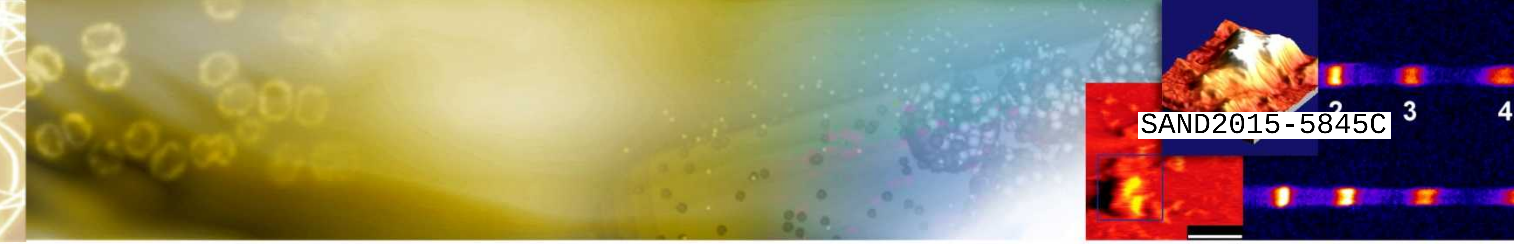


SAND2015-5845C

Sandia's Web-Based Bioinformatics Tools

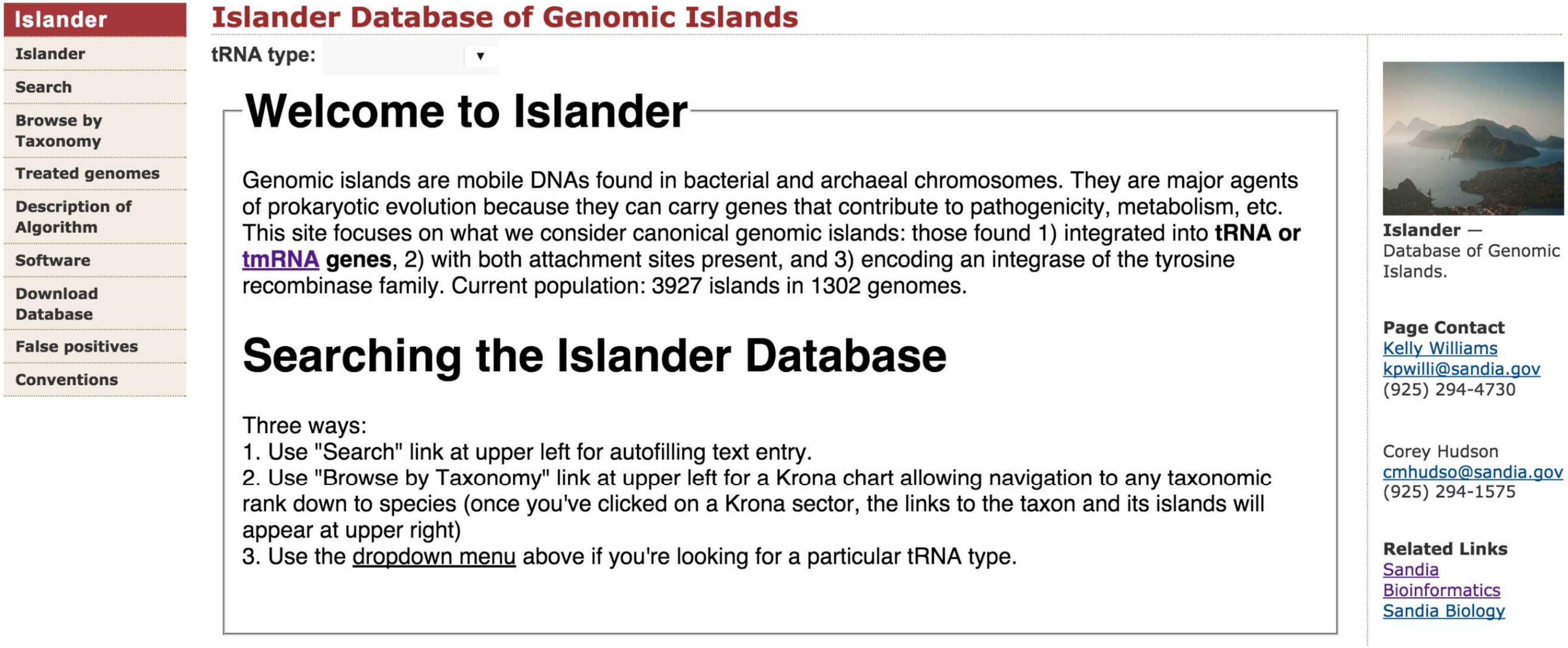
Corey Hudson

(08633)

