



Applying the GSID HIV Data Browser and Specimen Repository to Identify Epitopes Recognized by Broadly Neutralizing Antibodies

Faruk Sinangil, PhD

3rd Annual CAVD Meeting

Seattle, WA

December 2-5, 2008



Project Goals

- Acquire and disseminate information that will contribute to the development of a safe and effective HIV vaccine
- Establish a consortium to characterize and evaluate antigenic variation of viruses that mediate new infections



GSID Consortium



GENOMA

- Sequence Analysis
- Phylogenetics
- Alignments
 - PI: K. Crandall



- Consortium management
- HIV Data Browser
- Specimen repository
 - PI: F. Sinangil

UC SANTA CRUZ



- Bioinformatics
 - PI: J. Kent
- Protein expression and epitope mapping
 - PI: P. Berman



- Biostatistics and clinical data
 - PI: E. Li



- Pseudotype virus construction
- Evaluation of neutralization sensitivity
 - PI: B. Schweighardt and T. Wrin



Objectives

- Establish an AIDS VAX Phase III clinical specimen repository
- Establish a web accessible clinical and sequence database from the AIDS VAX Phase III studies (VAX003 and VAX004)



Specimen Repository

	VAX004		VAX003		Total
<u>Specimen inventory</u>	<u>Collection Time Points</u>	<u>Sample #</u> (tubes)	<u>Collection Time Points</u>	<u>Sample #</u> (tubes)	<u>Sample #</u> (tubes)
<u>Pre-Infection</u>					
Serum	73,137	135,042	34,484	68,966	204,008
<u>Post-infection</u>					
Plasma	2,527	29,282	1,792	15,320	44,602
Serum	2,852	2,852	2,171	4,255	7,107
Plasmid DNA library					
Full length gp120 plasmids (3 clones per individual sample)		1,050 (350 subjects)		600 (200 subjects)	1,650



Specimen Repository Activities

- **D. Forthal, UC, Irvine School of Medicine**

- Fc receptor genotyping
- Antibody-dependent cell-mediated virus inhibition activity of serum from AIDSVAX-immunized subjects

- **D. Goldstein, Duke University Medical Center**

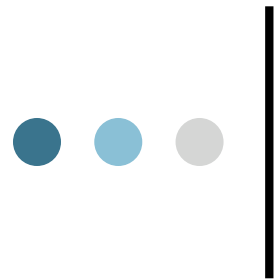
- Human genetic factors and B cell functions that govern HIV vaccine-elicited neutralizing antibody responses

- **P. Lusso, NIAID/NIH**

- Study of the epitopes of sera from human subjects vaccinated with recombinant HIV-1 gp120 within the framework of the VAX003 and VAX004 clinical trials

- **D. Montefiori, Duke University Medical Center**

- Neutralization studies



Specimen Repository Activities cont.

- **R. Wyatt, NIAID/NIH**

- Testing the ability of the gp120 monomer to elicit CD4-induced antibodies

- **S. Ahuja, Univ. of Texas Health Science Center**

- Genetic determinants of transmission, immune response to HIV-1 infection and vaccination in the VaxGen cohort

- **D. Barouch, Harvard Medical School**

- Testing baseline serum samples from VAX003 and VAX004 to assess if Ad5 and non-Ad5 adenovirus serotype seropositivity is a predictor of HIV infection

GSID HIV Data Browser

<http://www.gsidhivdatabrowser.org>

A Unique
Research Tool
Providing Access
to AIDS VAX Trial
Data and
Specimens

About the GSID HIV Data Browser Site

Welcome to the GSID HIV Data Browser, which is hosted by Global Solutions for Infectious Diseases (GSID). The GSID HIV Data Browser is a customized version of the UCSC Genome Browser, which was developed and is maintained by the Genome Bioinformatics Group at the University of California Santa Cruz (UCSCBG), a cross-departmental team within the Center for Biomolecular Science and Engineering (CBSE).

Under the guidance of Jim Kent and Fan Hsu, UCSCBG developed a relational database containing the significant clinical data and viral sequence information pertaining to the infected subjects participating in the VAX004 Phase III clinical trial conducted by VaxGen between 1998 and 2003. GSID, through a license and material transfer agreement with VaxGen and with funding provided by the **Bill & Melinda Gates Foundation**, is making this valuable resource of information and access to serological samples available to the HIV research community.

Three primary views are currently available on the GSID HIV Data Browser. **Subject View** provides the user with demographic and clinical information pertaining to the infected subjects who became HIV infected during the VAX004 Phase III clinical trial. **Table View** provides convenient access to the underlying database by enabling users to view multiple subjects sorted and displayed by the filter controls contained in **Select Subjects**. **Sequence View** contains tools and the ability to align sequences with each other, with reference sequences or with consensus sequences.

GSID Sequence View on HIV (HXB2) Oct. 2002 Assembly

position search: chr1:6,000-8,000 jump clear size 2,001 bp configure

Click on a feature for details. Click on base position to zoom in around cursor. Click gray/blue bars on left for track options and descriptions.

Use drop down controls below and press refresh to alter tracks displayed. Tracks with lots of items will automatically be displayed in more compact modes.

Mapping and Sequencing Tracks

Base Position: VAX004 Short Match
dense hide

Genes and Gene Prediction Tracks

Genes/Regions: InterPro
pack hide

Comparative Genomics

Conservation: MSA Protein Conservation Protein MSA
full hide hide

Subject View search for another subject [input] Go!

