LA-UR-16-25426 (Accepted Manuscript)

Panels of HIV-1 Subtype C Env Reference Strains for Standardized Neutralization Assessments

Hraber, Peter Thomas; Korber, Bette Tina Marie; Rademeyer, Cecilia; Seaman, Michael S; Ochsenbauer, Christina; Kappes, John; deCamp, Allan; Gottardo, Raphael; Edlefsen, Paul; Tang, Haili; Greene, Kelli; Gao, Hongmei; LaBranche, Celia; Mascola, John R; Williamson, Carolyn; Montefiori, David C

To be published in: Journal of Virology

DOI to publisher's version: 10.1128/JVI.00991-17

Permalink to record: http://permalink.lanl.gov/object/view?what=info:lanl-repo/lareport/LA-UR-16-25426

Panels of HIV-1 Subtype C Env Reference Strains for Standardized Neutralization Assessments

Peter Hraber,[a#] Cecilia Rademeyer,[b] Carolyn Williamson,[b] Michael S. Seaman,[c] Raphael Gottardo,[d] Haili Tang,[e] Kelli Greene,[e] Hongmei Gao,[e] Celia LaBranche,[e] John R. Mascola,[f] Lynn Morris,[g] David C. Montefiori[e], Bette Korber[a,h#]

Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, New Mexico, USA[a]; Division of Medical Virology & Institute of Infectious Diseases and Molecular Medicine, University of Cape Town and NHLS, Cape Town, South Africa[b]; Center for Virology and Vaccine Research, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, USA[c]; Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA[d]; Department of Surgery, Duke University Medical Center, Durham, North Carolina, USA[e]; Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA[f]; National Institute for Communicable Diseases, Johannesburg, South Africa[g]; New Mexico Consortium, Los Alamos, New Mexico, USA[h]

Running Head: HIV-1 Clade C Neutralization Panels

#Address correspondence to Peter Hraber and Bette Korber, pth@lanl.gov and btk@lanl.gov.

Abstract: 250 words. Text: 4,778 words (excludes references, table footnotes, figure legends).

ABSTRACT

In the search for effective immunologic interventions to prevent and treat HIV-1 infection, standardized reference reagents are a cost-effective way to maintain robustness and reproducibility among immunological assays. To support planned and ongoing studies where clade C predominates, here we describe three virus panels, chosen from 200 well-characterized clade C envelope (Env)-pseudotyped viruses from early infection. All 200 Envs were expressed as single-round of replication pseudoviruses, and tested to quantify neutralization titers by 16 broadly neutralizing antibodies (bnAbs) and sera from 30 subjects with chronic clade C infections. We selected large panels of 50 and 100 Envs to characterize cross-reactive breadth, either for sera identified as having potent neutralization activity based on initial screening, or to evaluate neutralization magnitude-breadth distributions of newly isolated antibodies. We identified these panels by down-selection after hierarchical clustering of bnAb neutralization titers. Resulting panels represent diversity of neutralization profiles throughout the range of virus sensitivities identified in the original panel of 200 viruses. A small 12-Env panel was chosen to screen sera from vaccine trials or natural-infection studies for neutralization responses. We considered panels selected by previously described methods, but favor a computationally informed method that enabled selection of viruses representing diverse neutralization sensitivity patterns, given that we do not *a priori* know what the neutralization-response profile of vaccine sera will be, relative to sera from infected individuals. The resulting 12-Env panel complements existing panels. Use of standardized panels enables direct comparisons of data from different trials and study sites testing HIV-1 clade C-specific products.

45 **IMPORTANCE**

46

47 HIV-1 M group includes nine clades and many recombinants. Clade C is the most common line-
48 age, responsible for roughly half of current HIV-1 infections, and a focus for vaccine design and
49 testing. Standard reference reagents, particularly virus panels to study neutralization by antibod-
50 ies, are crucial for developing cost-effective yet rigorous and reproducible assays against this di-
51 verse and variable virus. We developed clade C-specific panels for use as standardized reagents
52 to monitor complex polyclonal sera for neutralization activity, and to characterize potency and
53 breadth of cross-reactive neutralization by monoclonal antibodies, whether engineered or isolat-
54 ed from infected individuals. We chose from 200 southern African, clade C envelope-
55 pseudotyped viruses with neutralization titers against 16 broadly neutralizing antibodies and 30
56 sera from chronic clade C infections. We selected panels to represent diversity of bnAb neutrali-
57 zation profiles and Env neutralization sensitivities. Use of standard virus panels can facilitate
58 comparison of results across studies and sites.

INTRODUCTION

The quest to induce and understand protective immune responses by vaccination against HIV-1 remains a high priority. Passive administration of broadly neutralizing antibodies (bnAbs) is also being evaluated for its ability both to prevent and treat HIV-1 infection. Use of standardized reference reagents facilitates the comparison of results from different cohorts or trials (1). The demand for reagents that reflect global diversity of HIV-1 is offset by the overwhelming regional burden of specific forms of the virus. This regional burden is acutely clear for clade C viruses in southern Africa.

Clade C is far more common than any other HIV-1 lineage. For the period 2004-2007, nearly half (48%) of all HIV-1 infections were clade C, an estimated 15.8 million people (2). It is the dominant clade in southern Africa and India, and circulating recombinants that include C clade Env regions are very common in China (3). Although those prevalence estimates were current a decade ago, as of March 2017, sequences collected in the HIV database (http://hiv.lanl.gov/components/sequence/HIV/geo/geo.comp) indicate C clade predominates in South Africa (98% of 32,826 sequences are C clade) and India (95% of 13,475 sequences are C clade), and C clade or BC recombinants are present in roughly half of 30,188 sequences from China. Furthermore, multiple lines of evidence suggest that clade C is more transmissible (4-6) and may have greater replicative fitness than other subtypes (7, 8), so its prevalence is unlikely to have decreased in the past 10 years. The next most abundant non-recombinant forms are clades A (12%) and B (11%), present in 3.9 and 3.7 million individuals, respectively. Recombination is

also very common, with circulating recombinant forms (CRFs) and unique, non-circulating recombinants (URFs) together constituting 20%, or 6.7 million infections (2).

Here we describe development of standard clade C virus panels for two anticipated uses. Sets of 50 and 100 Envs are intended to enable detailed characterization of magnitude-breadth distributions for neutralizing antibodies and sera. A smaller, more manageable set of 12 Envs is intended to screen newly isolated antibodies or sera from vaccinees. The 12-Env C clade panel was selected to include informative examples of neutralization specificities that arise during the course of natural C clade infections. Envs for these neutralization panels were chosen from a set of 200 well-characterized clade C Envs, which we recently described elsewhere (9). The Envs represent HIV-1 clade C genetic and antigenic diversity of C clade in southern Africa, and do not include other geographic regions, such as India (9). A primary goal was to enable detection of neutralization responses in the new HVTN 702 vaccine efficacy trial that has recently begun in South Africa (10), wherein immune responses to a clade C vaccine will be monitored for their capacity to prevent infection in a clade C epidemic (11).

In related work, we recently described development of a 12-virus global panel that captures average neutralization responses across M (main) group diversity, including common subtypes and CRFs (12). The global panel and virus panels we develop in the present study are both intended to provide standardized reagents to investigators working to characterize adaptive immune responses to HIV. The panels developed here differ from the global reference panel in that the Envs are all from clade C, whereas the 12-virus global panel contains more genetic diversity by including clades A, B, C, and G, plus the recombinants CRF01 and CRF07. Second, a main se-

lection criterion for the global panel was ability to infer typical (median) serum potency. To that end, we identified nine viruses that satisfied the criterion optimally, then added three viruses deliberately to include patterns of neutralization response diversity that were not otherwise included (12). In contrast, here we describe a clade C panel of 12 Envs intended to detect relatively weak or potentially clade-specific Tier 2 neutralization responses. Vaccine sera that yield any detectable responses could be identified for further evaluation. Ultimately, both the clade C and M group panels are intended for use in vaccine trials and in other settings.

METHODS

The CAVIMC-CAVD HIV-1 Clade C Virus Neutralization Phenotype Study was reviewed and approved by the research ethics committee of the Faculty of Health Sciences of the University of Cape Town (168/2007; 513/2012). All participants provided written informed consent for study participation (9).

Neutralization titers were determined with the TZM-bl luciferase assay previously described (13, 14), to test 200 recently described Envs against 16 bnAbs and plasma samples from 30 chronic infections (9). Antibodies studied include five CD4 binding-site (CD4bs) bnAbs: VRC01 (15, 16), VRC07 (17), VRC07-523 (18), VRC13 (19), 3BNC117 (20); four V3-glycan (V3g) bnAbs: PGT121 (21), PGT128 (21), 10-1074 (22), and 10-1074V (22); five V1/V2-glycan (V2g) bnAbs: PGT145 (21), CAP256-VRC26.08 (23), CAP256-VRC26.25 (24), PG9 (25), PGDM1400 (26);

and two MPER bnAbs: 10E8 (27) and 4E10 (28). We will sometimes refer to CAP256-VRC26.08 and CAP256-VRC26.25 as VRC26.08 and VRC26.25, for brevity.

