HIV Database Workshop

www.hiv.lanl.gov

seq-info@lanl.gov

Presenters: Brian Foley, Elizabeth-Sharon Fung

Database PIs:
Karina Yusim (Immunology), Brian Foley (Sequence)

Additional database staff:
Werner Abfalterer, Kumkum Ganguly, Elena Giorgi, Elena Romero, James Szinger, Hyejin Yoon

Contract Officer Representative: Anjali Singh, NIAID, NIH

Theoretical Biology and Biophysics, T-6
Los Alamos National Laboratory

HIV DB Workshop slides:
https://tinyurl.com/2020-IEDB

Los Alamos HIV Databases

- The 2020 update includes all sequences through Dec 2019

- HIV Immunology Database: Searchable annotated T cell epitopes and Antibody entries

  10,691 CD8+ epitope entries from 1,353 papers

  1,609 CD4+ epitope entries from 389 papers

  3,579 distinct monoclonal antibody entries

  Neutralization data accessible through CATNAP
  - For 427 Abs, 40 antibody mixtures, and 40 polyclonal sera

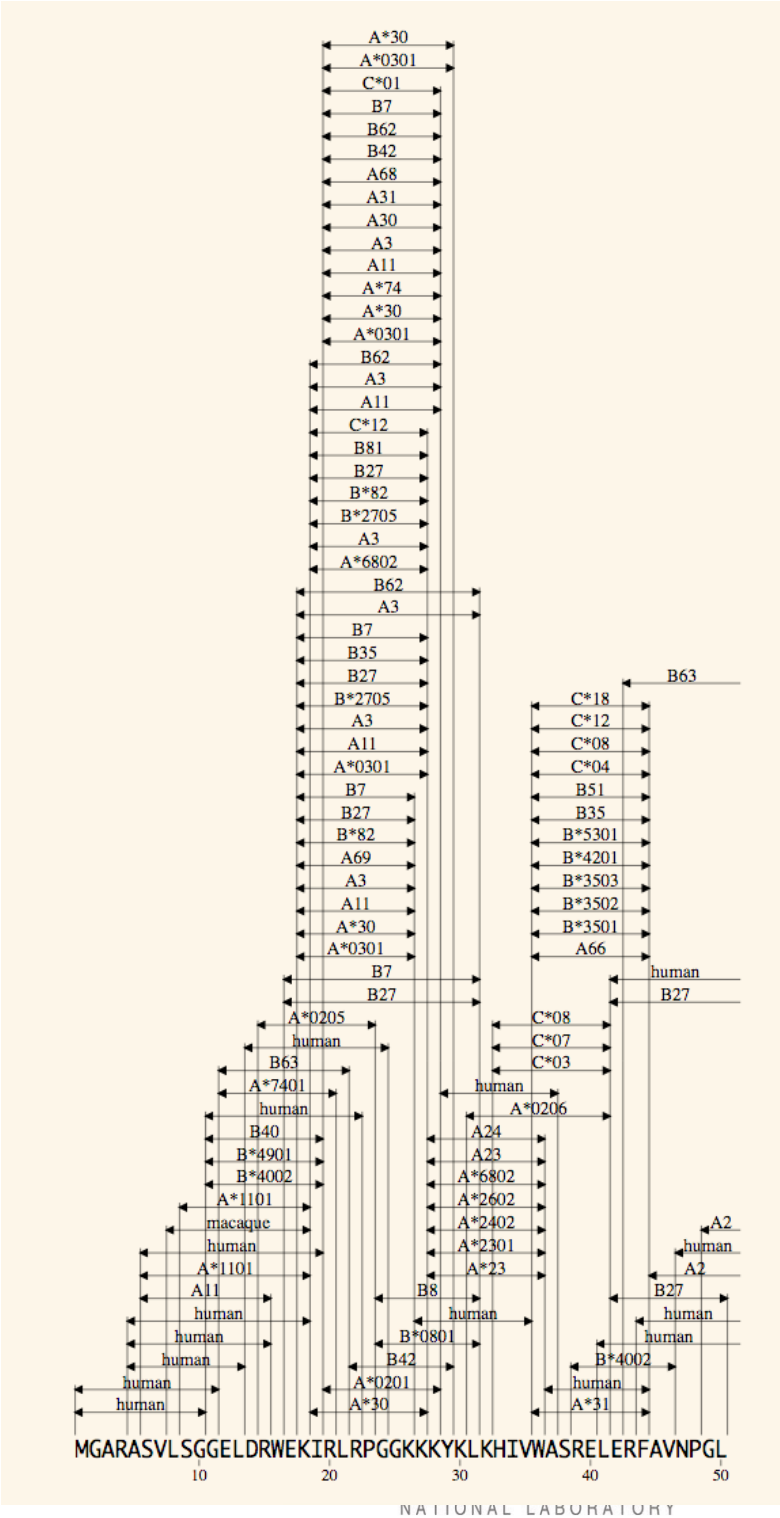
  1191 pseudoviruses tested, including 956 with sequences

- 68 bioinformatics tools with simple web interfaces
- Links to external tools, including IEDB's
- multiple search interfaces

  Tools split ~ 1/3rd between HIV-specific and 2/3rds general-use

- HIV Sequence Database: Over 935,458 searchable annotated HIV/SIV sequences total.

  Stored metadata enables us to provide custom made alignments or pre-made 1-sequence-per-person alignments.



Los Alamos HIV Databases

- Database update:

The 2020 update includes all sequences through Dec 2019

- HIV Immunology Database: Searchable annotated T cell epitopes and Antibody entries

10,691 CD8+ epitope entries from 1,353 papers

1,609 CD4+ epitope entries from 389 papers

3,579 distinct monoclonal antibody entries

Neutralization data accessible through CATNAP

  - For 427 Abs, 40 antibody mixtures, and 40 polyclonal sera
  - 1191 pseudoviruses tested, including 956 with sequences

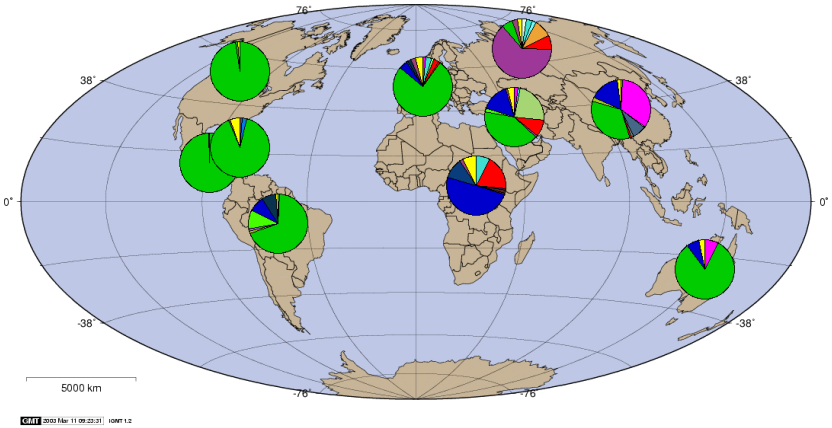
- 68 bioinformatics tools with simple web interfaces
- multiple search interfaces

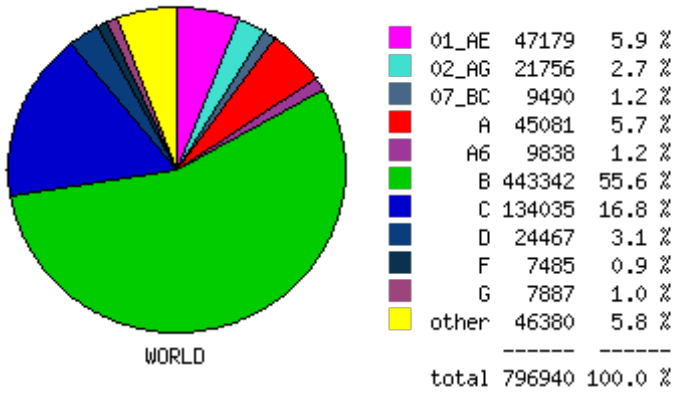
Tools split ~ 33% between HIV-specific and 66% general-use

- HIV Sequence Database: Over 935,458 searchable annotated HIV/SIV sequences total.

Stored metadata enables us to provide custom made alignments or pre-made 1-sequence-per-person alignments.

Global Clade and CRF distribution





Los Alamos NATIONAL LABORATORY

Integration of HIV Sequence and Immunology databases

- Los Alamos HIV Database: the first pathogen-specific database
  - HIV Sequence Database – founded in 1986 by G. Myers
  - HIV Immunology Database – founded in 1994 by B. Korber

- Integration of HIV sequence and immunological data via multiple tools, for example:

  HIV **Genome Browser** provides an interactive detailed view of the HIV genome or proteome with HIV sequence variability, functional domains and antibody and T cell epitopes marked by genome position

  CATNAP superimposes Ab neutralization data with the virus data, and links to structures, germline V/D/J genes, Ab sequences, Ab contact residues, Env alignments, positions associated with neutralization sensitivity …

  AnalyzeAlign shows the diversity and HIV variability of epitopes

  Multiple tools tap into the **Donors (or patient) database**, containing available donor HIV sequences, Ab sequences, monoclonal and polyclonal Ab data, HLAs, and T-cell epitopes

Los Alamos
NATIONAL LABORATORY

Beyond HIV

- **Twenty two of our computational tools (33%) are strictly HIV-specific. The remaining 66% are partially or fully applicable to other organisms**

- A striking example of successful extension beyond HIV is Mosaic/Epigraph vaccine design:
  - Rabies in bats (Stading *et al*, Plos Negl Trop Dis, 2017)
  - Filoviruses (Theiler *et al*, Sci Rep. 2016, Fenimore, PLoS One, 2012)
  - Chlamydia trachomatis (Badamchi-Zadeh *et al*, Front Immunol, 2016)
  - Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) in pigs (Cui *et al*, Vaccine reports, 2016)
  - Hepatitis C (Yusim *et al*, Clin Vaccine Immunol, 2013)
  - Foot-and-Mouth Disease in livestock (Devendra *et al*, in preparation)
  - Hepatitis B (Yusim *et al*, in preparation)
  - **... and the HIV-1 mosaic designs moved into Phase III human trials**

- The database structure and tools are transferrable to other pathogens. We have created several pathogen databases prototyped on the HIV database, and translating multiple tools: (***https://www.hiv.lanl.gov/content/otherviruses.html***):

  HCV Sequence (Kuiken et al, Nucleic Acid Res, 2008) and Immunology (Yusim et al, Appl Bioinformatics, 2005) Databases

  Hemorrhagic Fever Viruses (HFV) Sequence Database (80 viral species, found in 10 different genera comprising five different families: arena-, bunya-, flavi-, filo- and togaviridae) (Kuiken et al, Nucleic Acid Res, 2012)

  Filovirus Sequence and Immunology Database (Yusim et al, Database, 2016) (hfv.lanl.gov)

  COVID-19 Genome Analysis Pipeline *

  Because of lack of individual funding, only the sequence portions of these databases are automatically updated

Los Alamos NATIONAL LABORATORY

HIV Immunology Database Workshop

- Today's Outline

Overview of the HIV Immunology and HIV Sequence Databases

T Lymphocytes

- T cell epitope data and search interface
- Peptide tools

Antibodies

- Neutralizing Antibody Resources

CATNAP

- neutralization exploration
- tailored for HIV but pathogen-agnostic
- Integration of Antibody and Sequence Data

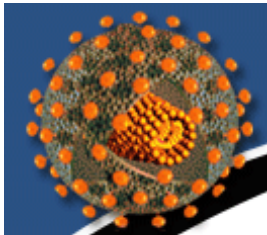
Sequence Database Organization

Sequence Database Search and Outputs

Sequence Database Alignments

Los Alamos NATIONAL LABORATORY

Antibody search interface and integration



HIV DATABASES

http://hiv.lanl.gov

The **HIV databases** contain comprehensive data on HIV genetic sequences and immunological epitopes. The website also gives access to a large number of tools that can be used to analyze and visualize these data. This project has been funded in whole or in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Interagency Agreement No. AAI12007-001-00000. Our content is reviewed by an Editorial Board.

SEQUENCE DATABASE ▸

IMMUNOLOGY DATABASE ▸

OTHER VIRUSES ▸

News:

Archived News ▸

HIV Molecular Immunology 2018-19

HIV Molecular Immunology 2018-19 is now available online. The PDF version is hypertext enabled and features clickable table-of-contents, indexes, references and links to external web sites. *14 September 2020*

Questions or comments? Contact us at seq-info@lanl.gov

Los Alamos NATIONAL LABORATORY

HIV Immunology Database Entries and Annotation

- **HIV T cell epitopes and Antibody data organization**

  T Cells (CTL and Helper epitopes)
  - One reference per entry, epitope/HLA combinations are often repeated
  - CTL and T-helper database organization is identical

  B Cells (Antibodies)
  - One entry for each monoclonal antibody
  - Many references per entry (> 800 for some well studied mAbs)

- **Descriptions of HIV T cell epitopes and Antibodies with associated data are harvested from regular periodic literature searches:**

  Epitope sequence, location, immunogen, vaccine details, subject details…

  Epitope Variants (escape, reduced binding, etc.)

  Host HLA or MHC, binding region, germline genes, etc

  Neutralizing Antibody Resources, contact residues, positions related to neutralization sensitivity or resistance, etc.