Demographic Information

subject ID: GSID4012
gender: Male age: 36 risk factor: Low
race: White/Non-Hispanic weight(kg): 68 location: Southwest

Vaccine and HIV Status

Vaccine/Placebo: Placebo Days of infection relative to first injection date: 541
HIV Status: Infected
Injections: 6

Clinical Information

Estimated Day of Infection*	HIV-1 RNA copies/ml	CD4 cells/microliter
196	5596	650
217	7567	594
245	53103	645
309	< 400	564

GSID Table View

subject	group	HIV-1 RNA	CD4	sex	age	race	geography	risk
GSID4382	Vaccine	30434	707	M	30	White/Non-Hispanic	South	Low
GSID4381	Vaccine	147777	635	M	38	White/Non-Hispanic	South	High
GSID4380	Vaccine	3837	654	M	39	White/Non-Hispanic	South	High
GSID4379	Vaccine	1736	950	M	44	White/Non-Hispanic	Midwest	High
GSID4378	Vaccine	338373	443	M	31	White/Non-Hispanic	Midwest	High
GSID4377	Vaccine	103923	375	M	33	White/Non-Hispanic	Midwest	High
GSID4376	Placebo	2807	518	M	31	White/Non-Hispanic	Northeast	High
GSID4375	Vaccine	N/A	N/A	M	22	White/Non-Hispanic	Northeast	High
GSID4374	Vaccine	6037	595	M	36	White/Non-Hispanic	Northeast	High
GSID4373	Vaccine	N/A	N/A	M	35	White/Non-Hispanic	Northeast	High
GSID4372	Vaccine	1096	856	M	32	White/Non-Hispanic	Northeast	High
GSID4371	Placebo	48778	345	M	41	White/Non-Hispanic	Northeast	High
GSID4370	Vaccine	11254	387	M	38	White/Non-Hispanic	Northeast	High
GSID4369	Vaccine	2082	477	M	34	Asian/Pacific Islander	Northeast	High
GSID4368	Vaccine	3812	646	M	36	White/Non-Hispanic	Northeast	High
GSID4367	Vaccine	N/A	N/A	M	33	White/Non-Hispanic	Northeast	High
GSID4366	Vaccine	6495	455	M	35	White/Non-Hispanic	Northeast	High
GSID4365	Vaccine	N/A	N/A	M	33	White/Non-Hispanic	Northeast	High
GSID4364	Placebo	6241	398	M	35	Hispanic	Northeast	High
GSID4363	Vaccine	20585	267	M	25	White/Non-Hispanic	Northeast	High

On this page you can restrict which subjects appear in the main table based on the values in any column. Click the **submit** button to return to the main Table View page with the current filter settings applied.

Filter Controls for Displayed Columns:

subject - GSID identification number
subject search (including * and ? wildcards): [input]
Include if [any/all] words in search term match.
Limit to items (no wildcards) in list: [input] paste list upload list

group - Immunization Status
group search (including * and ? wildcards): [input]
Include if [any/all] words in search term match.
Limit to items (no wildcards) in list: [input] paste list upload list