Magnitude-Breadth Panels (50 and 100 Envs)

Large virus panels are useful to characterize the magnitude and breadth of neutralizing antibodies, but panel size limits the rate at which results can be obtained. Using large neutralization panels is very expensive and may consume excessive reagent resources. The tradeoff is that excessively small panels may not contain sufficient information needed to make fair assessments across different bnAbs. We therefore down-selected representative sets of 50 and 100 Envs to facilitate studies of antibody magnitude and breadth.

We used a simple strategy to select subsets of viruses that represent the diversity of responses in the full set. To compare Env profiles, we used Euclidean distance between vectors of 16 bnAb IC50 neutralization titers, then hierarchically clustered the 200 Envs. We weighted the resulting dendrograms by geometric mean IC50 to obtain a gradient from most to least sensitive Env (within constraints of the dendrogram branching structure). We used Ward's method (29) for hierarchical clustering, but also considered other methods. A simple down-selection procedure alternated through rows of the dendrogram-ordered neutralization heatmaps, by including one Env and excluding the next. We repeated this procedure to down-select from the full panel of 200 Envs and obtain smaller panels comprised of 100 or 50 Envs. We kept the same row and column order in neutralization panels during down-selection, rather than recluster and reorder.

149 For each of 16 bnAbs, we compared magnitude-breadth distributions of the full panel of 200
150 clade C Envs with the down-selected panels. The area between curves (ABC) quantified the dif-
151 ference between the two cumulative distribution functions. We used resampling to evaluate fur-
152 ther the ABC values from down-selected panels. Random panel selection characterized the null
153 distribution of ABC values, to understand whether dendrogram-based down-selection gave sig-
154 nificantly lower values than could be obtained by chance. We randomly sampled 100-Env pan-
155 els from all 200 Envs (without replacement) 10^4 times. From each of these, we also sampled a
156 random 50-Env panel. We computed resampled ABCs against the distribution from 200 Envs,
157 and compared these with values from the down-selected panels.
158
159 We repeated the down-selection procedure to obtain an even smaller panel of 12 Envs.
160
161 **Serum Screening Panel (12 Envs)**
162
163 For the purpose of screening sera from vaccinees, we tried several approaches to select a small
164 panel of viruses, intended to include Envs sensitive to a variety of neutralizing antibodies and
165 sera. This smaller "candidate" panel includes 12 pseudoviruses chosen to detect neutralization
166 responses in vaccinees and suggest possible antibody specificities therein. Virus selection was
167 guided by neutralization titers from assays against bnAbs *and* chronic sera from each of 200
168 Envs. Tier phenotyping (30) of these Envs demonstrated 1.3% Tier 1A (n=2), 8.5% Tier 1B
169 (n=17), 75% Tier 2 (n=150), and 15.5% Tier 3 (n=31) Envs. We excluded the two Tier 1A Envs
170 and three highly sensitive Tier 1B Envs (geometric mean ID50 titers above 250 reciprocal dilu-

171 tions) from panel selection because they seemed unlikely to be useful in distinguishing protective
172 responses from those that are non-protective (31, 32).
173
174 Our strategy was to select Envs using bnAb IC50s, to ensure all specificities were included, and
175 to compare with ID50s from chronic plasmas. We used principal components analysis (PCA) to
176 simplify high-dimensional data from neutralization assays by projecting them onto fewer dimen-
177 sions. The overall effect of dimension reduction is achieved by decomposing correlations among
178 the data into principal components (33). This approach has recently been used for unsupervised
179 learning to characterize high-dimensional immunological data from HIV Env antigens (34).
180
181 In a computationally guided procedure, we iteratively selected candidate Env panels, then re-
182 viewed their distribution in lower-dimensional projections of bnAb IC50s. Where the candidate
183 panel contained clusters, rather than dispersed Envs, different Envs were chosen to increase the
184 separation between them, to increase coverage of known specificity profiles with the least over-
185 lap possible. This approach enabled us to select 12 candidate Envs that capture the diversity of
186 known bnAb specificities, while ensuring low redundancy among specificity profiles. We think
187 it is important to sample the diversity of natural antibody responses to heterologous virus isolates
188 because we do not know *a priori* the nature of neutralizing antibodies that may be elicited and
189 correlate with vaccine-mediated protection. We compared this PCA-guided strategy to automat-
190 ic selection using lasso (12, 35, 36) and a *k*-medoids clustering strategy (via the pam package in
191 R, version 2.0.5), in addition to the down-selection procedure developed for larger panels.
192

All analysis was done using R (versions 3.3.0 through 3.4.0). We computed Env hypervariable loop lengths and net charges as described previously (9, 37).

RESULTS

Antibody Neutralization. Neutralization titers are typically determined as point values (e.g. IC50, IC80) to summarize distributions from a series of reagent concentrations. The antibody concentration ranges tested in neutralization assays often produce censored neutralization IC50 titers, where the range of concentrations does not yield 50% neutralization. Censored outcomes are represented as "$>x$", where x is the greatest concentration used, or "$<y$", where y is the lowest concentration used. These cutoffs can differ across assays, generally due to practical constraints of limited serum or antibody availability. Such censoring is an issue for quantitative analysis, because standard practice would use a constant placeholder value for censored outcomes, e.g. an IC50 above 50 (">50") is replaced with 100. Censoring thresholds of 10, 20, 25, and 50 µg/ml were used for different bnAbs (**Figure 1**) and it was sometimes necessary to use different thresholds for even one bnAb, such as 3BNC117. Most of the IC50 titers by 3BNC117 were not censored (n=158 Envs). However, for 38 Envs the 3BNC117 values were reported as >20 µg/ml, and for 4 Envs this was >50 µg/ml. To standardize comparisons, and to compare different bnAbs against the 200-virus panel, we used a consistent censoring cutoff of 10 µg/ml across all assays, and IC50s below 0.01 µg/ml were censored at 0.01.

Magnitude-Breadth Panels

Envs down-selected for magnitude-breadth characterization sampled the spectrum of bnAb reactivity patterns from the full set of 200 Envs (**Figure 2**). Heatmaps show IC50 titers for the full neutralization panel (**Figure 2a**) and down-selected panels of 100 (**Figure 2b**) and 50 Envs (**Figure 2c**). Histograms show similar IC50 distributions at the top of each panel, combined for all 16 bnAbs.

For each of the bnAbs, **Figure 3** compares neutralization magnitude-breadth distributions of the full panel of 200 clade C Envs with the down-selected panels. In most cases the magnitude-breadth distributions show a high degree of overlap, which means the down-selected panels represent properties of the full set well. A slight shift towards greater neutralization sensitivity is apparent for some bnAbs, where distributions of selected Envs are biased towards slightly lower IC50 values than the excluded Envs. This small bias resulted from favoring more sensitive viruses when choosing alternate rows in the heatmap, i.e. by starting with the most sensitive virus, rather than skipping it for the next most sensitive.

Concordance of the breadth-potency curves was very high and consistent across bnAbs for 100 and 50 Envs (**Table S1**). Down-sampling further to obtain a 12-Env panel increased the bias in favor of some bnAbs and against others, and gave only a rough approximation to the full set of 200 Envs (**Figure S1**). Also, down-sampling to 12 Envs greatly increased the area between magnitude-breadth curves versus the full set (**Figure S2**), and is part of our rationale not to use down-sampling to select a 12-Env panel. We instead considered other approaches.

Serum Screening Panel (12 Envs)

Figure 4a summarizes Env sensitivity to neutralization by plasma, calculated as geometric mean ID50 among 30 chronic plasmas, together with the number of bnAbs that neutralized each Env. This coarse measure of sensitivity across all bnAbs was significantly associated with sensitivity to plasma (Kendall's τ, τ=0.338, $p=3.34\times10^{-11}$). We used this association to select Envs from PCA of bnAb neutralization data via computational guidance.

Informed by the results from testing each Env against multiple bnAbs, we sought to represent the diversity of different bnAb specificities, to reduce the risk of missing neutralization signal by over-representing the most common bnAb specificities. For this reason, we selected 12 Envs to represent a range of neutralization sensitivity to polyclonal plasma and monoclonal antibodies. **Figure 4b** shows the cumulative distribution of Env sensitivities to plasma. Env colors indicate the number of bnAbs with IC50 below 10 µg/ml from **Figure 4a**. Where plasma and bnAb sensitivities are closely associated, the progression of Envs appears in an order consistent with the progression of rows in **Figure 4a**. An overall trend is apparent for an association between serum and bnAb sensitivity, though small inconsistencies across Envs reflect wide variation in neutralization titers against sera and the number of bnAbs to which each Env is sensitive.