  Notes summarizing main findings

- **Multiple search interfaces and database products:**

  5 search interfaces for T cell epitopes, epitope variants and antibodies

  Computational tools for immunologists

  Epitope maps and summary tables that can also serve as search interfaces

  HLA typing and epitope mapping data sets

  Neutralizing antibody resources:
  - Neutralization, germline and antibody sequence data through CATNAP
  - Links to Germline Antibody Reconstruction tools
  - Search interface and a table for Ab contact residues, positions related to neutralization sensitivity or resistance, etc.
  - Assay protocols and neutralization serotype discovery data

Los Alamos
NATIONAL LABORATORY

Upcoming News: JSON Application Programming Interface for epitope data

- New tool will allow you to programmatically retrieve curated epitope data in JSON (JavaScript Object Notation) format

- Purpose
  - Alternative to HTML format presently available
  - Full contents of Immunology database available in structured form
  - Full power of search interface available, with data combinations possible
  - Fully documented via OpenAPI
  - Data once downloaded may be manipulated with user's choice of programming language
  - Tailor-made extraction of data, automated multiple searches possible

- **Example:**

Get the epitope sequence and notes for CTL record 42

```
wget -q -O-
'http://localhost/mojo/immunology/api/v1/epitope/ctl?id=42' \
    | jq '.epitopes|.[0]|(.epitope,.note)'

"GGKKKYKLK"
[
  {
    "note": "Study of an individual with partially defective
antigen processing.",
    "note_no": 0
  }
]
```

Created by: Jim Szinger
szinger@lanl.gov

Los Alamos NATIONAL LABORATORY

HIV Molecular Immunology Database: Tools & Links

https://www.hiv.lanl.gov/content/immunology/tools-links.html

Tools Produced by the Los Alamos HIV Databases

- Tools specific for HIV/SIV
- General use tools with some HIV/SIV-specific features
- General use tools

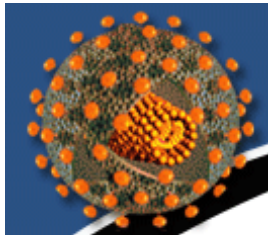
- CATNAP: Compile, Analyze and Tally NAb Panels Download or analyze neutralization data
- CombiNAber Predict the neutralization of combinations of antibodies
- HIV Genome Browser Display HIV genome and proteome
- QuickAlign Align amino acids or nucleotides against our alignments
- Analyze Align Show weblogos, calculate frequency by position, and find variants in an alignment
- Alignment Slicer Cut vertical slices from sequence alignments
- PeptGen Generate overlapping peptides for any protein
- PepMap Generate peptide maps in Fasta, HTML and PDF formats
- Motif Scan Scan alignments for HLA binding motifs
  - HLA genotype/serotype dictionary
  - HLA genotype/motif dictionary
  - HLA supertype dictionaries
- Hepitope Search for hopeful epitopes based on HLA enrichment
- HLA Frequency Analysis Tools Calculate HLA frequencies or HLA linkage disequilibrium in a population
- ELF Epitope location finder
- Sequence Locator Tool Find the location of any HIV/SIV sequence
- SeqPublish Produce pretty alignments for publication
- Heatmap Display a table of numbers using colors to represent the numerical values
- Epigraph Vaccine Suite Design and assess Epigraphs for vaccine design
- Mosaic Vaccine Suite Design and assess polyvalent protein sequences for T-cell vaccines
- N-Glycosite Find N-linked glycosylation sites
- Highlighter Highlight matches and mismatches in a set of aligned sequences
- Protein Feature Accent View 3D graphics of HIV proteins
- Variable Region Characteristics analyzes Env variable loops and reports length, glycosolations, and net charge
- Neutralization Index computes a tier-like score for neutralizing antibodies

- All Tools List of all software and tools in **both** the HIV sequence and immunology databases

LINKS TO EXTERNAL TOOLS : 25 Epitope Prediction; 6 Germline Ab Reconstruction; 10 Immunology; 2 Vaccine Studies; and more

Los Alamos NATIONAL LABORATORY

Antibody search interface and integration



HIV DATABASES

http://hiv.lanl.gov

The **HIV databases** contain comprehensive data on HIV genetic sequences and immunological epitopes. The website also gives access to a large number of tools that can be used to analyze and visualize these data. This project has been funded in whole or in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Interagency Agreement No. AAI12007-001-00000. Our content is reviewed by an Editorial Board.

SEQUENCE DATABASE ▸

IMMUNOLOGY DATABASE ▸

OTHER VIRUSES ▸

News:

Archived News ▸

HIV Molecular Immunology 2018-19

HIV Molecular Immunology 2018-19 is now available online. The PDF version is hypertext enabled and features clickable table-of-contents, indexes, references and links to external web sites. *14 September 2020*

Questions or comments? Contact us at seq-info@lanl.gov

Los Alamos NATIONAL LABORATORY

HIV Molecular Immunology Database

The HIV Molecular Immunology Database is an llection of HIV-1 cytotoxic and helper T-cell epitopes and antibody binding sites.

Search Interfaces

- CTL/CD8+ search
- T Helper/CD4+ search
- Antibody search
- CTL variant search
- T Helper variant search

- Search help
- Variant search help

Antibody Search

Multiple ways to database products and tools

Database Products

- All Database products and publications
- Epitope maps
- Epitope tables
- Epitope alignments
- T cell epitope variants and escape mutations
- Neutralizing antibody resources & CATNAP
- The HIV Molecular Immunology Compendium
- About the HIV Molecular Immunology Database
- How to cite this database
- Frequently-asked Questions (FAQ)

T cell epitope variants and escape mutations

Neutralizing Antibody Resources

Tools and Data Sets

- Tools & Links for immunologists
- SIV Epitopes (PDF) review article summarizing known SIV epitopes
- Identifying HLA-Associated Polymorphisms in HIV-1 (PDF) review article summarizing HIV polymorphism associated with escape mutations. Also a table of polymorphisms.
- HLATEM HLA Typing and Epitope Mapping Data Sets
- Standardized Assessments of Neutralizing Antibodies for HIV/AIDS Vaccine Development Assay protocols from Duke Central Reference Laboratory

https://www.hiv.lanl.gov/content/immunology/index.html

Los Alamos NATIONAL LABORATORY

HIV Molecular Immunology Database

The HIV Molecular Immunology Database is an annotated, searchable collection of HIV-1 cytotoxic and helper T-cell epitopes and antibody binding sites.

Search Interfaces

- CTL/CD8+ search
- T Helper/CD4+ search
- Antibody search
- CTL variant search
- T Helper variant search

- Search help
- Variant search help

Database Products

- All Database products and publications
- Epitope maps
- Epitope tables
- Epitope alignments
- T cell epitope variants and escape mutations
- Neutralizing antibody resources & CATNAP
- The HIV Molecular Immunology Compendium
- About the HIV Molecular Immunology Database
- How to cite this database
- Frequently-asked Questions (FAQ)

Epitope Tables

These tables summarize the epitopes from our database. HIV-1 epitope data may also be obtained in the form of downloadable maps or alignments.

- CTL epitopes
- Best-defined ("A-list") CTL epitopes
- CTL epitope variants and escape mutations
- T-helper epitopes
- T Helper epitope variants and escape mutations

- Antibody epitopes
- Best Neutralizing Antibodies
- Antibody-Dependent Cell-Mediated Cytotoxicity (ADCC)
- Antibody index by name
- Antibody index by binding type

- SIV epitopes

- Neutralizing antibody resources

Epitope alignments: epitopes aligned to HIV subtype Reference sequences in Fasta format

Tools and Data Sets

- Tools & Links for immunologists
- SIV Epitopes (PDF) review article summarizing known SIV epitopes
- Identifying HLA-Associated Polymorphisms in HIV-1 (PDF) review article summarizing HIV polymorphism associated with escape mutations. Also a table of polymorphisms.
- HLATEM HLA Typing and Epitope Mapping Data Sets
- Standardized Assessments of Neutralizing Antibodies for HIV/AIDS Vaccine Development Assay protocols from Duke Central Reference Laboratory

Reactive peptide maps and tables (with HLA and other subject data) from several large-scale studies scanning HIV proteins.

https://www.hiv.lanl.gov/content/immunology/index.html

CTL/CD8+ Epitope Summary (B-list)

- A comprehensive list of all unique epitopes in the database (including with unknown HLA, boundaries not fully defined…)
- Similar lists for Helper epitopes and linear Ab binding sites
- Unlike epitope maps that show epitope locations, each epitope sequence is shown

Epitope	Protein	HXB2 Location	Subtype	Species	HLA
MGARASVLSG	p17	1-10	CRF01_AE	human	
ASVLSGGEL	p17	5-13	B	human	
ASILRGGKLDK	p17	5-15	C	human	
SVLSGGQLDR	p17	6-15	B	human	A11
LSGGELDRWEK	p17	8-18		macaque	
GELDRWEKI	p17	11-19	B	human	B*4002, B40
GQLDRWEKI	p17	11-19	B	human	
GKLDSWEKIRLR	p17	11-22	A, CRF01_AE, CRF02_AG	human	

www.hiv.lanl.gov/content/immunology/tables/ctl_summary.html

Best-defined CTL/CD8+ Epitope Summary (A-list)

- Experimentally validated optimal epitopes with known HLA presenting molecules
- Defined/curated by Christian Brander and colleagues

Epitope	Protein	HXB2 Location	Subtype	Species	HLA
GELDRWEKI	p17	11-19		human	B*4002
KIRLRPGGK	p17	18-26		human	A*0301
IRLRPGGKK	p17	19-27	B	human	B*2705
RLRPGGKKK	p17	20-28		human	A*0301
RLRPGGKKKY	p17	20-29	B	human	A*0301
GGKKKYKLK	p17	24-32	B	human	B*0801
KYKLKHIVW	p17	28-36	B	human	A*2402
HLVWASREL	p17	33-41		human	Cw*0804

www.hiv.lanl.gov/content/immunology/tables/optimal_ctl_summary.html

Epitope variants and escape mutations

- Experimental epitope variants from the literature
  - Search interfaces
  - Summary tables (~3500 CTL epitope variants)
- HLA associated HIV polymorphisms (Zabrina Brumme, Bruce Walker)
  - Database review and a table

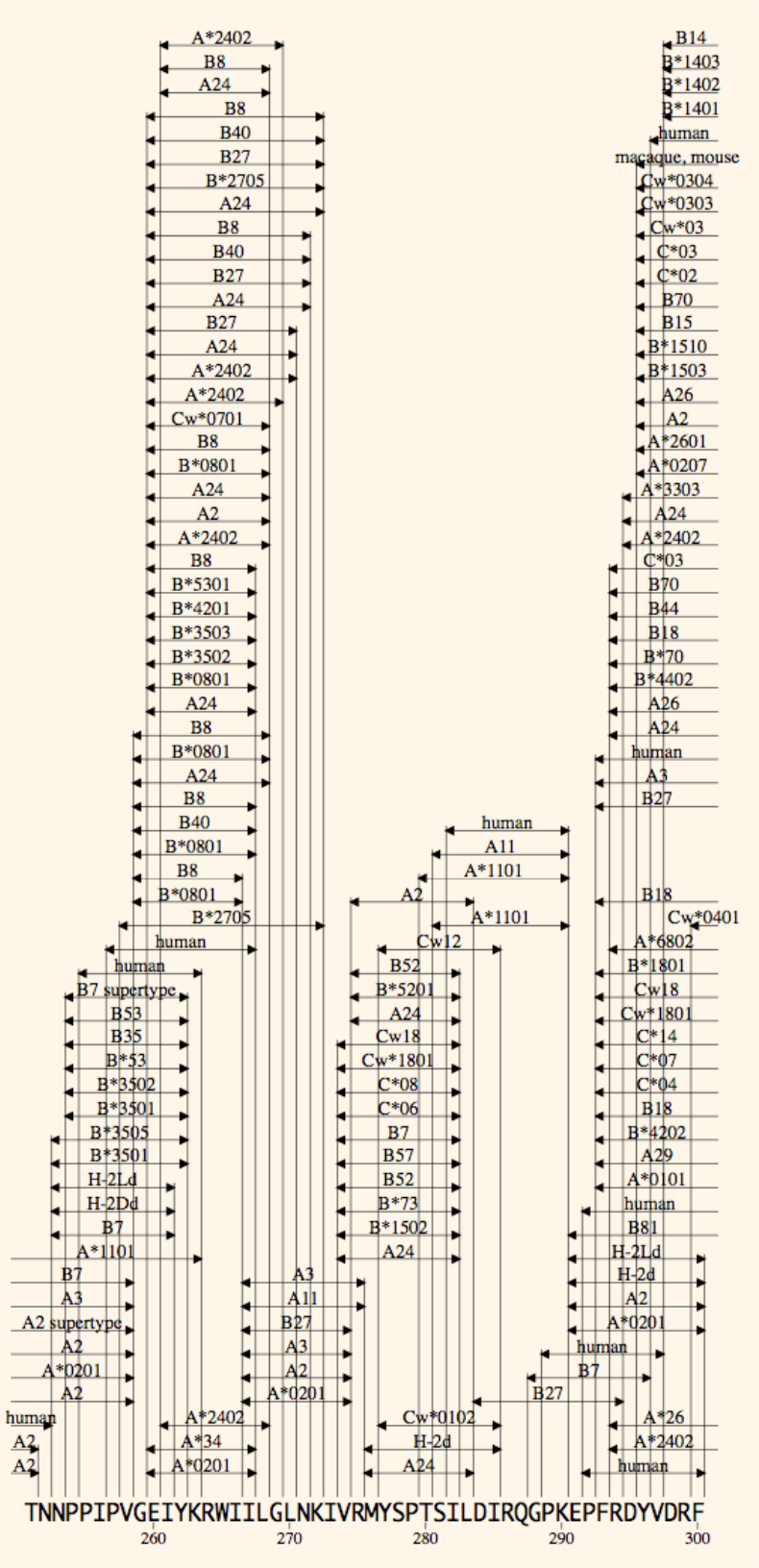
www.hiv.lanl.gov/content/immunology/pdf/2010/escape_article_supplement.html

Los Alamos NATIONAL LABORATORY

p17 CTL/CD8+ Epitope Map

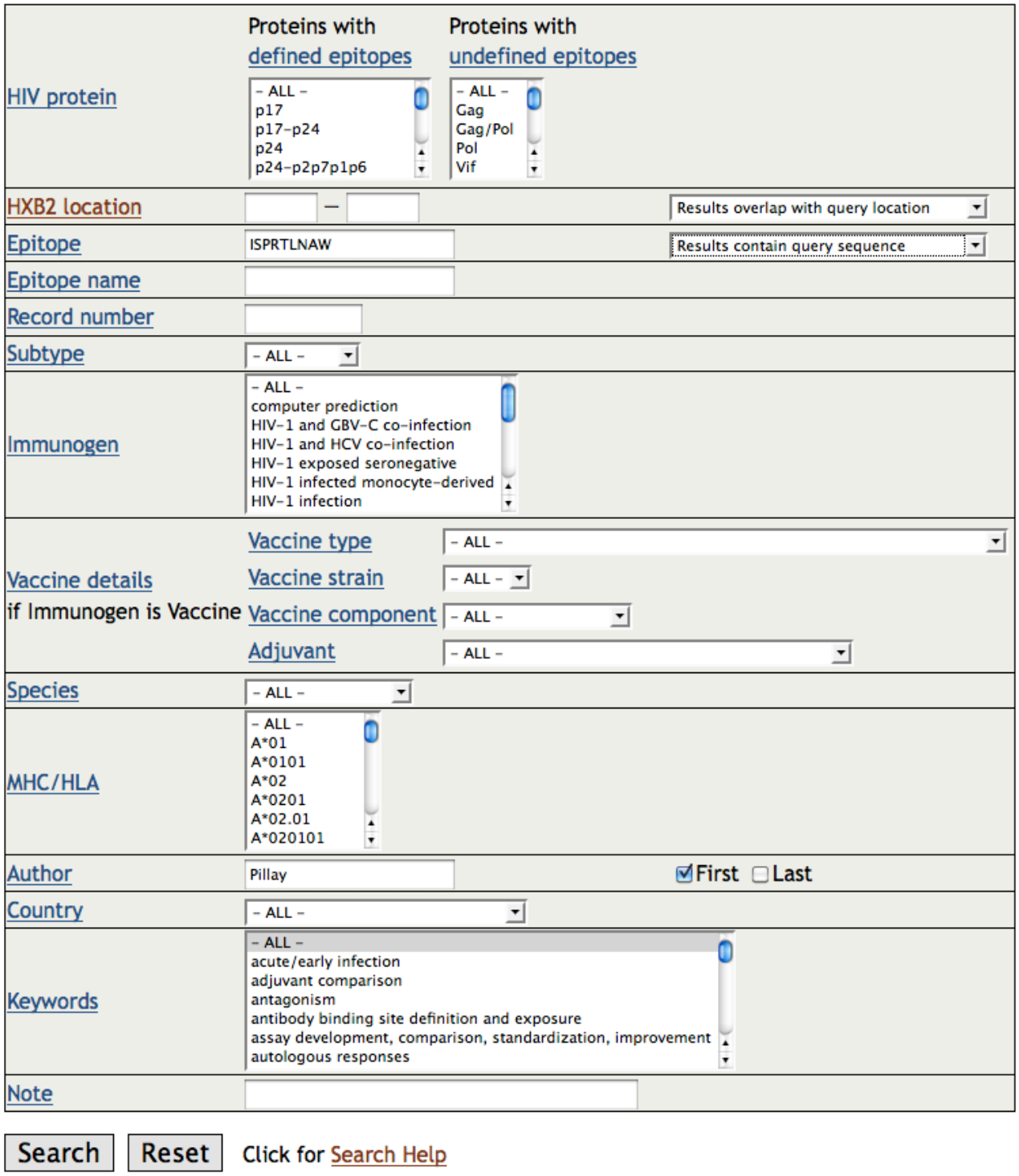
- Epitopes up to 14 aa long are mapped on HXB2
- HXB2 sequence may differ
- Epitopes with identical boundaries and HLA fields are included in the maps only once
- The epitope maps are interactive!