Available Values:
Placebo
Vaccine

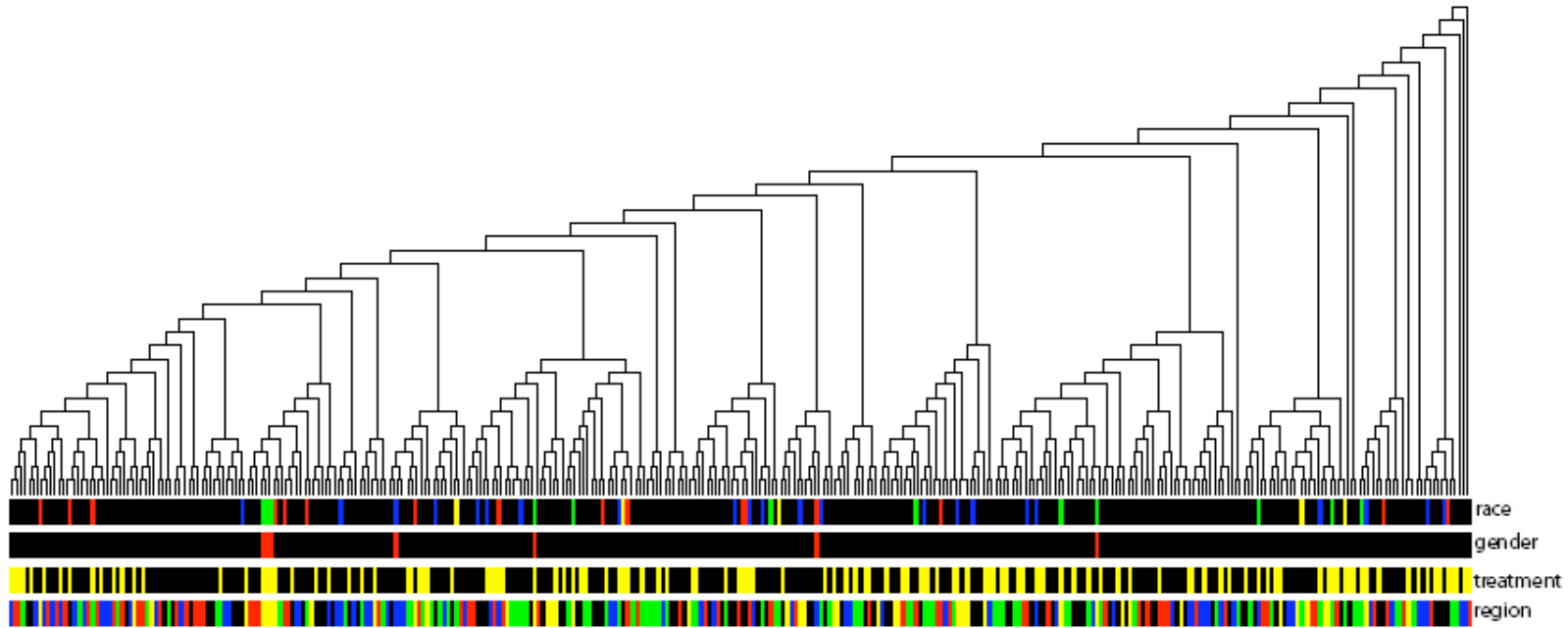


Unique Properties of the GSID HIV Data Browser

- Contains largest collection of full-length gp120 sequences from current time period
- Fully annotated with demographic and clinical information
- GSID HIV data browser offers rapid correlation of nucleotide and protein sequences from VAX004 and VAX003 viruses with multi-sequence alignment (MSA) tools



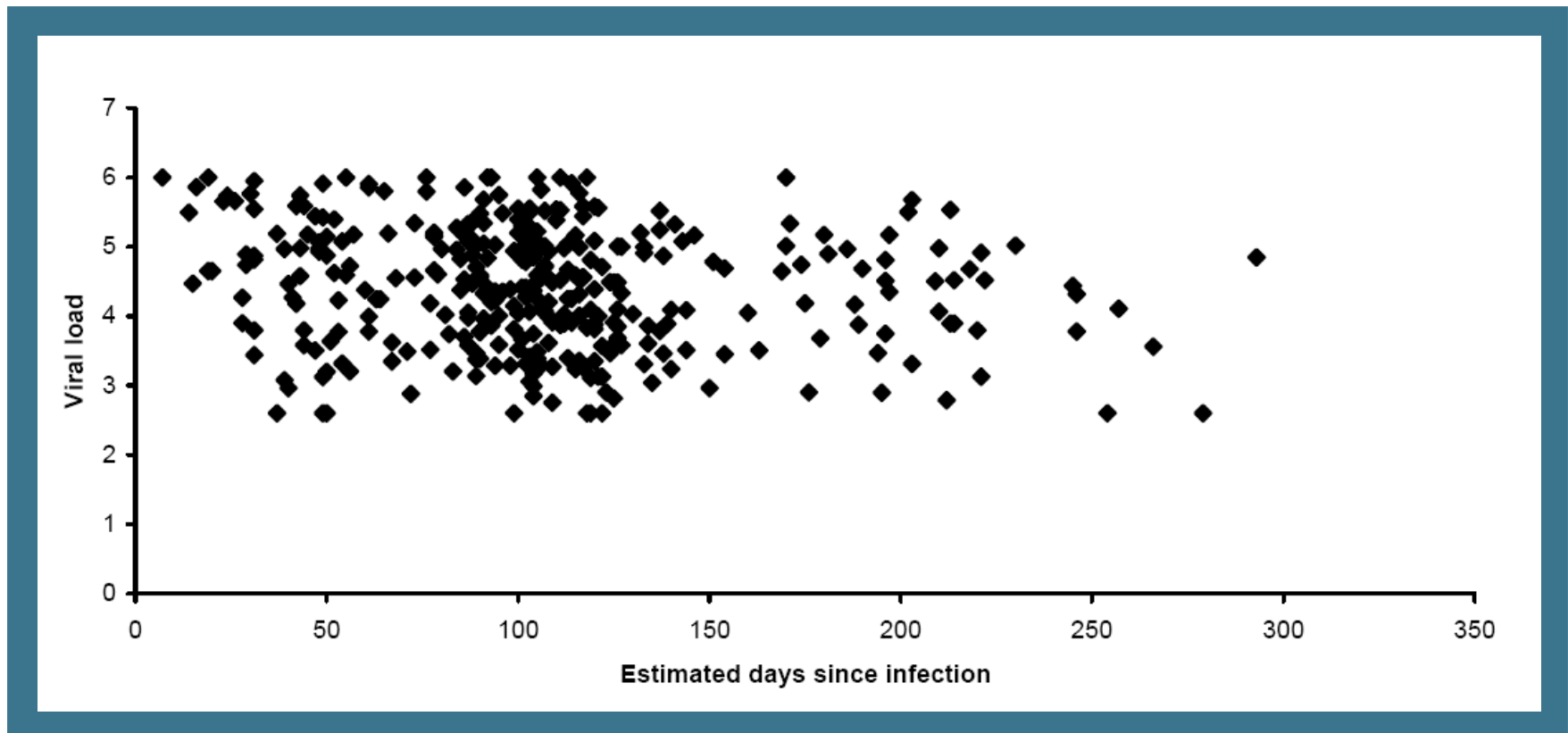
Phylogenetic Analysis of VAX004 Sequences: Subgroup Distributions By Treatment, Race, Gender, and Geographic Location