Sandia Bioinformatics has Developed a Front-Facing Web Presence

- 1) **Islander**: Database of nucleotide precision genomic islands.
  - http://bioinformatics.sandia.gov/islander/
- 2) ***Klebsiella pneumoniae*** **BAA-2146 genome browser**.
  - http://bioinformatics.sandia.gov/klebs/jbrowse/
- 3) **tmRNA** website: Source of tmRNA annotation, which serves a 3rd party for International Nucleotide Sequence Database archives (INSD), serving NCBI/Genbank, EMBL-EBI and DDBJ.
  - http://bioinformatics.sandia.gov/tmrna/index.html
- 4) **Software** tools for Island prediction, genome assembly, tDNA prediction and RapTOR (Rapid Threat Organism Recognition) Biodetection System for Public Health and Biodefense.
  - http://bioinformatics.sandia.gov/software/

Islander



Islander Database of Genomic Islands

Welcome to Islander

Genomic islands are mobile DNAs found in bacterial and archaeal chromosomes. They are major agents of prokaryotic evolution because they can carry genes that contribute to pathogenicity, metabolism, etc. This site focuses on what we consider canonical genomic islands: those found 1) integrated into **tRNA or tmRNA genes**, 2) with both attachment sites present, and 3) encoding an integrase of the tyrosine recombinase family. Current population: 3927 islands in 1302 genomes.

Searching the Islander Database

Three ways:
1. Use "Search" link at upper left for autofilling text entry.
2. Use "Browse by Taxonomy" link at upper left for a Krona chart allowing navigation to any taxonomic rank down to species (once you've clicked on a Krona sector, the links to the taxon and its islands will appear at upper right)
3. Use the dropdown menu above if you're looking for a particular tRNA type.

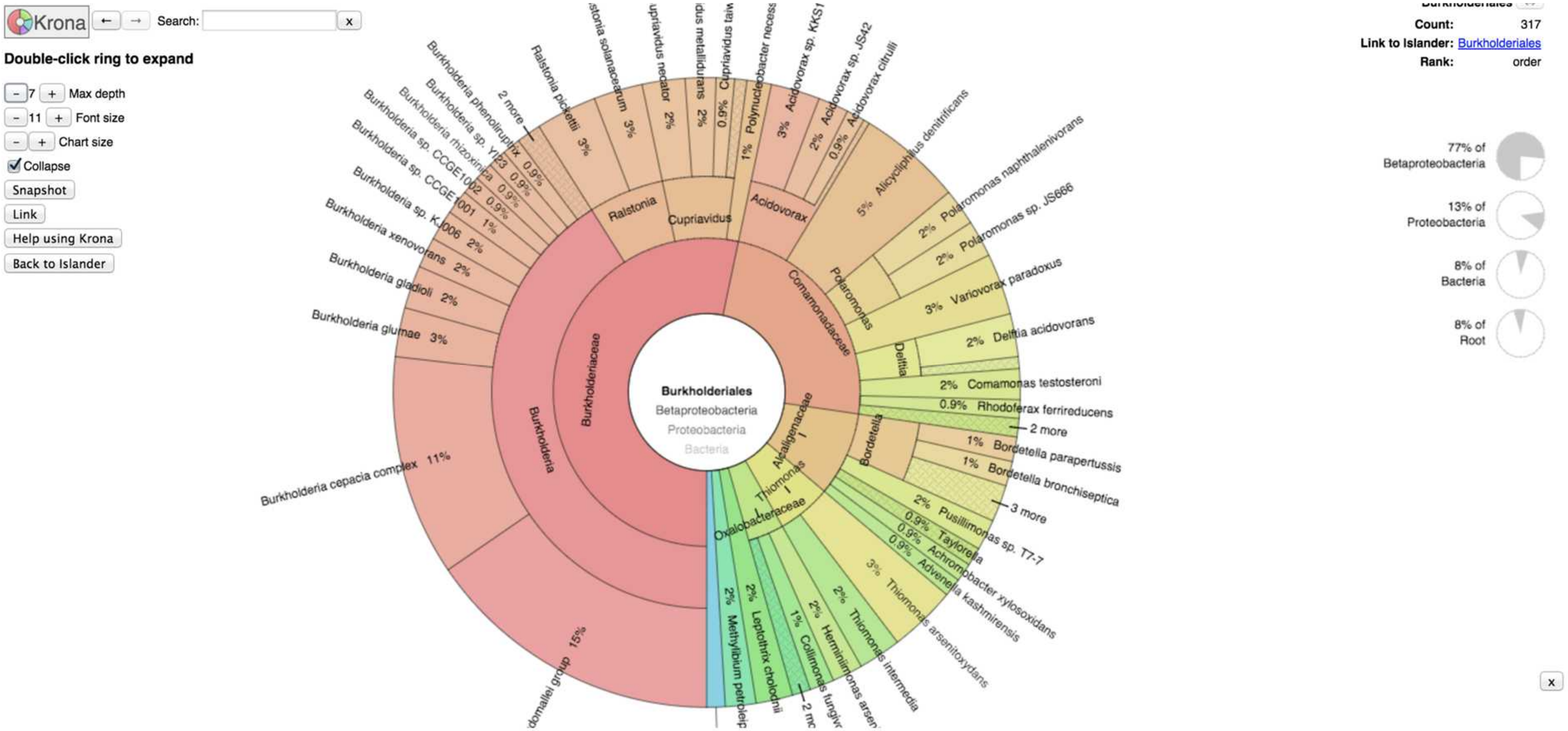
Islander — Database of Genomic Islands.