Figure 4c compares plasma ID50 distributions between the candidate panel and remaining Envs. The candidate panel was intentionally chosen to avoid extremely high or low geometric mean ID50 titers among chronic plasmas, both to reduce false negatives, and to exclude Tier 1 neutralization responses, which are readily obtained induced and do not correlate with immune protec-

tion (31, 32). We found no evidence that geometric mean ID50s of the selected panel (n=12) and the remaining Envs (n=183) were sampled from different distributions (two-sided, two-sample Kolmogorov Smirnov test, p=0.53). The candidate panel Envs were neutralized by different numbers (**Figure 4**) and subsets (**Figure 5**) of bnAbs, rather than Envs being sensitive to all the bnAbs studied, and we confirmed that multiple Envs that were well-targeted by each major monoclonal antibody epitope specificity tested were included.

To simplify the diverse outcomes of Env sensitivity to neutralization by different antibodies, and to facilitate the selection of 12 Envs that covered a range distinctive neutralization profiles to the 16 bnAbs tested, we used PCA, which flattens the neutralization data into orthogonal (minimally correlated) sets of linear combinations of bnAbs (**Figure S3**). The first two principal components together explain about half the variance (47.6%) in the bnAb IC50 data. Adding the third principal component accounts for 64.6% of the total variance. As detailed in Supplemental Materials, the first three principal components are strongly associated with combinations of CD4bs, V2 glycan, and V3 glycan bnAb specificities.

After comparing the alternative clustering methods, we favored Ward's method (29) with squared Euclidean distances (ward.D2) for clustering. Ward's method was best able to cluster distinctive patches of serum and virus specificities within the broader gradient of plasma neutralization sensitivities. The resulting clustered heatmap of serum neutralization ID50 titers (**Figure 6**) is annotated to identify the candidate panel of 12 Envs. The panels identified automatically (lasso and *k*-medoids), as described in Supplemental Materials are also shown, for comparison. All three sets of Envs represent a range of average neutralization sensitivities, as reflected by

their dispersal from the top to the bottom of the heatmap, which correspond to more resistant and more sensitive Envs, respectively. The candidate panel, chosen with computational guidance, covers a more limited range of sensitivities than the automatically chosen Envs. This was done intentionally during the iterative procedure described above, to avoid both highly sensitive and very resistant viruses.

Other clustering methods can yield quite different outcomes, and the correlation coefficient between cophenetic distances (38) summarizes similarity among clusters obtained using alternative algorithms (**Figure S4**).

Ordering ID50s by geometric mean titer reveals the continuum of neutralization responses (**Figure S5**), which is characteristic of the polyclonal mixture of antibody potencies and/or specificities found in plasma samples (39). This continuum further emphasizes the benefit of using bnAb sensitivities, rather than plasma responses, for computationally guided panel selection, given that we do not know whether a range of antibody sensitivities or varied antibody potencies dominates the neutralization response of any plasma sample.

Figure S6 summarizes serum neutralization responses among the 12-Env panels identified by 3 automated methods (down-selection, lasso, and *k*-medoids), versus computationally guided selection. Because computationally guided selection avoided individual Envs that were sensitive to all bnAb specificities, the candidate panel does not merely reflect the continuum of neutralization responses, as do the panels identified by automated methods. Sensitive and informative de-

tection of Tier 2 neutralization responses, not modeling the full distribution of Env plasma sensitivities, is the main purpose intended for the candidate 12-Env panel.

Information about the 12 candidate Envs, including the geographic region and year sampled, is summarized in **Table 1**. Other information is tabulated to summarize genetic attributes of these sequences, including glycosylation state (presence or absence of a potential N-linked glycosylation motif) at sites relevant to antibody binding susceptibility, hypervariable loop lengths and net charges, and the infection stage from which the virus was sampled.

Table 2 summarizes the IC50 neutralization titers by 16 bnAbs. In the candidate panel, ZM233M and Ce703010010_C4 are resistant only to PGT128. Another Env, Ko243, is sensitive to all bnAbs shown. The selection of Envs sensitive to specific bnAb families is evident in the last three rows (**Table 2**).

Dataset S1 lists the properties summarized in **Table 1** and **Table 2** for all 200 Envs.

Comparison with Earlier Panels. Earlier work, published in 2006, described a panel of 12 clade C Envs from South Africa and Zambia, selected from among 18 viruses which were all acquired by heterosexual transmission and represented acute or early infections (40). Their median collection date was June, 2001 (range: June, 1998 through June, 2005; there is always an inevitable lag between sample collection and publication). Median collection date among viruses in the current clade C panel was 2007 (range: October, 2002 through 2010; month of sample collection was not reported for these data). Average pairwise distance (APD) on trees from aligned *env*

330 nucleotide sequences is 9.2% greater for this panel (0.250) than for the 2006 clade C panel
331 (0.229), which are both lower than for the global multi-clade panel (0.330), as expected. Phylo-
332 genetic distances are significantly greater for the current panel than for the 2006 panel (n=66 in
333 both; two-sided Wilcoxon, p=0.00018), and reflect the more challenging conditions of the cur-
334 rent epidemic and test conditions for vaccine efficacy trials. These trees (not shown) were com-
335 puted by PhyML version 3 (41, 42) with the GTR+Γ4+I substitution model and rooted on HXB2,
336 though distances to HXB2 were excluded from panel APD calculations. Both panels were de-
337 signed to represent acute and early infection following heterosexual transmission. Because in-
338 creasing southern African clade C diversity is associated with reduced cross-reactive neutraliza-
339 tion between sera and circulating HIV strains (9), a more divergent, more contemporary clade C
340 panel better reflects the modern state of the epidemic. Such samples are difficult to obtain and it
341 takes years to acquire and evaluate them experimentally, so an even more a recent sampling to
342 assess vaccine trials that are currently underway is infeasible.
343
344 We also compared bnAb neutralization titers from viruses in each panel, and summarize neutral-
345 ization data for the 2006 clade C (**Figure 5a**) and global panels (**Figure 5b**) from the CATNAP
346 database (43). For the previously published panels, we extracted data available from CATNAP
347 as of May, 2017 (http://hiv.lanl.gov/catnap). Envs with multiple published results are summa-
348 rized as the geometric mean IC50 among unique values. That is, if an assay were published
349 three times with the same value and once with another value, only the two distinct neutralization
350 values were averaged. This was done to avoid biased estimates, where papers reproduce results
351 from earlier papers without repeating the experiment. One Env (ZM233M) was included in both
352 clade C panels, identified by an asterisk. The candidate 2017 clade C panel (**Figure 5c**) we have

described above is no less sensitive to known bnAbs, and is intended as an update to the 2006 clade C panel, for sensitive and informative plasma screening.

By design, several Envs in this new clade C panel share reactivity patterns to distinct bnAb classes. For example, B005582 is particularly sensitive to V3g bnAbs, Ce2103 to V2g bnAbs, and 2969249 to CD4bs bnAbs (**Figure 5c**). Detecting neutralization in plasmas that have responses to one or more of these viruses would provide clues about antibody specificities therein, and provide information for follow-up experiments that map specificities or isolate monoclonal antibodies.

DISCUSSION

To enhance scientific rigor, improve reproducibility, and unify efforts against HIV diversity, the use of standardized reference reagents for immunological assays is highly beneficial. Standardized reagents enable comparisons between different studies. We have described selection of standardized virus panels from HIV-1 clade C for several anticipated types of investigation, which include screening large numbers of sera from vaccinees for immune-induced neutralization responses, and to characterize the magnitude and breadth of neutralization responses by newly isolated monoclonal antibodies.

Guided by the anticipated uses for these panels, we have described practical selection criteria, which utilize available information to obtain appropriately representative Env panels. We have

described use of hierarchical clustering and a simple but elegant down-selection method to identify subsets of 100 and 50 clade C Envs from a panel of 200 well-characterized viruses. The panels performed better than randomly selected panels at characterizing magnitude-breadth distributions in aggregate across 16 bnAbs. For particular bnAbs, rather than the overall aggregate, moderate to almost no deviation appeared between the magnitude-breadth distributions reported by our down-selected panels and the full set of 200 Envs. This suggests that the smaller virus panels can be used in place of the full set to characterize bnAb magnitude-breadth distributions. Consequently, the use of smaller virus panels will accelerate the rate at which bnAbs can be characterized. Use of even smaller, 12-Env panels is not recommended in magnitude-breadth studies, to avoid bias in favor of some bnAbs and against others.

We used PCA on 16 bnAb IC50 neutralization titers to project 200 Envelope-pseudotyped viruses onto simplified coordinate systems, for computationally guided Env selection. Using this representation, we identified a panel 12 viruses that covered diverse bnAb sensitivity profiles on reduced dimensions. During panel selection, iterative refinement ensured the 12 had a representative range of sensitivity to 30 chronic plasmas.

We also tried automated methods (down-selection, lasso, *k*-medoids) but favored the panel identified with computational guidance, because it does not merely reiterate the plasma neutralization continuum. The diversified detection strategy embodied by the candidate panel may therefore utilize limited sample materials more effectively than the automatically chosen Env sets, which each contain closely related, and therefore redundant, neutralization profiles.