- 1047 sequences (349 infections)
- Each group of 3 clones forms distinct clusters
- Data available on GSID HIV Sequence Database

Race	Gender	Treatment	Region
Black	Male	Vaccine	West Coast
Other	Female		Southwest
Hispanic			South
Black			Northeast
Asian		Placebo	Midwest

VAX004 Estimated Dates of Infection and Viral Loads for Samples Sequenced



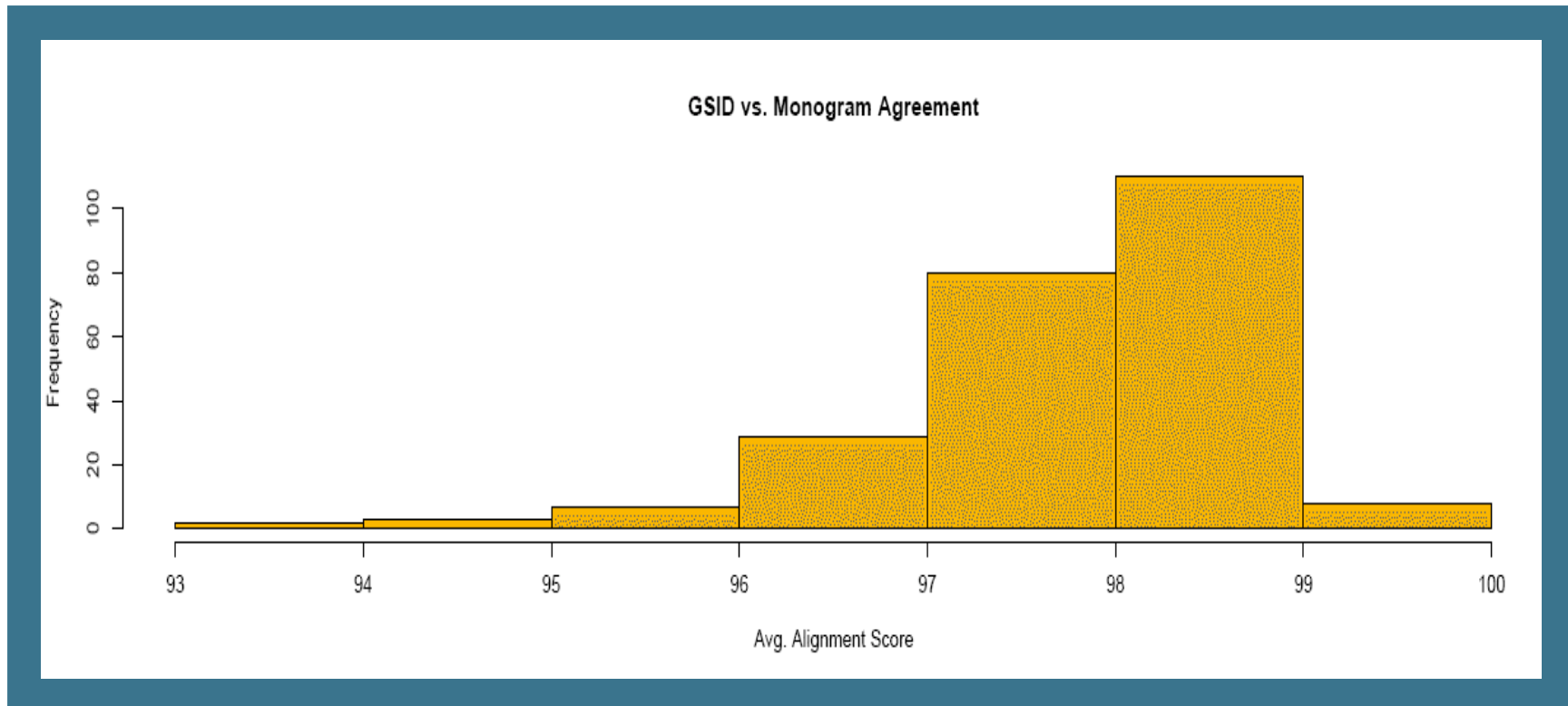


Quality Assessment – GSID HIV Data Browser & Specimen Repository

- Recover plasma samples from GSID specimen repository
- Clone and sequence envelope genes from 28 cryopreserved plasma samples
- Compare sequence data of plasma samples sequenced in 1998-2003 time frame with sequences of newly thawed specimens



Quality Assessment: Comparing gp120 sequences from 2003 with new sequences generated in 2008



28 plasma specimens were selected from the GSID specimen repository and gp160 sequences were determined. Sequences from 2008 were compared to sequences from 1998 – 2003.