  Clicking on an epitope leads to the epitope entry



www.hiv.lanl.gov/content/immunology/maps/maps.html

CTL/CD8+ Search (*www.hiv.lanl.gov/content/immunology/ctl_search*)



- Search by HIV protein, Epitope Sequence, Subtype, Immunogen, Vaccine Details, Species, presenting MHC/HLA, Author, Country, Keywords
- Search on epitope location and find fuzzy matches, overlaps and embedded epitopes
- Search examples:

  Example:

  - SLYNTVATL – 285 entries
  - Narrow the search with keyword “escape” – 35 entries

Search for ISPRTLNAW
With the first author Pillay

Los Alamos NATIONAL LABORATORY

Search CTL/CD8+ T-Cell Epitope Database

Found 1 matching record:

Displaying record number 53832

Immunological, virological, and epidemiological contexts:

Link to Epitope Maps → p24 Epitope Map

Link to Epitope Alignment → Epitope Alignment

Show epitope variants

Variant details with annotator's notes

HXB2 Location	p24(15-23)
Author Location	Gag(147-155)
Epitope	ISPRTLNAW
Subtype	C
Species (MHC/HLA)	human(B57)
Immunogen	HIV-1 infection
Donor MHC/HLA	A*3001, A*66, B*4201, B*5802, Cw*0602, Cw*1701; A*66, A*68, B*57, B*5802, Cw*0602, Cw*0701
Country	South Africa
Experimental methods	CD8 T-cell Elispot - IFNy
Keywords	epitope processing, responses in children, mother-to-infant transmission, escape, acute/early infection

Notes

- HIV-specific CTLs in infants were shown to be able to select for viral escape variants early in life, despite a lack of escape with the same CTL specificity in the mother. Infant CTL responses may be compromised by transmission of escape variants that arose in the mother and also those that arose in the father, if the father was the source of the mother's infection.
- ISPRTLNAW is the C consensus form of the epitope and was the autologous form in the mother, and was transmitted to her infant. By 33 weeks a new dominant form of the epitope had emerged in the infant, mSPRTLNAW, and two additional variants had arisen, one with a substitution proximal to the epitope, plSPRTLNAW, and lSPRTLNAW.

References

Pillay2005 Thillagavathie Pillay, Hua-Tang Zhang, Jan W. Drijfhout, Nicola Robinson, Helen Brown, Munira Khan, Jagadesa Moodley, Miriam Adhikari, Katja Pfafferott, Margaret E. Feeney, Anne St. John, Edward C. Holmes, Hoosen M. Coovadia, Paul Klenerman, Philip J. R. Goulder, and Rodney E. Phillips. Unique Acquisition of Cytotoxic T-Lymphocyte Escape Mutants in Infant Human Immunodeficiency Virus Type 1 Infection. *J. Virol.*, 79(18):12100-12105, Sep 2005. PubMed ID: 16140787. Show all entries for this paper.

Additional information provided in the entry:

- Location, Donor MHC/HLA, experimental methods, Notes
- Link to all entries for a reference
- PubMed links to papers
- Link to Epitope Maps
- Link to Epitope Alignment (aligned to large set of seq.)
- Epitope variants if studied in the paper

Los Alamos NATIONAL LABORATORY

Epitope Alignments

Genome map:

Query location(s) shown as colored bar(s) in map.

Summary & analysis:

Query: epitope

Query sequence	ISPRTLNAW
Query length	9
HXB2 Location	genome: 1228→1254, region: Gag 147→155
Alignment used	LANL HIV1 Gag Amino acid Filtered web

Summarize All | Summarize By Subtype (major subtype only) | Find Other Matches | Submit

Alignment slice:

Download alignment below in format fasta

"-" = identity to query sequence
"." = gap in sequence
"Red name" = perfect identity to query sequence

```
                                        epitope ISPRTLNAW
                    B.FR.83.HXB2_LAI_IIIB_BRU.K03455 ---------
          A1.CA.x.BCCFE_HOMER_HIV_GAG_3062.EU242119 L--------
                      A1.CD.02.02CD_KTB035.AM000055 ---------
                        A1.CD.97.97CD_KCC2.AM000053 L--------
                       A1.CD.97.97CD_KTB13.AM000054 L--------
           A1.CH.03.HIV_CH_BID_V3538_2003.JQ403028 L--------
                            A1.CH.04.pBV23.KJ689262 L--------
                            A1.CH.05.pBV26.KJ689264 F--------
                            A1.CH.08.pBV20.KJ689259 L--------
                            A1.CH.09.pBV32.KJ689270 M--------
                            A1.CH.10.pBV17.KJ689256 M--------
                            A1.CH.11.pBV13.KJ689253 ---------
                            A1.CH.11.pBV22.KJ689261 L--------
                            A1.CH.11.pBV48.KJ689279 L--------
                            A1.CH.12.pBV58.KJ689285 L--------
                             A1.CM.06.BS02.JX244900 L--------
                            A1.CM.07.46_10.KP718918 L--------
                             A1.CM.07.BS10.JX244906 L--------
                           A1.CM.08.886_24.KP718928 L--------
                          A1.CN.00.00CNLN14.EF122512 V--------
                            A1.CY.04.CY009.EU673416 L--------
                            A1.CY.05.CY012.EU673418 L--------
                            A1.CY.05.CY021.FJ388892 -A-KA-EG-
                            A1.CY.05.CY051.FJ388903 L--------
```

Also available as a separate tool **QuickAlign**

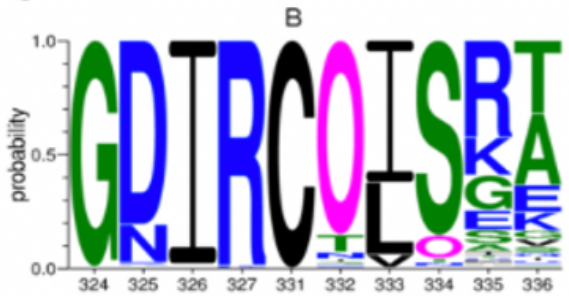
www.hiv.lanl.gov/content/sequence/QUICK_ALIGNv2/QuickAlign.html

OR **AnayzeAlign** Tool

Groups
[Download combined logs PDF EPS]
B A1 A2 C D F1 F2 G 01_AE 02_AG

Group B

Logo



Download: PNG PDF EPS

Submit alignment to get variant positions, Position frequencies, and WebLogos.
* Discontinuous positions allowed.

Los Alamos NATIONAL LABORATORY

www.hiv.lanl.gov/content/sequence/ANALYZEALIGN/analyze_align.html

Variant details

Displaying record number 53832

HXB2 Location	p24(15-23)	p24 Epitope Map
Epitope	ISPRTLNAW	Epitope Alignment
	mSPRTLNAW	escape documented in this paper
Variants	lSPRTLNAW	diminished response
	p|lSPRTLNAW	not determined
Species (MHC/HLA)	human(B57)	

Link back to epitope entry

Variant Details

Showing all 3 variants.

Variant ID.	1413
Epitope Seq.	ISPRTLNAW
Variant Seq.	mSPRTLNAW
Mutations	I/M
Epitope Location	I1M
HXB2 Location	I15M
Mutation Type	**E**: escape documented in this paper
Method	CD8 T-cell Elispot - IFNy, Sequence
Note	This is de novo variant seen in infant by week 33 of age. The index peptide was recognized, but not the variant.

Variant ID.	1414
Epitope Seq.	ISPRTLNAW
Variant Seq.	lSPRTLNAW
Mutations	I/L
Epitope Location	I1L
HXB2 Location	I15L
Mutation Type	**DR**: diminished response
Method	CD8 T-cell Elispot - IFNy, Sequence

Mutation type

Note describing why the variant was designated as a particular mutation type

Mutation type examples:

- ❑ E escape
- ❑ IE inferred escape
- ❑ DR diminished response
- ❑ SF susceptible form
- ❑ etc…

Los Alamos NATIONAL LABORATORY

Antibody search example: subject CH505

Patient Detail

Patient Code	Donor CH505 (703010505)
Patient Sex	Male
Risk Factor	Heterosexual (SH)
Infection Country	MW
Infection City	
Infection Year	2008
HLA Type	A*30, A*30, B*4202, B*570301, Cw*17, Cw*18
Patient Ethnicity	African
Progression	
Species	human
Patient Note	African donor enrolled approximately 4 weeks after infection and followed for over 6 years. During this time viral load ranged from 14,460 to 847,279 copies/ml (median = 173,667 copies/ml), and CD4 counts ranged from 69 to 431 cells/mm3 (median = 294 cells/mm3). A single founder virus is estimated to have established HIV-1 clade C with development of autologous neutralizing antibodies at 14 weeks; Abs CH103,CH104,CH105,CH106 isolated 136 weeks post-infection. Antibody CH235 was isolated from the patient's week 41-peripheral blood memory B cells in culture.
CTL CD8+ Records	59059, 59060
T-Helper CD4+ Records	
Antibody Records	CH103 (2861), CH104 (2862), CH105 (2863), CH106 (2864), IA1 (3176), IA2 (3177), IA3 (3178), IA4 (3179), IA5 (3180), IA6 (3181), IA7 (3182), IA8 (3183), CH103 UCA (3184), CH235 (3185), CH236 (3186), CH239 (3187), CH240 (3188), CH241 (3189), CH186 (3190), CH187 (3191), CH188 (3192), CH200 (3193), DH151 (3234), DH228 (3235), CH235.9 (3291), CH235.12 (3292), CH243 (3374), CH244 (3375), CH245 (3376), CH247 (3377), CH248 (3378), 1AH92U (3380), CH235.7 (3381), CH235.10 (3382), CH235.11 (3383), CH235.13 (3384)
Sequence Database Patient ID Record	56552

CATNAP and Sequence DB: CH235.9

DATABASES SEARCH ALIGNMENTS TOOLS PUBLICATIONS GUIDES search site Search

Go to CATNAP main page

Antibody information

Number of antibodies: 1

Download heavy and light aa na sequences in Fasta

Download table below

Expand table below to show heavy and light chain sequences and sources for germline data

Antibody	Antibody binding type	Structure	Donor	Clonal lineage	Isolation paper	Neutralizing antibody feature	Heavy V (IGHV)	Heavy D (IGHD)	Heavy J (IGHJ)	Light V (IGKV or IGLV)	Light J (IGKJ or IGLJ)	Light chain type	GenSig analysis	Aliases	LANL comments
CH235.9	gp120 CD4BS	EMD-8080 EMD-8081 5F9O	Donor CH505	CH235	Bonsignori2016	Antibody-driven selection in donor CH505; Electrostatic interactions with D368	-46*01	3-10*01	4*02	3-15*01	1*01	K	IC_{50}	CH493	

Assay

Analyze assay data in CATNAP Submit

Number of data: 199

Download table below with additional virus info

Expand table below to show virus infomation

Antibody	Virus	Reference	IC50	Mean IC50	IC80	Mean IC80	ID50	Mean ID50
CH235.9	0013095_2_11	Bonsignori et al. Cell 165:449 (2016)	>50	UD				
CH235.9	001428_2_42	Bonsignori et al. Cell 165:449 (2016)	0.417	0.417				
CH235.9	0077_V1_C16	Bonsignori et al. Cell 165:449 (2016)	41.7	41.7				
CH235.9	00836_2_5	Bonsignori et al. Cell 165:449 (2016)	>50	UD				
CH235.9	0260_V5_C36	Bonsignori et al. Cell 165:449 (2016)	10.5	10.5				
CH235.9	0330_V4_C3	Bonsignori et al. Cell 165:449 (2016)	1.88	1.88				

Los Alamos NATIONAL LABORATORY

CATNAP and Sequence DB expanded view: CH235.9

HIV sequence database

DATABASES SEARCH ALIGNMENTS TOOLS PUBLICATIONS GUIDES search site Search

Go to CATNAP main page

Antibody information

Number of antibodies: 1

Download heavy and light aa na sequences in Fasta

Download table below

Collapse

Antibody	Antibody binding type	Structure	Donor	Clonal lineage	Isolation paper	Neutralizing antibody feature	Germline paper	Germline software & DB	Heavy V (IGHV)	Heavy D (IGHD)	Heavy J (IGHJ)	Heavy CDR3 length	Heavy CDR3 seq	Light V (IGKV or IGLV)	Light J (IGKJ or IGLJ)	Light CDR3 length	Light CDR3 seq	Light chain type	Heavy chain	Light chain	GenSig analysis	Aliases	LANL comments
CH235.9	gp120 CD4BS	EMD-8080 EMD-8081 5F9O	Donor CH505	CH235	Bonsignori2016	○ Antibody-driven selection in donor CH505 ○ Electrostatic interactions with D368	Bonsignori2016	Cloanalyst	1-46*01	3-10*01	4*02	15	CVRNVGTAGSLLHYDHW	3-15*01	1*01	8		K	CH235.9 immunoglobulin heavy chain QVRLLQYGGGVKRPGASMTISCVASGYNFNDYYIHWVRQAPGQGLELMGW IDPSGGRTDYAGAFGDRVSMYRDKSMNTLYMDLRSLRSGDTAMYYCVRNV GTAGSLLHYDHWGLGVMVTVSS KU570037 CAGGTGCGACTACTACAATATGGGGGTGGAGTGAAGAGGCCTGGGGCCTC AATGACGATTTCCTGCGTGGCGTCTGGATACAACTTCAACGACTACTATA TACACTGGGTGCGACAGGCCCCTGGACAAGGCCTCGAATTGATGGGATGG ATCGACCCTAGTGGTGGTCGCACAGATTACGCAGGGGCGTTTGGGGACAG AGTGTCCATGTACAGGGACAAGTCCATGAACACACTCTACATGGACCTGA GGAGCCTGAGATCTGGCGACACGGCCATGTATTATTGTGTTAGAAATGTG GGAACGGCTGGCAGCTTGCTCCACTATGACCACTGGGGCCTGGGAGTTAT GGTCACCGTCTCCTCA		IC_{50}	CH493	