399 Clade-specific panels may be better able to detect relevant neutralization responses than non-
400 specific panels. In a previous study that tested South African plasma samples, from individuals
401 with C clade infections from the CAPRISA cohort, a panel of Tier 2 clade C viruses showed
402 greater sensitivity to neutralization than Tier 2 virus panels from clades A and B (44). Similar
403 findings have been reported in other studies (37, 40). We will not know how the two panels will
404 compare with vaccinee sera until there is a vaccine that generates some measurable activity
405 against Tier 2 viruses. The earliest success at generating Tier 2 virus neutralization could reflect
406 partially matured bnAbs, and it is not known how these immature bnAbs might be differentially
407 detected with clade-specific versus global-virus panels.
408
409 On the other hand, we do not necessarily expect the candidate panel to perform "better" than the
410 2006 panel with HIV-1 sera. In fact, some of our previous data suggest the panels could perform
411 similarly (12). The underlying scientific question concerns potential differences in panel per-
412 formance with vaccine-elicited antibodies, which cannot be assessed at the moment, because no
413 vaccine yet elicits sufficient tier 2 virus neutralization responses. With this in mind, our goal
414 was to design a panel of clade C viruses that is more contemporary, and selected based on more
415 robust analysis methods, to assure the best possible representation of the current epidemic in
416 southern Africa. The 2006 panel did not use neutralization phenotype data to guide its selection,
417 but rather included what was known and available at the time regarding Env genetic variation,
418 and reported neutralization assay results for the selected panel. We incorporated neutralization
419 phenotypes throughout panel selection, and selected from a very large, clade-specific neutraliza-
420 tion panel. We expect the useful phenotypic characteristics of this new panel to emerge in sub-
421 sequent work.

422

423 Our panel of 12 C clade Envs is intended as an update to the panel reported in 2006 (40). The
424 2006 panel was selected from a small subset through convenience sampling, whereas the 2017
425 panel was rationally selected from a much larger collection of viruses. The 2017 clade C panel
426 contains more recently sampled Envs, deliberately includes sensitivity profiles that are character-
427 istic of the currently known bnAb families, and includes greater genetic diversity than the earlier
428 panel. This is important, because within-clade cross-reactive neutralization tends to decrease as
429 genetic distance increases (37). Also, the candidate panel includes a range of plasma sensitivities
430 and favors neutralization-sensitive Envs without including known Tier 1, to help identify weak
431 clade-specific responses without detecting non-specific antibody neutralization that is typical of
432 a Tier 1 response (30). Consequently, the 2017 clade C panel should be more informative and
433 may be more sensitive than the 2006 clade C panel.

434

435 While we think the candidate screening panel might provide hints about antibody specificities in
436 plasmas, it is intended for screening, and not for epitope mapping, which would be performed to
437 characterize samples that test positive for Tier 2 neutralization activity. Further analysis would
438 be needed to differentiate between possible specificities in a serum, and "next-generation" fin-
439 gerprinting methods (45) could be useful for such purposes.

440

441 Unlike the 12-virus global panel of multi-clade viruses described in an earlier publication (12),
442 we planned these panels to be used for screening sera and bnAbs from vaccinees where clade C
443 infections predominate and clade C vaccines are being tested. We did not formulate a single
444 quantitative metric to choose the virus panels proposed here for standardization. Instead, we

considered a range of current needs for standardized reagents and selected sets of Envs that together satisfied these needs as we thought best. An extremely large number (6×10^{18}) of alternative 12-Env panels is possible. We have described several methods to select useful sets of sequences that are intended to represent diversity in a large neutralization assay panel (6,000 plasma ID50s and 2,600 antibody IC50 titers). The Env panels we propose are reasonably representative of the diversity of the population from which they were chosen, by several different criteria. They represent distinctive bnAb sensitivity patterns and generally reflect the diversity of neutralization responses seen among sera from infected individuals.

HIV-1 clade C, which constitutes about half of all infections worldwide at present, represents formidable genetic diversity. As long as virus evolution continues, the ability to induce and detect immune responses against this highly diverse pathogen will be of sustained significance.

ACKNOWLEDGMENTS

For providing plasma samples, we thank the following networks/institutions: the SANBS (South African National Blood Services); CHAVI (Centre for HIV/AIDS Vaccine Immunology); HVTN (HIV Vaccine Trials Network); CAPRISA (Centre for the AIDS Programmed of Research in South Africa); the Zambia-Emory HIV Research Project; HISIS (HIV super infection study); MHRP (US Military HIV Research Program); Malawi and HIV-1 in Pregnancy Program; and participants and study team of the Tshedimoso study in Botswana (Botswana/Harvard Partnership). For providing serum samples, we thank the following institutions and individuals: Des-

mond Tutu HIV foundation (DTHF), Heidi Freislich; CAPRISA and the Perinatal HIV Research Unit (PHRU). Feedback from two anonymous reviewers greatly improved an earlier draft.

FUNDING INFORMATION

This study is part of the Comprehensive Antibody Vaccine Immune Monitoring Consortium supported by the Bill & Melinda Gates Foundation as part of the Collaboration for AIDS Vaccine Discovery (Grant IDs 38619 and 1032144). CAPRISA samples have been supported by the National Institute of Allergy and Infectious Disease (NIAID), National Institutes of Health (NIH) (grant AI51794), the National Research Foundation (grant 67385), the Columbia University-Southern African Fogarty AIDS International Training and Research Programme (AITRP) funded by the Fogarty International Center, NIH (grant D43TW00231), and the South African Government Department of Science and Technology. The opinions expressed herein are those of the authors and do not purport to reflect the official views of the networks that contributed samples.

FIGURE LEGENDS

Figure 1. Cumulative distributions of neutralization IC50 titers from 16 bnAbs. Each line shows the proportion of 200 Envs with IC50s given by the value along the *x*-axis. Grey lines at the lower and upper range of IC50s indicate where censoring cutoffs differed among assays. Asterisks are intended to help locate the example of 3BNC117 censored at 20 and 50 µg/ml discussed in the text.

Figure 2. Heatmaps of IC50 neutralization titers from assaying 200 clade C envelopes against 16 bnAbs. (**a**) Hierarchically clustered heatmap of IC50 titers of 200 Envs against 16 bnAbs. The Env dendrogram is shown; the bnAb dendrogram is not shown. Leaf colors indicate 100 viruses included (red) or excluded (blue) by down-selection. The histogram (black line) above the heatmap summarizes the distribution of assay results, with histogram breakpoints at 10, 4.64, 2.15, 1.00, 0.464, 0.215, 0.10, 0.0464, 0.0215, and 0.01 µg/ml. Low IC50s were censored at 0.01 µg/ml to standardize censoring thresholds across bnAbs. (**b**) 100-Env panel, down-selected from alternating rows, i.e. red branches on the dendrogram. (**c**) 50-Env panel down-selected by alternating over 100 Envs.

Figure 3. Magnitude-breadth curves (cumulative distribution functions) compare IC50s from 100- and 50-Env panels (colored and grey lines, respectively) with all 200 Envs (black lines). Axes are scaled the same across all panels. Specificities for each bnAb are also listed. The curves are continuous distribution functions, which indicate the proportion of viruses with IC50 neutralization titers (µg/ml) no less than the corresponding *x*-axis value. Because the emphasis

here is on comparisons within the same bnAb, differences in censoring noted for **Figure 1** across different bnAbs are not relevant here. The differences are quantified in Supplemental Materials.

Figure 4. Comparison of Env sensitivity to neutralization by plasmas and bnAbs. (**a**) Geometric mean plasma ID50 per Env, stratified by the number of bnAbs with IC50s below 10 µg/ml, from among the 16 tested. (**b**) Cumulative distribution of geometric mean ID50s from 30 chronic plasmas. Vertical lines indicate the 12 selected Envs. (**c**) Distribution of geometric mean ID50s between 12-virus panel and non-panel Envs. Five Envs with geometric mean ID50s above 250 µg/ml are not shown. Colors in (b) and (c) summarize the number of neutralizing bnAbs shown in (a).

Figure 5. Comparison of neutralization IC50 titers between 12-Env panels. Comparison of (**a**) clade C panel from 2006, (**b**) global panel (12), and (**c**) candidate clade C panel from this manuscript. As noted in **Figure 1**, all assay results were censored above 10 and below 0.001 µg/ml, to standardize dilution ranges across different experiments. NA indicates no data. Data for historical panels (a**)** and (b) were computed as geometric means from the CATNAP database, as detailed in the text. Env names in (a**)** and **(**b) are shortened as in CATNAP, and (c) lists short names from **Dataset S1**.

Figure 6. Hierarchically clustered dendrogram of 200 Tier 2 envelopes with heatmap of neutralization ID50s. The dendrogram was computed from squared Euclidean distance using Ward's clustering method. Leaves (rows) were weighted by geometric mean neutralization titer for dendrogram layout. Colors indicate viruses selected for the candidate 12-Env panel (black). Panels

528 defined by the automatic methods are also indicated, for lasso (red), and by *k*-medoids (blue)

529 with *k*=12. Other virus names are grey.

