-EHTT-V---------S--NK--Q---T-Q-I--EVQY-S---AQ-T--L----R--L-DMCPA-ALINM-Y-EL--TT--N--.HNFSD-K----KGA----Q--  962
TAN.UG.x.SIVagmTAN1         --I-I--------FL-----AR---K---K----IIS---ITKL-----P--VSQE-QTI---GQVEHTT----------S-----RQ--ET-EKI-EDCAF-E---L--CH------------MTPA--LINM-T-QLEIQHI-T-Q.Q--S--K----EG-E-V----  972
VER.DE.x.AGM3               ----I--------F-------R---K---K----ILS---ITQL-----P---SQE-A-M---GK-EHTT-V---------SI-----Q--E---KI--DCQYTE---L--CH------------LTPA--LINM-T-QLELQH--TK-.Q--L-------EG-D-V----  969
VER.KE.x.9063               --IVI--------F-------R---R---K----ILS---IVQL-----P---SQE-A-I---GK-EHTT-V---------S------Q--E---KI--DCQYTE---L--CH------------LTPA--LINM-T-QLELQT--TK-.Q--L-------EG-D-V----  969
VER.KE.x.AGM155             --IVI--------F-------R---K---K----II----ITHL-----P---SQE-A-M---GKVEHTT-V---------SI-----Q--E---KI--DCQYTE---L--CH------------LT-A--LINM-T-QLEINT--TK-.Q--L-------EG-D-V----  972
VER.KE.x.TYO1_patent        ---VI--------F-------R---K---K----ILS---ITQL-----P---SQE-A-I---GK-EHTT-----------SI-----Q--E---KI--DCQYTE---L--CH------------QTSA--LIN--T-QLEIQH--TK-.Q--L-------EG-D-V----  972
WRC.CI.97.97CI14            --I--N--------L--KI-S--DAYQ--L-T-Q--S---I-K------T--KSQKMQ-I---L--E--------------I--K-RR--D--HSI-ED--K-E--LA--L------E---L--.-PA--L-NM-TS-LE-QQT-Q-K.L-FK--Q----TGA-QQ-Q--  976
WRC.CI.98.98CI04            --I--N--------L--KI----DAYQ--L---QI-----M-K------T--KSQKMQTI---L--E--------------I--K-RR--E--HRI-ED--K-E--LA--L------E---L--.-PA--L-NM-TAELE-NQT-Q-K.L-FK--Q----TGA-QQ-Q--  975
WRC.GM.05.Pbt_05GM_X02      --I--N--------L--KI--T-DAYQ--L-TIQ--------K------T--KSQKME-I---L--E--------------I--K-RV--Q--QRI-ED--K-E--LA--IY-----E---L--.-PA----NM-QS--E-NQT-Q-K.F-FK--Q----TGASQQ-Q--  967
```

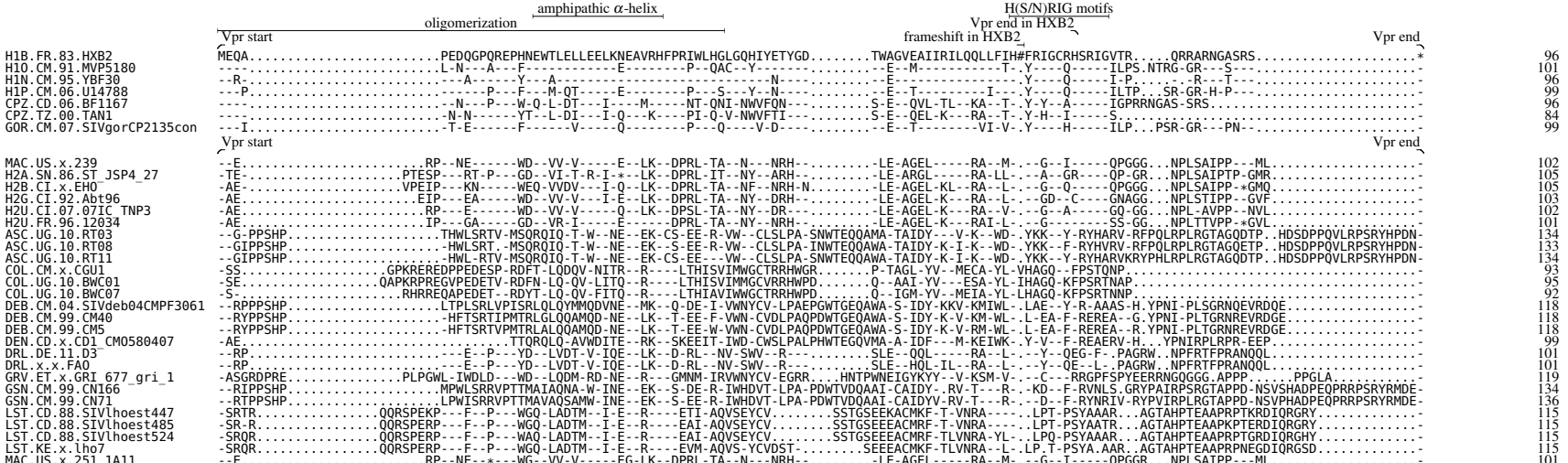
```
                                                                       p31 Integrase end
                                                                                 Pol end
H1B.FR.83.HXB2           AKLLWKGEGAVVIQDNS.DIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED...............*  1003
H1O.CM.91.MVP5180        -Q-------------KG.---------------------T-SM-N--T-SESMEQPGEIP.....-  1010
H1N.CM.95.YBF30          ----------------G.-----------------------G----G---NQEME...........-  1015
H1P.CM.06.U14788         -T-------------KG.--------------N------------DT-R-SESLEQSG........-  1008
CPZ.CD.06.BF1167         -R-----------KEGE.E------------K-----V--AGG--D--N--...............-  1007
CPZ.TZ.00.TAN1           -R-----------KEGE.-------------KE-------AGGMDD--N-T...............-  1000
GOR.CM.07.SIVgorCP2135con -T-------------KG.---------------------------DT-G-SESLEQSS........-  1003
                                                                       p31 Integrase end
                                                                                 Pol end
MAC.US.x.239             GE---------ILKVGT.-------------K---GGKEV-SSSHMEDTGEAREVA..........-  1020
H2A.SN.86.ST_JSP4_27     GE-----D---IVKVGA.---II--------K---GRQEM-SGSNLEGAREDGEVA..........-  1033
H2B.CI.x.EHO             GD---------I-KVGT.E---I---------N--GGKEL-CSADVEDTMQAREVAQSN.......-  1036
H2G.CI.92.Abt96          GE---------I-KVGT.E------------K---GGKEV-SSTNMEDTRQTGEVA..........-  1016
H2U.CI.07.07IC_TNP3      GE---------I-KVGT.----I--------K---GGKEV-SGSNLEGARQDGEMAQLGEIP....-  1022
H2U.FR.96.12034          -D---------L-KVGT.----I------------GGKEL-SSSNLEGA-KVREMALPDQTP.....  1034
ASC.UG.10.RT03           -Q-----------KTTTGELLT--K------KP--ASQNVGNKEGMDSPRKEDTSLD.........-  1018
ASC.UG.10.RT08           -Q-----------KTTTGELLT--K------KP--ANQNVG-TENMDSPRKEDTSLD.........-  1021
ASC.UG.10.RT11           -Q-----------KTTTGELLT---------KP--ASQSVG-KENMDSPRKEDTSLD.........-  1021
COL.CM.x.CGU1            GE-V---------RNSQGTLF-K----V--T-LQYGEDV-SENLL-NGQKEAETVKGMD.......-   991
COL.UG.10.BWC01          GQ-V---------KNQEGSLF-K----V--TKVNYGERMVSKALLPNEQ-EAGEVKEVD.......-   993
COL.UG.10.BWC07          GQ-V---------RSTEGTLF-K----V--S-ITHGTDQ-VVSEDLLPNGKEKPGETETLD.....-   995
DEB.CM.04.SIVdeb04CMPF3061 -T-----------KLDTG-L---------V-K----DVGSKIDPQDTHE.................-  1011
DEB.CM.99.CM40           -T----------CKTEVG-----------V-KE---D-DSKVNTQNSNE.................-  1008
DEB.CM.99.CM5            -T----------CKTEVG-----------V-K----GVDSKTDTQNPSE.................-  1008
DEN.CD.x.CD1_CMO580407   -E-----------KTQAG--FP---------KP-NAE.SRKGEQNKGM-SETDI............-  1032
DRL.DE.11.D3             -E-V---------KEGT.-L--I--------K----NVDSHTNME-....................-  1013
DRL.x.x.FAO              -E-V---------KEGT.-L--I--------K----NVDSHTNME-....................-  1013
GRV.ET.x.GRI_677_gri_1   -R-I---------KEGE.-------------K---ERKTM-SEGSMEGVREANKQMEGDSDLQDQE-  1041
GSN.CM.99.CN166          -V----------V-TQAGE-IT---------KP--EAKTE-VGG-THTNN-...............-  1016
GSN.CM.99.CN71           -V----------V-TQAGE-IT---------KP--EAK-E-VGS-AHTSN-REEGRMAD.......-  1017
LST.CD.88.SIVlhoest447   -Q----------VKEGD.T-FS-------LVK---EGPKDSEGGLHNN..................-  1020
LST.CD.88.SIVlhoest485   -QI---------VKEGD.T-FS-------LVK---EGPKDSEGSLDNN..................-  1020
LST.CD.88.SIVlhoest524   -Q----------VKEGE.N-FSI------LVK---EGPKDSQSSMDN-..................-  1021
LST.KE.x.lho7            -Q----------VKEGE.N-FS-------LVK---EGPKDSESSLDNN..................-  1020
MAC.US.x.251_1A11        GE---------ILKVGT.-------------K---GGKEM-SSSHMEDTGEAREVA..........-  1020
MND-1.GA.x.MNDGB1        GI----------LKYQE.E--I-----C---K---E.SGKNSQ-NLESV.................-  1010
MND-2.CM.98.CM16         -E-I---------KEGT.-L----------------TVDS-PN-EA....................-  1011
MND-2.GA.x.M14           -E-I---------KEGT.-L----------------TVDSNPHMED--ETA...............-  1017
MND-2.x.x.5440           -E-I---------KEGV.-L-----------K----TVDSNPHMGR....................-  1012
MNE.US.x.MNE027          GE---------ILKVGT.-------------K---GGKEV-SSSHMEDTGEAREVA..........-  1016
MON.CM.99.L1_99CML1      -R-----------KTEEG--LT---------KP--TTENVGGDTNQYNLRKQDGLAN.........-  1018
MON.NG.x.NG1             -H--------L--KTDQGEVIT----------P--SKEDVGSKPSAH-IREVDGMAD.........-  1020
MUS-1.CM.01.CM1239       -A----------VKTEQGQVIT---------KP--AKENVGSKSNTGDHRKEDGLDN.........-  1012
MUS-1.CM.01.SIVmus_01CM1085 -V---R-------KTD-GEVIT---------KP--AKEDVGSKSDTGDLRKEDRLDN.........-  1012
MUS-2.CM.01.CM1246       -R--------L--KTKEGE-VT---------K---TRQNV-SEPD-VHVRKEDGLAD.........-  1017
MUS-2.CM.01.CM2500       -Q--------I--KTDQGE-IT---------K---TKENV-SDPNPVY-RKEDGLAD.........-  1014
MUS-3.GA.09.09GabOI81    -T-----------KTDLGEV-T---------KP-NA-KNVGSKSNTGDIREEDGLVDQGN......-  1016
MUS-3.GA.11.11GabPts02   -A-----------KTDLGE-LT-------V-KP--PKKDVGGIPD-GD-GKTPGLVD.........-  1016
OLC.CI.97.97CI12         -E---------LLKEGD.KYFS--L-----VKP--VDSSANV........................-   995
RCM.CM.00.SIVagi_00CM312 -T-I-----------GQ.-L-------C--VK---RKNVDSETNMEGGQ-EN..............-  1018
RCM.CM.02.SIVrcm_02CM8081 -T-I-----------GQ.-L-------C---K---RKDVDSETSMEGRQ-KD..............-  1024
RCM.GA.x.SIVRCMGAB1      ---I----------EETG-L-----------KE--RKDVDSEANLAGRQ-EN..............-  1019
RCM.NG.x.NG411           ---I-E---------QE.-L-------C---K---RKEVDRETNMEGRQ-ES..............-  1023
SAB.SN.x.SAB1            ---I----------EQG.EL-TI--------K-----..-L-SQAPLEGNGRTAGEVD.........-  1040
SMM.CI.79.SIVsmCI2       -N---------I-KVGT.----I--------K---GGKELGNSPYLENP-EDGKMAQPD.......-  1028
SMM.LR.89.SIVsmLIB1      GE---------ILKVGT.-------------K---GGKEV-SGSHLEDTREIREMA..........-  1010
SMM.SL.92.SIVsmSL92A     GE---------I-KVG-.----I--------K---GGKEM-SSADMENTTQVRETAQLSEISKV..-  1023
SMM.SL.92.SIVsmSL92B     -E-----------KVGT.E------------K---GRQEMGSSASMEDQ.................-  1006
SMM.SL.92.SL92B          -E-----------KVGT.E------------K---GRQEMGSSASMEDQ.................-  1006