Quality Assessment – GSID HIV Data Browser & Specimen Repository

- Results demonstrate:
 - Viral RNA is still intact
 - Proper specimen storage conditions were maintained
 - Despite complex coding and blinding system from AIDSVAX clinical trials, specimens are traceable



GSID HIV Data Browser

- Relational database encompassing significant AIDSVAX clinical trial data
 - VAX004 and VAX003 data available
- Primary views for accessing data
 - **Subject View**
 - **Table View**
 - **Sequence View**
 - **Protein View**

Subject View

Subject View search for another subject:

Demographic Information

subject ID: gsid4336
 gender: Male age: 34
 race: White/Non-Hispanic weight(kg): 68

Vaccine and HIV Status

Vaccine/Placebo: Vaccine Estimated Stud
 HIV Status: Infected
 Injections: 4

Clinical Information

Days After Estimated Infection (DAEI)*	HIV-1 RNA copies/mL	CD4 cells/microliter
126	4879	592
161	1386	922
229	6059	684
258	22449	591
370	10373	608
515	8832	N/A
622	4645	493
727	3160	514
853	25527	N/A

Subject View search for another subject:

Demographic Information

subject ID: gsid3005
 gender: Male age: 24
 race: Asian/Pacific Islander weight(kg): 6

Vaccine and HIV Status

Vaccine/Placebo: Vaccine Estimated St
 HIV Status: Infected
 Injections: 7

Clinical Information

Days After Estimated Infection (DAEI)*	HIV-1 RNA copies/mL	CD4 cells/microliter
177	98544	622
188	97491	921
209	164725	901
265	109992	630
377	187185	541
489	231908	651



Table View

Table View

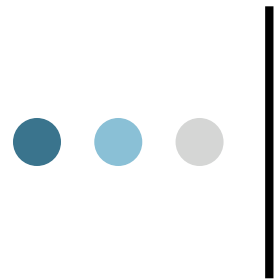
display

subject	group	HIV-1 RNA	CD4	DAEI*	ESDI*	gender	age	race	geography	risk	DNA Sequences	Protein Sequences
GSID3219	Vaccine	300493	575	98	253	F	26	Asian/Pacific Islander	Thailand	High	ss.T-131c2 ss.T-131c3 ss.T-131c4	p1.T-131c2 p1.T-131c3 p1.T-131c4
GSID3103	Placebo	15089	720	101	927	F	34	Asian/Pacific Islander	Thailand	High	ss.T-201c1 ss.T-201c2 ss.T-201c3	p1.T-201c1 p1.T-201c2 p1.T-201c3
GSID3075	Vaccine	47010	625	202	169	F	26	Asian/Pacific Islander	Thailand	Low	ss.T-064c4 ss.T-064c5 ss.T-064c6	p1.T-064c4 p1.T-064c5 p1.T-064c6
GSID4145	Placebo	7288	771	134	881	F	29	Black/Non-Hispanic	Midwest	High	ss.U-333-2c1 ss.U-333-2c3 ss.U-333-2c8	p1.U-333-2c1 p1.U-333-2c3 p1.U-333-2c8
GSID4142	Vaccine	23649	621	85	99	F	29	Black/Non-Hispanic	Midwest	High	ss.U-025c4 ss.U-025c5 ss.U-025c8	p1.U-025c4 p1.U-025c5 p1.U-025c8

Displayed 5 out of 5 subjects passing filter.

* Estimated Study Day of Infection (ESDI), click [here](#) for further explanation.

* Days After Estimated Infection (DAEI), click [here](#) for further explanation.



Sequence View

- Rapid visual correlation of nucleotide and protein sequences from viruses with multi-sequence alignment (MSA) tools
- DNA and protein conservation from MSA results

Sequence View

- VAX004 viral sequences from selected subjects, aligned with base genome (here, MN gp120)

Home Sequence View Gateway Select Subjects Blat Table View DNA PDF/PS Help

GSID Sequence View on HIV MN (GP120) Oct. 2007 Assembly

move <<<< << < > >> >>>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x

position/search chr1:799-855 jump clear size 57 bp. configure

chr1: | 805 | 810 | 815 | 820 | 825 | 830 | 835 | 840 | 845 | 850 |

TGTACACATGGARATTAGGCCAGTAGTATCAACTCAACTGCTGTTAARTGGCAGTCTA

VAX004 HIV-1 gp120 Sequences

HIV Genes and Regions

gp120 ENV

VAX004 DNA Sequence Conservation

Conservation 0 1

VAX004 Protein Sequence Conservation

Protein Conservat 0 1

Protein Multiple Sequence Alignment

Gaps	C	T	H	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L
HIV_MN_(GP120)	C	T	H	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L
U-238c1	C	T	H	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L
U-238c2	C	T	H	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L
U-238c3	C	T	H	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L
U-183c1	C	T	H	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L
U-183c2	C	T	H	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L
U-183c3	C	T	H	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L
U-896c5	C	T	H	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L
U-896c6	C	T	H	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L
U-896c7	C	T	H	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L
U-333-2c1	C	T	H	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L
U-333-2c3	C	T	R	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L
U-333-2c8	C	T	H	G	I	R	P	V	V	S	T	Q	L	L	L	N	G	S	L

move start < 2.0 > Click on a feature for details. Click on base position to zoom in around cursor. Click gray/blue bars on left for track options and descriptions. move end < 2.0 >

default tracks hide all add custom tracks configure refresh

Use drop down controls below and press refresh to alter tracks displayed. Tracks with lots of items will automatically be displayed in more compact modes.

