

Genome-wide prediction of vaccine targets for human herpes simplex viruses using Vaxign reverse vaccinology

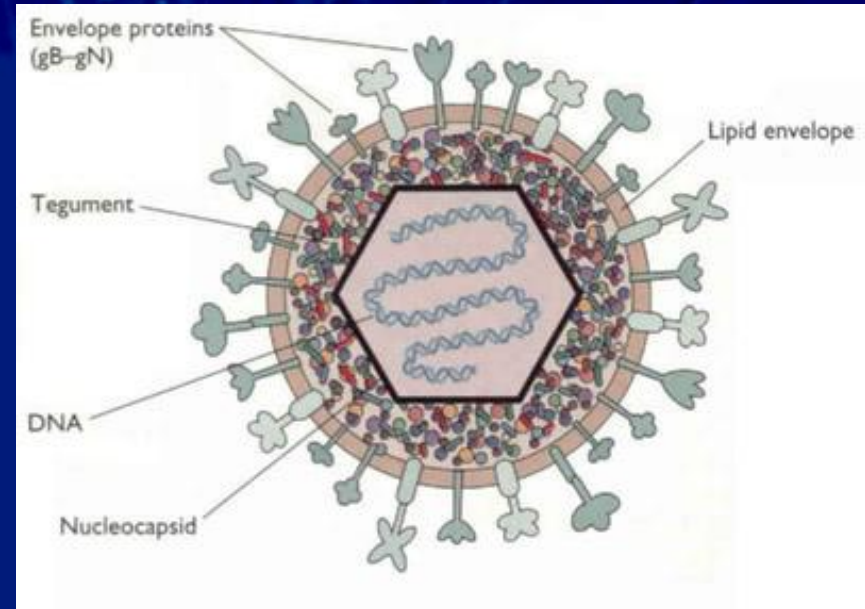
Zuoshuang “Allen” Xiang, Yongqun “Oliver” He

**Unit for Laboratory Animal Medicine
Department of Microbiology and Immunology
Center for Computational Medicine and Bioinformatics
Comprehensive Cancer Center
University of Michigan Medical School
Ann Arbor, MI 48109**



Human Herpes Simplex (HSV) Viruses

- Herpesviruses are a family of DNA viruses that cause diseases in humans and various animals.
- All herpesviruses are species-specific.
- Human herpesviruses (HHVs) have eight members, including: HSV-1 and HSV-2, the most common infectious agents of humans.
- Infectious virions are spherical.



Virion structure: a linear, double-stranded DNA molecule densely packaged into a protein cage called **capsid**. The capsid is surrounded by an amorphous protein layer, called the **tegument**, consisting of viral proteins and viral mRNAs and a lipid bilayer membrane (envelope).