Assay

Analyze assay data in CATNAP Submit

Number of data: 199

Download table below with additional virus info

Expand table below to show virus infomation

Antibody	Virus	Subtype	Tier	Infection stage	Coreceptor	Country	Year	Accession	Alias	Reference	IC50	Mean IC50	IC80	Mean IC80	ID50	Mean ID50
CH235.9	0013095_2_11	C	2	intermediate	CCR5	INDIA	2000	EF117267	0013095, 0013095-2.11, 0013095.2.11, HIV-0013095-2.11, HIV_0013095_2_11	Bonsignori et al. Cell 165:449 (2016)	>50	UD				
CH235.9	001428_2_42	C	2	intermediate	CCR5	INDIA	2000	EF117266	001428, 001428-2.42, HIV-001428-2.42, HIV_001428_2_42	Bonsignori et al. Cell 165:449 (2016)	0.417	0.417				

Los Alamos NATIONAL LABORATORY

Antibody Contacts and Features DB

ID 85

Description Antibody-driven selection in donor CH505

Antibody class CD4BS

Reference Hraber2015

Type resistance

MAb name CH103 CH235 CH235.12 CH235.9 (Click MAb name to get to Immunology DB notes)

Env pos.	Feature	HXB2 AA	Entropy Group M	Entropy Subtype B	Entropy Subtype C	Annotation
4	Signal peptide	K	1.292	1.114	1.115	Site associated with at least 80% loss of transmitted-founder sequence in CH505 (donor of mAbs CH103 - CH106 and CH235 lineage); these sites are likely to represent antibody-driven selection.
130	gp120	K	1.274	0.883	1.495	Site associated with at least 80% loss of transmitted-founder sequence in CH505 (donor of mAbs CH103 - CH106 and CH235 lineage); these sites are likely to represent antibody-driven selection.
132	gp120, V1-hypervariable, V1	T	1.450	0.849	1.573	Site associated with at least 80% loss of transmitted-founder sequence in CH505 (donor of mAbs CH103 - CH106 and CH235 lineage); these sites are likely to represent antibody-driven selection.
144	gp120, V1-hypervariable, V1	S	2.255	2.072	2.207	Site associated with at least 80% loss of transmitted-founder sequence in CH505 (donor of mAbs CH103 - CH106 and CH235 lineage); these sites are likely to represent antibody-driven selection.
145	gp120, V1-hypervariable, V1	G	2.331	2.228	2.255	Site associated with at least 80% loss of transmitted-founder sequence in CH505 (donor of mAbs CH103 - CH106 and CH235 lineage); these sites are likely to represent antibody-driven selection.
147	gp120, V1-hypervariable, V1	M	2.618	2.513	2.375	Site associated with at least 80% loss of transmitted-founder sequence in CH505 (donor of mAbs CH103 - CH106 and CH235 lineage); these sites are likely to represent antibody-driven selection.
151	gp120, V1-hypervariable, V1	K	2.576	2.285	2.605	Site associated with at least 80% loss of transmitted-founder sequence in CH505 (donor of mAbs CH103 - CH106 and CH235 lineage); these sites are likely to represent antibody-driven selection.

ID 84

Description Electrostatic interactions with D368

Antibody class CD4BS

Reference Bonsignori2016

Type binding

MAb name CH235.12 CH235.9 VRC01 (Click MAb name to get to Immunology DB notes)

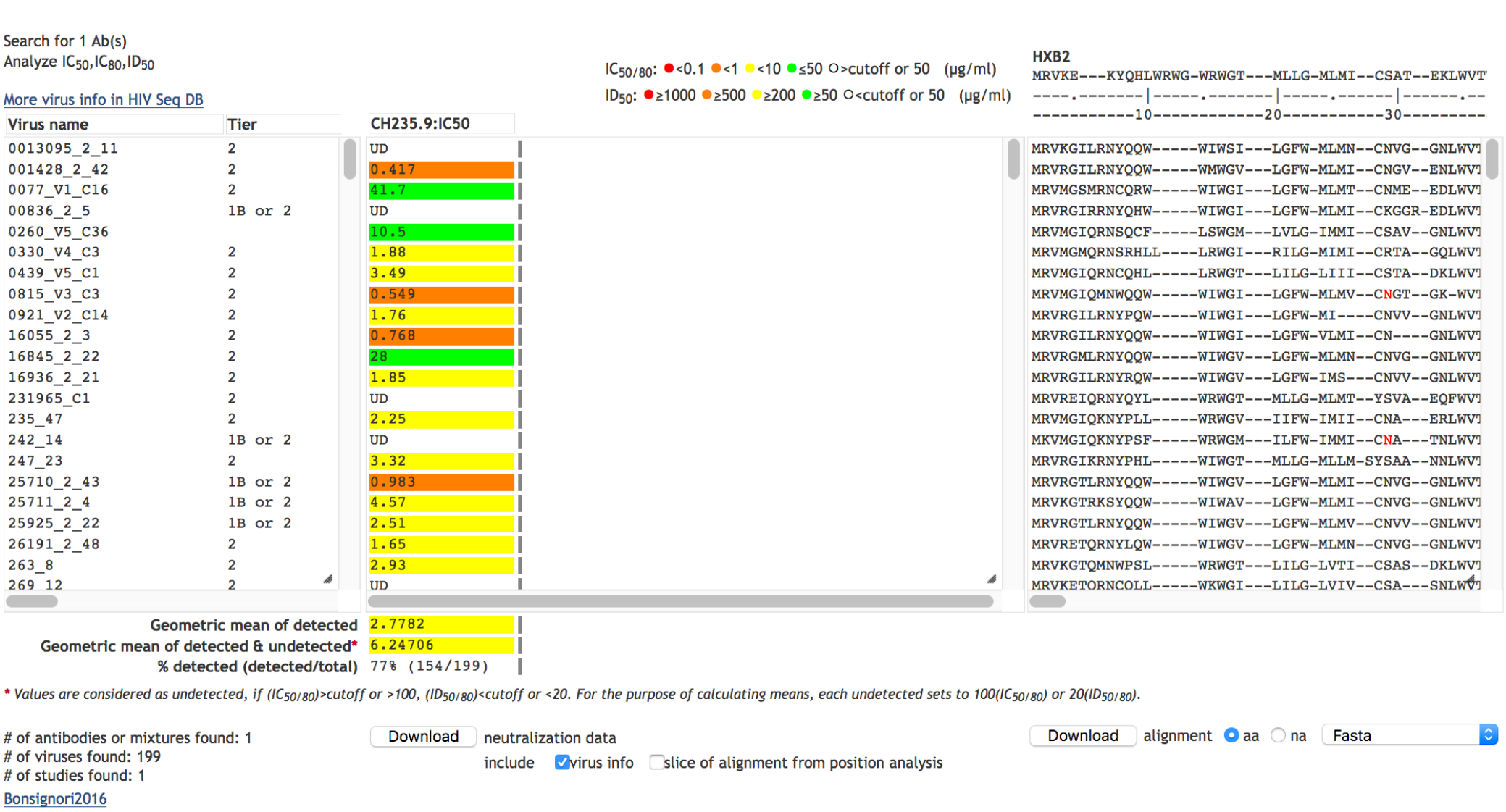
Env pos.	Feature	HXB2 AA	Entropy Group M	Entropy Subtype B	Entropy Subtype C	Annotation
368	gp120, CD4 binding loop	D	0.024	0.023	0.029	D368 contacts the CDR H2 loop of VRC01, CH235.9, and CH235.12 by electrostatic interactions.

Important position(s) with Hxb2 amino acid: D368

Los Alamos NATIONAL LABORATORY

Neutralization Data: CH235.9

- Antibodies with neutralization data are linked to CATNAP
  - Detailed antibody information including Ab sequences and germlines
  - Inhibition assay results against virus panels



DATABASES SEARCH ALIGNMENTS TOOLS PUBLICATIONS GUIDES Search Site

http://hiv.lanl.gov/catnap

CATNAP

Compile, Analyze and Tally NAb Panels

The CATNAP family of tools has been designed to facilitate the analysis of neutralizing antibodies (NAbs) through the identification of potential genetic signatures resulting from a NAb's interaction with a protein. While interactions between NAbs and HIV-1 Env are the emphasis, the Custom Input version can accomodate other types of data, including other proteins and organisms.

CATNAP

Purpose: Analyze our database of HIV-1 IC_{50} and IC_{80} neutralization data from publicly-available sources, in conjunction with HIV-1 Envelope sequences. Access our extensive databases of information about neutralizing antibodies and viruses used in published neutralization studies. Alignments of Env sequences for these viruses are also provided.

Help: CATNAP Help.

CATNAP: Custom Input

Purpose: Find potential genetic signatures based on your own numerical data in association with protein sequences. In addition to neutralization data, this tool is flexible enough to accomodate almost any kind of data in conjunction with almost any protein sequence.

Help: Custom CATNAP Help.

CATNAP: Hybrid

Purpose: Compare and analyze your HIV-1 IC_{50} and IC_{80} neutralization data with published data. This tool will display your data side-by-side with data from our database of published HIV-1 neutralization data.

Help: Hybrid CATNAP Help.

- Custom Input requires
  - Numerical data (IC50, ID50, AUC, any phenotypic data)
  - Aligned sequences associated with the data

- You can also combine your own HIV data with the published HIV data (Hybrid CATNAP)

Reference

Yoon et al. CATNAP: a tool to compile, analyze and tally neutralizing antibody panels. Nucleic Acid Res 2015 Jul 1;43(W1):W213-9. PMID 26044712.

Los Alamos NATIONAL LABORATORY

CATNAP

Compile, Analyze and Tally NAb Panels

Purpose: To provide easy analysis of data associated with HIV-1 neutralizing antibodies, including neutralization panel data, sequences, and structures.

See also: Help | Other CATNAP tools | How to Cite

Download CATNAP data

New! Click "Attributes" to select antibodies based on donor, germline genes, or binding type. Or select viruses based on tier, subtype, infection stage, or coreceptor. Details...

Select by: Antibody and Virus | Study

Antibodies by: Names | Attributes

Viruses by: Names | Attributes | Panels

# of Abs = 307, # of Ab mixtures = 25

# of Viruses = 1011 (785 seqs available)

Select	Name	Donor	# of viruses tested
☐	10-1074	Donor 17	420
☐	10-1074-IgG3C	Donor 17	119
☐	10-1074V	Donor 17	200
☐	10-996	Donor 17	121
☑	10E8	Donor N152	433
☐	10E8-0fH/0fL	Donor N152	21
☐	10E8-0fH/4fL	Donor N152	21
☐	10E8-10fH/10fL	Donor N152	21
☐	10E8-10fH/16fL	Donor N152	21
☐	10E8-10fH/4fL	Donor N152	21
☐	10E8-19fH/10fL	Donor N152	21
☐	10E8-19fH/16fL	Donor N152	21
☐	10E8-2fH/0fL	Donor N152	21
☐	10E8-2fH/10fL	Donor N152	21
☐	10E8-2fH/4fL	Donor N152	21
☐	10E8V1.1/P140	Donor N152	118

Select	Name	Subtype	# of Abs tested	Seq
☐	0013095_2_11	C	147	Yes
☐	001428_2_42	C	146	Yes
☐	0041_V3_C18	C	23	Yes
☐	0077_V1_C16	C	72	Yes
☐	00836_2_5	C	71	Yes
☐	0260_V5_C1	A1	11	Yes
☐	0260_V5_C36	A1	163	Yes
☐	0301_BM_A12	C	12	Yes
☐	0301_BM_A2	C	12	Yes
☐	0301_BM_A6	C	12	Yes
☐	0330_V4_C3	A1	122	Yes
☐	0404_BM_B9	C	12	Yes
☐	0404_BM_D4	C	12	Yes
☐	0404_BM_F3	C	12	Yes
☐	0404_BM_G3	C	12	Yes
☐	0404_BM_H4	C	12	Yes

Multiple ways to pick antibodies and viruses

Options

Retrieve: Antibody details | Virus details | Assay

Or

Analyze along with virus sequences: IC_{50} | IC_{80} | Both

Large sets of data run slowly. Limit the number of antibodies or viruses for quicker response.

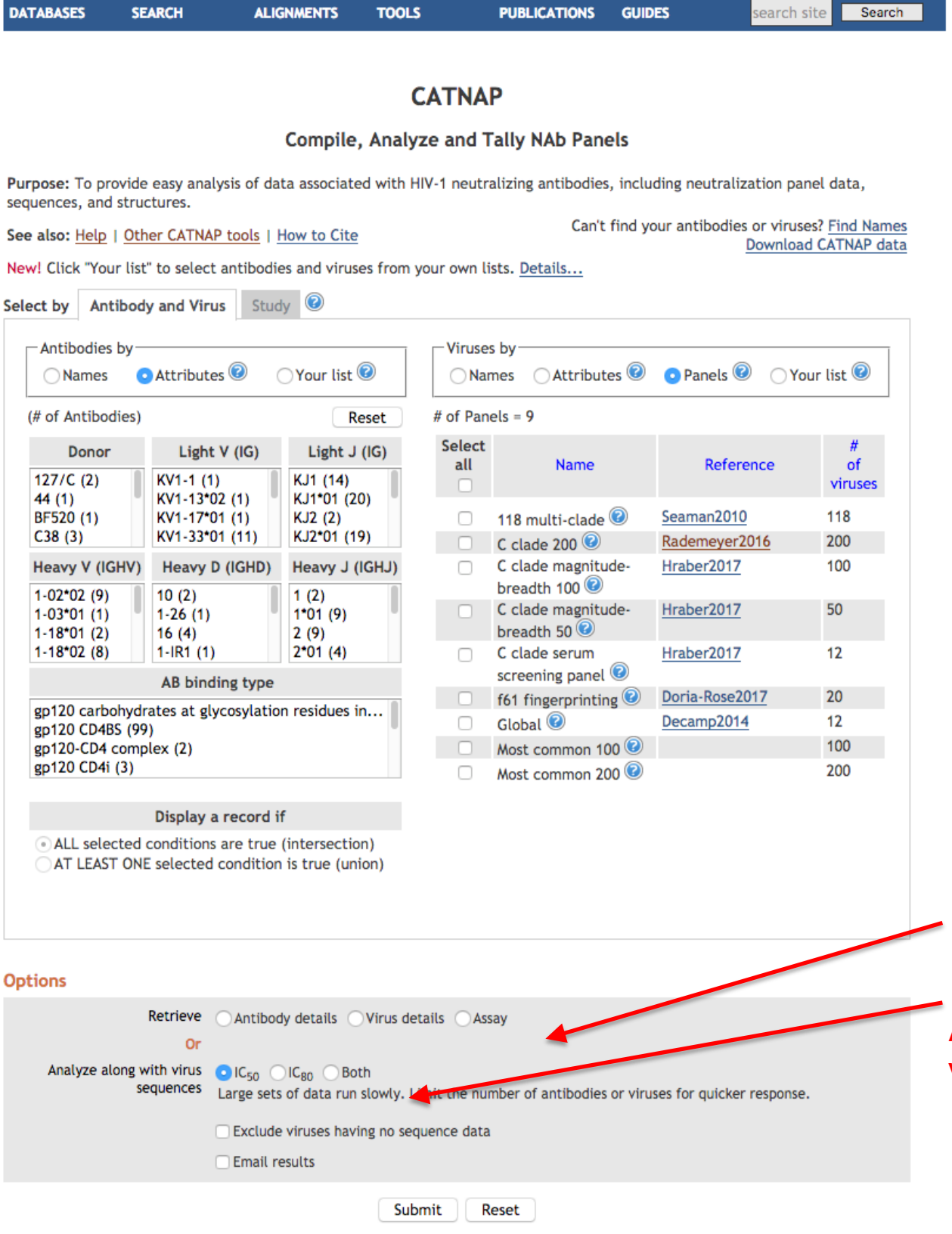
☐ Exclude viruses having no sequence data