Page Contact
Kelly Williams
kpwilli@sandia.gov
(925) 294-4730

Corey Hudson
cmhudso@sandia.gov
(925) 294-1575

Related Links
Sandia
Bioinformatics
Sandia Biology

Browse by Taxonomy



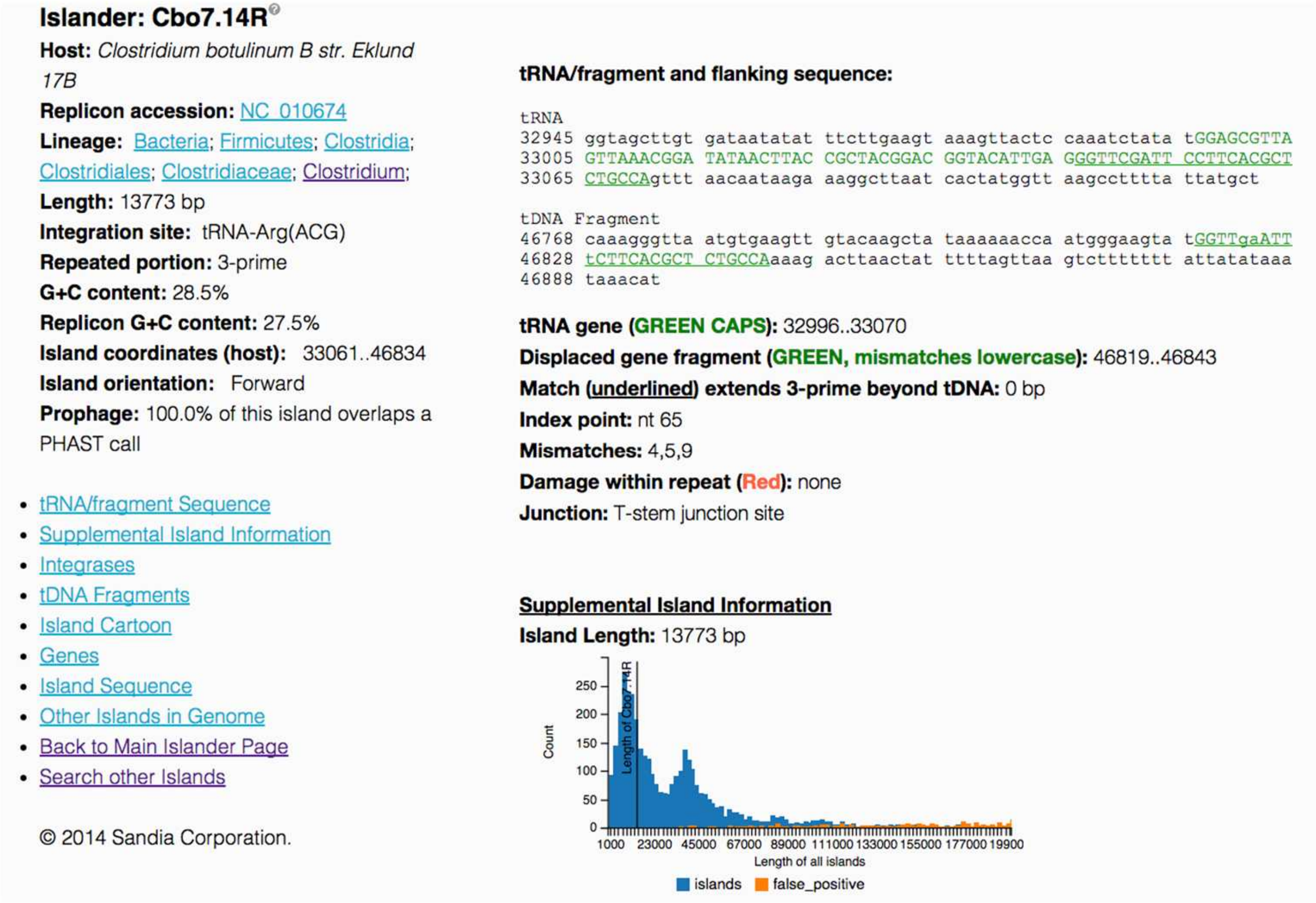
Burkholderiales

Count: 317

Link to Islander: Burkholderiales

Rank: order

Rich set of genomic island features



Islander: Cbo7.14R

Host: *Clostridium botulinum B str. Eklund 17B*

Replicon accession: NC_010674

Lineage: Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium;

Length: 13773 bp

Integration site: tRNA-Arg(ACG)

Repeated portion: 3-prime

G+C content: 28.5%

Replicon G+C content: 27.5%

Island coordinates (host): 33061..46834

Island orientation: Forward

Prophage: 100.0% of this island overlaps a PHAST call

- tRNA/fragment Sequence
- Supplemental Island Information
- Integrases
- tDNA Fragments
- Island Cartoon
- Genes
- Island Sequence
- Other Islands in Genome
- Back to Main Islander Page
- Search other Islands

© 2014 Sandia Corporation.

tRNA/fragment and flanking sequence:

```
tRNA
32945 ggtagcttgt gataatatat ttcttgaagt aaagttactc caaatctata tGGAGCGTTA
33005 GTTAAACGGA TATAACTTAC CGCTACGGAC GGTACATTGA GGGTTCGATT CCTTCACGCT
33065 CTGCCAgttt aacaataaga aaggcttaat cactatggtt aagcctttta ttatgct

tDNA Fragment
46768 caaagggtta atgtgaagtt gtacaagcta taaaaaacca atgggaagta tGGTTgaATT
46828 tCTTCACGCT CTGCCAaaag acttaactat ttttagttaa gtcttttttt attatataaa
46888 taaacat
```