Protein View

Protein p1.U-191c1 - GSID Proteome Browser - Microsoft Internet Explorer

Address: <https://hiv1.cse.ucsc.edu/cgi-bin/pbGsid?proteinID=p1.U-191c1>

Home PDF/PS New Query Help

GSID Protein View

protein: p1.U-191c1

Move <<< << < > >> >>> Current scale: FULL Rescale to 1/6 1/2 FULL

AA Scale | 1 | 50 | 100 | 150 | 200

AA Sequence V P V A K E A T T T L F C A S D A K K Y D T E V H N H A T H A C V P T D F N P Q E V K L E N T E N F N H K N D M V E Q R H E D I I S L N D G S L K F C V K L T F L C V T L N C T D A K F G N N T G N S S T L E E P G E M K N C S F N V T T L S D K K K E Y A L F Y R L D V V P I E E N S S G N K S G S Y R L I S C H T S A I T A Q C K V S F E P I P I H Y C T P A G F A L L K C N D K K F N G K G F C K N V S T V G C T H G I K F V V S T G L L L N G S R A E D E V V

Polarity ↑

Hydrophobicity

Cysteines Predicted

Glycosylation

AA Anomalies

AA Scale | 1 | 50 | 100 | 150 | 200

Explanation of Protein Tracks

pI 8.9

Amino Acid Frequencies

Hydrophobicity -0.5

Number of Cysteines 18

Amino Acid Anomalies

M C H H Y N F I D Q K R T V P G E A S L

Explanation of Protein Property Histograms

Total amino acids: 472

FASTA record:

Done Internet



Specimens and Sequence Data from AIDSVAX Trials Provide a New Approach to the Understanding of bNABs Required for an Effective HIV Vaccine

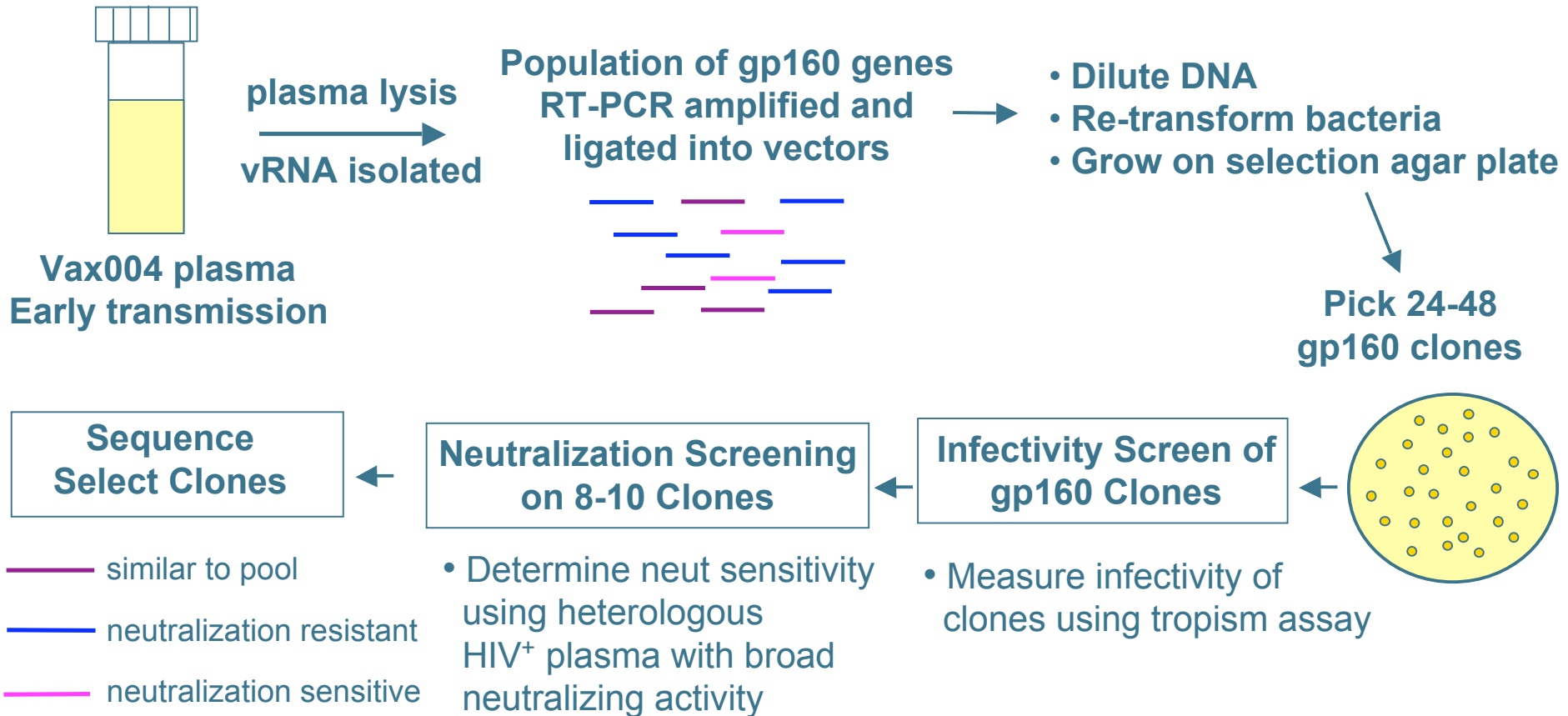
- Sequence data from the VAX004 allow us to define the sequences of viruses responsible for new infections, and the range of virus variation that must be controlled
- Specimens allow us to create panels of representative viruses that can be used in neutralization assays to evaluate performance of candidate vaccines
- Cloned genes represent a library of 1047 antigens available for inclusion in multivalent vaccines
- Longitudinal specimens allow us to define the co-evolution of viruses and the immune response over time and the finite opportunities for immune escape available to the virus