☐ Email results

Retrieve Antibody, Virus or Assay details

Analyze IC_{50}, IC_{80} or Both along with the viral sequences

Los Alamos NATIONAL LABORATORY



Select Antibodies and Viruses in Several Ways:

- Individual or all antibody and viruses
- Select by study
- Select antibodies by attributes (germline and binding region)
- Select viruses by attributes (Tier, Subtype, Infection stage)
- Select viruses by a virus panel

Example: 10E8 and PG9

Retrieve Antibody, Virus or Assay details

Analyze IC_{50}, IC_{80} or Both along with the viral sequences

Los Alamos NATIONAL LABORATORY

Collapse or expand details from individual studies

CATNAP: IC_{50} & IC_{80}/HIV-1 alignment

4 antibodies & 1007 viruses selected to search

More virus info in HIV Seq DB

*: Geometric means (Expand to individual values)

●<0.1 ●<1 ●<10 ●≤50 ○>cutoff or 50 (µg/ml)

Virus name	Tier	10E8 IC50	10E8 IC80	PG9 IC50	PG9 IC80	PG1
001428_2_42	2	1.48106*	6.44387*	0.01010*	0.01221*	0.0
0041_V3_C18	2	0.44000	3.70700	0.00100	0.00400	1.8
0077_V1_C16	2	1.45938*	9.19287*	0.09106*	0.35301*	UD:>
00836_2_5	1B or 2	0.50193*	1.90470*	49.00000*	UD:>50*	31.
0260_V5_C36		9.80180*	30.00417*	1.57709*	13.80908*	0.0
0301_BM_A12		–	–	0.66000	–	–
0301_BM_A2		–	–	0.91000	–	–
0301_BM_A6		–	–	0.91000	–	–
0330_V4_C3	2	1.04674*	3.83465*	0.01468*	0.05294*	0.0
0404_BM_B9		–	–	0.08000	–	–
0404_BM_D4		–	–	6.87000	–	–
0404_BM_F3		–	–	0.04000	–	–
0404_BM_G3		–	–	0.58000	–	–
0404_BM_H4		–	–	0.01000	–	–
0439_V5_C1	2	1.16920*	4.69871*	UD:>50*	UD:>50*	UD:>
0702_BM_B4		–	–	0.83000	–	–
0702_BM_B9		–	–	0.38000	–	–
0702_BM_D1		–	–	0.18000	–	–
0702_BM_H12		–	–	0.37000	–	–
0815_V3_C3	2	0.31354*	1.86820*	UD:>10, >25,...*	UD:>25, >50*	0.0
0907_V4_C12		–	–	–	–	–
Geometric mean of detected		0.35352	2.15327	0.18501	0.38293	0.0
Geometric mean of all (undetected set to 100)		0.46515	2.43935	0.92140	2.20904	0.9
% detected (detected/total)		95% (411/432)	97% (387/400)	74% (543/729)	69% (285/416)	64%

HXB2

```
MRVKE----KY-QHLW-RWG---WRWGTMLLG---MLMI---CSAT---
----.---------|--------.----|-------.-------|----
----------10-----------20--------------30----
MRVRGILR-NY-QQW--------WMWGVLGFW---MLMI---CNGV----
MRVRGILR-NW-QLW--------WTWGILGFW---MVMN---CNVR----
MRVMGSMR-NC-QRW--------WIWGILGFW---MLMT---CNME----
MRVRGIRR-NY-QHW--------WIWGILGFW---MLMI---CKGGR---
MRVMGIQR-NS-QCF--------LSWGMLVLG---IMMI---CSAV----
MRVRGMMR-NW-QQW--------WIWGILGFW---MLMI---CSVL----
MRVRGMMR-NW-QQW--------WIWGILGFW---MLMM---CSVL----
MRVRGMMR-NW-QQW--------WIWGILGFW---MLMM---CSVL----
MRVMGMQR-NS-RHL--------LLRWGIRILG--MIMI---CRTA----
MRVRGILR-NC-PQW--------WTWGILGFW---MLMI---CSVW----
MRVRGILR-NC-PQW--------WTWGILGFW---MLMI---CSVW----
MRVRGILR-NC-PQW--------WTWGILGFW---MLMI---CSVW----
MRVRGILR-NC-PQW--------WTWGILGFW---MLMI---CSVW----
MRVRGILR-NC-PQW--------WTWGILGFW---MLMI---YSVW----
MRVMGIQR-NC-QHL--------LRWGTLILG---LIII---CSTA----
MRVRGILR-NW-ELW--------WIWGILGFW---MFMI---CNML----
MRVRGILK-NW-KLW--------WIWGILGFW---MFMI---CNML----
MRVRGILR-NW-ELW--------WIWGILGFW---MFMI---CNML----
MRVRGILR-NW-KLW--------WIWGILGFW---IFMI---CNTL----
MRVMGIQM-NW-QQW--------WIWGILGFW---MLMV---CNGT----
MRARGMKR-NC-QNL--------WKWGIMLLG---ILMI---CSAA----
```

Potency and Breadth of neutralization over multiple studies

# of viruses found: 799

Download neutralization data with virus information

Go to antibody information section

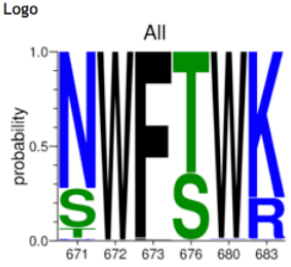
Download ● aa ○ na in Fasta

Antibody context and feature position(s) (based on HXB2)

(See Spreadsheet of neutralizing antibody contexts and features (.xls) for more information)

- **10E8 contacts** (LogoByAll LogoBySubtype): N671 W672 F673 T676 W680 K683
- **PG9-like contacts** (LogoByAll LogoBySubtype): N156 N160 I165 G167 K168 V169 Q170 K171 Y173

Logo



Frequency by position

See full raw counts

	Percentage and raw count of non-gap
671	N: 72.14% (3947) S: 21.09% (1154) T: 5.63% (308) other: 1.13% (62)
672	W: 99.80% (5460) other: 0.20% (11)
673	F: 98.83% (5407) other: 1.17% (64)
676	T: 64.17% (3511) S: 35.48% (1941) other: 0.35% (19)
680	W: 98.92% (5412) other: 1.08% (59)
683	K: 77.82% (4257) R: 21.39% (1170) other: 0.79% (43)

Position analysis

Analyze HXB2 position 160 for Ab PG9

Pick Ab and click on contact position to analyze, or enter your own position

Run CombiNaber

Submit

# of viruses tested							
10E8 IC50: 432	10E8 IC80: 400	PG9 IC50: 729	PG9 IC80: 416	PGT121 IC50: 634	PGT121 IC80: 393	VRC01 IC50: 781	VRC01 IC80: 444

388 virus(es) tested against all antibodies retrieved will be submitted to CombiNaber.

More tools for Immunologists

Most tools are applicable to any organism and some to any numerical data

- **CATNAP**: Compile, Analyze and Tally published and your own NAb Panels
- **CombiNAber**: Predict and analyze neutralization by antibody combinations
- **Sequence Locator:** Find epitope location on the reference genome
- **PepMap:** Map an input set of peptides on the reference sequence (Fasta, PDF and HTML)
- **PeptGen:** Generate sets of overlapping peptides for epitope mapping.
- **QuickAlign** and **AnalyzeAlign:** Align query sequences or discontinuous positions to an alignment, create WebLogos, calculate frequency by position, tally variants in an alignment
- **ELF:** Epitope Location Finder. Search query sequence for
  - Known epitopes from our HIV immunology databases
  - HLA binding motifs
  - Epitopes predicted by the IEDB binding algorithm.
- **N-Glycosite:** Find potential N-linked glycosylation sites in an alignment
- **Mosaic** and **Epigraph:** Generate candidate vaccine protein cocktails with optimized potential epitope coverage, calculate and visualize coverage, using a suite of tools
- **Gen Sig:** picks out signatures or correlations wrt a given sequence feature
- **Heatmap:** Display and organize neutralization or other quantitative data. **And more …**