SMM.US.04.G078           GE---------ILKVGT.E------------K--AGGKEVGSGSHLEDPGEAREMA..........-  1016
SMM.US.05.D215           GE---------ILKVGT.---------V--VK--AGGKEV-SSSHLEDTRETGEVA..........-  1016
SMM.US.06.FTq            GE---------I-KVGT.-------------K--AGGKEL-SGTHVEDTREAGEMA..........-  1020
SMM.US.86.CFU212         GE---------ILKVGT.-------------K---GGKEV-SGSHLEDTREAGEMA..........-  1016
SMM.US.x.H9              GE---------ILKVGT.E------------K---GGKEL-SGSHLEDTGEAREVA..........-  1016
SMM.US.x.PGM53           GE---------ILKVGT.E------------K---GGKEL-SGPHLEDTGEAREVA..........-  1016
SMM.x.x.pE660.CG7G       GE---------ILKVGT.E------------K---GGKEL-SGSHLEDTGEAREVA..........-  1016
STM.US.89.STM_37_16      GE---------IVKVGT.-------------K---GGKEV-SGSHLEDTREAGEVA..........-  1016
SUN.GA.98.L14            -------------KQGE.N-L--------LVK---GESSSVEM-G.....................-  1017
SYK.KE.x.KE51            -S-------------TPDQQVIA---------SSD-ERVDSGTHLEITSKSN..............-  1041
SYK.KE.x.SYK173_COMGNM   -R----------V-TEEG--FA---------T-H-ER-DSGSH-ENDPKT-...............-  1037
TAL.CM.00.266            -H-I---------LRTDEGEVIT---------KP--QALGNKIDLE-SKEQ-AEMGRDN........-  1020
TAL.CM.01.8023           -H-V---------LRTDEGEVIT---------KP--QA-GNKTDLEGSKEQ-AEMGRDN........-  1021
TAN.UG.x.SIVagmTAN1      -T-----------KEGE.EL-----------KE--ERKTVG-KTNMEG..................-  1020
VER.DE.x.AGM3            GQ-I---------KGGV.EL-EY--------K--EPRKRMG-ESNLEGAGGADN............-  1023
VER.KE.x.9063            -Q-I---------KEGE.EL-I---------K--EPRKRMG-EGNMEGLRGPDNQMARNSQILDD.-  1034
VER.KE.x.AGM155          -R-I--------LKEGE.EL-----------K--EPRKTLG-ETHLEGAGGSDHQMAG.DS.....-  1032
VER.KE.x.TYO1_patent     -Q-I--------LK.DGS-L-----------K--EPKQRVGNEGDVEGTRGSDN............-  1026
WRC.CI.97.97CI14         G--P------L--ETPE.G-IT-------L-KVWNGEGMDRSSSDKN...................-  1023
WRC.CI.98.98CI04         G---------L-VETPE.G-IT--K------KVWNGEGMDRSSSNKD...................-  1022
WRC.GM.05.Pbt_05GM_X02   G--V------L-V-TPEG-LLT---------KTWDGKGM#..........................-  1007
```

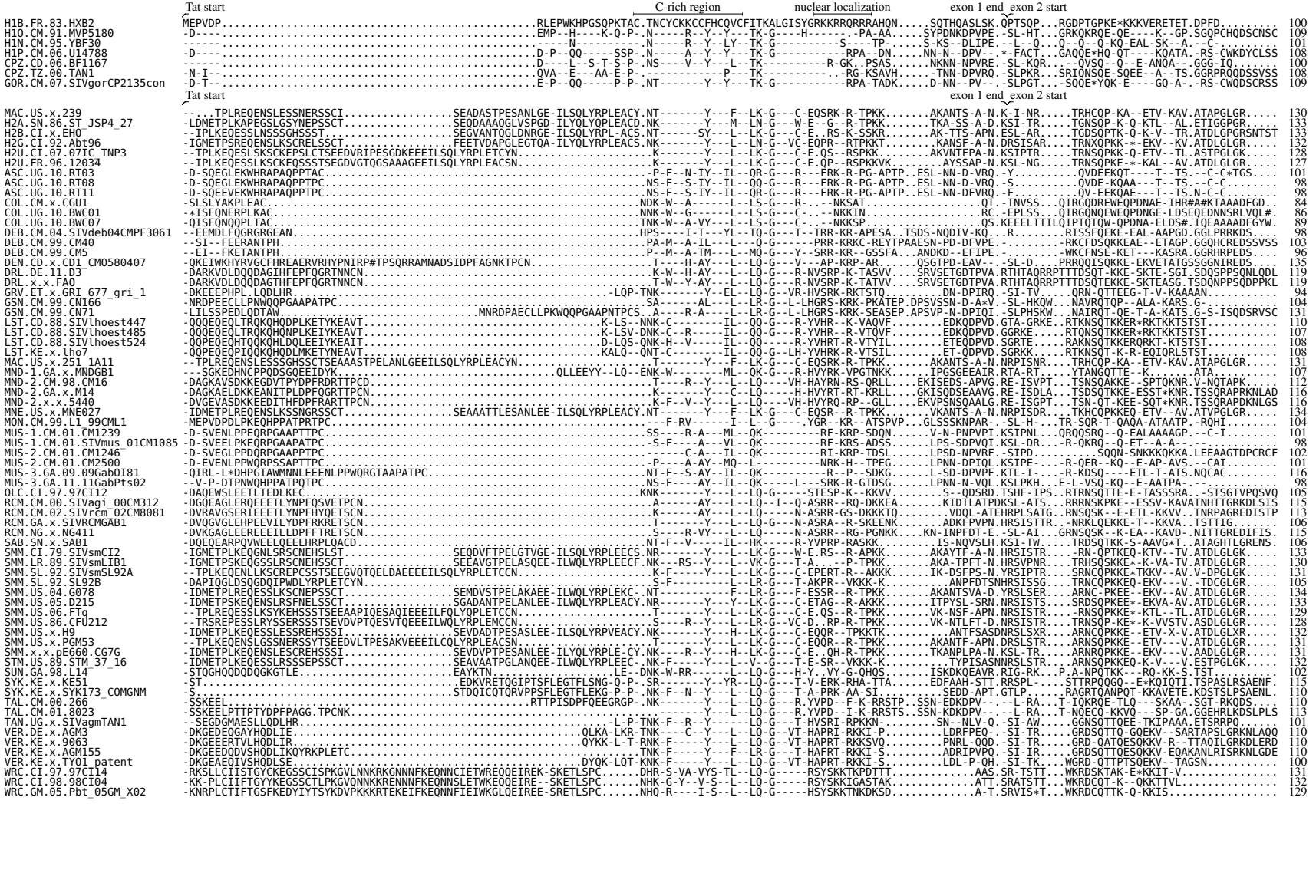


Draft of 30 September 2015, not for public release
plvprot_main.tex r8876–10546

	Vpx start	Vpx end	
MAC.US.x.239	MSD..PRERIP..PGNSGEETIGEAFEWLNRTVEEINREAVNHLPRELIFQVWQRSWEYWHDEQGMSPSYVKYRYLCLIQKALFMHCKKGCRCLGEGHGAGGW.RPGPPPPPPPG.LA*		112
H2A.SN.86.ST_JSP4_27	-AG..---T-..----------D----A----------R---------I--T-------M---M-I-S-R--T---G---P----.-S---------.-V-		113
H2B.CI.x.EHO	-.-..----V-..--------V--------ET-L-HL--V----------K--A--RE-----I--T------M---M-I-FA---G--R----P----.-S---------.---		112
H2G.CI.92.Abt96	-.-..----V-..-E--D----------E---T----I----------R----------A--T-----L-M---M-V------T--QK---P----.-Q---------.---		112
H2U.CI.07.07IC_TNP3	----.--------..-----D--VE------ES--TA--Q----------G--------T---T----------Y----R--Q--R----P----.----------.---		113
H2U.FR.96.12034	-G-..--------.----E---V----S--E---RD---A--Q----------Q--R------E--T----------Q-------P----.S---------.-I-		113
DRL.DE.11.D3	-AEGQ-A--A-AE-VGA-----V-LE---Q-SLLR--Q--RL-FHP-FL-RL-NTCM-HY--ALQL-FI-S-----L-L---M---FQQ--S-VQGR-PPPLRPAGDRL-----PG-V-		116
DRL.x.x.FAO	-AERQSV--A-AE-MGA-----V-LE---Q-SLLR--Q--RL-FHP-FL-RL-NTCM-HY--ALQL-FT-S-----L-L---M---FQQ--S--QGR-PPPLRPAGDRL-----PP..-		114
MAC.US.x.251_1A11	----.--------.----------.----------.----------.----------.----------.----------.----------.----------.----------.---		113
MND-2.CM.98.CM16	-AE..GAPE--E...GA----.VDLNT--E-SL-K--Q--RL-FHP-FL-RL-NACI-H---RHQR-L--A-----L-MN--M-T-MQQE-P-RSG.-P.....-...-----.MV-		100
MND-2.GA.x.M14	-AE.....-A-EA-QGA----.VGLEQ--E-SL-QV----QL-FHP-FL-RL-NTCV-H---RLRRTLN-A-----L-M---M-V-MQQ--P-RSG.-S.....-...-----.---		100
MND-2.x.x.5440	-AE.....-A-EA-EGA----.VGLEQ--ETSL-R-----RL-FHP-FL-RL-NTCV-H---RHQR-LD-A-----L-MH--MYT-MQQ--P-RNG.RPR.....-...-----.M--		100
MNE.US.x.MNE027	----.--------.----------.----------.----------.----------.----------.----------.----------.----------.----------.---		113
RCM.CM.00.SIVagi_00CM312	-AE.....-A-..E-PT-AGEVEFQPWLARMLY-INQEA.RL-F-G-F--HL-RTCV--------RTLE-AG----L-M-----T-MRS--KLR...--DPPR.QR-ERV-IL---.MQ-		106
RCM.CM.02.SIVrcm_02CM8081	-AE.....-A-..EVPT-AGEVEFQPWLARMLY-INQEA.RL-FHP-F--RL-RTCV-H----L-R-LE-AG----L-M-----T-MRS---LR...--DPPR.QR-ERV-IL---.MQ-		106
RCM.GA.x.SIVRCMGAB1	-AE.....-A-EV-TGA----.A-FQP--RDML-KV-L--RL-FHP-F--RL-RTCV-H---VHQR-LE-AA----L-M-----I--QT--SQR...--PNPR.AV-ERITIL---.M.-		105
RCM.NG.x.NG411	-AE..G---V-..EAPT-AGDVEF-PWLHRMLT-VNLEA.RL-FHP-F--RL-RTCV-H---RL-R-LE-AG----L-M-----I--QS--SQR...--Q-QAREA-ERIQIL---.M.-		109
SMM.CI.79.SIVsmCI2	-GS..A-------.--T-------------YS---AV--G--K----------R----F--------QA-TQ----L-------C-V----T--R----S----.-A---------.---		113
SMM.LR.89.SIVsmLIB1	----.--------.----------.----------E---------Q----------R----F--------A--T----------M-L----------.S----.S---------.---		112
SMM.SL.92.SIVsmSL92A	-.-.----Q---.-----D-------D--E-----V--T----------R----------AA-T----------P----------.F--		111
SMM.SL.92.SIVsmSL92B	-T-..--------..----------HN---AL-QT--Q-------S---R-C----V----Y----A----VQ-M---M-Q-FR---T-R----SQ----.T---------.---		113
SMM.SL.92.SL92B	-T-..--------..----------HN---AL-QT--Q----------R-C----V----Y----A----VQ-M---M-Q-FR---T-R----SQ----.T---------.---		113
SMM.US.04.G078	----..--------..----------H---A----------R------Y----R-E--T----------W---M------R--------.----------.---		113
SMM.US.05.D215	--N..--------..----------E----D---V----------R----------T----------YI----------D-.----------.---		113
SMM.US.06.FTq	----..--------..----------E----D---A----------R----------R--A--T----------V---------R--------.T---------.---		113
SMM.US.86.CFU212	-A-..--------..----------L--V----------R-C--F--------A--T-------M---M----R---N--R---E----.-Q---------.---		113
SMM.US.x.H9	--X..--------..--------X----D--H--X-----A----------R-X-------M---V--T-X----------GE------.----------.---		113
SMM.US.x.PGM53	----..--------..----------A----------R------R--M---E--T----------V---R-------E------.-S---------.---		113
SMM.x.x.pE660.CG7G	----..--------..---------E----------G---A----------R---------M---E--T----------V----------E------.-T---------.---		113
STM.US.89.STM_37_16	----..--------..-------A-E------H----D----------R----------PG----------S-R----------.----------.---		113

Draft of 30 September 2015, not for public release
plvprot_main.tex r8876–10546

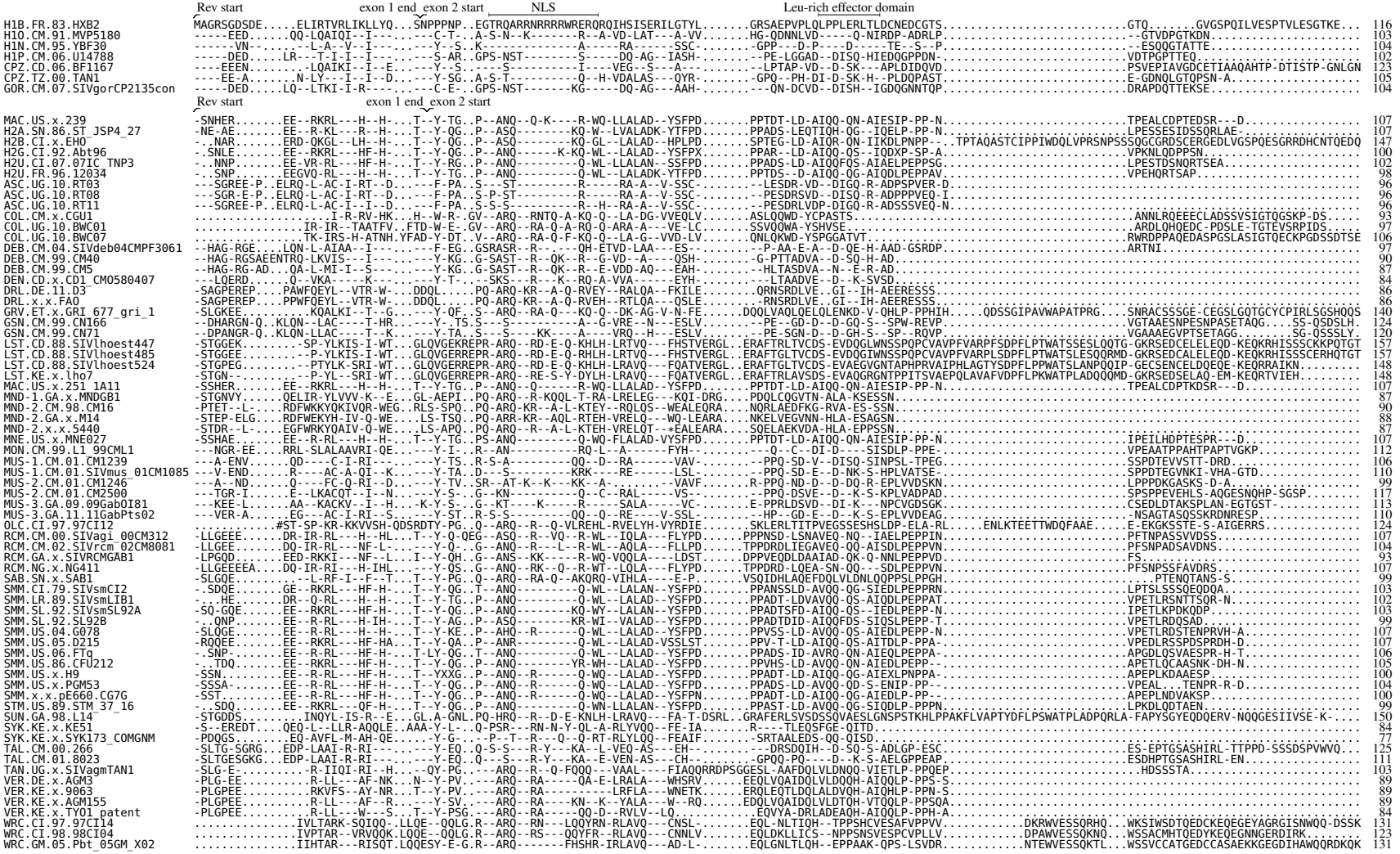

oligomerization
amphipathic α-helix
H(S/N)RIG motifs
Vpr start
Vpr end in HXB2
frameshift in HXB2
Vpr end
H1B.FR.83.HXB2
H1O.CM.91.MVP5180
H1N.CM.95.YBF30
H1P.CM.06.U14788
CPZ.CD.06.BF1167
CPZ.TZ.00.TAN1
GOR.CM.07.SIVgorCP2135con
MAC.US.x.239
H2A.SN.86.ST_JSP4_27
H2B.CI.x.EHO
H2G.CI.92.Abt96
H2U.CI.07.07IC_TNP3
H2U.FR.96.12034
ASC.UG.10.RT03
ASC.UG.10.RT08
ASC.UG.10.RT11
COL.CM.x.CGU1
COL.UG.10.BWC01
COL.UG.10.BWC07
DEB.CM.04.SIVdeb04CMPF3061
DEB.CM.99.CM40
DEB.CM.99.CM5
DEN.CD.x.CD1_CMO580407
DRL.DE.11.D3
DRL.x.x.FAO
GRV.ET.x.GRI_677_gri_1
GSN.CM.99.CN166
GSN.CM.99.CN71
LST.CD.88.SIVlhoest447
LST.CD.88.SIVlhoest485
LST.CD.88.SIVlhoest524
LST.KE.x.lho7
MAC.US.x.251_1A11
MND-1.GA.x.MNDGB1
MND-2.CM.98.CM16
MND-2.GA.x.M14
MND-2.x.x.5440
MNE.US.x.MNE027
MON.CM.99.L1_99CML1
MON.NG.x.NG1
MUS-1.CM.01.CM1239
MUS-1.CM.01.SIVmus_01CM1085
MUS-2.CM.01.CM1246
MUS-2.CM.01.CM2500
MUS-3.GA.09.09GabOI81
MUS-3.GA.11.11GabPts02
OLC.CI.97.97CI12
RCM.CM.00.SIVagi_00CM312
RCM.CM.02.SIVrcm_02CM8081
RCM.GA.x.SIVRCMGAB1
RCM.NG.x.NG411
SAB.SN.x.SAB1
SMM.CI.79.SIVsmCI2
SMM.LR.89.SIVsmLIB1
SMM.SL.92.SIVsmSL92A
SMM.SL.92.SIVsmSL92B
SMM.SL.92.SL92B
SMM.US.04.G078
SMM.US.05.D215
SMM.US.06.FTq
SMM.US.86.CFU212
SMM.US.x.H9
SMM.US.x.PGM53
SMM.x.x.pE660.CG7G
STM.US.89.STM_37_16
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173_COMGNM
TAL.CM.00.266
TAL.CM.01.8023