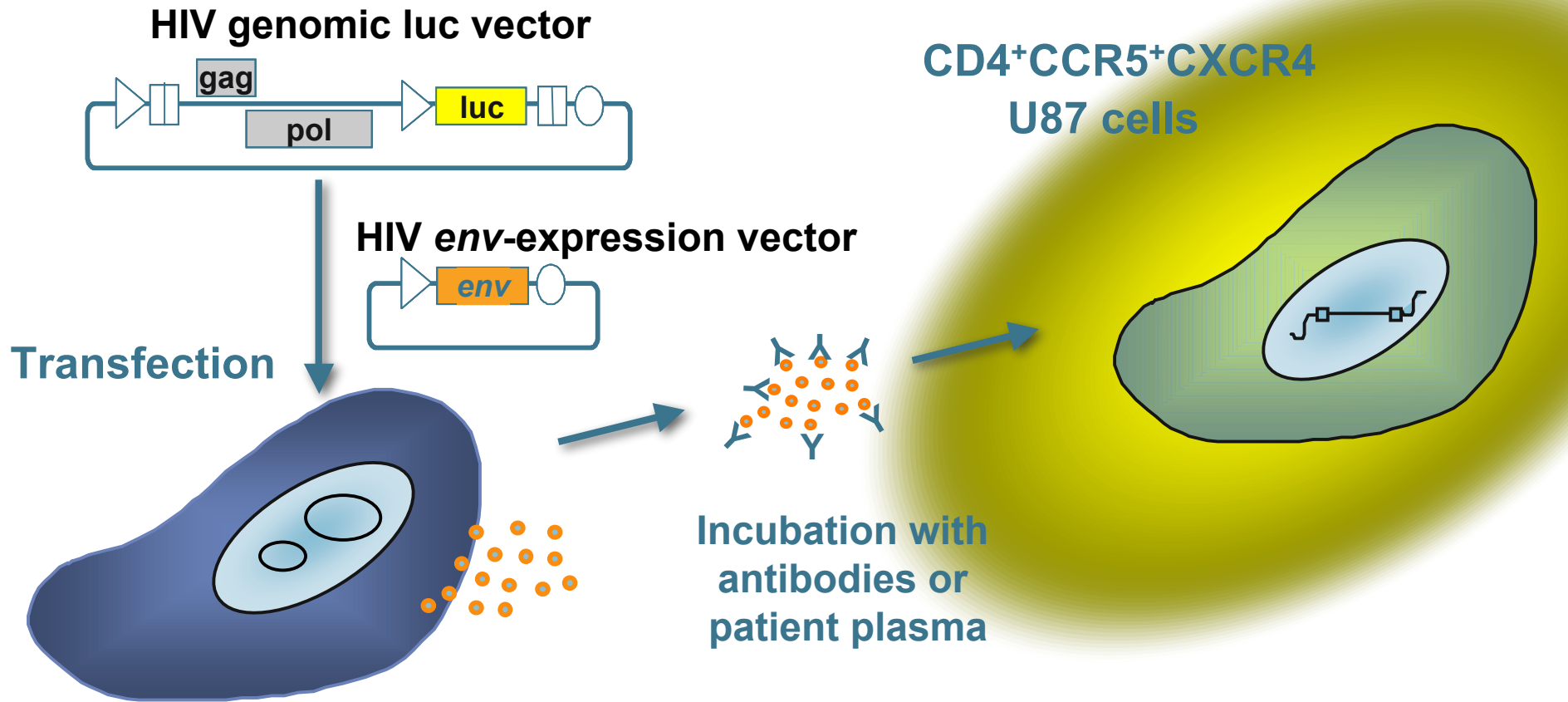
Defining VAX004 Sequence Variation in New Infections: Virus Selection for a Panel of New Immunogens

<u>Characteristic</u>	<u>Specimens Analyzed</u>
Randomly selected	27
Close to consensus sequence	10
Distant from consensus sequence	10
Early / acute infections	10
High virus loads (transmitters)	10
Unusual disulfide structures	5
Long or short V regions	5
<hr/>	
Total:	77

Defining Important Antigenic Regions within gp160 Recognized by Neutralizing Antibodies