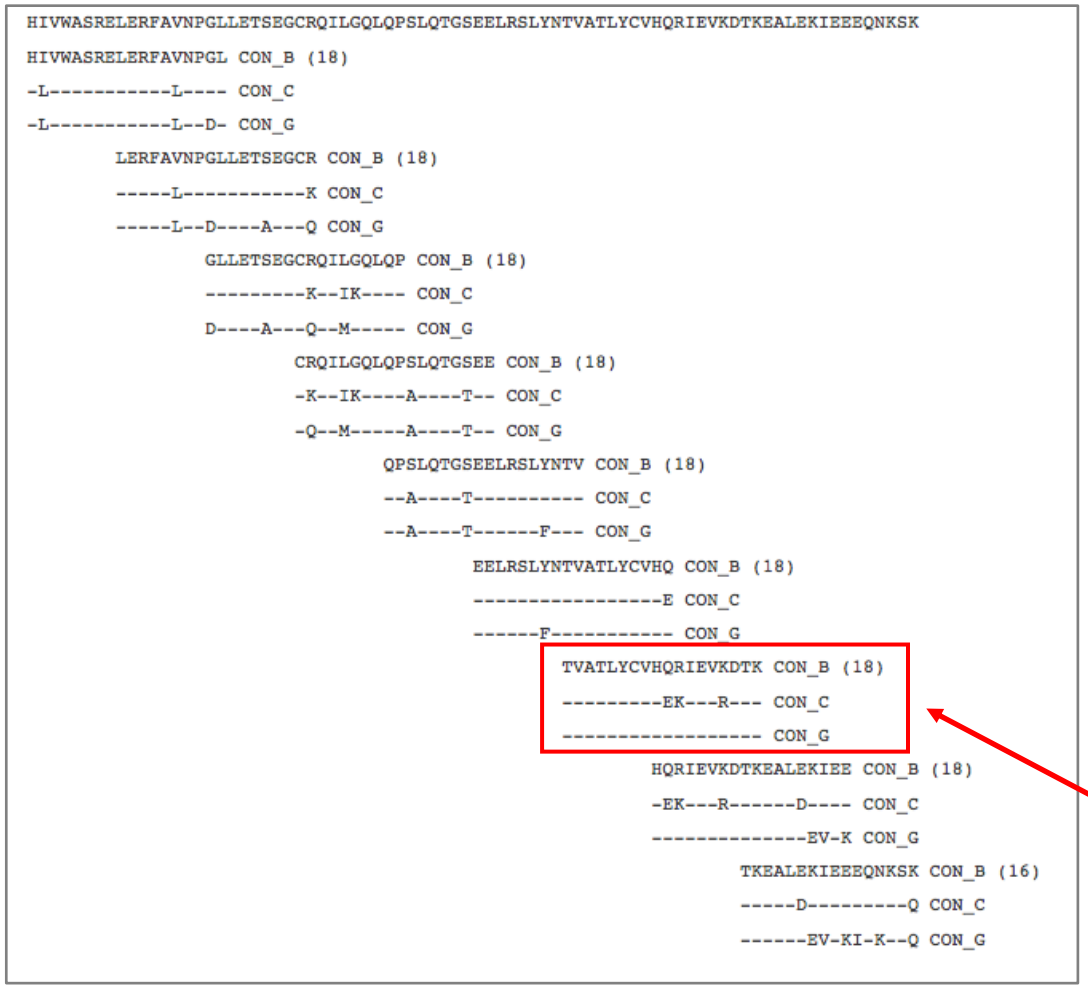
Los Alamos NATIONAL LABORATORY

PeptGen

https://www.hiv.lanl.gov/content/sequence/PEPTGEN/peptgen.html

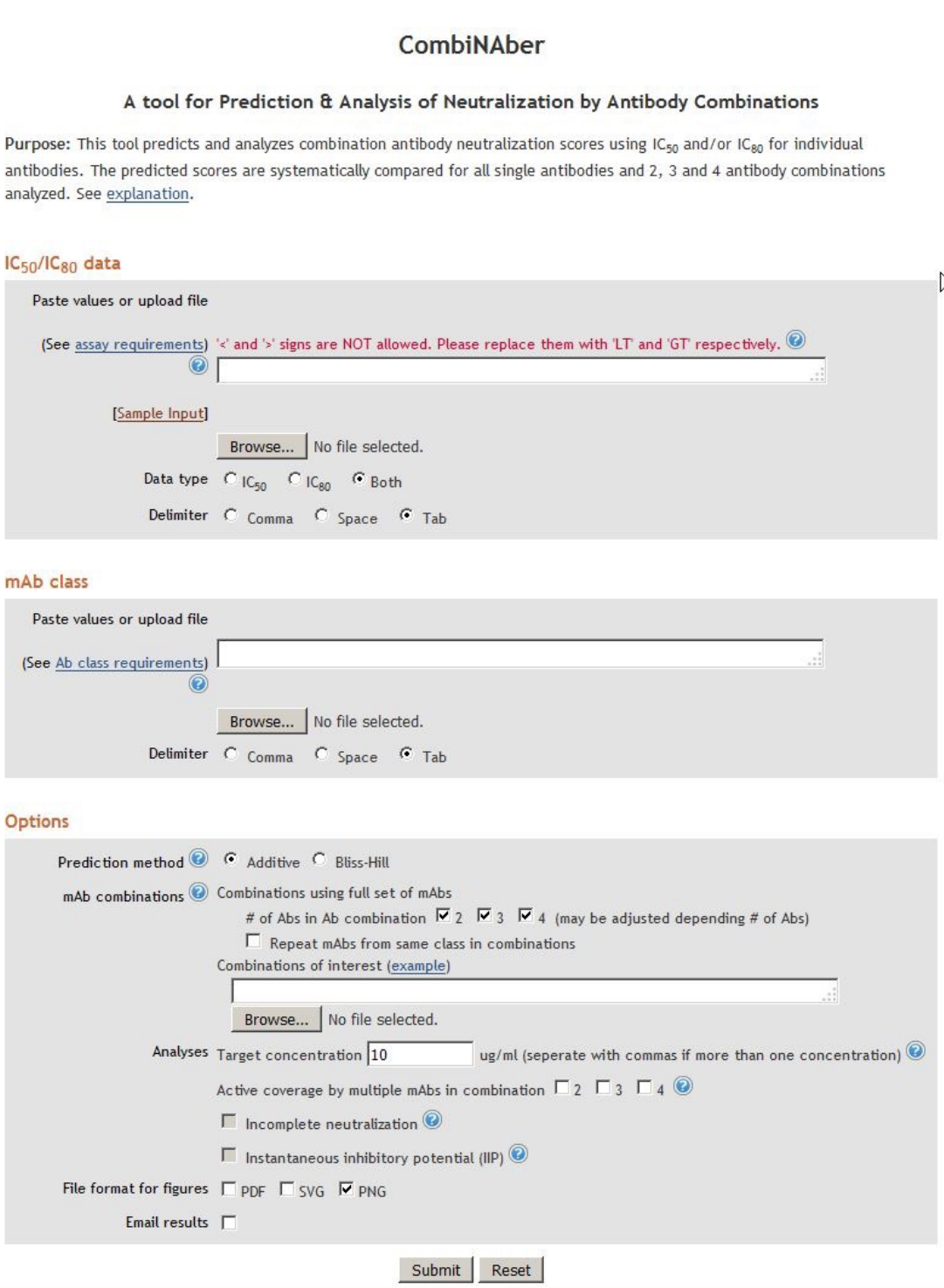
- Generates overlapping peptides for any protein sequence
- Takes alignment as an input and removes duplicate peptides

```
Seq1        HIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEVKDTKEALEKIEEEQNKSK
Seq2        HLVWASRELERFALNPGLLETSEGCKQIIKQLQPALQTGTEELRSLYNTVATLYCVHEKIEVRDTKEALDKIEEEQNKSQ
Seq3        HLVWASRELERFALNPDLLETAEGCQQIMGQLQPALQTGTEELRSLFNTVATLYCVHQRIEVKDTKEALEEVEKIQKKSQ
```




```
1 HIVWASRELERFAVNPGL 1 s1 1  s1 - -
2 HLVWASRELERFALNPGL 1 s2 1  - s2 -
3 HLVWASRELERFALNPDL 1 s3 1  - - s3

4 LERFAVNPGLLETSEGCR 2 s1 1  s1 - -
5 LERFALNPGLLETSEGCK 2 s2 1  - s2 -
6 LERFALNPDLLETAEGCQ 2 s3 1  - - s3

7 GLLETSEGCRQILGQLQP 3 s1 1  s1 - -
8 GLLETSEGCKQIIKQLQP 3 s2 1  - s2 -
9 DLLETAEGCQQIMGQLQP 3 s3 1  - - s3

10 CRQILGQLQPSLQTGSEE 4 s1 1  s1 - -
11 CKQIIKQLQPALQTGTEE 4 s2 1  - s2 -
12 CQQIMGQLQPALQTGTEE 4 s3 1  - - s3

13 QPSLQTGSEELRSLYNTV 5 s1 1  s1 - -
14 QPALQTGTEELRSLYNTV 5 s2 1  - s2 -
15 QPALQTGTEELRSLFNTV 5 s3 1  - - s3

16 EELRSLYNTVATLYCVHQ 6 s1 1  s1 - -
17 EELRSLYNTVATLYCVHE 6 s2 1  - s2 -
18 EELRSLFNTVATLYCVHQ 6 s3 1  - - s3

19 TVATLYCVHQRIEVKDTK 7 s1&s3 2  s1 - s3
20 TVATLYCVHEKIEVRDTK 7 s2 1  - s2 -

21 HQRIEVKDTKEALEKIEE 8 s1 1  s1 - -
22 HEKIEVRDTKEALDKIEE 8 s2 1  - s2 -
23 HQRIEVKDTKEALEEVEK 8 s3 1  - - s3
```

CombiNAber

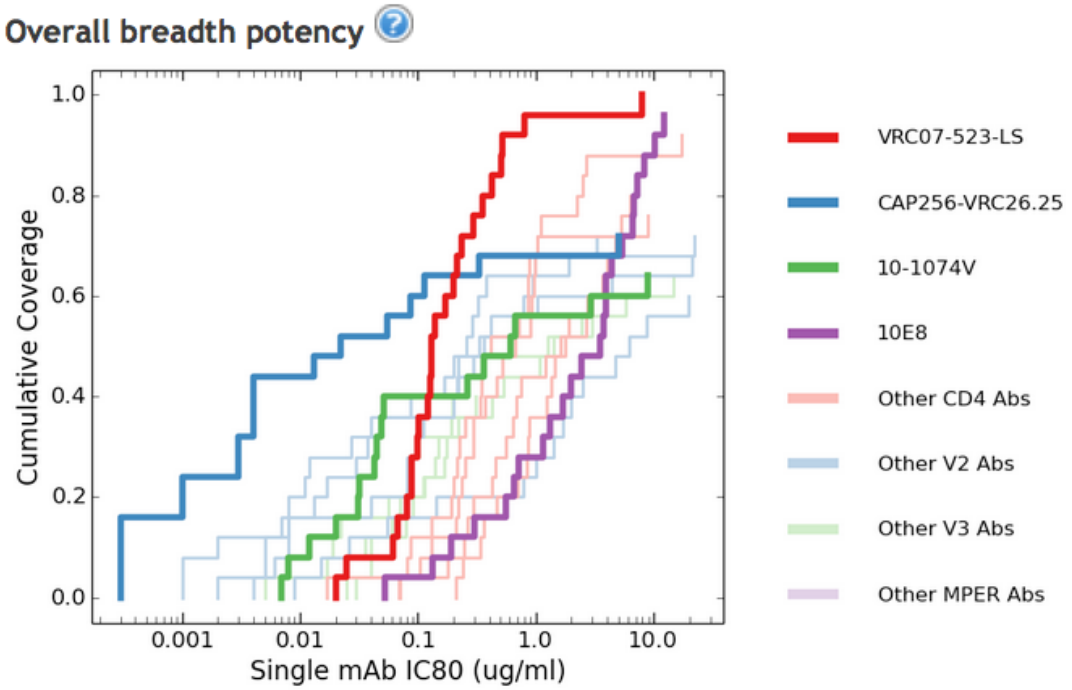


- Background
  - Kong *et al*, 2015, *J Virol*
  - Wagh *et al*, 2016, *PLOS Pathogens*
  - Questions: Kshitij Wagh, kshitij@lanl.gov
- Purpose: predict neutralization by antibody combinations (to optimize immunotherapy options)
- Input:
  - Neutralization data (IC50 and / or IC80) with antibody and virus names
  - Antibody type (i.e. binding region)

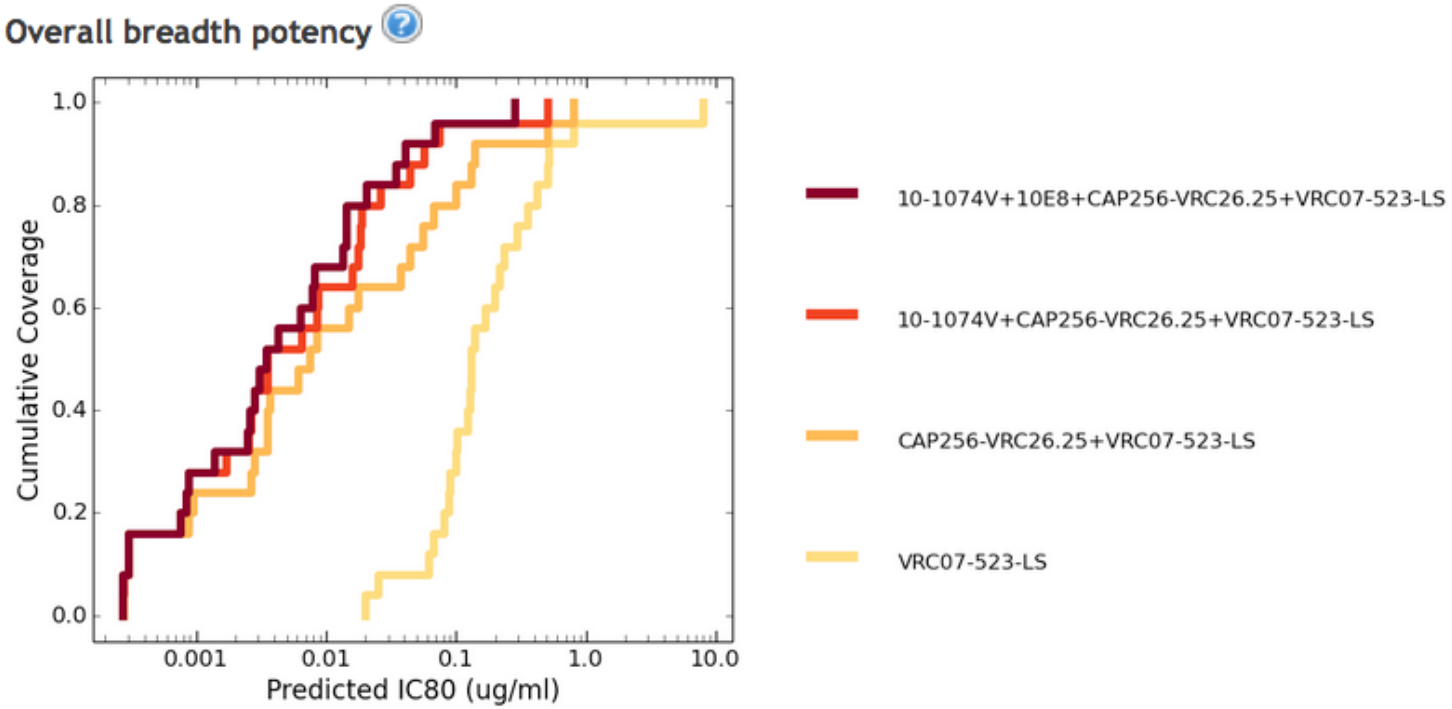
www.hiv.lanl.gov/content/sequence/COMBINABER/combinaber.html

Los Alamos NATIONAL LABORATORY

CombiNAber



Single mAbs



Best 4, 3, 2, 1 Combinations

Los Alamos NATIONAL LABORATORY

HIV sequence database

All kinds of basic information about HIV and about our database

DATABASES | SEARCH | ALIGNMENTS | TOOLS | PUBLICATIONS | GUIDES | Search Sit

- Tutorials
- CRFs
- HIV-1 Gene Map
- In-depth Annotation
- Neutralizing Antibody Resources & CATNAP
- 3D Structure
- Data Dictionary
- How to Cite this Database
- HIV Database News
- FAQs
- Links

Tutorials and Basic Informatic

Previous workshop presentations

Tutorials

Keystone 2014 HIV sequence database workshop

Keystone 2014 HIV Immunology database workshop

Sequence quality control explains several common problems with sets of viral sequences

How to make a phylogenetic tree explains how to build a phylogenetic tree

How to use these databases summaries of workshops given at conferences

HIV numbering relative to reference strain HXB2

SIV numbering relative to reference strain SIVmm239

Articles

3D views of HIV macromolecular structures provides links to 3D views of HIV proteins

Stalking the AIDS Virus [PDF] article from LANL Research Quarterly (Fall 2003) about HIV Database research on the HIV-immune system interaction as a step toward an AIDS vaccine

Reference Inform

Circulating recombinan etails about all documented CRFs

HIV-1 gene map illustra of HIV-1, including HXB2 breakpoints

HXB2 annotated spreadsheet (.xls) provides a fully-annotated sequence of HXB2 with base-by-base detail

HIV and SIV subtype nomenclature gives an overview of HIV and SIV subtype nomenclature, particularly HIV-1 groups and subtypes

Primate immunodeficiency virus nomenclature lists SIV species and nomenclature

How the HIV database classifies sequences explains how recombinants are named and annotated

Common sequence formats for alignments shows examples of common sequence formats for alignments

How to cite this Database explains how to cite this website and the printed HIV compendia

Codes and symbols in sequences decodes the symbols and IUPAC codes that appear in sequences and alignments

Codon table gives the translation of nucleotides into amino acids

FAQs answers basic questions about the HIV Sequence Database

Links HIV/AIDS resources and bioinformatics tools on other websites

Yes! We do respond to this e-mail address!