TAN.UG.x.SIVagmTAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TYO1_patent
WRC.CI.97.97CI14
WRC.CI.98.98CI04
WRC.GM.05.Pbt_05GM_X02



Draft of 30 September 2015, not for public release
plvprot_main.tex r8876–10546

		Tat end	
H1B.FR.83.HXB2	.....	*	100
H1O.CM.91.MVP5180	TRISGQ.....	-	116
H1N.CM.95.YBF30	.....	-	102
H1P.CM.06.U14788	GTT.....	-	112
CPZ.CD.06.BF1167	.....	-	101
CPZ.TZ.00.TAN1	GRTSGTSSSGYTRPFKTSSGSSGSACKH	-	137
GOR.CM.07.SIVgorCP2135con	GTISEP.....	-	116

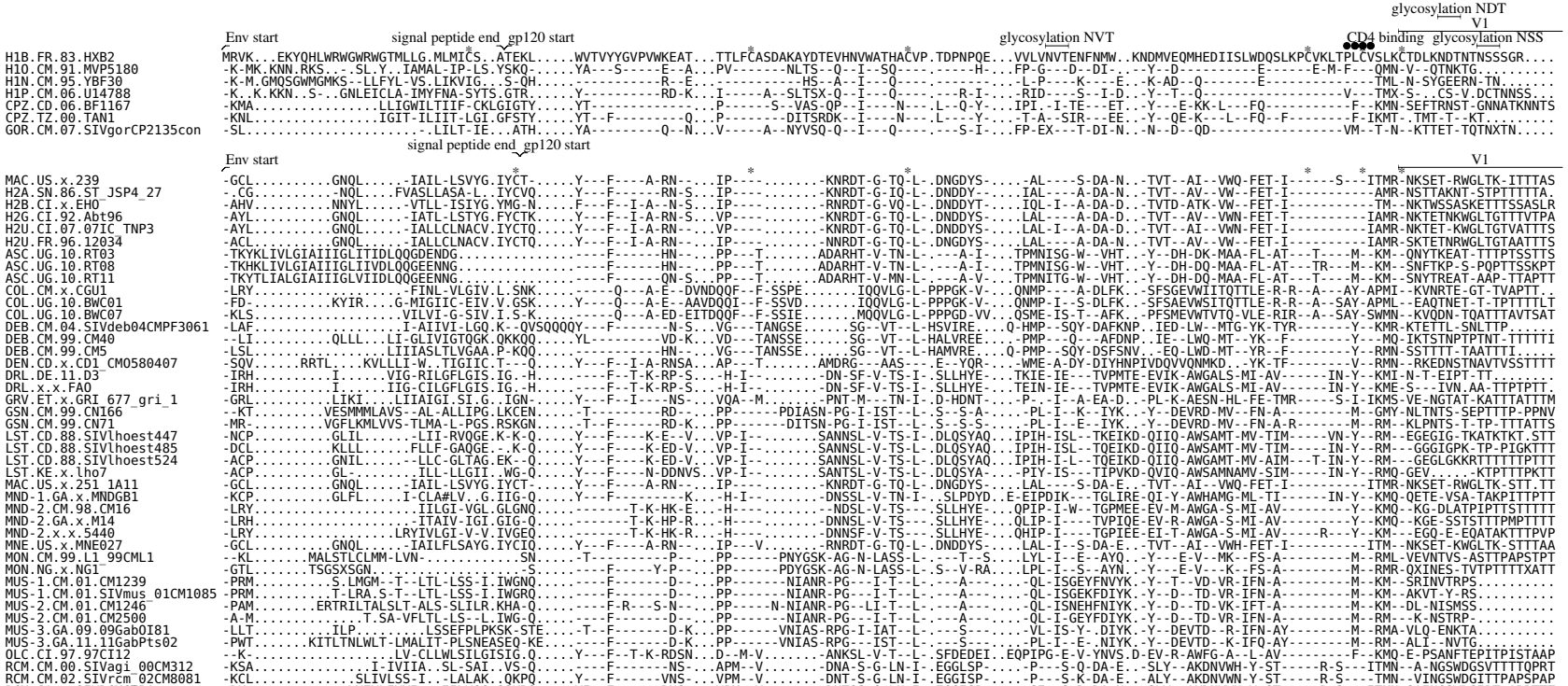
		Tat end	
MAC.US.x.239	.....	-	131
H2A.SN.86.ST_JSP4_27	.....	-	134
H2B.CI.x.EHO	SRFAN.....	-	139
H2G.CI.92.Abt96	.....	-	133
H2U.CI.07.07IC_TNP3	.....	-	129
H2U.FR.96.12034	.....	-	128
ASC.UG.10.RT03	.....	-	102
ASC.UG.10.RT08	.....	-	99
ASC.UG.10.RT11	.....	-	99
COL.CM.x.CGU1	.....	-	85
COL.UG.10.BWC01	.....	-	87
COL.UG.10.BWC07	.....	-	90
DEB.CM.04.SIVdeb04CMPF3061	.....	-	99
DEB.CM.99.CM40	GRISNNC.....	-	111
DEB.CM.99.CM5	.....	-	97
DEN.CD.x.CD1_CMO580407	.....	-	136
DRL.DE.11.D3	GAKK.Q.....	-	125
DRL.x.x.FAO	GAKK.Q.....	-	125
GRV.ET.x.GRI_677_gri_1	.....	-	95
GSN.CM.99.CN166	.....	-	105
GSN.CM.99.CN71	GTT.....	-	135
LST.CD.88.SIVlhoest447	.....	-	111
LST.CD.88.SIVlhoest485	.....	-	108
LST.CD.88.SIVlhoest524	.....	-	109
LST.KE.x.lho7	.....	-	109
MAC.US.x.251_1A11	.....	-	132
MND-1.GA.x.MNDGB1	.....	-	108
MND-2.CM.98.CM16	.....	-	113
MND-2.GA.x.M14	LGGQSK.....	-	123
MND-2.x.x.5440	LGSQSKSGAG.....	-	127
MNE.US.x.MNE027	.....	-	135
MON.CM.99.L1_99CML1	.....	-	105
MUS-1.CM.01.CM1239	.....	-	102
MUS-1.CM.01.SIVmus_01CM1085	.....	-	99
MUS-2.CM.01.CM1246	QKTSTT.....	-	109
MUS-2.CM.01.CM2500	.....	-	102
MUS-3.GA.09.09GabOI81	.....	-	117
MUS-3.GA.11.11GabPts02	.....	-	99
OLC.CI.97.97CI12	GH.....	-	108
RCM.CM.00.SIVagi_00CM312	.....	-	116
RCM.CM.02.SIVrcm_02CM8081	.....	-	114
RCM.GA.x.SIVRCMGAB1	.....	-	107
RCM.NG.x.NG411	.....	-	116
SAB.SN.x.SAB1	.....	-	107
SMM.CI.79.SIVsmCI2	.....	-	134
SMM.LR.89.SIVsmLIB1	.....	-	131
SMM.SL.92.SIVsmSL92A	.....	-	132
SMM.SL.92.SL92B	.....	-	106
SMM.US.04.G078	.....	-	135
SMM.US.05.D215	.....	-	134
SMM.US.06.FTq	.....	-	130
SMM.US.86.CFU212	.....	-	129
SMM.US.x.H9	.....	-	133
SMM.US.x.PGM53	.....	-	132
SMM.x.x.pE660.CG7G	.....	-	132
STM.US.89.STM_37_16	.....	-	133
SUN.GA.98.L14	.....	-	103
SYK.KE.x.KE51	.....	-	116
SYK.KE.x.SYK173_COMGNM	.....	-	111
TAL.CM.00.266	.....	-	111
TAL.CM.01.8023	GRTSTASSATVTRP.....	-	128
TAN.UG.x.SIVagmTAN1	.....	-	102
VER.DE.x.AGM3	SGRATGASD.....	-	120
VER.KE.x.9063	KREAVGANA.....	-	120
VER.KE.x.AGM155	TRGPVGAGN.....	-	120
VER.KE.x.TYO1_patent	.....	-	101
WRC.CI.97.97CI14	.....	-	132
WRC.CI.98.98CI04	.....	-	133
WRC.GM.05.Pbt_05GM_X02	.....	-	130


Rev start
exon 1 end
exon 2 start
NLS
Leu-rich effector domain
H1B.FR.83.HXB2
H1O.CM.91.MVP5180
H1N.CM.95.YBF30
H1P.CM.06.U14788
CPZ.CD.06.BF1167
CPZ.TZ.00.TAN1
GOR.CM.07.SIVgorCP2135con
MAC.US.x.239
H2A.SN.86.ST_JSP4_27
H2B.CI.x.EHO
H2G.CI.92.Abt96
H2U.CI.07.07IC_TNP3
H2U.FR.96.12034
ASC.UG.10.RT03
ASC.UG.10.RT08
ASC.UG.10.RT11
COL.CM.x.CGU1
COL.UG.10.BWC01
COL.UG.10.BWC07
DEB.CM.04.SIVdeb04CMPF3061
DEB.CM.99.CM40
DEB.CM.99.CM5
DEN.CD.x.CD1_CMO580407
DRL.DE.11.D3
DRL.x.x.FAO
GRV.ET.x.GRI_677_gri_1
GSN.CM.99.CN166
GSN.CM.99.CN71
LST.CD.88.SIVlhoest447
LST.CD.88.SIVlhoest485
LST.CD.88.SIVlhoest524
LST.KE.x.lho7
MND-1.GA.x.MNDGB1
MND-2.CM.98.CM16
MND-2.GA.x.M14
MND-2.x.x.5440
MNE.US.x.MNE027
MON.CM.99.L1_99CML1
MUS-1.CM.01.CM1239
MUS-1.CM.01.SIVmus_01CM1085
MUS-2.CM.01.CM1246
MUS-2.CM.01.CM2500
MUS-3.GA.09.09GabOI81
MUS-3.GA.11.11GabPts02
OLC.CI.97.97CI12
RCM.CM.00.SIVagi_00CM312
RCM.CM.02.SIVrcm_02CM8081
RCM.GA.x.SIVRCMGAB1
RCM.NG.x.NG411
SAB.SN.x.SAB1
SMM.CI.79.SIVsmCI2
SMM.LR.89.SIVsmLIB1
SMM.SL.92.SIVsmSL92A
SMM.SL.92.SL92B
SMM.US.04.G078
SMM.US.05.D215
SMM.US.06.FTq
SMM.US.86.CFU212
SMM.US.x.H9
SMM.US.x.PGM53
SMM.x.x.pE660.CG7G
STM.US.89.STM_37_16
SUN.GA.98.L14
SYK.KE.x.KE51
SYK.KE.x.SYK173_COMGNM
TAL.CM.00.266
TAL.CM.01.8023
TAN.UG.x.SIVagmTAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TYO1_patent
WRC.CI.97.97CI14
WRC.CI.98.98CI04
WRC.GM.05.Pbt_05GM_X02

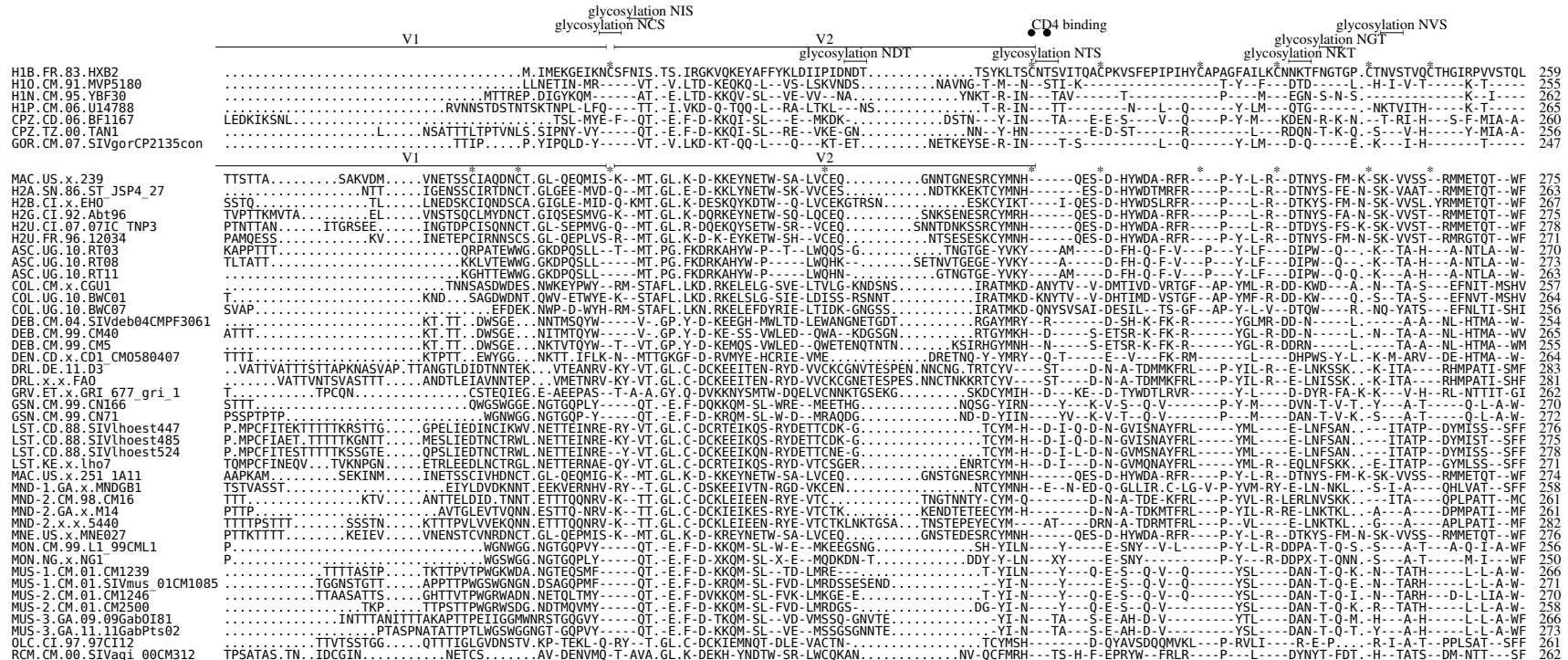
	Rev end	
H1B.FR.83.HXB2	.....*	116
H1O.CM.91.MVP5180	.....-	104
H1N.CM.95.YBF30	.....-	105
H1P.CM.06.U14788	.....-	103
CPZ.CD.06.BF1167	SN.....-	126
CPZ.TZ.00.TAN1	.....-	106
GOR.CM.07.SIVgorCP2135con	.....-	105
	Rev end	
MAC.US.x.239	.....-	108
H2A.SN.86.ST_JSP4_27	.....-	108
H2B.CI.x.EHO	TRG.....-	151
H2G.CI.92.Abt96	.....-	101
H2U.CI.07.07IC_TNP3	.....-	103
H2U.FR.96.12034	.....-	99
ASC.UG.10.RT03	.....-	97
ASC.UG.10.RT08	.....-	97
ASC.UG.10.RT11	.....-	97
COL.CM.x.CGU1	.....-	94
COL.UG.10.BWC01	.....-	98
COL.UG.10.BWC07	VATICDECRA.....-	117
DEB.CM.04.SIVdeb04CMPF3061	.....-	98
DEB.CM.99.CM40	.....-	91
DEB.CM.99.CM5	.....-	88
DEN.CD.x.CD1_CMO580407	.....-	85
DRL.DE.11.D3	.....-	87
DRL.x.x.FAO	.....-	87
GRV.ET.x.GRI_677_gri_1	KKSAARP.....-	148
GSN.CM.99.CN166	.....-	125
GSN.CM.99.CN71	.....-	121
LST.CD.88.SIVlhoest447	KEEMEIQIQRKKWVSQ-	174
LST.CD.88.SIVlhoest485	KEKMEIQIQREKWVSQ-	174
LST.CD.88.SIVlhoest524	.....-	149
LST.KE.x.lho7	.....-	149
MAC.US.x.251_1A11	.....-	108
MND-1.GA.x.MNDGB1	.....-	88
MND-2.CM.98.CM16	.....-	91
MND-2.GA.x.M14	.....-	89
MND-2.x.x.5440	.....-	88
MNE.US.x.MNE027	.....-	108
MON.CM.99.L1_99CML1	.....-	113
MUS-1.CM.01.CM1239	.....-	107
MUS-1.CM.01.SIVmus_01CM1085	.....-	111
MUS-2.CM.01.CM1246	.....-	100
MUS-2.CM.01.CM2500	.....-	118
MUS-3.GA.09.09GabOI81	.....-	114
MUS-3.GA.11.11GabPts02	.....-	111
OLC.CI.97.97CI12	.....-	124
RCM.CM.00.SIVagi_00CM312	.....-	108
RCM.CM.02.SIVrcm_02CM8081	.....-	105
RCM.GA.x.SIVRCMGAB1	.....-	94
RCM.NG.x.NG411	.....-	108
SAB.SN.x.SAB1	.....-	100
SMM.CI.79.SIVsmCI2	.....-	104
SMM.LR.89.SIVsmLIB1	.....-	103
SMM.SL.92.SIVsmSL92A	.....-	104
SMM.SL.92.SL92B	.....-	100
SMM.US.04.G078	.....-	108
SMM.US.05.D215	.....-	108
SMM.US.06.FTq	.....-	107
SMM.US.86.CFU212	.....-	106
SMM.US.x.H9	.....-	101
SMM.US.x.PGM53	.....-	105
SMM.x.x.pE660.CG7G	.....-	101
STM.US.89.STM_37_16	.....-	100
SUN.GA.98.L14	.....-	151
SYK.KE.x.KE51	.....-	85
SYK.KE.x.SYK173_COMGNM	.....-	78
TAL.CM.00.266	.....-	126
TAL.CM.01.8023	.....-	112
TAN.UG.x.SIVagmTAN1	.....-	104
VER.DE.x.AGM3	.....-	90
VER.KE.x.9063	.....-	90
VER.KE.x.AGM155	.....-	90
VER.KE.x.TYO1_patent	.....-	85
WRC.CI.97.97CI14	Q.....-	133
WRC.CI.98.98CI04	.....-	124
WRC.GM.05.Pbt_05GM_X02	QQHQGNYCQTLHKP..-	146

Draft of 30 September 2015, not for public release

	phos phos transmembrane domain | cytoplasmic domain Vpu start | α-helix | α-helix | Vpu end	
H1B.FR.83.HXB2	TQP..IPIVAIVALVVAIIIAIVVWSIVIIE....YRKILRQRK.IDRLIDRLIERAEDSGNESEGEIS.....ALVEMGVEMGH...HAPWDVDDL....*	82
H1O.CM.91.MVP5180	MHQ..ENLL-LI--SALCL-NVLI-LFNLRIYLVQRKQDR-EQE.-LERLR-IK-IRD--DY--NE-EQ...QE.....-MELIHSHGF..AN-MFEL.....-	86
H1N.CM.95.YBF30	MLS.....LGFI--GA-VS--VI--ALLYR.....K--KL-E--.KHIRQ-IR--E-------D-DAEWLDGDEEW.---TLLSSSKLD..QGN-V.....-	84
H1P.CM.06.U14788	MH-...RDEAVLIIAG-LLLCC--X-GK-LLL....VL-ERERD-.FVQRLA-WR-GQ--E-Y--NE-EE...EQLRELG.....NLLGFDHV.....-	79
CPZ.CD.06.BF1167	MLWQFLQWLQYLGWGG--V-W-IALLVIRKAYLY.IKQVQEEDR.-EQA-IDRYQ-RGSTDSGI-EDEN...GLI.....NWMDPDNYFLL...-	82
CPZ.TZ.00.TAN1	MIKIVVGS-STNVIGILC-LL-LIGGGLL-GIG..I-RE-ERER.QHQRVLERLA-RLSIDSGV-EDEE.....FNWNNFDP..-NYNPR-WI....-	84
GOR.CM.07.SIVgorCP2135con	MH-..RD-IV-IIGITLLAATVI--LKALAL....-LRDR-E-R.FFNRLE--LSTK--E-Y--NE-EA...AELMERG.....NELGFDFN-H...-	80
DEN.CD.x.CD1_CMO580407	MLK...IKLGPIEY-CLVFAVVIT-IAA-GVGYL.AYRAYKSYR.EELRYI--RLWSI---Y--SQ-DP.....-	65
GSN.CM.99.CN166	MH-.....A-VWWWGA---TF-YLCVALLALYLAWDKWVKGKP-PTQVAVI----DE----IYDDASSE.....LTGFNGF..AN-GFEV.....-	79
GSN.CM.99.CN71	MSA.....A-LWWWGA-V-TF-YFCLAIFALYLAWDKW-KGKP-..PVAVI--V-DD-E--IFEDASSE.....PNAYGF..AN-GFEV.....-	77
MON.CM.99.L1_99CML1	MNY..WWSLVAITYSLIL-ALP-AAWAWWRY....-KITK-FKR.--QE-Q---QIH-RRRHD-GVDTE.....SES-QHE-THGF..VN-VFN--FGEWV-	88