530 **REFERENCES**

531 1. Mascola JR, D'Souza P, Gilbert P, Hahn BH, Haigwood NL, Morris L, Petropoulos CJ,
532 Polonis VR, Sarzotti M, Montefiori DC. 2005. Recommendations for the design and use of
533 standard virus panels to assess neutralizing antibody responses elicited by candidate human
534 immunodeficiency virus type 1 vaccines. J Virol 79:10103–7.
535 2. Hemelaar J, Gouws E, Ghys PD, Osmanov S, WHO-UNAIDS Network for HIV Isolation
536 Characterisation. 2011. Global trends in molecular epidemiology of HIV-1 during 2000-
537 2007. AIDS 25:679–89.
538 3. Korber BT, Letvin NL, Haynes BF. 2009. T-cell vaccine strategies for human
539 immunodeficiency virus, the virus with a thousand faces. J Virol 83:8300–14.
540 4. John-Stewart GC, Nduati RW, Rousseau CM, Mbori-Ngacha DA, Richardson BA,
541 Rainwater S, Panteleeff DD, Overbaugh J. 2005. Subtype C is associated with increased
542 vaginal shedding of HIV-1. J Infect Dis 192:492–6.
543 5. Iversen AK, Learn GH, Skinhoj P, Mullins JI, McMichael AJ, Rambaut A. 2005. Preferential
544 detection of HIV subtype C' over subtype A in cervical cells from a dually infected woman.
545 AIDS 19:990–3.
546 6. Walter BL, Armitage AE, Graham SC, de Oliveira T, Skinhoj P, Jones EY, Stuart DI,
547 McMichael AJ, Chesebro B, Iversen AK. 2009. Functional characteristics of HIV-1 subtype
548 C compatible with increased heterosexual transmissibility. AIDS 23:1047–57.
549 7. Rodriguez MA, Ding M, Ratner D, Chen Y, Tripathy SP, Kulkarni SS, Chatterjee R,
550 Tarwater PM, Gupta P. 2009. High replication fitness and transmission efficiency of HIV-1
551 subtype C from India: implications for subtype C predominance. Virology 385:416–24.

8. Abraha A, Nankya IL, Gibson R, Demers K, Tebit DM, Johnston E, Katzenstein D, Siddiqui A, Herrera C, Fischetti L, Shattock RJ, Arts EJ. 2009. CCR5- and CXCR4-tropic subtype C human immunodeficiency virus type 1 isolates have a lower level of pathogenic fitness than other dominant group M subtypes: implications for the epidemic. J Virol 83:5592–605.
9. Rademeyer C, Korber B, Seaman MS, Giorgi EE, Thebus R, Robles A, Sheward DJ, Wagh K, Garrity J, Carey BR, Gao H, Greene KM, Tang H, Bandawe GP, Marais JC, Diphoko TE, Hraber P, Tumba N, Moore PL, Gray GE, Kublin J, McElrath MJ, Vermeulen M, Middelkoop K, Bekker LG, Hoelscher M, Maboko L, Makhema J, Robb ML, Abdool Karim S, Abdool Karim Q, Kim JH, Hahn BH, Gao F, Swanstrom R, Morris L, Montefiori DC, Williamson C. 2016. Features of recently transmitted HIV-1 clade C viruses that impact antibody recognition: implications for active and passive immunization. PLoS Pathog 12:e1005742.
10. Cohen KW, Frahm N. 2017. Current views on the potential for development of a HIV vaccine. Expert Opin Biol Ther 17:295–303.
11. Wang Z, Lorin C, Koutsoukos M, Franco D, Bayat B, Zhang Y, Carfi A, Barnett SW, Porter F. 2016. Comprehensive characterization of reference standard lots of HIV-1 subtype C Gp120 proteins for clinical trials in Southern African regions. Vaccines (Basel) 4.
12. deCamp A, Hraber P, Bailer RT, Seaman MS, Ochsenbauer C, Kappes J, Gottardo R, Edlefsen P, Self S, Tang H, Greene K, Gao H, Daniell X, Sarzotti-Kelsoe M, Gorny MK, Zolla-Pazner S, LaBranche CC, Mascola JR, Korber BT, Montefiori DC. 2014. Global panel of HIV-1 Env reference strains for standardized assessments of vaccine-elicited neutralizing antibodies. J Virol 88:2489–507.

13. Montefiori DC. 2005. Evaluating neutralizing antibodies against HIV, SIV, and SHIV in luciferase reporter gene assays. Curr Protoc Immunol Chapter 12:Unit 12 11.

14. Sarzotti-Kelsoe M, Bailer RT, Turk E, Lin CL, Bilska M, Greene KM, Gao H, Todd CA, Ozaki DA, Seaman MS, Mascola JR, Montefiori DC. 2014. Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1. J Immunol Methods 409:131–46.

15. Wu X, Yang ZY, Li Y, Hogerkorp CM, Schief WR, Seaman MS, Zhou T, Schmidt SD, Wu L, Xu L, Longo NS, McKee K, O'Dell S, Louder MK, Wycuff DL, Feng Y, Nason M, Doria-Rose N, Connors M, Kwong PD, Roederer M, Wyatt RT, Nabel GJ, Mascola JR. 2010. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. Science 329:856–61.

16. Wu X, Zhou T, Zhu J, Zhang B, Georgiev I, Wang C, Chen X, Longo NS, Louder M, McKee K, O'Dell S, Perfetto S, Schmidt SD, Shi W, Wu L, Yang Y, Yang ZY, Yang Z, Zhang Z, Bonsignori M, Crump JA, Kapiga SH, Sam NE, Haynes BF, Simek M, Burton DR, Koff WC, Doria-Rose NA, Connors M, NISC Comparative Sequencing Program, Mullikin JC, Nabel GJ, Roederer M, Shapiro L, Kwong PD, Mascola JR. 2011. Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. Science 333:1593–602.

17. Wu X, Zhang Z, Schramm CA, Joyce MG, Kwon YD, Zhou T, Sheng Z, Zhang B, O'Dell S, McKee K, Georgiev IS, Chuang GY, Longo NS, Lynch RM, Saunders KO, Soto C, Srivatsan S, Yang Y, Bailer RT, Louder MK, NISC Comparative Sequencing Program, Mullikin JC, Connors M, Kwong PD, Mascola JR, Shapiro L. 2015. Maturation and diversity of the VRC01-antibody lineage over 15 years of chronic HIV-1 infection. Cell 161:470–85.

18. Rudicell RS, Kwon YD, Ko SY, Pegu A, Louder MK, Georgiev IS, Wu X, Zhu J, Boyington JC, Chen X, Shi W, Yang ZY, Doria-Rose NA, McKee K, O'Dell S, Schmidt SD, Chuang GY, Druz A, Soto C, Yang Y, Zhang B, Zhou T, Todd JP, Lloyd KE, Eudailey J, Roberts KE, Donald BR, Bailer RT, Ledgerwood J, NISC Comparative Sequencing Program, Mullikin JC, Shapiro L, Koup RA, Graham BS, Nason MC, Connors M, Haynes BF, Rao SS, Roederer M, Kwong PD, Mascola JR, Nabel GJ. 2014. Enhanced potency of a broadly neutralizing HIV-1 antibody in vitro improves protection against lentiviral infection in vivo. J Virol 88:12669–82.

19. Zhou T, Lynch RM, Chen L, Acharya P, Wu X, Doria-Rose NA, Joyce MG, Lingwood D, Soto C, Bailer RT, Ernandes MJ, Kong R, Longo NS, Louder MK, McKee K, O'Dell S, Schmidt SD, Tran L, Yang Z, Druz A, Luongo TS, Moquin S, Srivatsan S, Yang Y, Zhang B, Zheng A, Pancera M, Kirys T, Georgiev IS, Gindin T, Peng HP, Yang AS, NISC Comparative Sequencing Program, Mullikin JC, Gray MD, Stamatatos L, Burton DR, Koff WC, Cohen MS, Haynes BF, Casazza JP, Connors M, Corti D, Lanzavecchia A, Sattentau QJ, Weiss RA, West AP, Jr., Bjorkman PJ, Scheid JF, Nussenzweig MC, et al. 2015. Structural repertoire of HIV-1-neutralizing antibodies targeting the CD4 supersite in 14 donors. Cell 161:1280–92.

20. Scheid JF, Mouquet H, Ueberheide B, Diskin R, Klein F, Oliveira TY, Pietzsch J, Fenyo D, Abadir A, Velinzon K, Hurley A, Myung S, Boulad F, Poignard P, Burton DR, Pereyra F, Ho DD, Walker BD, Seaman MS, Bjorkman PJ, Chait BT, Nussenzweig MC. 2011. Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. Science 333:1633–7.