tRNA gene (GREEN CAPS): 32996..33070

Displaced gene fragment (GREEN, mismatches lowercase): 46819..46843

Match (underlined) extends 3-prime beyond tDNA: 0 bp

Index point: nt 65

Mismatches: 4,5,9

Damage within repeat (Red): none

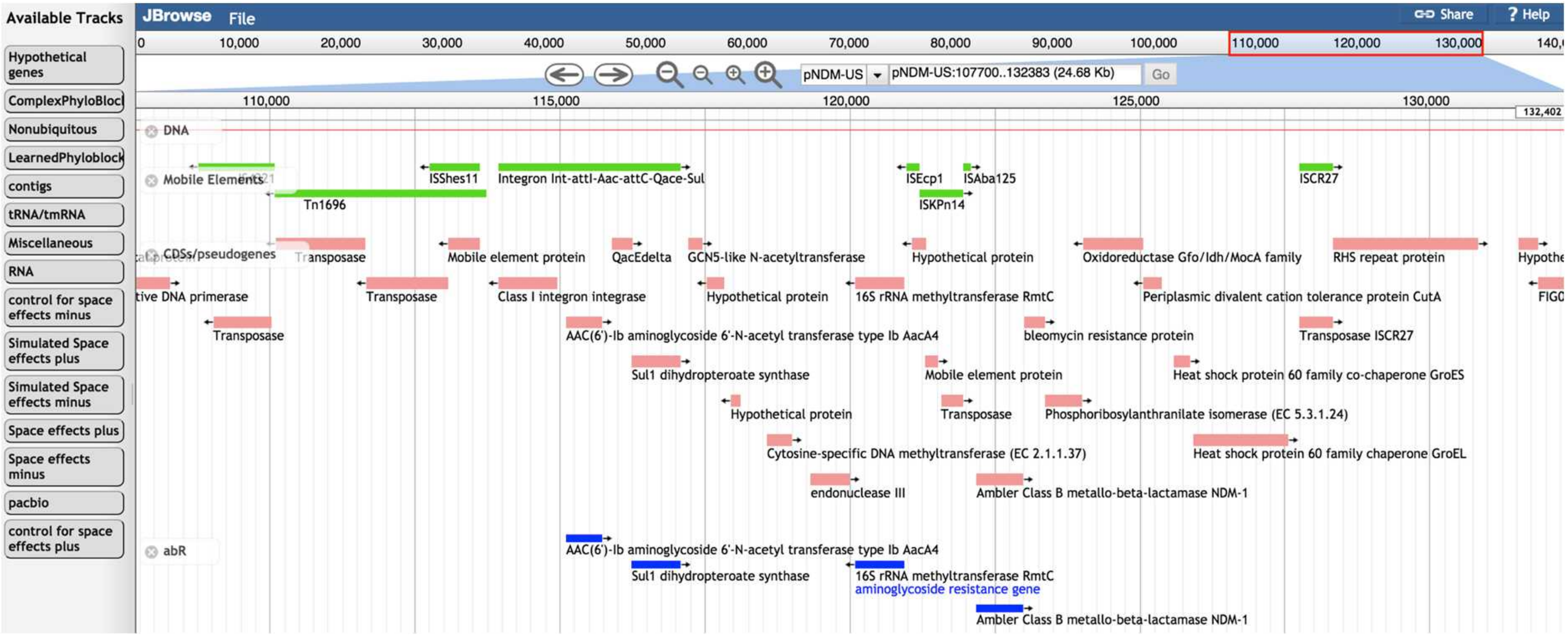