MON.NG.x.NG1	MNT..YW-A--SVWC-IVLL-PCTLYLLYQS....-KEHK-VSR.FEQDFQ--VQXYQEXDSGY-D-DEDHNSFDNPLFDDGDPDQWV*FRKQIMGHSILWSAGL-	98
MUS-1.CM.01.CM1239	MNY..WYLA-ALVTGIYFL--LFAFVLAYQRWCRP.....KVEVSVIRLLEEG--DSGIFEDAE.....DEP.....N-DGHHAF..AN-AFEQ.....-	76
MUS-1.CM.01.SIVmus_01CM1085	MNY..WYLA-VIVTGIYFV---FAFVLAYQRWCKP.....-..KVEVSVIRLLEEG-GDSGIFEDAE.....DDM.....AES-HHAF..AN-AFEQ.....-	77
MUS-2.CM.01.CM2500	MNW..WWFA-A-VTAIYFV--L-AFVLAYQR....WCQP...KG.QVEVNVIRLLEEG-TDSGIFEDAE.....DG.....SDDQRHGF..LN-AFEL.....-	78
MUS-3.GA.09.09GabOI81	MNV..WYLA-AIVTAIYXA-V--ALWLL.....WD-Y.VRKEPTVVSVI--L-EGDSGYE-VFEDAP.....QEP.....GDPLINAF..DN-AFEQ.....-	78



Draft of 30 September 2015, not for public release
plvprot_main.tex r8876–10546



Draft of 30 September 2015, not for public release
plvprot_main.tex r8876–10546

```
                                                       CD4 binding      glycosylation NNT
                                                                   glycosylation NCT         V3 tip
                                               glycosylation NFT                              V3
                            glycosylation NGS             glycosylation NTS                                          glycosylation NNT   glycosylation NKT                                    CD4 binding
H1B.FR.83.HXB2              LLNGSLAEEEVVIR........SVNFT.DNAKTIIVQLNTSVEINCTRPNNN.TRKRIRIQRGPGRAFV....TIGK.IGNMRQAHCNISRAKWNNTLKQIASKLREQFGNNKTII..............................FKQSSGGDPEIVTHSFNCGGEF  382
H1O.CM.91.MVP5180           I---T-SR-K.--.......IMGK-I-.ES--N---T---PINMT-I-EGIA.EVQD-YT..--M-WRSMTLKRSNNTSPRS-V-Y-TYNKTV-E-A-Q-T-I.RYL..NLVNQT.....ENVTIIF.....SRT..................----A-VSHLH---H---  383
H1N.CM.95.YBF30             I-----NTDGI---...........N.-SHSNLL--W-ET-P------G--.-GGQVQ-..--AMT-Y....N-E-IV-DI---Y--V-KEL-EPMWNRTREEIKKIL-K-N..............................ITFRARERNE--L-VTHLM---R---  381
H1P.CM.06.U14788            I---T--KG-PLVI.......TQ-VS.-TG-V---K--E--SLT-I--G--.--GQVQL..-V.MTWY....NMKHYV-DI-A----V--EN-KR--EWVSEAIWKAYPQHSNHTSK......................HNHTFV--N-T-----VSSLH-S-H---  394
CPZ.CD.06.BF1167            H----VE-G-TKAY........Y-EAKHNAPLL-KFNKA-VFN-T-V--G--.--GQVQ-.---MT-Y....N-RNV--DT-K-F-YVNKTL-E-ATEAAKVAIN-TLK-TTFKN....................TTYTFRTDGDL-WI----Q-TQ--W---Q---  393
CPZ.TZ.00.TAN1              H-----E---TKAY........FVN-.SVNTPLL-KF-V-INLT-E-TG--.--GQVQ-.---MT-Y....N-ENVV-DT-K-Y-SVNATT-YRN-DWAMAAINTTMRARNETV...........................QQT-QWQRD----VTSFW---Q---  380
GOR.CM.07.SIVgorCP2135con   I---T--K-HAT-I........-K-S-.-SG-D---K-AR----T-E-TS--.--GQ-QV.--MTIY....NSENIV--T-K-F-KYNKTN-Q-A--DTVRA-KANQIRFQ..............................N-------VTFLH---H---  363
                                                                                     V3 tip
                                                                                   V3
MAC.US.x.239                GF--TR--NRTY-Y........WHGR.--RTI-SLNKYYNLTMK-R--G-K.-VLPVT-..MS-LV-.....HSQPINDRPK--W-WF.GG--KDAI-EVKQTIVKHPRY...TGTNNT.DKINLTAPGG.......................----VTFMWT--R---  399
H2A.SN.86.ST_JSP4_27        GF--TR--NRTY-Y........WHGR.--RTI-SLNKFYNLTVH-K--G-K.-VVP-TL..MS-LV-.....HSQPINRRP---W-WF.KGE-KEAM-EVKLT-AKHPRY...KGTNDT.EKIRFIAPGER.....................S---VAYMWT--R---  388
H2B.CI.x.EHO                GF--TR--NRTY-Y........WHGK.--RTI-SLNSYYNLTMH-K--G-K.MVVP--T..VS-IL-.....HSQPINKRPK--W-WF.KGN-TEAIQEVKETIKNHPRYSG-TNI.........................SQIRLAEHARSS---VRYMWT--R---  392
H2G.CI.92.Abt96             GF--TR--NRTY-Y........WHGR.--RTI-SLNKYYNLT-X-R--G-K.-VLP-T-..MS-LV-.....HSQPINERP---W-WF.GG--KEAMQEVKETVVKHPSY...KGTNDTKXITFTTPGEG.......................S---VKFMWT--R---  400
H2U.CI.07.07IC_TNP3         GF--TR--NRTY-Y........WHGR.--RTI-SLNKYYNLTVR-R--G-K.-VLPVT-..MS-LV-.....HSQPINERPK--W-WF.GGS-KEAIQEVKET-VKHPRYTGINDTKKINLTQ.......................PG-----ATFMWT--R---  402
H2U.FR.96.12034             GF--TRS-NRTY-Y........WHGR.--RTI-SLNKYYNLTMY-R--G-K.-VLPVT-..MS-LV-.....HSQAINN-PK--W-WF.GGN-KGAIQEVKETIANHPRY...HGTKNI.SQINLAEPAG.......................--S-VKFMWT--R---  395
ASC.UG.10.RT03              Q---TYEKNTDEVQV.......MRKYNSNYSVAF-FSESQI-NLT-V--G-K.SIRNLQ-..-A-MT-Y....SQLIVG--T-R-Y-K-NGTQ-Y-A-NATYRAIKKEYNLTRDQPITWRSQPR.......................--R-VES-W-Q-Q---  396
ASC.UG.10.RT08              Q---TYEKNTDKVQV.......MRKYNSNYSVAF-FSESQI-NLT-V--G-K.SIRNLQ-..-A-MT-Y....SQLIVG--T-R-Y-K-NGTQ-YSA-NATYRAIQKEYNLTENQSITWRSQPR.......................--R-VES-W-Q-Q---  399
ASC.UG.10.RT11              Q---TYEKNTDEVQV.......MRKYNSNYSVAF-FSESQI-NLT-V--G-K.-IRNLQ-..-S-MT-Y....SQLIVG-DT-R-F-R-NGTQ-YHA-NATYRAIKKEYNLTQNQPITWRSQPR.......................--R-VES-W-Q-Q---  389
COL.CM.x.CGU1               -V-A-KELSDWAKD......REG-WKN.-SGTIEYYWFPKDIALG-I-RG-S.SHRNLNT..AN-AK-Y....YELIPYSKGIYGR-QF.VPMTGQNK-NKTQFAI-IKK-LTAWLERISRKNI...................TITPRNGNRTS---ATFTFVI-HRL-  392
COL.UG.10.BWC01             -V-A-EELK-WSKD......REG-WKN.ESGTFEYYWFPKDIALG-I-RG-S.SHRNLNT..AN-AK-Y....YEVRPEADRWKGR-QF.VPMNQKNMTKK.EFAGKIVE-LTNWLQNITEYGN...................ITIKPRNGNKTS---ATFTFLI-HRL-  399
COL.UG.10.BWC07             VV---KTLVDYAKT......REGIWAN.ESGNIEYYWFPKDLYLG-K-RG-V.-NRNLNT..AN-AV-Y....YELRGQDADWK-R-QF.VKIGM-KTTLAQEIKKNLTYWL-N-TSYDNITIK................PRNGNKTSDPEATYTFITFHGMFLY-NASS  395
DEB.CM.04.SIVdeb04CMPF3061  QF---DE-RAEE-HV....IRK..GLE.NKTI--RIPPEYNLRLD-V--G-K.-YRATHM..AT-LS-Y....-TFVERLRIKR---RM.-GN-SGAVSD-KQV-KGIY...NVTN.............................ITMHYP---R-VQRVW-Q-H--Y  375
DEB.CM.99.CM40              QF---DE-RAEELHI..IRKEVKGEVQ.NGSI--R-PAKYNLTLT-V--G-K.-YRA-HM..AT-LS-Y....-TFIQRLRIKR---RL.NGS-A-AT-EMRQ-IL-I--KANRTNN.............................LTIHYPK--R-VQSVW-Q-H---  394
DEB.CM.99.CM5               QF---DE-RAED-HI....IRK..GLE.NKTI-VRIPPKYNLSLD-V--G-K.-YRAVHM..AT-LS-Y....-TFIPRLRIKR---RM.-GN-TGAV-E-R-R-K-IY...NVTK.............................ITVHYP---R-VQNTW-Q-H---  376
DEN.CD.x.CD1_CMO580407      IM-STVELLDNHTQI.....IRNTSKD.NQTIA-LFRKSSVIKLK----G-V.-L-GLNM..AA-IVMR....ATIIPRRAI-R-A-K-VGGN-TR-RQE-EKAVK-E-NI-G-VKW..............................DVPRV-----TSSAR-Q-R---  390
DRL.DE.11.D3                GF--TKHDQDEL-ET.....RGRPE-V.-HKYVFR-NKKWNLK-M-R-KG-R.SVISTPS..AT-LL-Y....HGLEPGK-LNRGM-KF.EGQ-GKA-ESL-KE-KAVNASIWRNTTIKGKECKRLKETKD..TKTVGTGCMLKQLKVDNYTTK--LASENLMML----Y  437
DRL.x.x.FAO                 GF--TKHGQDEL-ET.....RGKPD-L.NHKYVFR-NKKWNLK-I-R-KG-R.SIISTPS..AS-LL-Y....HGLEPGK-LKRGM-KF.EGQ-GKA-GSL-KE-KSVS-SIWRNTT..NSTCISLKGKKI..QDRNVTGCMLKQLKVDNYTTK--LASENLMML----Y  433
GRV.ET.x.GRI_677_gri_1      G----RS-NRTE-W.......QKGGND.NDTVI-KLNKFYNLTVR-R--G-K.-VLPVT-..MA-LV-.....HSQ-YNTRLK--W-HF.QGD-KGAW-EVREEVKKVKNLTEVS-E.............................NIHLRRIW..----SANFW---Q---  386
GSN.CM.99.CN166             H---TYQPGNNTRIM.......MNGMK.NESIV-GFGEDYHLNLT-I--G-K.-IRNLQ-..-A-MT-Y....SQVIVG--T-K-Y-RLDPDR--RAIREAMKAMN-HWE-KTG.........................RNDTQIRWTSEPK--L-VQ--W-Q-Q---  398
GSN.CM.99.CN71              H---TYQPGNNTRVM.......MNGKK.NESIV-GFGEDYQLTLT-I--G-K.-L-NLQ-..-A-MT-Y....SQIIVG-DT-K-Y-KLNHT--DIAIR-AMKAMKNHWEKINNDTT.........................PNKTQIRWTSEPK--L-VQ--W-Q-Q---  403
LST.CD.88.SIVlhoest447      GF--TNHTDDEL-PL...TPKKMG-LL.EAKFVYK-AGKWGLI-R-I-KG-R.SEVSTIS..ST-LL-Y....YGLEHGSRL-L-Q-KF.EGQ-GRMFHNLGNM-KRLNATAMNYTHGNCIDNT...........KPCGRQLKGLPIANMTRR--NLATEMLMHT--E-M  423
LST.CD.88.SIVlhoest485      GF--TNHT-DKL-PL...TPKKMG-LL.-AKFVYK-SGKWGLI-R-I-KG-R.SEVSTIS..ST-LL-Y....YGLEHGSRL-L-Q-KF.EGQ-GRMFNNLGKM-QKLNATAMNYTNGQCPEGK...........QPCGRHLKGLPIANMTRR-EGLATEMLMHT--E-M  422
LST.CD.88.SIVlhoest524      GF--TNHT-DEL-PL...TPKKMGDLN.GAKFVYK-AGKWGLI-R-I-KG-R.SEVSTIS..ST-LL-Y....YGLEHGSRL-L-Q-KF.EGQ-GRMFNNLGRM-KKLNETAMNYKEGNCSTPG...........KPCGRQLKGLPIANMTRR-V-LATEMLMHT--E-M  425
LST.KE.x.lho7               GF--TNHTRDEL-PL...TPNKMEDLN.GAKFVYK-AGKWGLI-R-I-KG-R.SEVSTIS..ST-YL-Y....YGLEHGSRL-L-Q-KF.EGQ-GRMFNNLGKM-K-LNAEAMNYTEGTGTCDSKK.........TTCGRKLKGLPIANMTRH-A-LATEMLMHT--E-M  420