Neutralization Sensitivity Measured in Monogram Neutralization Assay



Pseudovirus Production

Virions pseudotyped with patient-derived envelope

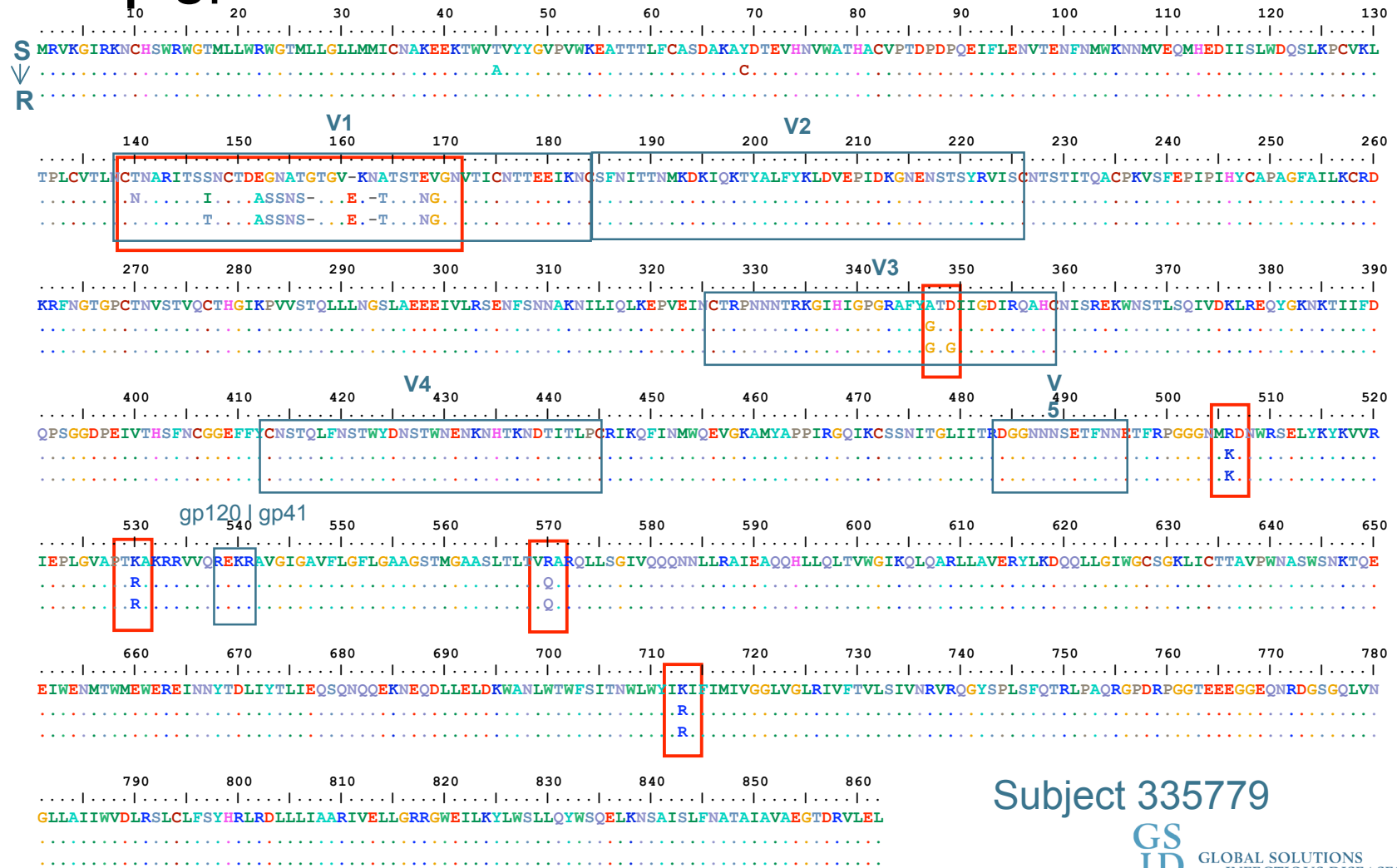
Neutralization Sensitivity of 8-10 Envelope Clones from Each Patient Tested with Heterologous HIV⁺ Plasma

HIV⁺ Plasma from Individuals with Broadly Neutralizing Activity

Pt ID	Viral Envelope	Z1679	Z1684	N16	Z23	
5435631	env pool	661	458	207	686	population pool
	clone 1	336	152	112	358	
	clone 5	390	233	114	330	
	clone 6	106	64	<40	107	more resistant
	clone 10	308	183	131	253	
	clone 12	392	245	160	555	similar to pool
	clone 14	373	185	148	265	
	clone 15	378	183	104	390	
	clone 18	386	200	151	325	
	clone 22	259	167	91	229	
	clone 24	989	595	402	1219	more sensitive



Amino Acid Differences Detected between Neutralization Sensitive and Resistant gp160 Clones



Subject 335779

● ● ● | GSID Consortium Strategy for Development of an Effective HIV Vaccine

- Define the sequences and range of virus variation among viruses responsible for new infections
- Assess the sensitivity/resistance of these viruses to neutralization by broadly neutralizing antibodies (bNAbs) in HIV+ patient sera
- Identify epitopes on envelope proteins recognized by bNAbs in HIV+ sera
- Identify the naturally occurring polymorphisms that occur at these epitopes and their effect on virus neutralization
- Purify populations of broadly neutralizing antibodies from HIV+ sera and define mechanism of action (specificity, affinity/avidity)
- Engineer antigens and multivalent vaccines able to elicit antibodies to the most common polymorphisms at each neutralizing site on the envelope glycoprotein



Acknowledgements

- **Bill and Melinda Gates Foundation**

- **GSID**

- Don Francis
- Carter Lee
- Michael Peterson
- Keith Higgins
- Evie Zaharias
- Melissa Daoust

- **UC Santa Cruz - Bioinformatics**

- Jim Kent
- Fan Hsu
- Ann Zweig
- Robert Kuhn
- Galt Barber
- Erich Weiler

- **UC Santa Cruz – Biomolecular Engineering**

- Phil Berman
- Sarah O'Rourke
- Dora Fonseca
- Bin Yu

- **Genoma LLC**

- Keith Crandall
- Marcos Pérez-Losada
- David Posada

- **PharmaStat**

- Elizabeth Li

- **Monogram Biosciences**

- Becky Schweighardt
- Terri Wrin
- Julie Goss