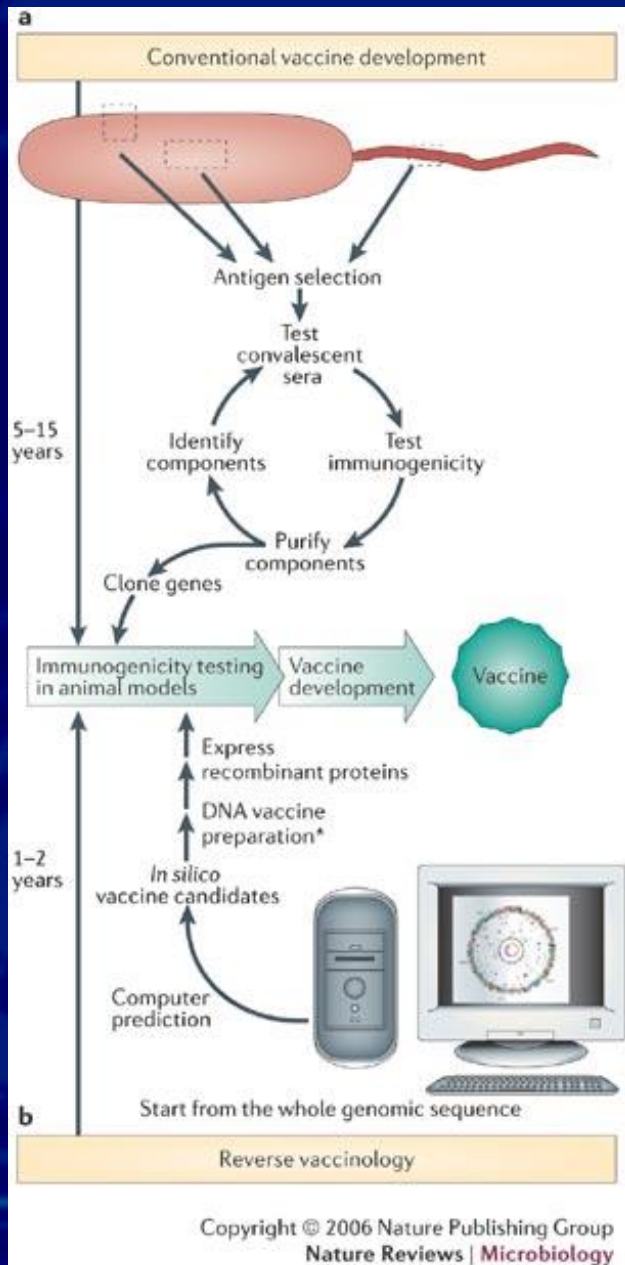
Human Herpes Simplex (HSV) Viruses

- HSV-1 and -2 are the most common infectious agents of humans.
- USA: Seroprevalence of HSV-1 and HSV-2 in adults is 68% and 21%, respectively
- USA: ~700-2000 cases of neonatal HSV infections per year occur in the US.
- No safe and effective HSV vaccines are available.
- Herpesvirus genomes available:
 - 52 herpesvirus genomes
 - 3 HHV-1 genomes: HHV-1 genome has 77 proteins
 - 12 HHV genomes
- **Question:** Can we predict vaccine candidates using these genome data?

Reverse Vaccinology (RV)



Dr. Rino Rappuoli
Pioneer of
Reverse Vaccinology

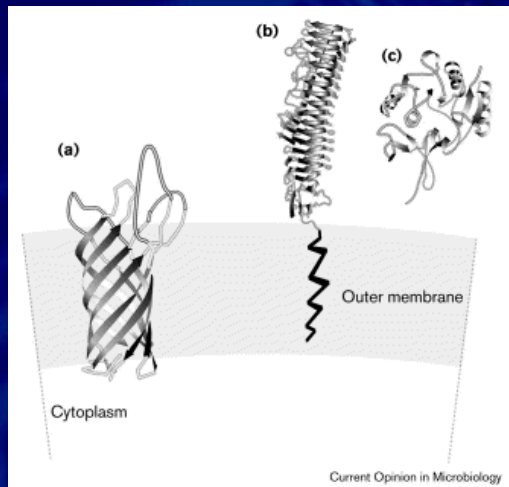


Definition: RV is a vaccine development strategy that starts with bioinformatics analysis to find potential vaccine candidates from pathogenic genomes. These candidate genes can then be tested in normal wet lab for protective immune responses.

- Johri *et al.* *Nat Rev Microbiol.* 2006 Dec; 4(12):932-42.
- Rappuoli R. *Curr Opin Microbiol.* 2000 Oct;3(5):445-50.

The First RV Success: MenB

MenB: Serogroup B meningococcus



First success: MenB

- 1) Genome sequence of *Neisseria meningitidis* serogroup B strain MC58 was obtained and used.
- 2) 570 genes predicted to code for surface-exposed or exported proteins.
- 3) 350 were successfully cloned to *E. coli*, expressed, and purified.
- 4) Mice were immunized.
- 5) 25 proteins induced bactericidal antibodies, which correlate with vaccine efficacy in humans.

Reference: Pizza M, et al. *Science*. 2000 Mar 10;287(5459):1816-20.