MAC.US.x.251_1A11           GF--TR--NRTY-Y........WHGR.--RTI-SLNKYYNLT-K-R--G-K.-VLPVT-..MSQLV-.....HSQPINDRPK--W-WF.-G--KDAI-EVKQTIVKHPRY...TGTNNT.DKINLTAPGG.......................----VTFMWT--R---  398
MND-1.GA.x.MNDGB1           GF--TMHK.-GELIPIDDKYRGPEE-H.QRKFVYK-PGKYGLK-E-H-KG-R.SVVSTPS..AT-LL-Y....HGLEPGK-LKKGM-TF.KGR-GLA-WSL-KE-NKLNDSI-VNQTCKNFTSTGEENKQNTDKQKEFAKCIKTLKIDNYTT-.--RAAEMMMMT-Q--M  417
MND-2.CM.98.CM16            GSS-TKHDYNEL-QT..NTKKGKEE-H.-HKYVYR-DEKYNLQVV-R-KG-R.SVISTPS..AT-LI-Y....SGLEPGK-LKKGM-QL.KGQ-GIAMHDL-IE--KIDSSIWRNVT.KGYCKGLKRK.......ENRTGCALKTIKVSDYTTK-E-GAE-IMLL----Y  412
MND-2.GA.x.M14              GF--TKHNYDEL-LV..NPQ...KE-H.-HKYVYR-NKKWGLKVV-D-KG-R.SIVSVPS..AS-LI-Y....HGLEPGR-LKKGM-RL.IGQ-GRA-NALSRE-KKVNASIYRGFNASGPCRGINNR.....GVVNKTGCALKTIEVSNYTTI-E-GAE-IMIL----Y  412
MND-2.x.x.5440              GF--TKHDYDEL-QT..NPRKGKDE-H.-HKYVYR-DKKWGLQVR-R-KG-R.SIISTPS..ST-LL-Y....HGLEPGK-L-KGK-QL.EG--GQA-HSLSLE--KINDSIYKDNH.NMTCKSSNNK.......KNTTGCHLKTISISESTVK-E-GAE-IMLL----Y  433
MNE.US.x.MNE027             GF--TR--NRTY-Y........WHSK.--RTI-SLNKYNNLTMK-R--G-K.-VLPVT-..MS-LV-.....HSQPINERPK--W-RF.EGN-KEAI-EVKQTIVKHPRY...TGTNNT.DKINLTAPRG.......................----VTFMWT--R---  400
MON.CM.99.L1_99CML1         Q--STGNAPNTTVMM.......NKQK.NESIVVRLAKHLH-N-T-I--G-K.-IRNLQ-..-A-MT-Y....SQLIVG--T-K-Y-KVNKTQ-ETA-QAVHEAVKTEWEKKNNGTN.........................VTTISWRFQPQ--K-VQ--W---Q---  384
MON.NG.x.NG1                Q--ATS-TNXTTVMM.......NGKR.NESIV-RLXKTL--N-T-I--G-K.-LRNLQ-..-A-MT-Y....SQIIVG--T-K-Y-QVNKTQ-DAA-HAAHXAIXXEWESKNNNVT.........................VANITWRYQPK--N-VQ--W---Q---  378
MUS-1.CM.01.CM1239          Q---TYL-GNNTAVM.......MNGDK.NESIG-KFGE-FR-NLT-L--G-K.-IRNLQ-..-A-MT-Y....SQLIVG-DT-R-Y-K-NKTQ-DVA-R-AMLAM-KHWHGRLEKQGE.NITEE...................KIKIRWTSQPK--N-VQ--W---A---  401
MUS-1.CM.01.SIVmus_01CM1085 Q---TYL-GNNTAVM.......MNGDK.NESIG-KFGE-FR-NLT-I--G-K.-IRNLQ-..-A-MT-Y....SQLIVD-DT-R-Y-K-NKTQ-DPA-R-AMLAM-KHWHGRLQKQNHSSINET...................KIKIRWRSEPK--Q-VQ--W-------  407
MUS-2.CM.01.CM1246          Q---TYLPGNDTAVM.......MNGDK.NESIA-KFGE-FR-NLT-I--G-K.-IRNLQ-..-A-MT-Y....SQLIVD-DT-R-Y-K-NKTQ-DPA-R-AMIAM-DHWH-RLQKANK.TINKT...................EIKIRWTSEPK--R-VQ--W---Q---  405
MUS-2.CM.01.CM2500          Q---TYLPGNDTAVM.......MNGKK.NESIA-QFGE-FR-NLT-L--G-K.-IRNLQ-..-A-MS-Y....SQLIVD-DT-R-Y-K-NKTQ-DPA-R-AMKAM-AHWA.KMD.....NRSES...................DIKIRWTSEPK--R-VQ--W-------  388
MUS-3.GA.09.09GabOI81       Q---T-QPGNDTAVV.......MNGKK.NESVV-RLGE-YK-NLT-I--G-K.-IRNLQ-..-A-MT-Y....SQLIVG-DT-R-Y-K-DKTQ-DAAIR-SMIAM-DHWI.KINKNDSTHNYTK..................ENITIRWISQPK----VQ--W-A-D---  402
MUS-3.GA.11.11GabPts02      Q---TYQ-GNNTAVM.......MNGKK.NESIVVRLGE-HR-QLT-I--G-K.-IRNLQ-..-A-MS-Y....SQLIVD-DT-R-Y-K-DKD--D-AIR-VTLAV-DHWV.KINNGTNA......................EKVTMRWTSQPR--R-VE--W-A-D---  404
OLC.CI.97.97CI12            GF--TRHKASRL-PI.TGDLNHKPREG.TGVYVMAIADKYKL--E-I-KG-R.SIVSVNS..AS-LVYY....AGLEPY--IDKGL-RF.KGQ-GWM-YALGKG-EKENNTWANYTGVNITGIGR.....NSTTFAKAIKR...FKFDE-LKR---DAAENMMMM---YM  413
RCM.CM.00.SIVagi_00CM312    G----IS-NRTW-Y........QRKQ.S-RTI-GLNSFYNLTVS-R--S-R.-I-G-SL..AT-.V-.....ISL-VEKRPKG-W-RF.YGN-KGAW-EVKE-VIQTK-YKG-NDT...........................TKINIRSVY---D-TKYFWL--D---  385
RCM.CM.02.SIVrcm_02CM8081   G----IS-NRTW-Y........QRKQ.S-RTV-GLNREYNLTVE-R--S-RTVKGIS.L..AT-.V-.....ISLRVEKRPKG-W-RF.YGN-TGAWEEVKKEVIKTK-YKG-NDT...........................KKIYIRSHY---D-AKYFWL--D---  394
RCM.GA.x.SIVRCMGAB1         G----INVN-TW-Y........QRRQ.S-RTV-GLNSFYNLSVT-R--S-R.-V-G-SL..AT-.V-.....ISLRVEKRPKG-W-RF.EGN-TDAW-EVKERVKTTK-YRG-SNT...........................DKIKIRTVY---D-ARYFWL--N---  396
RCM.NG.x.NG411              G----IS-NRTW-Y........QRKQ.S-RTV-GLNRHFNLTVT-R--S-RTVKGIS.L..AT-.V-.....ISLRVQKRPKG-W-RF.DGD-KGAW-EVKEAINKTK-YKG-RQP...........................ENITIRSVY---D-AKCFWL--D---  390
SAB.SN.x.SAB1               GI---YVANRTE-W.......QKNGNS.NDSVI-RLNRYFNLT-R-R--G-K.-VLPVT-..MA-LV-.....HSQ-YNTRLK--W-WF.GGN-RGAW-EVKETIVRLPPKKYSGTN.........................DTNKI-L-RQW----SEFFF---Q---  396
SMM.CI.79.SIVsmCI2          GF--TR--NRTY-Y........WHGR.--RTI-SLNKYYNLTMR-R--G-K.-VLPVT-..MS-LV-.....HSQPINERPK--W-WF.GG--KEA--EVKET-VKHPRY..KGGTNDT.RKINFTEPGRK.....................S---VAFMWT--R---  409
SMM.LR.89.SIVsmLIB1         GF--TR--NRTY-Y........WHGR.--RTI-SLNKYYNLT-R-R--G-K.-VLPVT-..MS-LV-.....HSQPINERPK--W-WF.GGS-RKAIQEVKET-VKHPRY...TGTNDT.NKINLTAPGG.......................----VTFMWT--R---  396
SMM.SL.92.SIVsmSL92A        GF--TR--NRTY-Y........WHGR.S-RTI-SLNKTYGL-LH-R--G-K.-VLPVT-..MS-LV-.....HSQPINERPK--W-WF.KGE-KKAIREVKETIVQHPRY...TGTNRT.EKIYLTAPGG.......................----VTFMWT--R---  401
SMM.SL.92.SL92B             GF--TR--NRTY-Y........WHGG.S-RTI-SLNKYYNLT-K-R--G-K.-VLPVT-..MS-LV-.....HSQPINERP---W-WF.GG--REAMQEVKKTIVKHPRY...TGTNDT.RKINLTAPGG.......................----VTFMWT--R---  402
SMM.US.04.G078              GF--TR--NRTY-Y........WHGR.--RTI-SLNKFYNLT-R-R--G-K.-VLPVT-..MS-LV-.....HSQPINERPK--W-WF.GGS-KKAIQEVKETIVKHPRY...TGTNDT.EKINLTAPRG.......................----VTFMWT--R---  406
SMM.US.05.D215              GF--TR--NRTY-Y........WHGK.S-RTI-SLNKYYNLT-R-R--G-K.-VLP-T-..MS-LV-.....HAQPINERP---W-WF.GG--KQAI-EVKET-VKHPRY...TGTNDT.SKIKLTKPGG.......................----VTFMWT--R---  400
SMM.US.06.FTq               GF--TR--NRTY-Y........WHGR.--RTI-SLNKYYNLT-K-R--G-K.-VMPVT-..MS-LV-.....HSQPINDRP---W-WF.GG--KSAMQEVKQT-VKHPRY...TGTNDTKNINFTEPGKT.......................S---VAFMWT--R---  401
SMM.US.86.CFU212            GF--TR--NRTY-Y........WHGH.R-STI-SLNSYYNLT-R-R--G-K.-VLPVT-..MS-LV-.....HSQPINERP---W-WF.GG--KSAI-EVKET-VKHPRY...TGTNDT.SKINLTAPGG.......................----VTFMWT--R---  402
SMM.US.x.H9                 GF--TR--NRTY-Y........WHXR.S-RTI-SLNKYXNLTMR-R--G-K.-VLPVT-..MS-LV-.....HSQPINERPK--W-WF.GGE-KEAIREVKXX-VKHPRY...KGTNXX.XQIKLTAPGG.......................----VTFMWT--R---  402
SMM.US.x.PGM53              GF--TR--NRTY-Y........WHSR.S-RTI-SLNKYYNLT-R-R--G-K.-VLPVT-..MS-LI-.....HSQPINERPK--W-WF.GGS-REAIQEVKETIVKHPRY...TGTNET.KKINLTAPGG.......................----VTFMWT--R---  404
SMM.x.x.pE660.CG7G          GF--TR--NRTY-Y........WHGK.S-RTI-SLNKYYNLTMR-R--G-K.-VLPVT-..MS-LV-.....HSQPINERPK--W-WF.GGS-KEAIQEVKET-VKHPRY...TGTNDT.KKINLTAPAG.......................----VTFMWT--R---  401
STM.US.89.STM_37_16         GF--TR--NRTY-Y........WHGR.--RTI-SLNKYYNLTMS-R--G-K.-VLPVT-..MS-LV-.....HSQPINERPK--W-WF.GGE-RGAI-EVKET-VKHPRY...TGTNDT.AKIRIVAPGG.......................----VTFMWT--R---  398
SUN.GA.98.L14               GF--TRHK-DEL-PI....NNKVGQGA.EGEYVWK-AAKWGLV-Q-I-KG-R.SQVSTIS..ST-LL-Y....YGLEPGSKL-L-Q-KF.VGQ-GRAFATLGKL--QVEPLANMSNG..CTFDNITKT....CRFTNGTDFKKMIKFKPMEKH-A-AATEMLMMT--E-M  426
SYK.KE.x.KE51               Q---TYKAKDK-RFI.......RQKDK.NESVI-L-PEKLGLQLV-E--G-E.SI-N-QL..AA-FFLPVI.QGRL-TGKAAKR-F-KV.KGE-GKFFE-VHNESIKVWK-..............................VTSTSWRSQPQ--L-VR--W-Q-----  385
SYK.KE.x.SYK173_COMGNM      Q---TYKAKDK-RFI.......KQKDK.NESVI-L-PEALRLQ-I-E--G-E.SI-N-QL..AA-YFLPVI.QGKL-TGRDAKR-F-RV.TGN-TEFF--VHEQATKTWK-..............................VTNTTWRSQP---L-VR--W-Q-----  378
TAL.CM.00.266               I---TK.-NHTQVI........HQAPK.NESAI-LLGKEHNLT-T-I--G-R.-I-DLQ-..AA-LM-H....SQIIAGK-LNR-Y-KL.EGN-TKAMESANEEIAKRYKEHINNTD.........................KVNITWTAVRR--L-VQLFWHA-Q---  378
TAL.CM.01.8023              V---TK.-NQTQ-I........HQAPK.NESAI-LLGKEHNIT-Q-I--G-R.-I-DLQ-..AA-LM-H....SQIIAGKDLKR-Y-K-.EGN-TRAMEDVNE-IAKAYKDHINSSS........................GTAQIRWQGVRK--L-VSLFWHA-Q---  395
TAN.UG.x.SIVagmTAN1         GI---Q--NRTE-W.......QKHGVS.N-SVI-KLNKHYKLK-V-R--G-K.-VLPVT-..MA-LV-.....HSQQYNTKL---W-HF.QGD-KGAWREVRKTIV-LPKEKYRGTN.........................NTRQIWLSRQW---AANIWL--Q---  397