618 21. Walker LM, Huber M, Doores KJ, Falkowska E, Pejchal R, Julien JP, Wang SK, Ramos A,
619 Chan-Hui PY, Moyle M, Mitcham JL, Hammond PW, Olsen OA, Phung P, Fling S, Wong
620 CH, Phogat S, Wrin T, Simek MD, Protocol GPI, Koff WC, Wilson IA, Burton DR,
621 Poignard P. 2011. Broad neutralization coverage of HIV by multiple highly potent
622 antibodies. Nature 477:466–70.
623 22. Mouquet H, Scharf L, Euler Z, Liu Y, Eden C, Scheid JF, Halper-Stromberg A,
624 Gnanapragasam PN, Spencer DI, Seaman MS, Schuitemaker H, Feizi T, Nussenzweig MC,
625 Bjorkman PJ. 2012. Complex-type N-glycan recognition by potent broadly neutralizing HIV
626 antibodies. Proc Natl Acad Sci U S A 109:E3268–77.
627 23. Doria-Rose NA, Schramm CA, Gorman J, Moore PL, Bhiman JN, DeKosky BJ, Ernandes
628 MJ, Georgiev IS, Kim HJ, Pancera M, Staupe RP, Altae-Tran HR, Bailer RT, Crooks ET,
629 Cupo A, Druz A, Garrett NJ, Hoi KH, Kong R, Louder MK, Longo NS, McKee K, Nonyane
630 M, O'Dell S, Roark RS, Rudicell RS, Schmidt SD, Sheward DJ, Soto C, Wibmer CK, Yang
631 Y, Zhang Z, NISC Comparative Sequencing Program, Mullikin JC, Binley JM, Sanders RW,
632 Wilson IA, Moore JP, Ward AB, Georgiou G, Williamson C, Abdool Karim SS, Morris L,
633 Kwong PD, Shapiro L, Mascola JR. 2014. Developmental pathway for potent V1V2-directed
634 HIV-neutralizing antibodies. Nature 509:55–62.
635 24. Doria-Rose NA, Bhiman JN, Roark RS, Schramm CA, Gorman J, Chuang GY, Pancera M,
636 Cale EM, Ernandes MJ, Louder MK, Asokan M, Bailer RT, Druz A, Fraschilla IR, Garrett
637 NJ, Jarosinski M, Lynch RM, McKee K, O'Dell S, Pegu A, Schmidt SD, Staupe RP, Sutton
638 MS, Wang K, Wibmer CK, Haynes BF, Abdool-Karim S, Shapiro L, Kwong PD, Moore PL,
639 Morris L, Mascola JR. 2016. New member of the V1V2-directed CAP256-VRC26 lineage
640 that shows increased breadth and exceptional potency. J Virol 90:76–91.

641 25. Walker LM, Phogat SK, Chan-Hui PY, Wagner D, Phung P, Goss JL, Wrin T, Simek MD,
642 Fling S, Mitcham JL, Lehrman JK, Priddy FH, Olsen OA, Frey SM, Hammond PW, Protocol
643 GPI, Kaminsky S, Zamb T, Moyle M, Koff WC, Poignard P, Burton DR. 2009. Broad and
644 potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target.
645 Science 326:285–9.
646 26. Sok D, van Gils MJ, Pauthner M, Julien JP, Saye-Francisco KL, Hsueh J, Briney B, Lee JH,
647 Le KM, Lee PS, Hua Y, Seaman MS, Moore JP, Ward AB, Wilson IA, Sanders RW, Burton
648 DR. 2014. Recombinant HIV envelope trimer selects for quaternary-dependent antibodies
649 targeting the trimer apex. Proc Natl Acad Sci U S A 111:17624–9.
650 27. Huang J, Ofek G, Laub L, Louder MK, Doria-Rose NA, Longo NS, Imamichi H, Bailer RT,
651 Chakrabarti B, Sharma SK, Alam SM, Wang T, Yang Y, Zhang B, Migueles SA, Wyatt R,
652 Haynes BF, Kwong PD, Mascola JR, Connors M. 2012. Broad and potent neutralization of
653 HIV-1 by a gp41-specific human antibody. Nature 491:406–12.
654 28. Zwick MB, Labrijn AF, Wang M, Spenlehauer C, Saphire EO, Binley JM, Moore JP,
655 Stiegler G, Katinger H, Burton DR, Parren PW. 2001. Broadly neutralizing antibodies
656 targeted to the membrane-proximal external region of human immunodeficiency virus type 1
657 glycoprotein gp41. J Virol 75:10892–905.
658 29. Ward JH. 1963. Hierarchical grouping to optimize an objective function. J Am Stat Assoc
659 58:236–44.
660 30. Seaman MS, Janes H, Hawkins N, Grandpre LE, Devoy C, Giri A, Coffey RT, Harris L,
661 Wood B, Daniels MG, Bhattacharya T, Lapedes A, Polonis VR, McCutchan FE, Gilbert PB,
662 Self SG, Korber BT, Montefiori DC, Mascola JR. 2010. Tiered categorization of a diverse

panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. J Virol 84:1439–52.

31. Gilbert P, Wang M, Wrin T, Petropoulos C, Gurwith M, Sinangil F, D'Souza P, Rodriguez-Chavez IR, DeCamp A, Giganti M, Berman PW, Self SG, Montefiori DC. 2010. Magnitude and breadth of a nonprotective neutralizing antibody response in an efficacy trial of a candidate HIV-1 gp120 vaccine. J Infect Dis 202:595–605.
32. Montefiori DC, Karnasuta C, Huang Y, Ahmed H, Gilbert P, de Souza MS, McLinden R, Tovanabutra S, Laurence-Chenine A, Sanders-Buell E, Moody MA, Bonsignori M, Ochsenbauer C, Kappes J, Tang H, Greene K, Gao H, LaBranche CC, Andrews C, Polonis VR, Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, Self SG, Berman PW, Francis D, Sinangil F, Lee C, Tartaglia J, Robb ML, Haynes BF, Michael NL, Kim JH. 2012. Magnitude and breadth of the neutralizing antibody response in the RV144 and Vax003 HIV-1 vaccine efficacy trials. J Infect Dis 206:431–41.
33. Wall ME, Rechtsteiner A, Rocha LM. 2003. Chapter 5: Singular value decomposition and principal component analysis. In D.P. Berrar, W. Dubitzky, M. Granzow, eds. A Practical Approach to Microarray Data Analysis. Kluwer, Norwell, MA.
34. Choi I, Chung AW, Suscovich TJ, Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, O'Connell RJ, Francis D, Robb ML, Michael NL, Kim JH, Alter G, Ackerman ME, Bailey-Kellogg C. 2015. Machine learning methods enable predictive modeling of antibody feature:function relationships in RV144 vaccinees. PLoS Comput Biol 11:e1004185.
35. Tibshirani R. 1996. Regression shrinkage and selection via the lasso. J R Stat Soc Series B Stat Methodol 1:267–88.

686 36. Friedman J, Hastie T, Tibshirani R. 2008. Regularization paths for generalized linear models
687 via coordinate descent. J Stat Softw 33.
688 37. Hraber P, Korber BT, Lapedes AS, Bailer RT, Seaman MS, Gao H, Greene KM, McCutchan
689 F, Williamson C, Kim JH, Tovanabutra S, Hahn BH, Swanstrom R, Thomson MM, Gao F,
690 Harris L, Giorgi E, Hengartner N, Bhattacharya T, Mascola JR, Montefiori DC. 2014. Impact
691 of clade, geography, and age of the epidemic on HIV-1 neutralization by antibodies. J Virol
692 88:12623–43.
693 38. Sneath PHA, Sokal RR. 1973. Numerical taxonomy: the principles and practice of numerical
694 classification. Freeman, San Francisco, CA.
695 39. Hraber P, Seaman MS, Bailer RT, Mascola JR, Montefiori DC, Korber B. 2014. Prevalence
696 of broadly neutralizing antibody responses during chronic HIV-1 infection AIDS 28:163–9.
697 40. Li M, Salazar-Gonzalez JF, Derdeyn CA, Morris L, Williamson C, Robinson JE, Decker JM,
698 Li Y, Salazar MG, Polonis VR, Mlisana K, Karim SA, Hong K, Greene KM, Bilska M, Zhou
699 J, Allen S, Chomba E, Mulenga J, Vwalika C, Gao F, Zhang M, Korber BT, Hunter E, Hahn
700 BH, Montefiori DC. 2006. Genetic and neutralization properties of subtype C human
701 immunodeficiency virus type 1 molecular env clones from acute and early heterosexually
702 acquired infections in Southern Africa. J Virol 80:11776–90.
703 41. Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large
704 phylogenies by maximum likelihood. Syst Biol 52:696–704.
705 42. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New
706 algorithms and methods to estimate maximum-likelihood phylogenies: assessing the
707 performance of PhyML 3.0. Syst Biol 59:307–21.