Reverse Vaccinology Criteria

Original Criteria:

1. Subcellular localization
 - ✓ Outer membrane proteins
 - ✓ Secreted proteins

New Criteria:

1. Transmembrane domains
 - ✓ >2 α -helix domains \rightarrow difficult to isolate
2. Adhesin probability
 - ✓ Adhesin is important for pathogen invasion
3. MHC-Epitope binding
 - ✓ MHC class I epitope \rightarrow cell-mediated immunity
 - ✓ MHC class II epitope \rightarrow antibody response
4. Sequence conservation and exclusion
 - ✓ Shared genes in pathogens but not in avirulent strains
5. Similarity to host proteins
 - ✓ Avoid autoimmunity or immune tolerance

**It is challenging to apply reverse vaccinology
without a comprehensive pipeline**

→ To address this challenge, we developed

Vaxign: <http://www.violinet.org/vaxign>

- Vaxign vaccine design for *Brucella*:

Citations: -- Xiang Z, He Y. 2009. Vaxign: a web-based vaccine target design program for reverse vaccinology. *Procedia in Vaccinology*. Volume 1, Issue 1, Pages 23-29.

-- He Y, Xiang Z. Bioinformatics analysis of *Brucella* vaccines and vaccine targets using VIOLIN. *Immunome Res*. 2010 Sep 27;6 Suppl 1:S5.

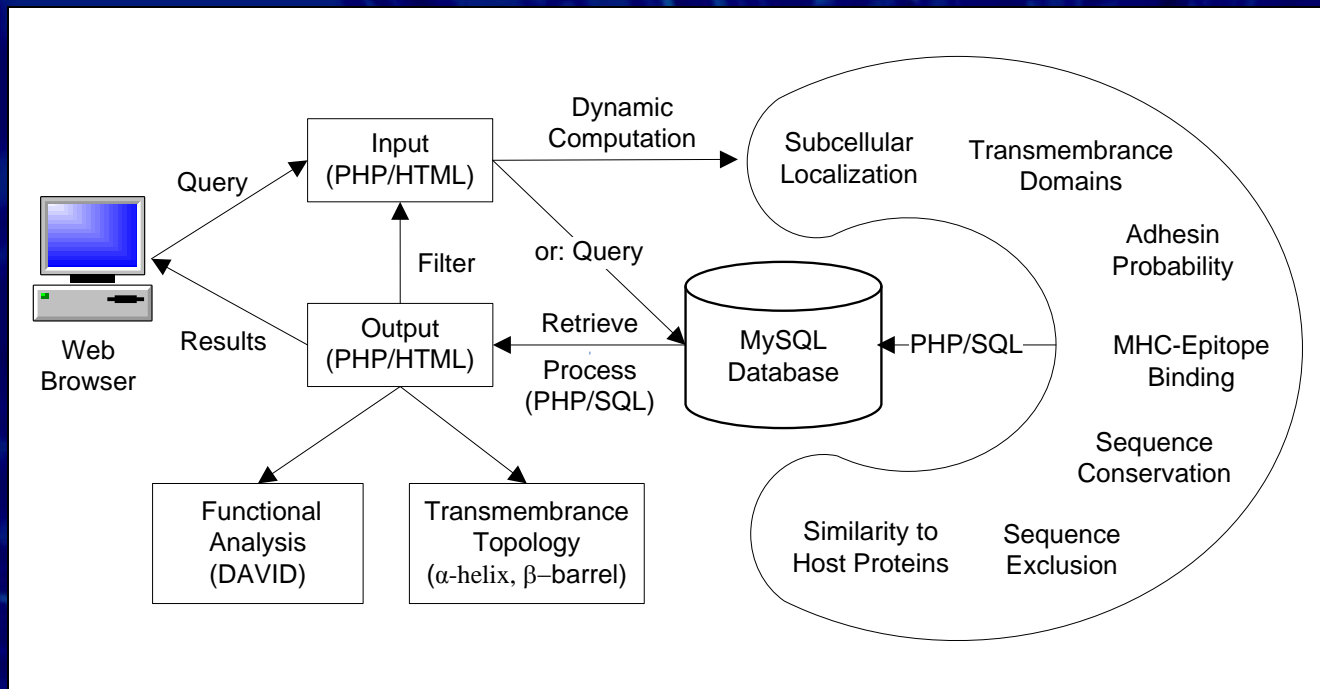
- Vaxign vaccine design for uropathogenic *E. coli*:

Citation: He Y, Xiang Z, Mobley HLT. Vaxign: the first web-based vaccine design program for reverse vaccinology and an application for vaccine development. *Journal of Biomedicine and Biotechnology*. Volume 2010 (2010), Article ID 297505, 15 pages.