VER.DE.x.AGM3               -----YS-NRTQ-W.......QKHRVS.NDSVLVLFNKHYNLTVT-K--G-K.-VLPVT-..MA-LV-.....HSQRYNTRL---W-HF.QGN-RGAW-EVKNEIVKLPKDRYQGTN.........................DTEEIYL-RLF----AANLW---Q---  406
VER.KE.x.9063               -----YS-NRTQ-W.......QKHRVN.NDSVIVLFNKYYNLTVT-K--G-K.-VLPVT-..MA-LV-.....HSQ-YNTKL---W-HF.QGR-REAW-EVKEEIIKLPKDRYRGTN.........................DTEKIYL-RQF----AANLW---Q---  406
VER.KE.x.AGM155             -----YS-NRTQ-W.......QKHGVS.NDSVL-LLNKHYNLTVT-K--G-K.-VLPVT-..MA-LV-.....HSQ-YNTRL---W-HF.QGN-KGAW-EVQEEIVKLPKERYQGTN.........................DTNKI-L-RQF----AANLW---Q---  401
VER.KE.x.TYO1_patent        -----YH-NRTQ-W.......QKHRVN.N-TVL-LFNKHYNLSVT-R--G-K.-VLPVT-..MA-LV-.....HSQ-YNMKL---W-HF.EGN-RGAWREVKQ-IV-LPKDRYKGTN.........................NTEHIYL-RQW----ASNLW---Q---  399
WRC.CI.97.97CI14            GF--TRHK-GEL-EV.IGNINHKPKEG.SG-FVYK-PDKYN-K-E-H-KG-R.SLVSVNS..AS-LIYY....AGLEPHR-I-KGL-KF.TGL-GKM-YDLGKN-EKYNSSWVNYTEGDCNSKN......NSADFAKCIKR...FQIRNFTRQ-A-RATENIMMI----M  415
WRC.CI.98.98CI04            GF--TRYKKGELVEI.IGNINHRPHEG.TG-FVYK-PSRYK---E-N-KG-R.SVVSVNS..AS-LIYY....AGLEPHR-I-KGL-KF.KGL-GHM-YDLGKN-Q-YNNTWVNYTE.DCNQKN......DSREFAKCIKQ...FQIRNFTRK-A-RATENIMMI----M  425
WRC.GM.05.Pbt_05GM_X02      GF--TKYDKGEL-EV.IGNINHKPKDG.TG-FVYK-PAKYN-A-E-H-KG-R.SLVSVSS..AS-LIYY....AGLEPYK-I-K-L-KF.TGQ-GHM-YGLGKE-QAYNSSWVNYTE.ECNKKG......NSSAFATCIKQ...FQIRNFTKQ-A-RATENLMMI----M  417
```

Draft of 30 September 2015, not for public release
plvprot_main.tex r8876–10546

Draft of 30 September 2015, not for public release
plvprot_main.tex r8876–10546

fusion peptide

glycosylation NAS

glycosylation NHT

immunodominant region

glycosylation NKS

glycosylation NYT

```
H1B.FR.83.HXB2          LGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLL.RAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKS.............LEQIWNHTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWAS.LWNWFNITN  677
H1O.CM.91.MVP5180       -S---------ATA---RTHSV-K------D---.---Q------R-S----R--R--LQ-L-TLIQN--R-NL---K-----Y-S-K--T---GRY...........NDDS--DNL--QQ--QH---VS-I-YDE-QAA-D-----VKA-----E---.-----D--K  686
H1N.CM.95.YBF30         ----------I-------T----------I-.----------SI---------KV--I----R----I-SL-----T--Y-T-----ET---NTS...........YDT---NL--QQ--EKVR--SGV-FG---QA-E--NT--KS-------Q-D-.--S--G--K  666
H1P.CM.06.U14788        -S---------GI------K---H--------M-.-------E--R-S----R--R--L--I-T--R------L----QIX-Y-N----ST-T--N..........ETE-DG--GNL--Q---KLVD---DT-YLE-QKA-E--KE--RK-------Q.--S-MD--K  698
CPZ.CD.06.BF1167        -S---------IA-----G---------A---Q----------S----------L----K--R------L---AN----HSSLE--S--L--TLCGNKTNNYENNYEC.--SQY--QQ--E--R-ISDV-YLSL-TA-V---R-HKQ-------S-------D--  691
CPZ.TZ.00.TAN1          -S---------IA--A---G---------Q----Q----------S----------M----K-IR-----SL---AN--V-HSS----LT-AED-TKCNHSD...AKYYDC---NL--Q----LVE-S-GT-Y--L-KA-T-----K---------S----D--D--Q  678
GOR.CM.07.SIVgorCP2135con MS---------AV--------V-H------X-M-.---A---E--R-S----R--R--L--I-T--R------L----T-R---Y-N----KT-TG-N..........ETE-DD--GNM--QQ--KLVD---DT-FLE-QKA-E---V--KA--------D.--S-LD--Q  663

MAC.US.x.239            -AT---A-----L---A-S-T--A------QQ---.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NA---AFRQV-H-T---PNAS.................-TPK--NE--Q--E-KVDFLEEN-TA-L--A-I-----MY--QK-NS-DV.FG---DLAS  686
H2A.SN.86.ST_JSP4_27    -TT--AA-----L---SA-S-T--A------QQ---.DVVKR--EM-R-----T-N----VT-I-K-----AQ-NS---AFRQV-H-T---VNDT.................-TPD--NM--Q--EQR-R-LEAN-SESL-QA-I-----MY--QK-NS-DV.FG---DL-S  666
H2B.CI.x.EHO            -AT---A-----L---SA-S-T--A------QQ-V.DVVKR--E--R-----T-N----VT-I-K-----AQ-NS---AFRQV-H-T---VNES.................-KPD--NM--QQ-E-QVRFLDAN-TK-L--A-I-----MY--QK-NQ-DI.FS---DF-S  670
H2G.CI.92.Abt96         -AM---A--T--L---SA-S-T--A------QQ---.DVVKR--EM-R-----T-N--T-VT-I-K-----AR-NS---AFRQV-H-T---D-LGA--T.............--PQ--NM--Q--EKQ--FLEDN-TR-L--A-I-----MY--QK-NS-DV.FG---DL-S  691
H2U.CI.07.07IC_TNP3     -AT--TA--T--L---SA-S-T--A------QQ---.DVVKR--E--R-----T-N--T-VT-I-K-----AS-NA---AFRQV-H-T---VNNS.................-QPD--NM--Q--EVKVKDLEEN-TV-L--A-I-----MYQ-QK-NS-DV.FG--LDL-S  691
H2U.FR.96.12034         -AT---A-----L---SA-S-T--A------QQ---.DVVKR--EM-R-----T-N--T-VT-I-K-----AS-NA---AFRQV-H-T---INDT.................-TPN-DNM--Q--EEKV-YLEEN-TQ-L-AA-I-----MY--QK-NN-DI.FG---DL-S  681
ASC.UG.10.RT03          -ST--A---SVATA----S-S--E---A--KA---.E-V-H--Q----S----RN-N--LT-I-K-V---GI-ASY--QF-Q--Y-S----K-FA.............DGA-PAWH-.M---I--E-KVT-H-GI-NT-LV-A-RR--E-THK-QK-GE-DN.-----D-SK  693
ASC.UG.10.RT08          -ST--A---SVATA----S-S--E---A--KA---.E-V-H--Q----S----RN-N--LT-I-K-V---GI-ASY--QF-Q--Y-S----K-FA.............DGEALPE---NM--I--E-KVT-H-GI-NG-LV-A-RR--E-THK-QK-GE-DN.--S--D-SR  697
ASC.UG.10.RT11          -ST--A---SVATA----S-S--E---A--KA---.E-V-H--Q----S----RN-N--LT-I-K-V---GI-ASY--HF-Q--Y-S----I-FA.............DGGVLPD---NM--I--E-KVT-H-GI-NN-LV-A-RR--E-THK-QK-GE-DN.--S--D-SR  688
COL.CM.x.CGU1           -SG--AA--S--VA--I--QS-NGRASASS-RM-LKLV-T-SA--------V-N--V-VATI-G--EE-AK-ASI--ANMQ--R-I----KT-...............GE-DP-QNM--KQ-HERVR---DI-EADLV-AYDL--E--KK-A--GD-TN.WFSG-GLF-  673
COL.UG.10.BWC01         -SG--AA--S--VA--I--QS--E-L----RM---.KLV-T-SA--------V-N--V-VATI-S--EE-AK-ASI---NMQ--R-I----KT-...............GM-EP-QNY--KQ-YEQVA---NI-EADLV-AYDL--A--KK-A--GD-TN.WFSG-GLF-  679
COL.UG.10.BWC07         -SG--AA--S--VA--I--QS--D-L-R--RQ---.KLV-T-SA--------V-N--V-VATI----EE-AK-AVI--ANLQV-R-----ST-...............GD--P-MNM--KQ-HERVK---DV-QEELQ-AYEL--E--KK-A--GD-T-.WFSG-G-FS  665
DEB.CM.04.SIVdeb04CMPF3061 -ST--AA-----TA----S-S---------QE---.K-V--HG---T--A--V-N-NT-LT-L-K-----SK-NE---AF-Q--H-T----G-I.............GQPD--NM--Q--EQKVA---DIWEEALQKA-EK--Q-VHA-QS-QD-D------DLSR  656
DEB.CM.99.CM40          -ST--AA-----TA----S-S---------QE---.K-V--HGQ--T--A--VRN-NT-LT-L-K-----AK-NE---AF-Q--H-T----N-L.............EDPD-DNM--Q--EMKVA---DEWEGALQRA-E---R-VHA-QS-QD-D-.-----DLSR  679
DEB.CM.99.CM5           -ST--AA-----TA----S-S---------QE---.K-V--HG---S--A--VRN-NT-LT-L-K-----SK-NE---AF-Q--H-T----HT-...............G-PD--NM--Q--E-KVA---DEWEGALQRA-E---R-VHA-QS-TD-D------DLSR  659
DEN.CD.x.CD1_CM0580407  -S---TV---VGTG-A--S-D--A------QE-----.V-GHSA----S-----N-N--LT-I-K-----A--NE---AW-Q--H-S-E-SNQS..........IPGLALYQPDFQNM--QQ-ETD-AALFGNVSDALKKA-A---R-VH-VQS-YD-DN.-----DLSK  686
DRL.DE.11.D3            -SL-----SVATAM---SQA-----E--KQ---.LV-Q--E--K--I--V-N----LT-L-E-VG--AR-SL----FAQV-H-N-V-PNES.................VTPN-TSE-----QKRVDSISNN-TLDLQKAYE-EQ--IF--QK-GDLT-.WA---DF-W  726
DRL.x.x.FAO             -SL---A-CSVATAM---SQA--T-M-E--KQ---.LV-Q--E--K--I--V-N----LT-L-E-IG--AM-SL----FAQV-H-N-V-PNES.................VTPN-TSE-----QKRVDSISNN-TLDLQKAYE-EQ--IF--QK-GDLT-.WA---DF-W  722
GRV.ET.x.GRI_677_gri_1  -----TA----AT-----S-H--A--L---K----.A-V-Q--Q--K--I--V-N-N--VT-L-K--E--AR-NS---AW-QV-H-T---KYNN.............TPK-DNM--L--E-Q--ALEGN-TQ-L--A---ES--LDLYQK--D-SG.F-S--SLST  682
GSN.CM.99.CN166         -S---T-----ATA--E-S-S--A--M---E----.--V----S---PS--------T-LSSL-K--R--TI-QA---ANRP--H-I----T-.............WANGSLPD-ENM--QK-SMLVE-D-YT-QQ-L-QANQ--AS-LN--MK-S--D-.--S--D-SD  688
GSN.CM.99.CN71          -S---A-----ATA----S-S--A------E----.--V----S---S---------LSSL-K--R--TI-QA---ANQP--H-I----D-.............WAKNSTPD-E-M--Q--SKL-E-D-YT-QQ-L-NANH--S--MND--K-S--D-.--S--D-S-  697
LST.CD.88.SIVlhoest447  -S---TA--LV-TI-----QVVIQ--L---KQ---.VLV-K--E--R--I--V-N----LT-I-E-----T--AS---QW-QV-H-N-E--YN.............ITPN-TRD--I----QVGVLEAN-ST-LQ-AYTTELE-RNAFKK-QEFN..F--LD-LS  710
LST.CD.88.SIVlhoest485  -S---TA--LV-TI-----QVVIQ--L---KQ---.VLV-K--E--R--I--V-N----LT-I-E-----A--AS---QW-QV-H-N-E--YN.............ITPN-TRD--I----RVGVLEAN-ST-LQ-AYTTELE-RNAFKK-QEFN..F-S-LD-LS  709