last modified: Tue Aug 8 12:41 2017

Questions or comments? Contact us at seq-info@lanl.gov

Los Alamos NATIONAL LABORATORY

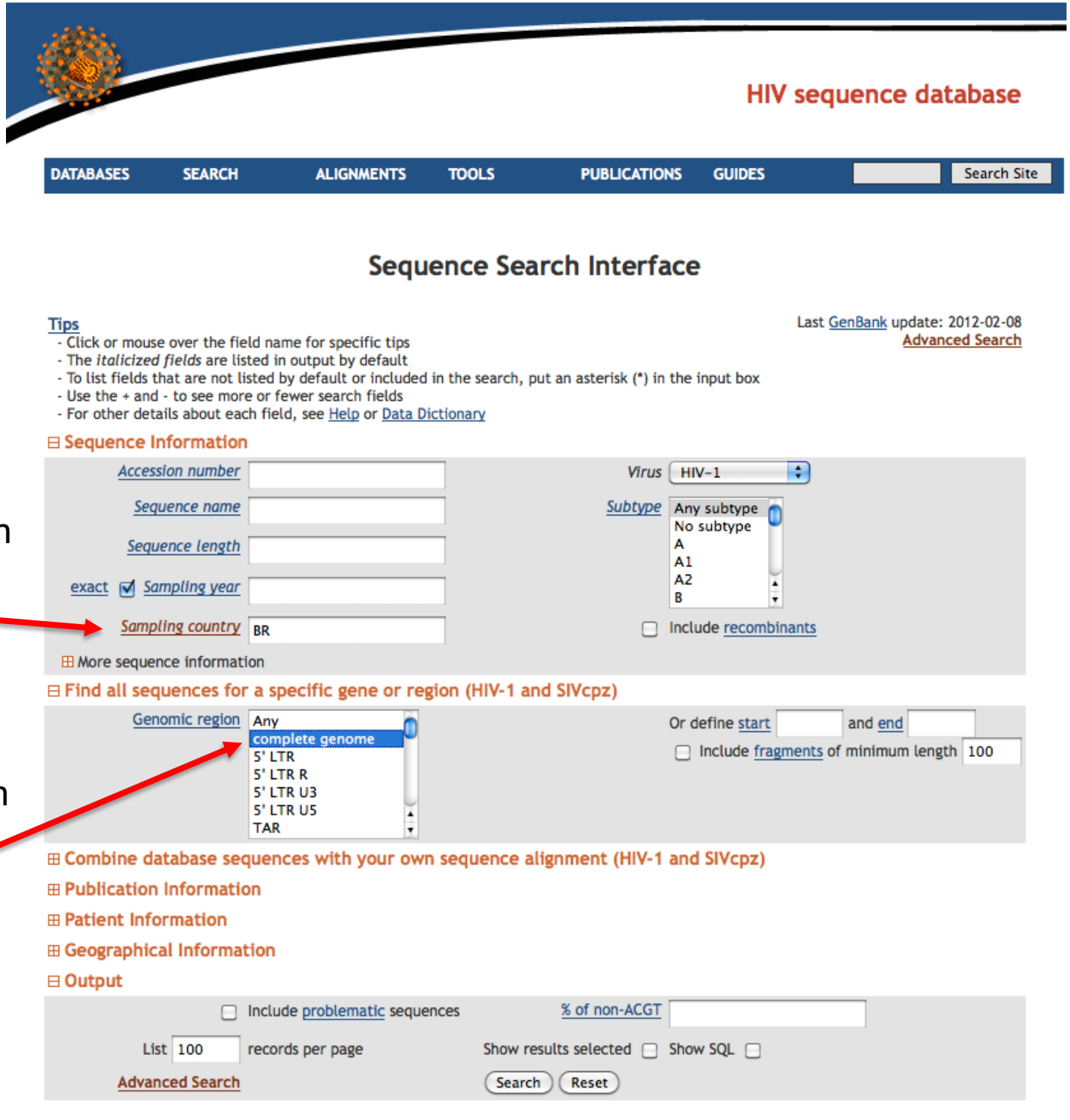
Search Interface

- **Results (what you want)**
  - Can download aligned or unaligned sequences
  - Alignments based on multiple pairwise alignments – alignments are good, but need hand editing for an optimal alignment
  - Select all or a subset of sequences for download
  - Sequences can be re-ordered by clicking on fields at the top of the page, and names customized
- **Searches (how you get it)**
  - Searches are case-insensitive
  - Records are searchable through sequence, patient, genomic region, or publication information and can be matched to the genomic region of a user-provided alignment
  - First seven fields will appear in search results page by default
  - A "*" in a textbox will cause that field to be included in the results page
  - Patient information (Infection year, Infection country) is different than sequence information (Sampling year and Sampling country)
  - Problematic sequence filters (hypermutation, frequent ambiguities, potential contamination)
- **Analysis (what you can do with it)**
  - Build a tree with user alignment, search results and subtype reference sequences combined
- **Help (if all else fails, read the instructions!)**
  - Tips at the top of the page are often overlooked
    - Ranges, operators, wildcards, logical groupings
  - Mouse-over provides brief descriptions; click field names for details in Help file

Los Alamos
NATIONAL LABORATORY

We will search for country = Brazil (BR)

We will search for complete genomes.

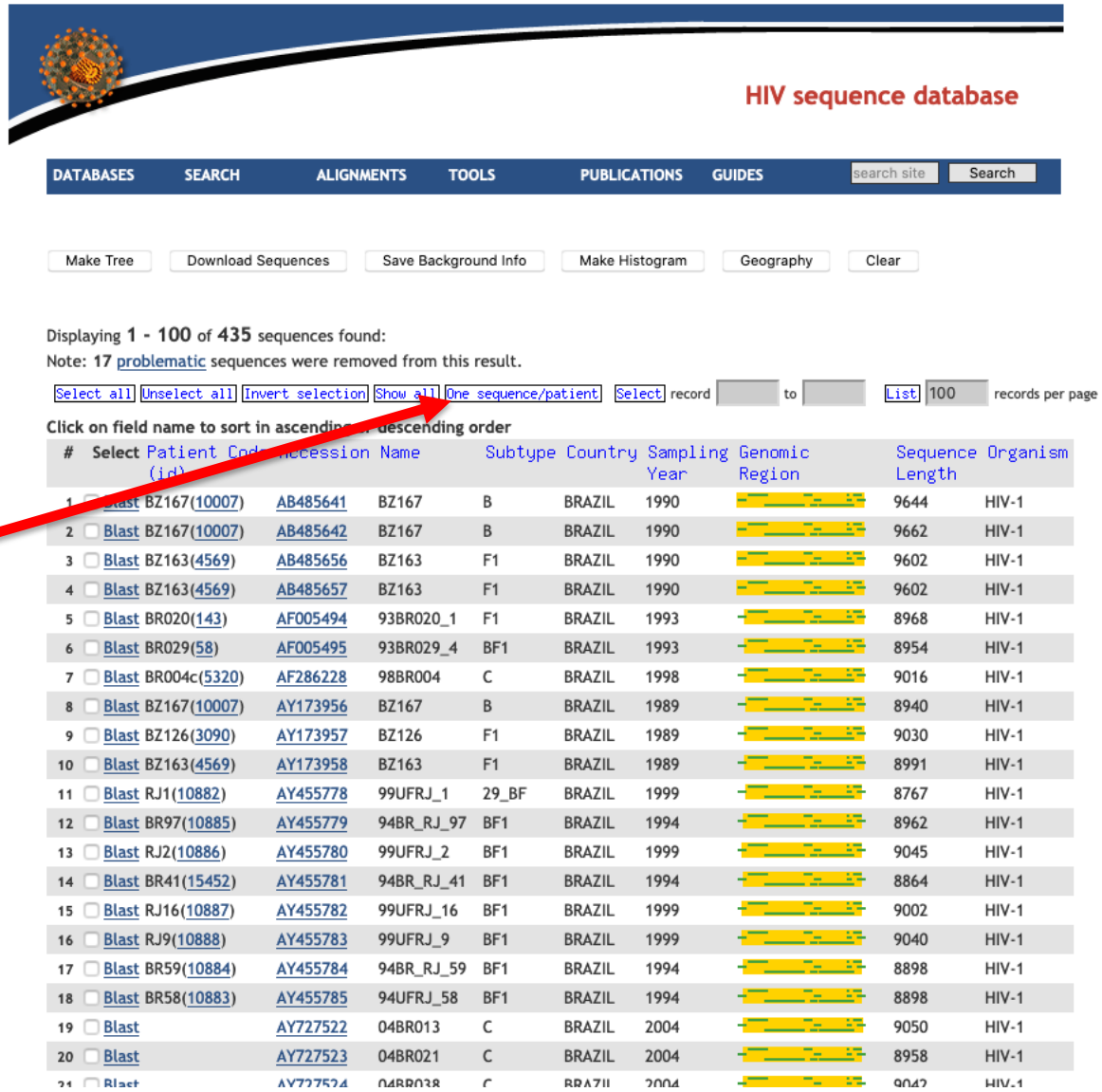


Los Alamos NATIONAL LABORATORY

Questions or comments? Contact us at seq-info@lanl.gov.

Results for HIV-1 complete genomes from Brazil

Choose "One sequence /patient" to remove very similar sequences (only available if a region is selected)



HIV sequence database

DATABASES SEARCH ALIGNMENTS TOOLS PUBLICATIONS GUIDES search site Search

Make Tree | Download Sequences | Save Background Info | Make Histogram | Geography | Clear

Displaying **1 - 100** of **435** sequences found:

Note: **17** problematic sequences were removed from this result.

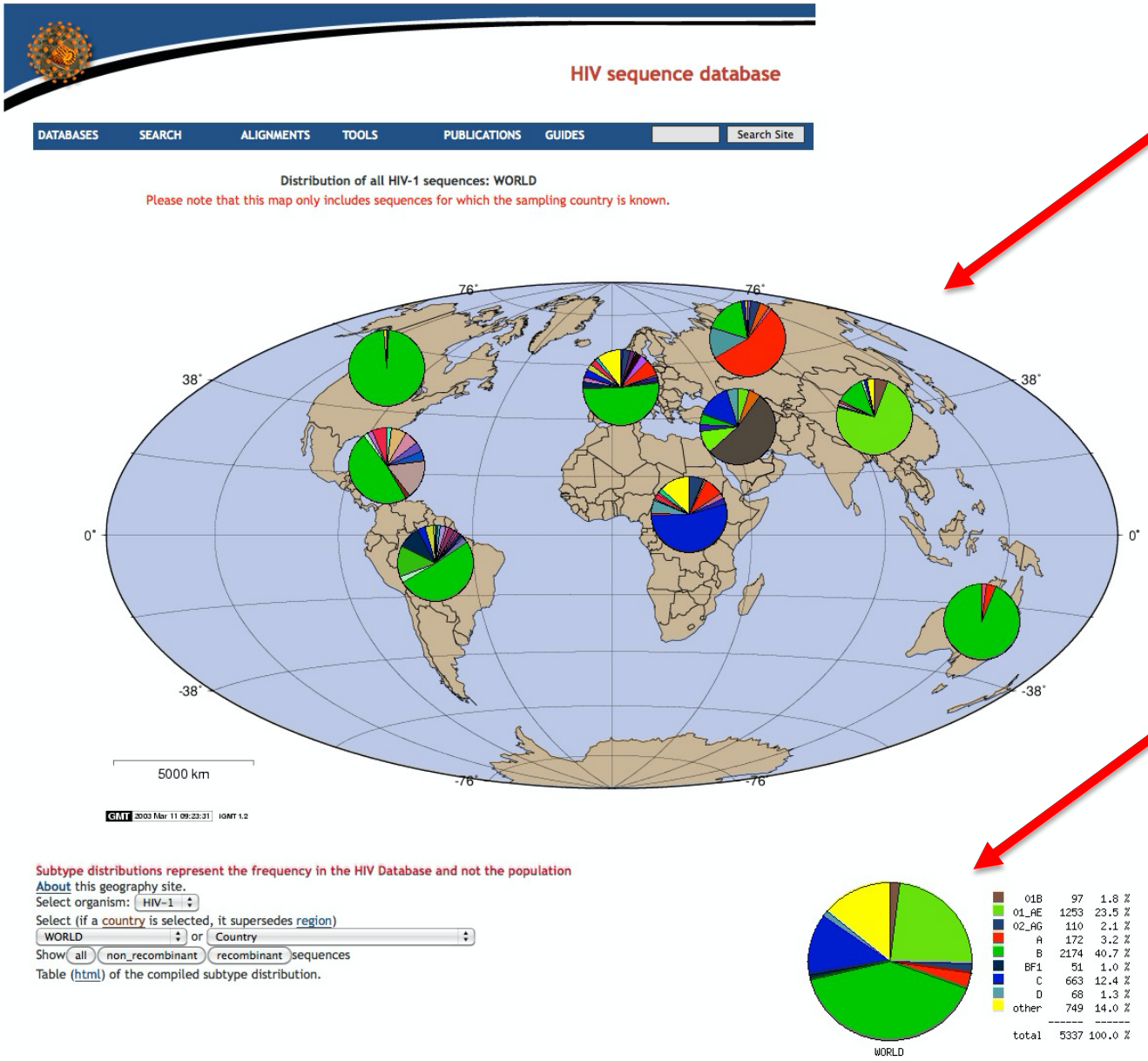
Select all | Unselect all | Invert selection | Show all | One sequence/patient | Select record to List 100 records per page

Click on field name to sort in ascending or descending order

#	Select	Patient Code (id)	Accession	Name	Subtype	Country	Sampling Year	Genomic Region	Sequence Length	Organism
1		Blast BZ167(10007)	AB485641	BZ167	B	BRAZIL	1990		9644	HIV-1
2		Blast BZ167(10007)	AB485642	BZ167	B	BRAZIL	1990		9662	HIV-1
3		Blast BZ163(4569)	AB485656	BZ163	F1	BRAZIL	1990		9602	HIV-1
4		Blast BZ163(4569)	AB485657	BZ163	F1	BRAZIL	1990		9602	HIV-1
5		Blast BR020(143)	AF005494	93BR020_1	F1	BRAZIL	1993		8968	HIV-1
6		Blast BR029(58)	AF005495	93BR029_4	BF1	BRAZIL	1993		8954	HIV-1
7		Blast BR004c(5320)	AF286228	98BR004	C	BRAZIL	1998		9016	HIV-1
8		Blast BZ167(10007)	AY173956	BZ167	B	BRAZIL	1989		8940	HIV-1
9		Blast BZ126(3090)	AY173957	BZ126	F1	BRAZIL	1989		9030	HIV-1
10		Blast BZ163(4569)	AY173958	BZ163	F1	BRAZIL	1989		8991	HIV-1
11		Blast RJ1(10882)	AY455778	99UFRJ_1	29_BF	BRAZIL	1999		8767	HIV-1
12		Blast BR97(10885)	AY455779	94BR_RJ_97	BF1	BRAZIL	1994		8962	HIV-1
13		Blast RJ2(10886)	AY455780	99UFRJ_2	BF1	BRAZIL	1999		9045	HIV-1
14		Blast BR41(15452)	AY455781	94BR_RJ_41	BF1	BRAZIL	1994		8864	HIV-1
15		Blast RJ16(10887)	AY455782	99UFRJ_16	BF1	BRAZIL	1999		9002	HIV-1
16		Blast RJ9(10888)	AY455783	99UFRJ_9	BF1	BRAZIL	1999		9040	HIV-1
17		Blast BR59(10884)	AY455784	94BR_RJ_59	BF1	BRAZIL	1994		8898	HIV-1
18		Blast BR58(10883)	AY455785	94UFRJ_58	BF1	BRAZIL	1994		8898	HIV-1
19		Blast	AY727522	04BR013	C	BRAZIL	2004		9050	HIV-1
20		Blast	AY727523	04BR021	C	BRAZIL	2004		8958	HIV-1
21		Blast	AY727524	04BR038	C	BRAZIL	2004		9042	HIV-1

Los Alamos NATIONAL LABORATORY

Geography output



Each continent's pie chart is clickable to "zoom in" on that continent.

Likewise for each country once you are zoomed in to the continent level.

Most complete genomes in the HIV database are subtype B. But subtype C is more prevalent in human infections. Beware of this type of sampling bias.

Los Alamos NATIONAL LABORATORY

Pre-Built Sequence alignments

- Based on both manual and HMM alignments
- Manually curated
- Alignments are in reading frame (codon aligned)
- Contain non-redundant data (one sequence per patient)
- Compendium alignments show a small “readable” subset
- Reference alignments contain up to four representatives of each subtype (CRFs optional).
  - Useful to provide context for newly generated sequences!
- Protein alignments with frameshifts compensated
- Subtype consensuses and “maximum likelihood ancestors” are available for reagent production
- Special interest alignments

  Sequence sets (“authors’ alignments”) of particular research interest

  Suggestions and additions welcome!

Los Alamos
NATIONAL LABORATORY

Thank you for attending!

Please send us comments, questions, and suggestions!