Many have cited Vaxign, esp. this year!

Vaxign: Vaccine Design System

- Aim: Vaccine target prediction for reverse vaccinology



- The 1st web-based reverse vaccinology system
- Freely available: <http://www.violinet.org/vaxign>

Reference: He Y, Xiang Z, Mobley HLT. Vaxign: the first web-based vaccine design program for reverse vaccinology and an application for vaccine development. *Journal of Biomedicine and Biotechnology*. Volume 2010 (2010), Article ID 297505, 15 pages. [PMID: [20671958](https://pubmed.ncbi.nlm.nih.gov/20671958/)]

Two Forms of Vaxign Usage

Pre-computed data query

Select a Genome(s), Query a Protein (Optional), and Set up Parameters (Optional)

Select a Genome Group (Required)

Select a Genome (Required)

Sequence ID(s)

Keywords

Sort by

Filter Options:

1. Select Subcellular Localization

2. Number of Transmembrane Helices (Note: check to include this filtering option)

3. Adhesion Probability (0-1.0) (Note: check to include this filtering option)

4. Have Orthologs in of the above selected genomes

5. Exclude Proteins having Orthologs in Any of Selected Genome(s)

6. Similarity to Human Proteins Yes No Do not use this option

7. Similarity to Mouse Proteins Yes No Do not use this option

8. Similarity to Pig Proteins Yes No Do not use this option

[help](#)

- Pre-computed results
- > 200 genomes in database
- Easy to query

Dynamic analysis

Input Protein Sequence(s), Set up Parameters, and Submit a Job

Protein Sequence(s):
(Examples: Gram- *B. abortus* SodC, Gram+ *Bacillus anthracis* PA, 62317454, or YP_016495.2)
(Note: Select FASTA or other format, up to 500 sequences)

Sequence Format

File Upload

Bacterium Gram +/- (NOT for Virus)

Include Analyses [Select all] [Unselect all]

- Subcellular Localization
- Transmembrane Helices
- Adhesion Probability
- MHC Class I Binding
- MHC Class II Binding
- Similarity to Host Proteins Human Mouse Pig
- MHC class I and II epitope prediction (Note: This function is available after the above analyses)