LST.CD.88.SIVlhoest524  -S---TA--LV-TI-----QVVIQ--L---KQ---.VLV-K--E--R--I--V-N----LT-I-E-----A--AS---QW-QV-H-N----YN.............VTPN-TRD--I--E-QVGSLEAN-TT-LQ-AYTTELE-RNNFKK-QDFN..F-S-MDL-T  712
LST.KE.x.lho7           -S---TA--LV-TI-----QAV-Q--L---KQ---.VLV-K--E--R--I--V-N----LT-L-E-V-H-A--AS---QW-QV-H-N-E-TYN.............ITPN-TKD--R--ESKVAI-DKN-T--LQ-AYTTELE-QNKFKK-QEFN..F-S-LD-SH  707
MAC.US.x.251_1A11       -AT---A-----L---A-S-T--A------QL---.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NA---AFRQV-H-T---PNAS.................-TPK--NE--Q--E-KVDFLEEN-TA-L--A-I-----MY--QK-NS-DV.FG---DLAS  685
MND-1.GA.x.MNDGB1       -SL---A--SV-VA----SQS-VT---E--KQ---.KL--Q-SE--K--I--V-N--T-LTSL-N-I---A--SQ----WAQV-H-S-E-TNTS.............ITPN-TSE--K--ETRTDYLQQN-TEMLKQAYDREQR-TY--QK-GDLT-.WAS--DF-W  704
MND-2.CM.98.CM16        --L------SVAVA----SQA--N---E--KV---.SL-DQHSE--K--I--V-N----LT-L-D-VA--AR-SM----FAQV-H-H---PNDS.............ITPN-TSE--L---KRVTAL-DNMTVNLQKAYELEQ--IY--EK-GD-T-.WAS--DF-W  695
MND-2.GA.x.M14          --L---A--SVAVA----SQA--N---E--KI---.SL-DQHSE--K--I--V-N----LT-L-D-VA--SR-AV----FSQV-H-N---PNES.............ITPN-TSE--L----RVTAI-NNMTIDLQRAYELEQ--MY--QK-GDLT-.WAS--DL-W  696
MND-2.x.x.5440          -L---A--SVAVA----SQT--N---E--KV---.SL-DQHSE--K--I--V-N--V-LT-L-E-VA--SR-SV----FSQV-H-S-K-PNNS.............ITPN-TSE--L----RV-SIVTNMTIDLQRAYELEQR-IF--QK-GDLNFHGLTG-DL-W  719
MNE.US.x.MNE027         -AT---A-----L---A-S-T--A------QQ---.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NA---AFRQV-H-T---PNAS.................-TPN--NE--Q--E-KVDFLEEN-TA-L--A-I-----MY--QK-NS-DV.FG---DLAS  689
MON.CM.99.L1_99CML1     -ST--G-----ATA----S-S--A------E----.--VT---S-------V-------LT---KFI---T--NA---AN-AV-H-T----N-.............WAKGHFPE-DNM--QQ-SELVD-D-MT-QQ-L-AA-E--G--QH--MKPGQ-DF.------D-SK  679
MON.NG.x.NG1            -ST--G-----ATA----S-S--A------E----.--VT---S----S---------LT---KFI---T--NS---ANRAV-H-Q-L--NT.............WAKGHFPE-DNM--QQ-SMLVD-D-A--QX-L--A-E--G--AH--MK-GQ-DW.-----D-SK  676
MUS-1.CM.01.CM1239      -S---G-----ATA----S-S--A------A----.--V---------S---------LT-L-KFI---A--NL---ANRQ--H-R----T-.............WANHTLPG-ENM--QQ-SKLVD-D-NT-QG-L-AA-K---E-QHK-QK-LE-D-.---E--D-SK  694
MUS-1.CM.01.SIVmus_01CM1085 -S---G-----ATA----S-S--A------A----.--V---------S---------LT-L-KFI---A--NL---ANRQ--H-R----D-.............WANHTQPG-ENM--QQ-S-LVD-D-TT-QE-L-LA-R---E-QHK-QK-LE-D-.---E--D-SK  700
MUS-2.CM.01.CM1246      -S---G-----ATA----S-S--A------A----.--V---------S---------LT-L-KFV---AI-NL---ANRQ--H-S----D-.............WANHTQPG-DNM--QQ-S-LVD-D-AV--G-L-MA-Q---E-QHK-QK-LE-D-.---E--D-SK  696
MUS-2.CM.01.CM2500      -S---G-----ATA----S-S--A------A----.--V-T-------S---------LT-L-KFI---A--NL---ANRQ--H-R----N-.............WANHTQPE-DNM--Q--S-LVD-D-AT-QG-L-LA-K---E-QHK-QK-LE-DA.---E--D-SK  682
MUS-3.GA.09.09GabOI81   -S---G---S-AIA----S-S--A------Q----.--VQT------S---------MT-I-K--R--T--NS---ANRA--H-NL---K-.............WANETMPG-DNM--Q--S-L-D-D-V--SN-L--AEK---R-NHD-MK-AE-D-.---S--D-SH  695
MUS-3.GA.11.11GabPts02  -S---G---S-AIA----S-S--A------Q----.--VQT------MS--------LT-I-K-----T--N----ANRA--H-N----------.............WANHTQPG-DNM--Q--S-LVD-D-LE-ER-L-LANK---E-NHK-MK-AE-D-.---S--D-SH  696
OLC.CI.97.97CI12        --L---A--TV-TI----SSH--A---R--KHM-.EM--K--E--K--I--V-N----L-G-K-----A--KQ---EFMQV-H-T----YT.............QSPD-GND--LK-Q-KVD-L-QE-DDRLTIAYQE-QR-QY--QRFQELSD.WTK---LA-  692
RCM.CM.00.SIVagi_00CM312 -S---TA----ATA----S-H--A--L---KQ---.DIV-Q-----K-----T-N----VT-I-K--A--S--NTF--AWRQV-H-S---SF.............NKTPD-QKQ--L--E-NVSYLEAN-TASLQ-A-D-H---VH--EK-NN-GD.AFS-LRLDW  665
RCM.CM.02.SIVrcm_02CM8081 -S---TA----ATA----S-H--A--L---KK---.DIV-Q-----R-----T-N----VT-I-K--A--S--NTF--AWRQV-H-T-Q-TF.............NATPE--KQ--L--E-N-SIIEGN-SLALQ-A-D-H-R-VH--EK-NN-GD.AFS-LRLDW  674
RCM.GA.x.SIVRCMGAB1     -S---TA----ATA----S-H--A--L---K----.DIVKR--N--K-----T-N----VT-I-K--A--S--NTF--AWRQV-H-V---TF.............NKTPE-QKES-LQ-E-N-SYLEAN-TIALQ-A-D-H---VH--EK-SN-GD.AFS-L-LDW  676
RCM.NG.x.NG411          -S---TA----ATA----S-H--A--L---KK---.DIV-Q--E--K-----T-N----VT-I-K--A--S--NTF--AWRQV-H-T-E-IY-Q.............TPE--KQ--L--E-N-SRLEGN-SVALQDA-E-H-R-VHD-EK-NS-GD.MLS-L-MDW  670
SAB.SN.x.SAB1           -----AA----ATA----SQ---A--L---K----.A-V-Q--QM-K--I--V-N-N--VT-L-K--E--AR-N----AFRQV-H-T-L-KYNN.............TPD-ENM--Q--E-Q-EK-EAN-SRIL-QAHE-EQ--LDSYQK-VS-SD.F-S--DL-K  687
SMM.CI.79.SIVsmCI2      -AT---A-----L--SA-S-T--A------QQ---.DVVTR--E--R-A---T-N--T-VT-I-K-----AR-NA---AFRQV-H-T---ENNT.............-QPD--NM--Q--EIK-RDLEAN-SESL--A-I-----MY--QK-NS-DV.FG---DF-S  696
SMM.LR.89.SIVsmLIB1     -AT---A-----L---A-S-T--A------QQ---.DVVKR--E--R-----T-N----VT-I-K-----AQ-NS---AFRQV-H-T-E-PNKT.............-QPN-ENM--Q--E-KVDFLEAN-TQ-L--A-I-----MY--QK-NS-DV.FG---DL-S  685
SMM.SL.92.SIVsmSL92A    -AT---A-----L---A-S-T--A------RQ---.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NS---AFRQV-H-T---PNDS.............-TPD--NM--Q--E-KVDFLEAN-TLML--A-I-----MY--QK-NS-DV.FG---DLAS  692
SMM.SL.92.SL92B         -AT---A-----L---A-S-T--A------QQ---.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NS---AFRQV-H-T-L-PNDS.............-VPD--NM--Q--E-KKKVEFLEAN-TQML--ARL-----MY--QK-NS-DV.FG---DL-S  693
SMM.US.04.G078          -AT---A-----L--SA-S-T--A------QQ---.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NS---AFRQV-H-T---PNDS.............-TPN--NM--Q--E-KVDYLEAN-TEML--A-I-----MY--QK-NS-SV.FG---DLAS  697
SMM.US.05.D215          -AT---A-----L--SA-S-T--A------QQ---.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NS---AFRQV-H-T---PNDT.............-TPN--NM--Q--E-QVDFLEAN-TQML--A-I-----MY--QK-NS-DV.FG---DL-S  691
SMM.US.06.FTq           -AT---A-----L--SA-S-T--A------QQ---.DVVKR--E--R-----T-N--T-VT-I-K--Q--AK-NS---AFRQV-H-T---VNDS.............-QPD--NM--Q--EKQVAYLEAN-TQQL--A-IL----MYD-QK-NS-DV.FG---DL-S  688
SMM.US.86.CFU212        -AT---A-----L---A-S-T--A------QQ---.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NS---AFRQV-H-T-E-P-AN.............VTPD--NM--Q--E-KVDYLEAN-TL-L--A-I-----MY--QK-NS-DV.FG---DL-S  694
SMM.US.x.H9             -AT---A-----L--SA-S-T--A------QQ---.DVVKR--E--R-----T-N--T-VT-I-K----APXNS---AFRQV-H-T---PNDT.............-TPX--NMX-Q--EKQV-FLEAN-TXXL--A-I--X-MY--QK-NXXDX.FG--XDL-X  693
SMM.US.x.PGM53          -AT---A-----V-RSA-S-T--A------QQ---.DVVKR--E--R-----T-N--T-VT-I-K-R---AQ-NS---AFRQV-H-T---PNAS.............-VPN--NM--Q--E-QVDDLEAN-TQAL--A-I-----MY--QK-NS-DI.FG---DL-S  695
SMM.x.x.pE660.CG7G      FAT--A-----L--SA-S-T--A------QQ---.DVVKR--E--R-----T-N--T-VT-I-K-----AH-NS---AFRQV-H-T---PNET.............-VPN-SNM--Q--E-QVDFLEAN-TQ-L--A-I-----MY--QK-NS-DI.FG---DL-S  693
STM.US.89.STM_37_16     -AT---A-----L---A-S-T--T------QQ---.DVVKR--E--R-----T-N--T-VT-I-K-----AQ-NS---AFRQV-H-T---PNDS.............-VPD--NM--Q--E-KVDFLEAN-TQ-L--A-V-----MY--QK-NS-DV.FG---DL-S  689
SUN.GA.98.L14           -S---TA--LV-TI----QAV-Q--L---KQ---.VLV-K--E--R--I--V-N----LT-L-E-VQ--S--AS---QW-QV-H-N----YN.............ITPN-TKD-------QVKM-DDN-TA-LQ-AYVTELE-QNKFKQ-QEFN..F-S-LDLSQ  713
SYK.KE.x.KE51           -S---TA--S-ATA--L-SQT--A------QK---.E-V-------G-----V-N-N--LT-L-T--R--AIMSN---AF-Q--H---T-QQACG-N-...........RCPTPQ-ENM--HT-E-QVD-L-DH-DN-LR-A-E-----VHD-TK-QE-D-.--S--DLSK  674
SYK.KE.x.SYK173_COMGNM  -S---TA--G-ATA--L-SQT--A------QK---.E-V-------G-----V-N-N--LT-L-T--R--AI-SN---AF-Q--H---T-EKACG-N-N.........FCPKPQ-KNM--HR-EQ-VD-L-DH-DG-LR-A-E---R-VHD-TK-QE-D-.---S--DLSK  669
TAL.CM.00.266           -ST--TA---VATA----S-S--------EH---.----H-------------N-N--LT-L-K--E--AK-NS---AW-Q--Y-S----KT.............WSNYTDPQ-QNM--Q--EMKVD-H-G--SQ-LQ-A-E---R-VH--QK-ND-D-.--S---LSA  682
TAL.CM.01.8023          -ST--TA---VATA----S-S--------EH---.----H-------------N-N--LT-L-K--E--AR-NS---AW-Q--Y-S----KT.............WTNSTNPD-QNM--Q--EKLVD-ASDT-TV-LQ-A-E---R-VH--QK-ND-D-.--S---LSA  700