708 43. Yoon H, Macke J, West AP, Jr., Foley B, Bjorkman PJ, Korber B, Yusim K. 2015.
709 CATNAP: a tool to compile, analyze and tally neutralizing antibody panels. Nucleic Acids
710 Res 43:W213–9.
711 44. Gray ES, Madiga MC, Hermanus T, Moore PL, Wibmer CK, Tumba NL, Werner L, Mlisana
712 K, Sibeko S, Williamson C, Abdool Karim SS, Morris L, Team CS. 2011. The neutralization
713 breadth of HIV-1 develops incrementally over four years and is associated with CD4+ T cell
714 decline and high viral load during acute infection. J Virol 85:4828–40.
715 45. Doria-Rose NA, Altae-Tran HR, Roark RS, Schmidt SD, Sutton MS, Louder MK, Chuang
716 G-Y, Bailer RT, Cortez V, Kong R, McKee K, O'Dell S, Wang F, Karim SSA, Binley JM,
717 Connors M, Haynes BF, Martin MA, Montefiori DC, Morris L, Overbaugh J, Kwong PD,
718 Mascola JR, Georgiev IS. 2017. Mapping polyclonal HIV-1 antibody responses via next-
719 generation neutralization fingerprinting. PLoS Pathog 13:e1006148.

Table 1. Properties of Tier 2 Envs selected for candidate 12-virus panel, chosen with computational guidance.

Accession[1]	Name[2]	Year[3]	Country/ Region[4]	is TF[5]	Stage[6]	N332 N334[7]	N293 N295	N130	N156 N160	V1V2 aas[8]	V1V2 charge[9]	V4 length	V4 charge	V5 length	V5 charge
FJ443533	Ce703010010	2006	MW	T	A1	N332	none	F	both	28	2	4	-1	9	-4
DQ388517	ZM233M	2002	ZM	NA	E	N332	N293	F	both	19	2	4	0	7	-1
DQ422948	ZM215F	2002	ZM	NA	E	N332	none	T	both	15	0	4	1	5	0
HM215307	3728	2004	TZ	F	A2	none	N295	F	both	22	-2	9	1	6	0
KF114892	Ko243	2009	ZA/nw	T	E	N332	N295	F	both	49	2	13	0	7	2
FJ444124	Ce704810053	2007	ZA/gp	T	A1	none	N295	T	both	34	-2	9	3	4	0
HQ615959	2759058	2006	ZA/kz	T	A1	none	none	T	both	23	-1	6	-1	10	-2
JN681246	So431	2007	ZA/gp	T	E	N332	none	F	both	31	-1	8	0	7	-1
KC154028	CAP382	2010	ZA/kz	T	E	N332	none	F	both	22	-5	5	1	11	-2
FJ444612	Ce2103	2005	MW	T	A1	N334	none	T	N156	34	-8	12	-1	6	0
KF114882	B005582	2007	BW	T	A2	N332	N295	F	both	31	-3	7	0	8	-1
HQ595766	2969249	2007	ZA/kz	T	A2	N334	N295	F	both	31	0	7	-1	8	-1

[1] A table with these data for all 200 Envs is available among Supplemental Materials.

[2] Common name of the sequence.

[3] Year in which the sequence was sampled.

[4] Country and region in which the sequence was sampled.

[5] Does the sequence represent transmitted/founder (TF) that established homogeneous infection?

[6] Infection stage at time of sampling. (E, early; A1, A2).

[7] Indicates presence of potentially N-linked glycosylation sequon motif at the site/s listed.

[8] Length, i.e. number of amino acids in the hypervariable region/s.

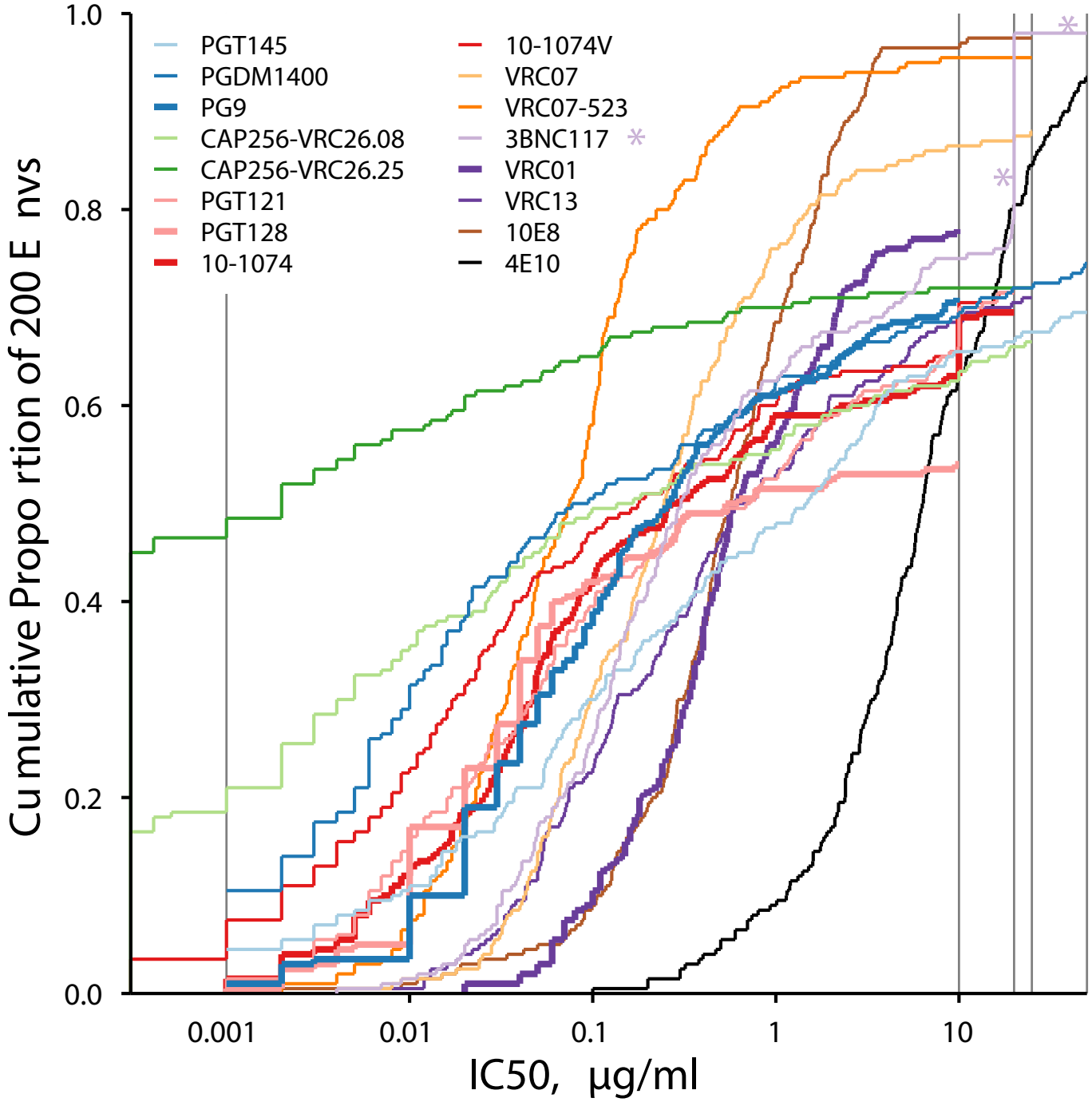
[9] Sum of amino acid charges in hypervariable region/s.

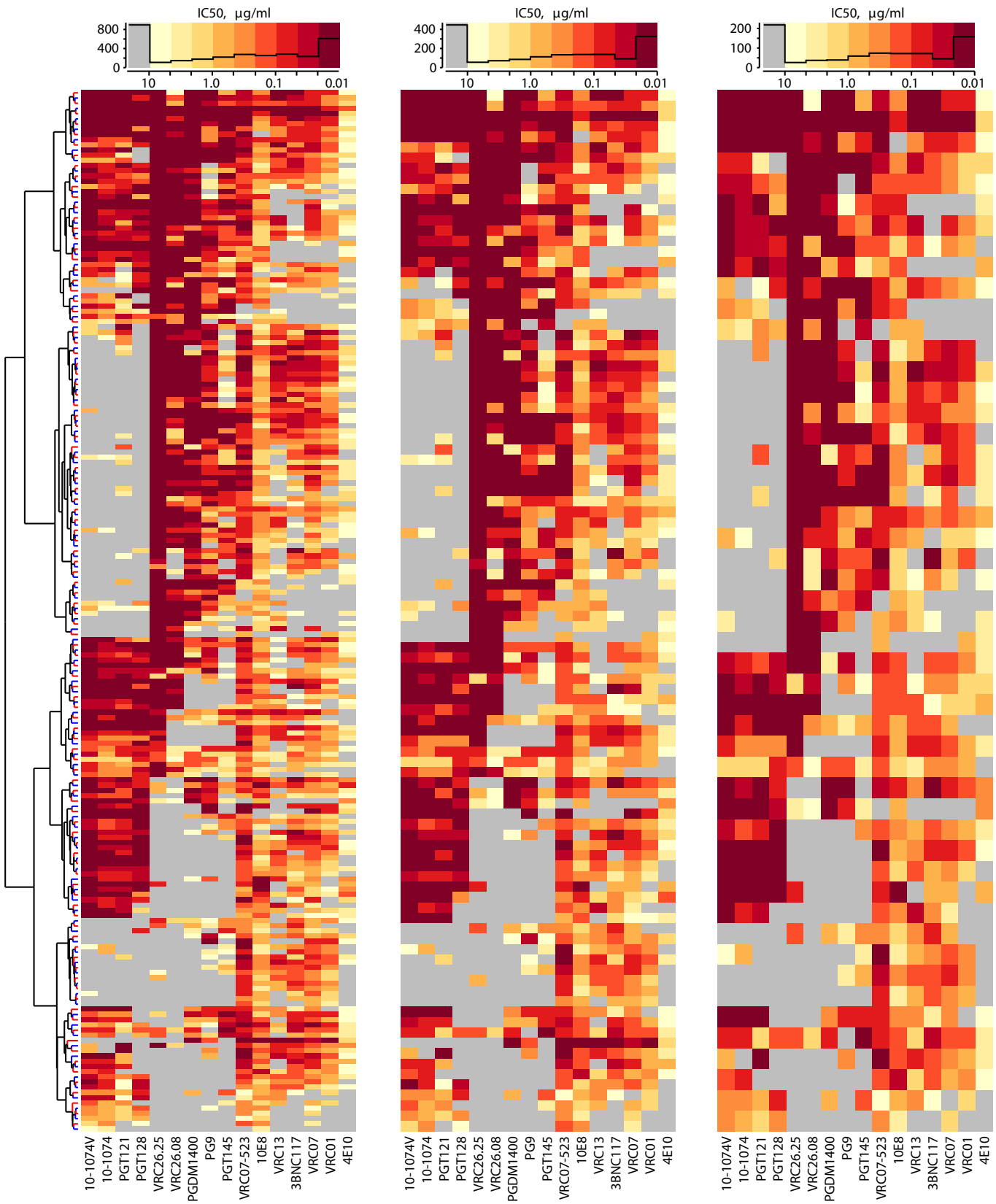
Table 2. Clade-C panel antibody neutralization IC50 titers[1] (µg/ml).