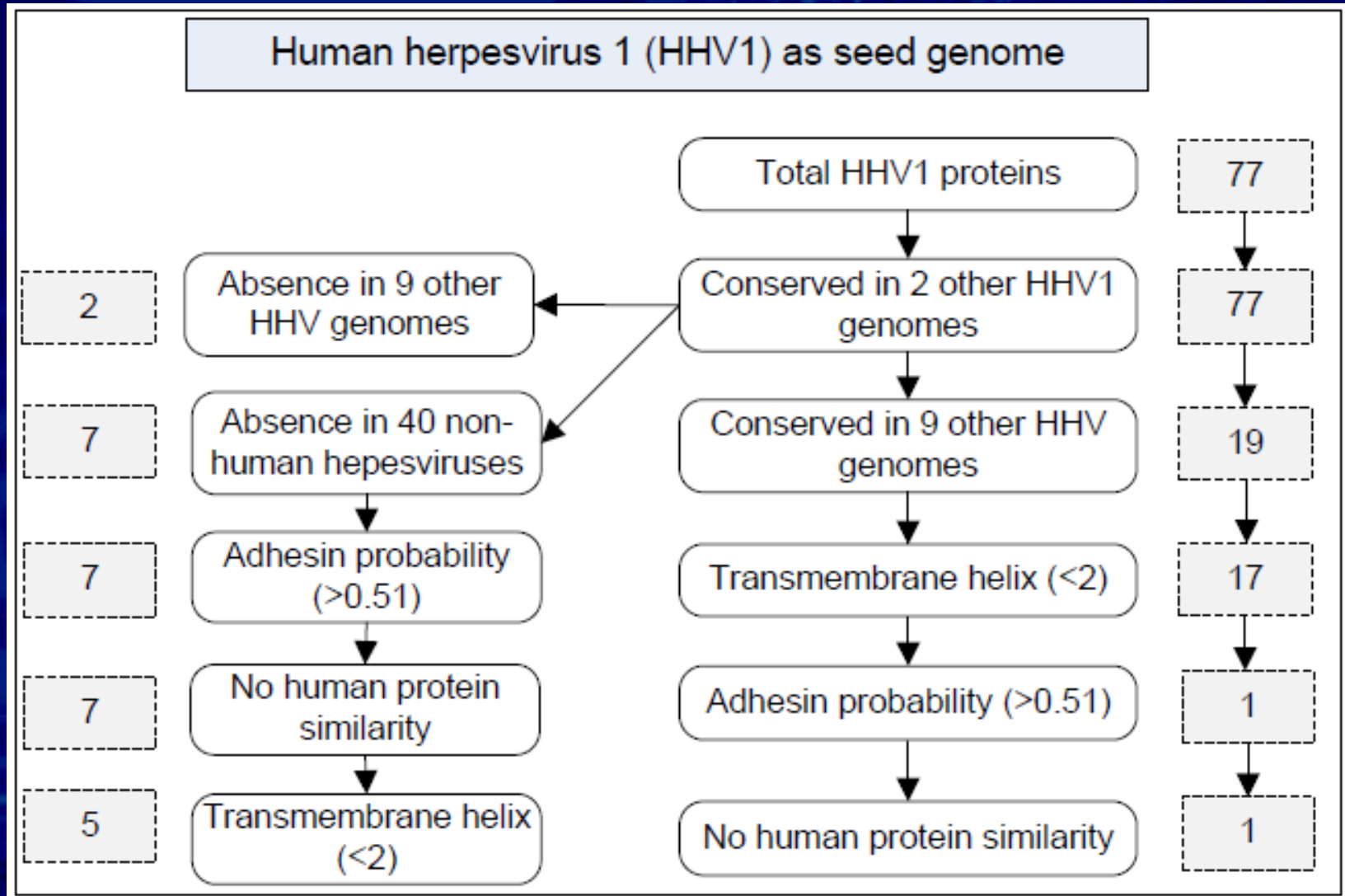
Your Note:

- Runtime execution
- Can analyze up to 500 proteins at one time

Human Herpesvirus Vaccine Design

- Downloaded from NCBI RefSeq database:
 - Three HHV-1 genomes:
 - ✓ 77 proteins in human herpesvirus 1
 - 12 HHV genomes
 - 52 herpesvirus genomes
- Vaxign pre-computation and results saved in Vaxign database
- User-friendly web interface for result query

Workflow and Result Summary



Seven HSV-1 proteins having No orthologs in all 40 non-human herpesviruses

#	<i>Protein Accession</i>	<i>Protein Note</i>	<i>Adhesin Probability</i>	<i>Trans-membrane helices</i>
1	NP_044675.1	TAP transporter inhibitor ICP47	0.079	0
2	NP_044674.1	tegument protein US11	0.245	0
3	NP_044667.1	envelope glycoprotein gJ	0.176	2
4	NP_044666.1	envelope glycoprotein gG	0.419	2
5	NP_044659.1	membrane protein UL56	0.238	1
6	NP_044661.1	neurovirulence protein ICP34.5	0.228	0
7	NP_044600.1	neurovirulence protein ICP34.5	0.228	0