Junction: T-stem junction site

Supplemental Island Information

Island Length: 13773 bp

Klebsiella pneumoniae BAA-2146 Genome Browser

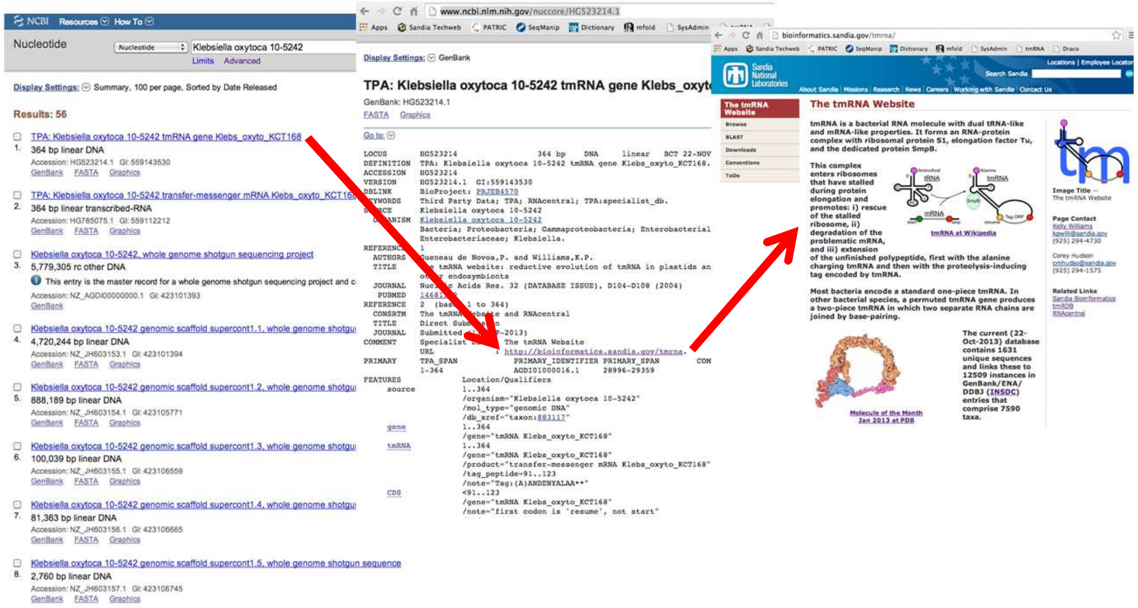
- Provides a full set of mobility and antibiotic resistance genes.