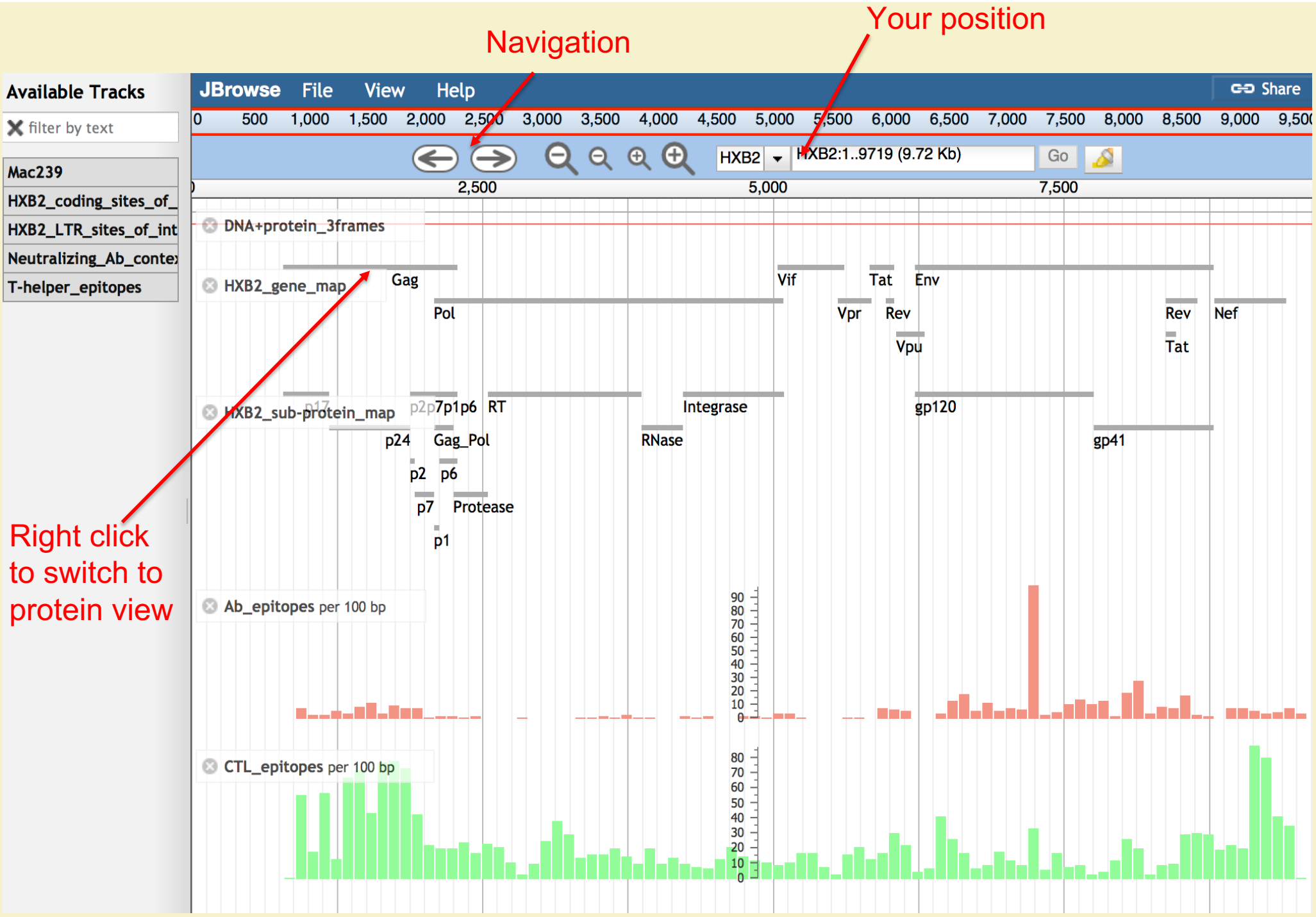
Your comments will help us provide future training and better tools.

Slides available at https://tinyurl.com/2020-IEDB

Contact us: seq-info@lanl.gov or immuno@lanl.gov

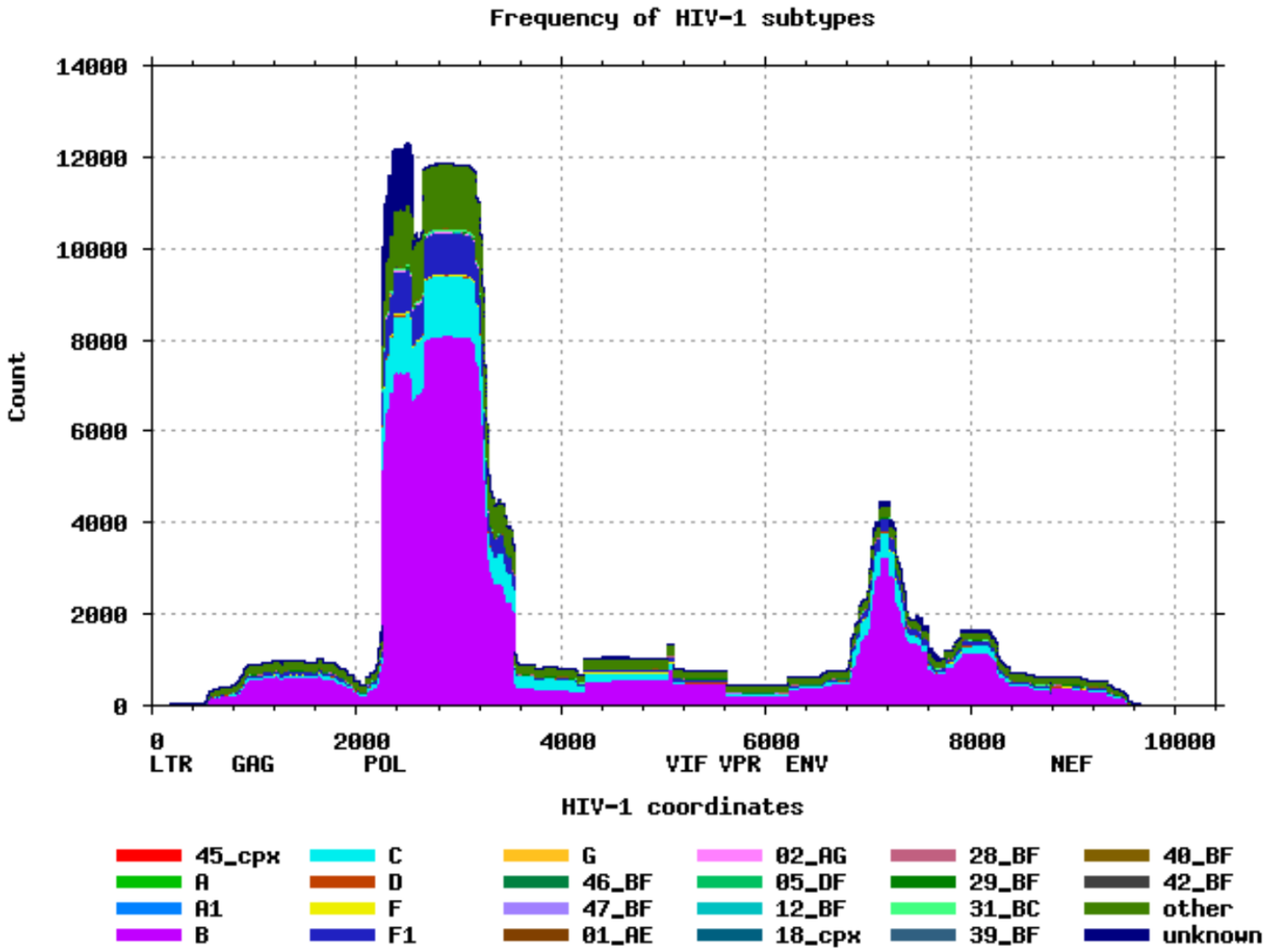
Los Alamos
NATIONAL LABORATORY

HIV Genome Browser: Nucleotide view



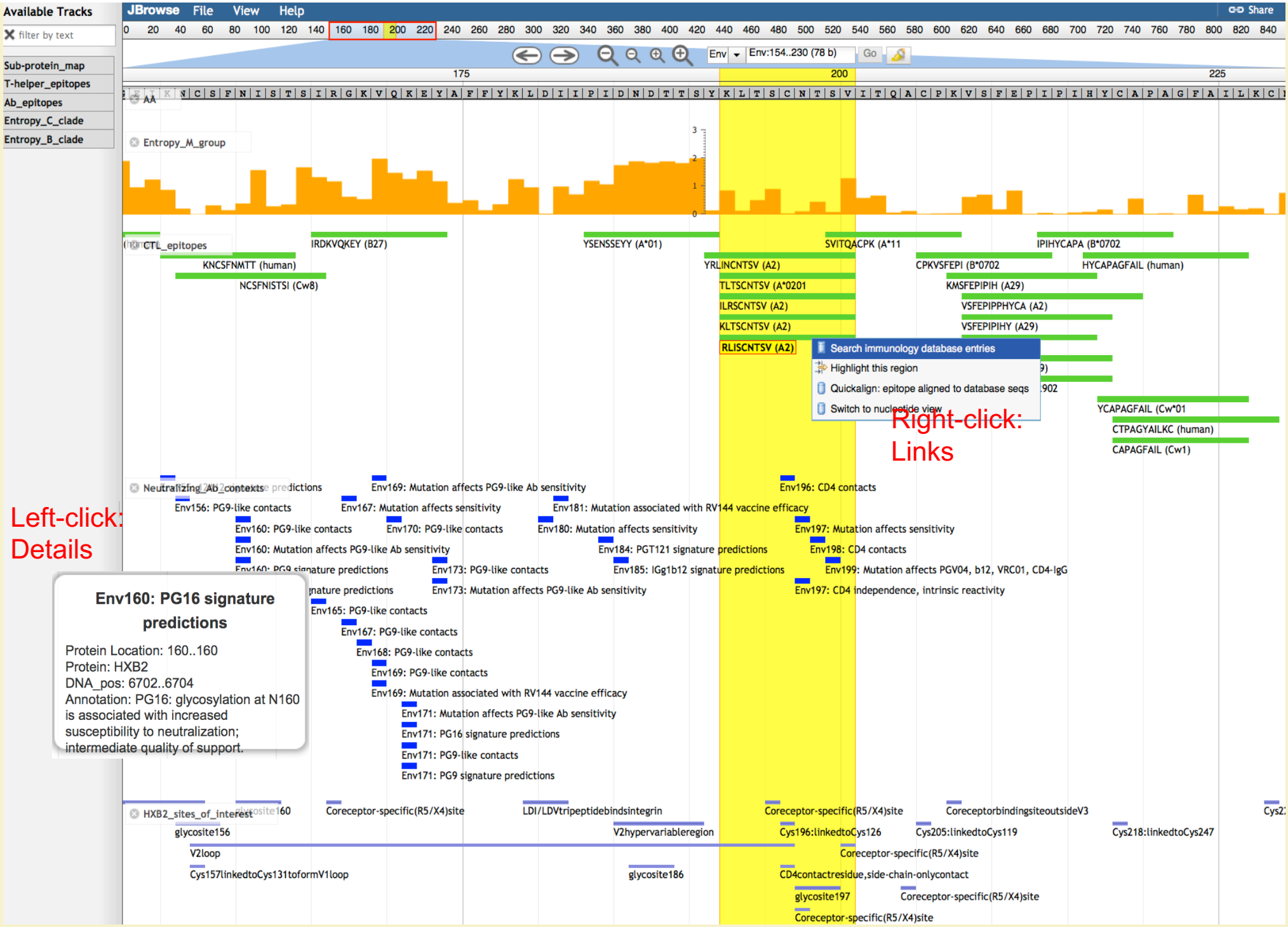
https://www.hiv.lanl.gov/content/sequence/genome_browser/browser.html

Histogram output



This histogram shows the distribution of sequences from your query across the entire HIV-1 genome. At each position across the genome, the number of sequences overlapping with that position is plotted. The colors represent different subtypes.

Los Alamos
NATIONAL LABORATORY


Available Tracks
filter by text
Sub-protein_map
T-helper_epitopes
Ab_epitopes
Entropy_C_clade
Entropy_B_clade
JBrowse File View Help
Share
Env:154..230 (78 b)
Go
AA
Entropy_M_group
CTL_epitopes
IRDKVQKEY (B27)
YSENSSEYY (A*01)
SVITQACPK (A*11
IPIHYCAPA (B*0702
KNCSFNMTT (human)
NCSFNISTSI (Cw8)
YRLINCNTSV (A2)
CPKVSFEPI (B*0702
HYCAPAGFAIL (human)
TLTSCNTSV (A*0201
KMSFEPIPIH (A29)
ILRSCNTSV (A2)
VSFEPIPPHYCA (A2)
KLTSCNTSV (A2)
VSFEPIPIHY (A29)
RLISCNTSV (A2)
Search immunology database entries
Highlight this region
Quickalign: epitope aligned to database seqs
Switch to nucleotide view
YCAPAGFAIL (Cw*01
CTPAGYAILKC (human)
CAPAGFAIL (Cw1)
Neutralizing_Ab_contexts
Env156: PG9-like contacts
Env169: Mutation affects PG9-like Ab sensitivity
Env167: Mutation affects sensitivity
Env196: CD4 contacts
Env181: Mutation associated with RV144 vaccine efficacy
Env160: PG9-like contacts
Env170: PG9-like contacts
Env180: Mutation affects sensitivity
Env197: Mutation affects sensitivity
Env160: Mutation affects PG9-like Ab sensitivity
Env184: PGT121 signature predictions
Env198: CD4 contacts
Env173: PG9-like contacts
Env185: IGg1b12 signature predictions
Env199: Mutation affects PGV04, b12, VRC01, CD4-IgG
Env173: Mutation affects PG9-like Ab sensitivity
Env197: CD4 independence, intrinsic reactivity
Env165: PG9-like contacts
Env167: PG9-like contacts
Env168: PG9-like contacts
Env169: PG9-like contacts
Env169: Mutation associated with RV144 vaccine efficacy
Env171: Mutation affects PG9-like Ab sensitivity
Env171: PG16 signature predictions
Env171: PG9-like contacts
Env171: PG9 signature predictions
HXB2_sites_of_interest
Coreceptor-specific(R5/X4)site
LDI/LDVtripeptidebindsintegrin
Coreceptor-specific(R5/X4)site
CoreceptorbindingsiteoutsideV3
glycosite156
V2hypervariableregion
Cys196:linkedtoCys126
Cys205:linkedtoCys119
Cys218:linkedtoCys247
V2loop
Coreceptor-specific(R5/X4)site
Cys157linkedtoCys131toformV1loop
glycosite186
CD4contactresidue,side-chain-onlycontact
glycosite197
Coreceptor-specific(R5/X4)site
Coreceptor-specific(R5/X4)site
Left-click: Details
Right-click: Links
Env160: PG16 signature predictions
Protein Location: 160..160
Protein: HXB2
DNA_pos: 6702..6704
Annotation: PG16: glycosylation at N160 is associated with increased susceptibility to neutralization; intermediate quality of support.