TAN.UG.x.SIVagmTAN1     -----AA---TATA----SQ---A--L---K----.A-V-Q--QM-K--I--V-N-N--VT-L-K--E--TR-NL---AF-QV-H-T---TFNN.............TPD-DNM--Q--ESQ-TALEGN-STTLVKAYE-EQ--MDTYQK-GD-T-.W--I-DVSS  693
VER.DE.x.AGM3           -----TA----ATA----SQH--A--L---K----.A-V----QM-K--I--V-N-N--VT-L-K--E--AR-NA---AW-QV-H-T---QWNN.............RTPD--NM--L--E-Q-SYLEGN-TTQL--ARA-E----LDAYQK-SS-SD.F-S--DFSK  704
VER.KE.x.9063           -----TA----ATA----SQH--A--M---K----.A-V----QM-K--I--V-N-N--VT-L-K--E--AR-NV---AW-QV-H-T---QWNN.............MTPN-QNM--L--E-Q-GELEGN-TEQLVKARE-E----LDAYQR-TS-SN.F-S--DFSK  704
VER.KE.x.AGM155         -----TA----ATA----SQH--A--L---K----.A-VG---QM-K--I--V-N-N--VT-L-K--A--AR-NA---AW-QV-H-T---TWNN.............-TPE--NM--L--EKQ-EGLEGN-TKQL-QARE-E---LDAYQK-SD-S--F-S--DFSK  698
VER.KE.x.TYO1_patent    -----TA----ATA----SQH--A--L---K----.A-V----QM-K--I--V-N-N--VT-L-K--E--AR-NS---AW-QV-H-T-E-PWTN.............RTPD-QNM--L--E-Q-ADLE-N-TRQLVKARE-E---LDAYQK-TS-SD.F-S--DFSK  696
WRC.CI.97.97CI14        --L---A--TV-TV--I-SQH--A------K----.ELV-K--E--R--I--V-N----LS-I-GF-E--LK-KQ---ELTQV-H------Q-.............WNATPN-ENE--QH-YKRTDELQAN-SAILVIAHD-EQA-MF--QK-SD-TA.WGS-LQWF-  696
WRC.CI.98.98CI04        --L---A--TV-TV--I-SQH--A--L---K----.ELV-K--E--R--I--V-N----LS-I-GF-E--LK-KQ---ELTQV-H-N---E-.............WNMTPA--NE--QH-Y-RTDELQAN-SV-LV-AH-TEQ--MF--QK-QD-TT.WGS-LQWF-  706
WRC.GM.05.Pbt_05GM_X02  --L---A--TV-TV--I-SQH--A------K----.ELV-K--E--R--I--V-N----LS---G--E--VK-KQ---ELTQV-H-----EEA.............WNMTPA--NE--QQ-HLRTDTLM-N-SDMLVQAHA-EQS-MF--QK-TD-TQ.WGS-LSWF-  698
```

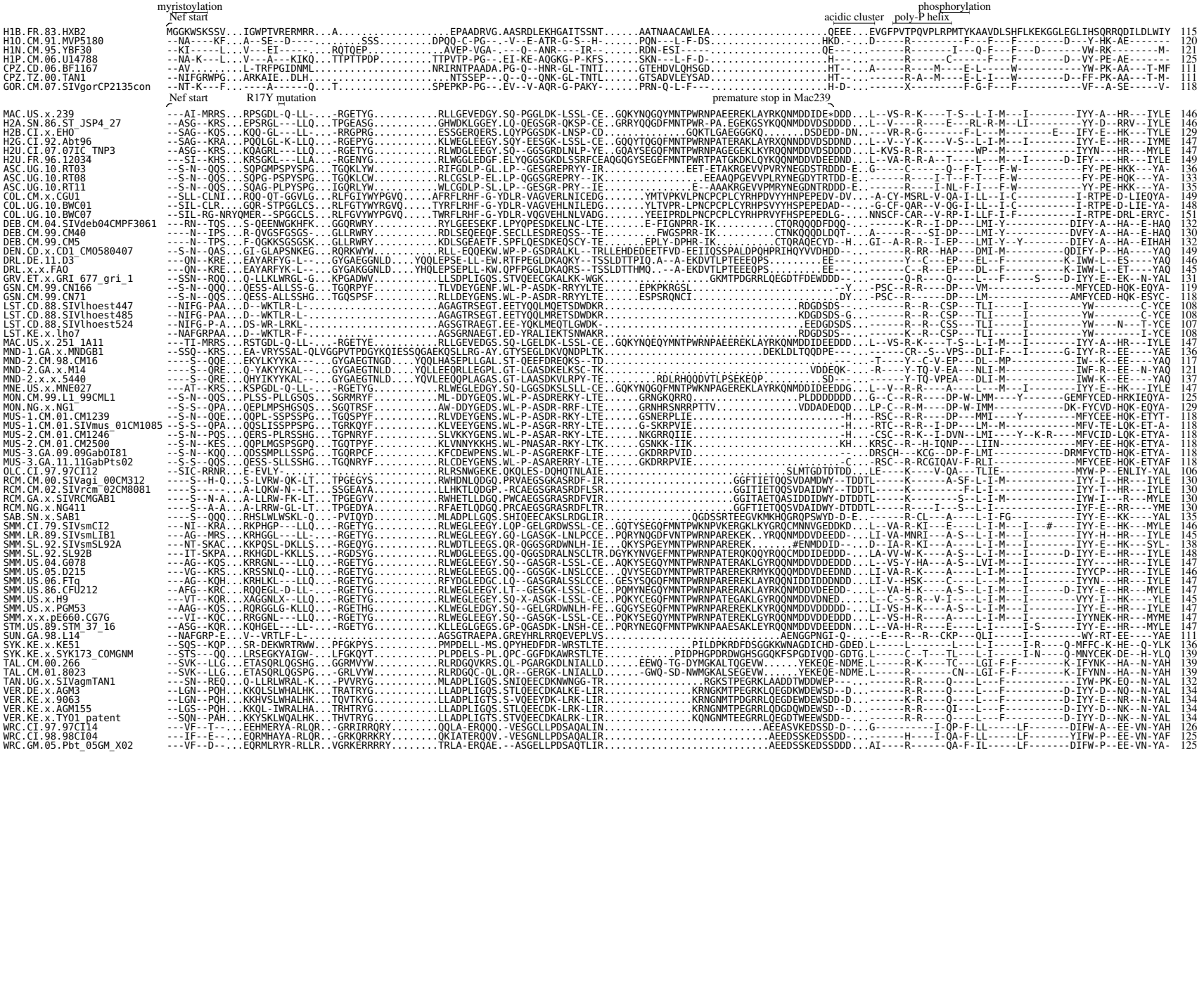
transmembrane domain
gp41 cytoplasmic tail start
glycosylation NGS
glycosylation NAT

Env end
gp41 end
cytoplasmic tail end

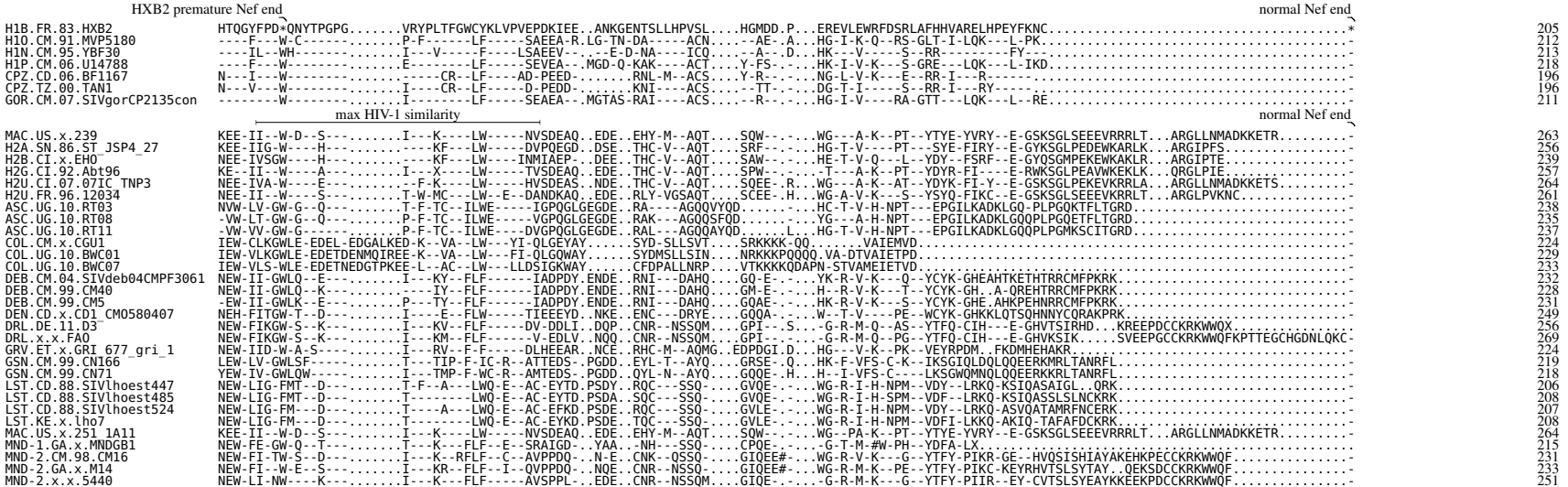
```
H1B.FR.83.HXB2              VAEGTDRVIEVV...QGACRAIRHIPRRIRQGLER..ILL*   856
H1O.CM.91.MVP5180           --NW--GI-LGL...-RIGQGFL---------A--...-V-   873
H1N.CM.95.YBF30             -------I--LA....RIG-G-L-----------..A-I-   853
H1P.CM.06.U14788            --NW--Q--A-G....-QIG-GFLN----L-------..S--   883
CPZ.CD.06.BF1167            TGWW--G---AL...RV-VDI-----T-------I..A-N-   885
CPZ.TZ.00.TAN1              TGWW--LI--G-...VYIA-G--N----------L..A-N-   872
GOR.CM.07.SIVgorCP2135con   T-DW--Q--AIA...-RIG-G-LN-----------..S---   850
```

Env end
gp41 end

```
MAC.US.x.239                L-GAWGDLW-TL...RRGG-W-LA----------L..T---   880
H2A.SN.86.ST_JSP4_27        LTNAWRGFWGTL...GQIG-G-LAV-------A-I..A---   860
H2B.CI.x.EHO                L-SAARTSWG-L...RR-AGE-IA--------A-L..A---   853
H2G.CI.92.Abt96             LTSTWRALWKTL...GRVG-G-LA----------L..T---   885
H2U.CI.07.07IC_TNP3         LGRVGRGILAIP.....------------F-L..A---      882
H2U.FR.96.12034             LSRTGRELW-TL...GRVG-WL-A--------F-L..A---   868
ASC.UG.10.RT03              TWNW-EA-LAYL...RRLAQG-IA--------A-I..L-N-   887
ASC.UG.10.RT08              TWNW-EA-LAAL...RQLAQG-IA--------A-I..L-N-   891
ASC.UG.10.RT11              TWNW-EA-LVAL...RRLAQEFIA--------A-I..L-N-   882
COL.CM.x.CGU1               IL-TLQ-WLRSA...ARGW-RAPEYL-GWIYDRPQ..GPA-   838
COL.UG.10.BWC01             IP-TLS-WLRSA...TRRW-RASEYL-GWLFDRPE..SPA-   843
COL.UG.10.BWC07             IPQTLQ-WLRSA...TS-G-RALESCG-WIR-DPQ..GPA-   835
DEB.CM.04.SIVdeb04CMPF3061  FQRWAEVALQ-L...GRGI-E-LA--A-----A-L..F-N-   847
DEB.CM.99.CM40              LQGIAEMALQGL...VR-GG--LRV-A-----A-L..--N-   871
DEB.CM.99.CM5               LQGIAQVA-QGL...VWGG-NLLA--A-----A-L..L-N-   851
DEN.CD.x.CD1_CMO580407      GRATAEV-LAAL...TR-A-EVVA----------I..V-N-   865
DRL.DE.11.D3                GRQFGLSAARAL...R-LAQGVAN--------A-I..L-N-   900
DRL.x.x.FAO                 GRQLGLSSARAL...RALAQEVAA--------A-V..LFN-   896
GRV.ET.x.GRI_677_gri_1      AR-AWG-LGAI-...RS-Y--VINS---V-----K..V-G-   855
GSN.CM.99.CN166             TWNW-EA-LHAC...RRVW-EFLA--------A-I..LFN-   882
GSN.CM.99.CN71              TWNW-EPLVQT-...GRVW-EFLA--------A-I..L-N-   887
LST.CD.88.SIVlhoest447      DISAAHARNLRL...GQKK-WRFRFRG-SGFPS-TTETA--   916
LST.CD.88.SIVlhoest485      DISAAHARDIRL...GQKK-WRFRFRG-SGFPS-TTETA--   915
LST.CD.88.SIVlhoest524      ELSK-NGTNIRL...GQNR-RW.RFRF-SGLPS-TTETT--   917
LST.KE.x.lho7               QLSS---KNIQL...GKKK-WRLRFGG-SGISS-ATETA--   913
MAC.US.x.251_1A11           L-GAWGDL*-AL...RRGG-W-LA----------L..T---   877
MND-1.GA.x.MNDGB1           IFT-CREA-IAA.....GTC-W--L-----SA---.P-N-    879
MND-2.CM.98.CM16            HQW-LSTAACFR....AIA-G-IN--------A-V..L-N-   867
MND-2.GA.x.M14              HQL-LSVIRGAA....AIG---GN--------A-V..L-N-   868
MND-2.x.x.5440              YQL-LSG-RGAA....AFG-G-WN--------A-A..L-N-   891
MNE.US.x.MNE027             L-GAWGDLW-TL...GRVG-W-LA-------E--L..T---   883
MON.CM.99.L1_99CML1         TIVW--ALLHAG...GRLW---VA--------A-I..F-N-   870
MON.NG.x.NG1                TVIW--A-LRLG...GRLW-GLVAV-------A-I..L-N-   867
MUS-1.CM.01.CM1239          TWNW-EPILQIG...RRVW-EFLA--------A-I..L-N-   892
MUS-1.CM.01.SIVmus_01CM1085 TWNW-ET-LQAG...RRVW-EFLA----V---A-I..L-N-   898
MUS-2.CM.01.CM1246          TWNW-EPILQFS...-KIW-KFLA--------A-I..L-N-   883
MUS-2.CM.01.CM2500          T-NW--A-LQ-S...KQL*K-FLA--K-----A-I..L-N-   879
MUS-3.GA.09_09GabOI81       TWNW-EA-LQ-L...RRVA-EFLA--------A-I..LFN-   893
MUS-3.GA.11.11GabPts02      TWNW-EPLFQAX...RRVW-EFLA--------A-I..L-N-   894
OLC.CI.97.97CI12            ...HGINLQQKK...-RRGSPLQKAQEQLGREGAE.TTGV-   844
RCM.CM.00.SIVagi_00CM312    AYYTWRGLQSLA...RQLAGWPATCG----------..L-N-   855
RCM.CM.02.SIVrcm_02CM8081   AYFVWRGLCVTA...-DIA-WPGTVC------F---.L-I-   862
RCM.GA.x.SIVRCMGAB1         AYYTWRGLCA-A...RDFAGWPAMVC----------..LCN-   851
RCM.NG.x.NG411              AYYTWRGLCA-C...RDPAGWPATLC------F---.F-N-   858
SAB.SN.x.SAB1               ATSAS-S-RHAC...RSIV--VIAH---M--E---..WFN-   868
SMM.CI.79.SIVsmCI2          L-SAGRDLW-TL...GRIG-GLLAA---------L..AF--   887
SMM.LR.89.SIVsmLIB1         L-SAWRDLWQTL...GRIG-GLWA----------L..T---   879
SMM.SL.92.SIVsmSL92A        LVSTGRTIWQTL...GHIG-G-AA----------L..T-N-   886
SMM.SL.92.SL92B             IVGAWGLIW-TL...GRVG-G-AA----------L..M-N-   887
SMM.US.04.G078              L-SAWRDIW-TL...GRVG-GLLA--------F-L..T---   891
SMM.US.05.D215              L-SAWRDLW-TL...GRVG-G-AA--------F-L..TM--   885
SMM.US.06.FTq               L-SAWRDLW-IL...RRVGGWMLA----------L..T---   882
SMM.US.86.CFU212            L-SAWRDLW-TL...GRVG-G-LN----------L..T---   888
SMM.US.x.H9                 L-SAWRDLW-TL...GRVG-G-LA--X--------L..T---   887
SMM.US.x.PGM53              L-SAWRDSW-TL...GRVG-WVLA----------L..A---   886
SMM.x.x.pE660.CG7G          L-SAWRDIW-TL...GRVG-G-LA--------F-L..A---   887
STM.US.89.STM_37_16         L-SAGRDLW-TL...GRVG-R-GA----------L..T---   883
SUN.GA.98.L14               RVSSSQKARSRTFSLGRKW-PKWNRT-GS-IPS-TTETT--   920
SYK.KE.x.KE51               MVAIWQALLAYA...RRVAENVAAL---L-----I..Y-N-   841
SYK.KE.x.SYK173_COMGNM      --AIWQAIYAAT...RRVVE-VAAL---L-----I..Y-N-   833
TAL.CM.00.266               TWWW-NG-LAAA...RW-GE--AA----T-----Y..CFT-   873
TAL.CM.01.8023              TWRW-SG-LAAA...RW-VQ--AA----T-----Y..C-A-   891
TAN.UG.x.SIVagmTAN1         AQSASRTLWNAC...RS-Y---LEH---M--E---..WFN-   880
VER.DE.x.AGM3               AQNAGYQIWLAC...RS-Y---INS---V-----G..--N-   878
VER.KE.x.9063               AQNAGHQIWLAC...RSTY-H-ISS---V-----E..--N-   878
VER.KE.x.AGM155             ARNAAHQIWLAC...RS-Y---INS---V-----E..V-N-   871
VER.KE.x.TYO1_patent        AQNACHQIWLAC...RS-Y-N-VNS---V-----E..--N-   864
WRC.CI.97.97CI14            S-T-SRKTAASS...RIRML-T-LCASFDKCRRSFSERRQ-   879
WRC.CI.98.98CI04            EK-ISENSNRET...T-SS-VWKL-T-LCSSLNKSRRR-F-   885
WRC.GM.05.Pbt_05GM_X02      EK-IYTPGSRET...SRSSSIRGITA-LCTNLNPCRRGQQ-   877
```



Draft of 30 September 2015, not for public release
plvprot_main.tex r8876–10546


HXB2 premature Nef end
normal Nef end
max HIV-1 similarity
normal Nef end