Resource for the CRE community

- Additionally allows users to input custom datasets
- Compare new genomes with the resistome and mobilome of a very well-annotated and well-assembled CRE genome.
- Further developing tools for experimental verification of mobile genetic elements that relied heavily of this browser, but made no preexisting bioinformatic assumptions.

tmRNA Website Increased Sandia’s Bioinformatics Presence



NCBI Resources How To

Nucleotide

Nucleotide: Klebsiella oxytoca 10-5242

Limits Advanced

Display Settings: Summary, 100 per page, Sorted by Date Released

Results: 56

1. TPA: Klebsiella oxytoca 10-5242 tmRNA gene Klebs_oxyto_KCT168
   364 bp linear DNA
   Accession: HG523214.1 GI: 559143530
   GenBank FASTA Graphics
2. TPA: Klebsiella oxytoca 10-5242 transfer-messenger mRNA Klebs_oxyto_KCT16
   364 bp linear transcribed-RNA
   Accession: HG785075.1 GI: 559112212
   GenBank FASTA Graphics
3. Klebsiella oxytoca 10-5242, whole genome shotgun sequencing project
   5,779,305 rc other DNA
   This entry is the master record for a whole genome shotgun sequencing project and c
   Accession: NZ_AGDI00000000.1 GI: 423101393
   GenBank
4. Klebsiella oxytoca 10-5242 genomic scaffold supercont1.1, whole genome shotgu
   4,720,244 bp linear DNA
   Accession: NZ_JH603153.1 GI: 423101394
   GenBank FASTA Graphics
5. Klebsiella oxytoca 10-5242 genomic scaffold supercont1.2, whole genome shotgu
   888,189 bp linear DNA
   Accession: NZ_JH603154.1 GI: 423105771
   GenBank FASTA Graphics
6. Klebsiella oxytoca 10-5242 genomic scaffold supercont1.3, whole genome shotgu
   100,039 bp linear DNA
   Accession: NZ_JH603155.1 GI: 423106559
   GenBank FASTA Graphics
7. Klebsiella oxytoca 10-5242 genomic scaffold supercont1.4, whole genome shotgu
   81,363 bp linear DNA
   Accession: NZ_JH603156.1 GI: 423106665
   GenBank FASTA Graphics
8. Klebsiella oxytoca 10-5242 genomic scaffold supercont1.5, whole genome shotgun sequence
   2,760 bp linear DNA
   Accession: NZ_JH603157.1 GI: 423106745
   GenBank FASTA Graphics