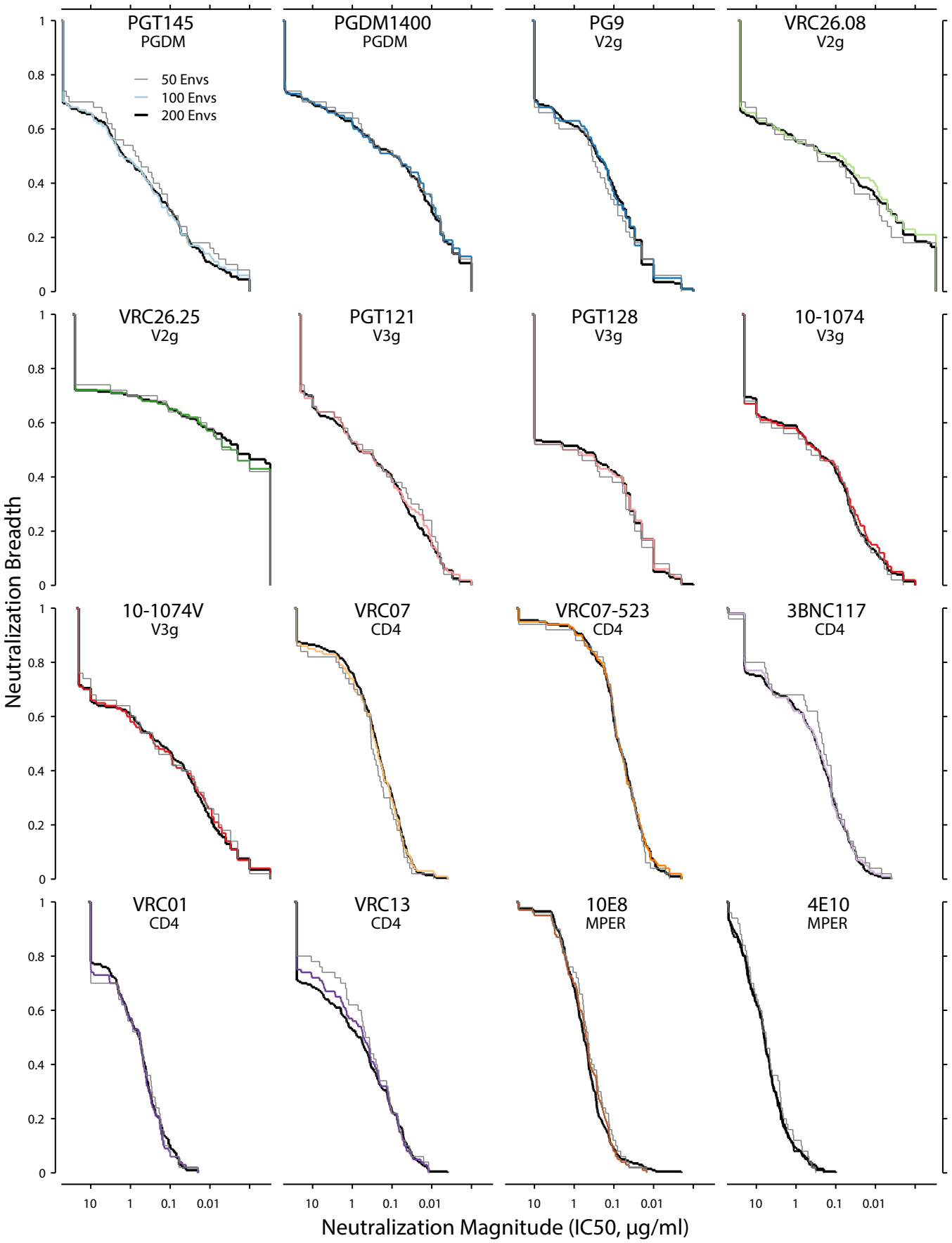
bnAb Specificity:		V2g	V2g	V2g	V2g	V2g	V3g	V3g	V3g	V3g	CD4bs	CD4bs	CD4bs	CD4bs	CD4bs	MPER	
Env Accession[2]	Env Name	PGT145	PGDM-1400	PG9	CAP256-VRC26.08	CAP256-VRC26.25	PGT121	PGT128	10.1074V	10.1074	VRC07	VRC07-523	3BNC117	VRC01	VRC13	10E8	4E10
FJ443533	Ce703010010	**0.003**	**0.001**	**0.01**	**<0.0003**	**<0.0003**	**0.008**	>10	**0.008**	**0.031**	**0.025**	**0.01**	**0.041**	**0.11**	**0.25**	**0.772**	**5.2**
DQ388517	ZM233M	**0.008**	**<0.001**	**0.002**	**<0.0003**	**<0.0003**	**2.809**	>10	**0.051**	**0.058**	**0.228**	**0.035**	**0.13**	**1.67**	**0.045**	**0.259**	**1.2**
DQ422948	ZM215F	>50	**0.008**	**0.02**	>25	>25	**0.01**	**0.06**	**0.006**	**0.028**	**0.058**	**0.041**	**0.018**	**0.17**	**0.365**	**0.067**	**0.4**
HM215307	3728	31.724	>50	**0.13**	**0.0004**	**<0.0003**	**3.783**	>10	>20	>20	**0.024**	**0.007**	**0.02**	**0.06**	**0.026**	**0.153**	**1.6**
KF114892	Ko243	**0.015**	**0.006**	**0.14**	**1.061**	**<0.0003**	**1.362**	**0.04**	**0.505**	**0.867**	**0.357**	**0.113**	**0.314**	**0.67**	**8.675**	**1.643**	18.62
FJ444124	Ce704810053	**3.398**	**0.003**	**0.16**	**<0.001**	**<0.0003**	>20	0.01	>20	>20	**1.41**	**0.069**	**1.582**	>10	**0.789**	**0.28**	**2.9**
HQ615959	2759058	**1.852**	**0.062**	**0.13**	**0.004**	**<0.0003**	**1.169**	>10	>20	>20	**0.537**	**0.155**	**1.141**	**2.64**	>25	**1.287**	**2.89**
JN681246	So431	35.309	>50	**0.34**	>25	>25	**0.021**	>10	**0.577**	>20	**0.354**	**0.012**	**0.198**	**0.56**	**0.072**	**0.161**	**2.5**
KC154028	CAP382	**1.272**	**0.021**	**0.07**	**<0.0003**	**<0.0003**	**9.911**	**6.42**	**0.8**	**5.885**	>25	**0.372**	>20	>10	>25	**1.511**	18.15
FJ444612	Ce2103	**2.533**	**0.012**	**2.6**	**0.002**	**<0.0003**	>20	>10	>20	>20	>25	>25	>20	>10	**0.221**	**2.506**	17.98
KF114882	B005582	>50	>50	>10	>25	>25	**6.839**	**0.03**	**0.027**	**0.034**	>25	**2.378**	>20	>10	**0.328**	**0.382**	23.64
HQ595766	2969249	>50	>50	>10	>25	>25	>20	>10	>20	>20	**0.93**	**0.222**	**0.64**	**1.37**	19.391	**0.921**	10.83

[1] Values below 10 µg/ml appear in bold text. See **Figure 5c** for the corresponding heatmap.

[2] A table containing these data for all 200 Envs is available among Supplemental Materials.







Journal of Virology