www.ncbi.nlm.nih.gov/nuccore/HG523214.1

Display Settings: GenBank

TPA: Klebsiella oxytoca 10-5242 tmRNA gene Klebs_oxyt

GenBank: HG523214.1

FASTA Graphics

Go to:

```
LOCUS       HG523214                 364 bp    DNA     linear   BCT 22-NOV
DEFINITION  TPA: Klebsiella oxytoca 10-5242 tmRNA gene Klebs_oxyto_KCT168.
ACCESSION   HG523214
VERSION     HG523214.1  GI:559143530
DBLINK      BioProject: PRJEB4570
KEYWORDS    Third Party Data; TPA; RNAcentral; TPA:specialist_db.
SOURCE      Klebsiella oxytoca 10-5242
  ORGANISM  Klebsiella oxytoca 10-5242
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacterial
            Enterobacteriaceae; Klebsiella.
REFERENCE   1
  AUTHORS   Gueneau de Novoa,P. and Williams,K.P.
  TITLE     The tmRNA website: reductive evolution of tmRNA in plastids an
            other endosymbionts
  JOURNAL   Nucleic Acids Res. 32 (DATABASE ISSUE), D104-D108 (2004)
   PUBMED   14681[illegible]
REFERENCE   2  (bases 1 to 364)
  CONSRTM   The tmRNA Website and RNAcentral
  TITLE     Direct Submission
  JOURNAL   Submitted ([illegible]-SEP-2013)
COMMENT     Specialist DB     : The tmRNA Website
            URL               : http://bioinformatics.sandia.gov/tmrna.
PRIMARY     TPA_SPAN            PRIMARY_IDENTIFIER PRIMARY_SPAN        COM
            1-364               AGDI01000016.1     28996-29359
FEATURES             Location/Qualifiers
     source          1..364
                     /organism="Klebsiella oxytoca 10-5242"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:883117"
     gene            1..364
                     /gene="tmRNA Klebs_oxyto_KCT168"
     tmRNA           1..364
                     /gene="tmRNA Klebs_oxyto_KCT168"
                     /product="transfer-messenger mRNA Klebs_oxyto_KCT168"
                     /tag_peptide=91..123
                     /note="Tag:(A)ANDENYALAA**"
     CDS             <91..123
                     /gene="tmRNA Klebs_oxyto_KCT168"
                     /note="first codon is 'resume', not start"
```

bioinformatics.sandia.gov/tmrna/

Sandia National Laboratories

Locations | Employee Locator

Search Sandia

About Sandia | Missions | Research | News | Careers | Working with Sandia | Contact Us

The tmRNA Website
- Browse
- BLAST
- Downloads
- Conventions
- ToDo

The tmRNA Website

tmRNA is a bacterial RNA molecule with dual tRNA-like and mRNA-like properties. It forms an RNA-protein complex with ribosomal protein S1, elongation factor Tu, and the dedicated protein SmpB.

This complex enters ribosomes that have stalled during protein elongation and promotes: i) rescue of the stalled ribosome, ii) degradation of the problematic mRNA, and iii) extension of the unfinished polypeptide, first with the alanine charging tmRNA and then with the proteolysis-inducing tag encoded by tmRNA.

tmRNA at Wikipedia

Most bacteria encode a standard one-piece tmRNA. In other bacterial species, a permuted tmRNA gene produces a two-piece tmRNA in which two separate RNA chains are joined by base-pairing.

Molecule of the Month Jan 2013 at PDB

The current (22-Oct-2013) database contains 1631 unique sequences and links these to 12509 instances in GenBank/ENA/DDBJ (INSDC) entries that comprise 7590 taxa.

Image Title -- The tmRNA Website

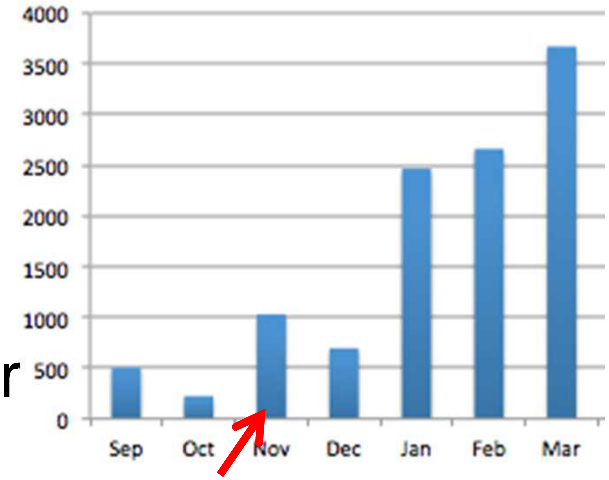
Page Contact
Kelly Williams
kpwilli@sandia.gov
(925) 294-4730

Corey Hudson
cmhudso@sandia.gov
(925) 294-1575

Related Links
Sandia Bioinformatics
tmRDB
RNAcentral

For any genome, top GenBank hit is our Third Party Annotation of tmRNA, linking to our Sandia tmRNA website --- Published with RNACentral

Web Hits



TPA goes live at NCBI

Distributing Sandia's Bioinformatics Software

- Five new genome assembly tools:
  - readStepper
  - extendMulti
  - getMates
  - denovo
  - indexReads
- Work to finish the assembly of partially assembled genomes.
- **Islander** --- produces nucleotide precision genomic islands.
- **tFind** --- identifies tDNA from raw sequence.
- **RapTOR** --- Rapid Threat Organism Recognition: Biodetection System for Public Health and Biodefense