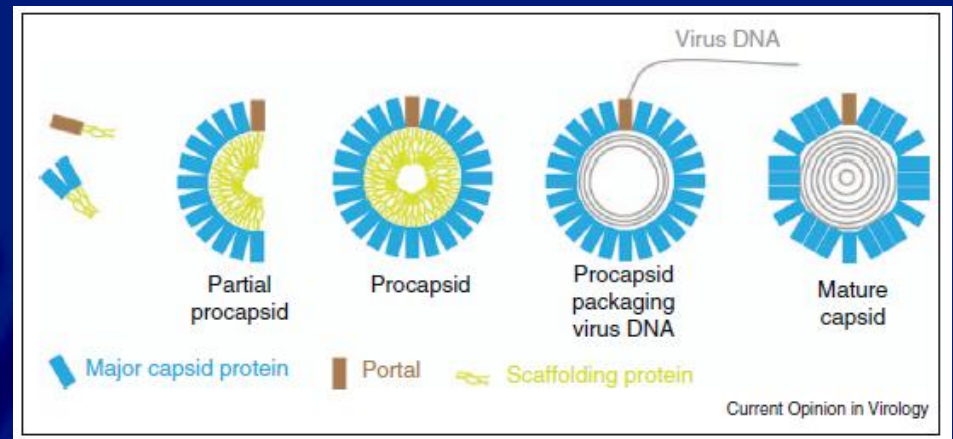
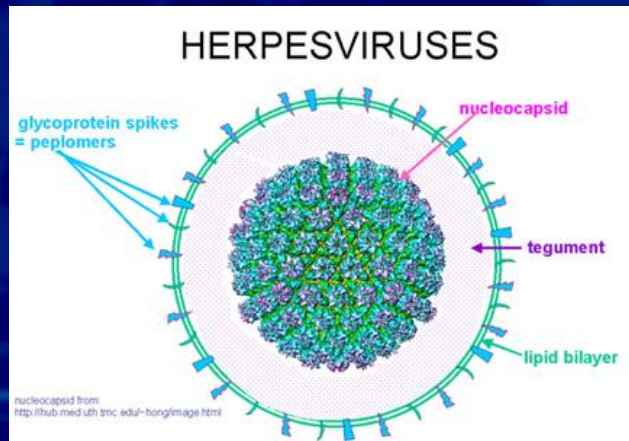
**So these proteins are human-specific.
Not good for animal test.**

19 HSV-1 proteins conserved in human herpesviruses

#	<i>Protein Accession</i>	<i>Protein GI</i>	<i>Protein Note</i>	<i>Adhesin Probability</i>	<i>Trans-membrane helices</i>
1	NP_044603.1	9629382	uracil-DNA glycosylase	0.262	0
2	NP_044606.1	9629385	helicase-primase helicase subunit	0.115	0
3	NP_044655.1	9629434	helicase-primase primase subunit	0.163	0
4	NP_044607.1	9629386	capsid portal protein	0.241	0
5	NP_044620.1	9629399	major capsid protein	0.113	0
6	NP_044627.1	9629406	capsid maturation protease (UL26)	0.386	0
7	NP_044628.1	9629407	capsid scaffold protein UL26.5	0.675	0
8	NP_044611.1	9629390	envelope glycoprotein gM	0.244	8
9	NP_044629.1	9629408	envelope glycoprotein gB	0.229	3
10	NP_044613.1	9629392	deoxyribonuclease	0.203	0
11	NP_044616.1	9629397	DNA packaging terminase subunit 1	0.165	0
12	NP_044630.1	9629409	DNA packaging terminase subunit 2	0.188	0
13	NP_044626.1	9629405	DNA packaging tegument protein UL25	0.210	0
14	NP_044634.1	9629413	DNA packaging protein UL32	0.185	0
15	NP_044635.1	9629414	DNA packaging protein UL33	0.264	0
16	NP_044625.1	9629404	nuclear protein UL24	0.195	0
17	NP_044631.1	9629410	single-stranded DNA-binding protein	0.168	0
18	NP_044632.1	9629411	DNA polymerase catalytic subunit	0.101	0
19	NP_044641.1	9629420	ribonucleotide reductase subunit 1	0.193	0

UL26.5 for HHV-1 Vaccine Development?

- U26.5 capsid scaffold protein is important for virus capsid formation
- Has not been reported for vaccine development
- U26.5 capsid scaffold protein has adhesin-like characteristics? Why?



<http://www.ncbi.nlm.nih.gov/pubmed/21927635>

Scaffold protein lost in mature virus

<http://microbiologybook.org/mhunt/dna1.htm>

MHC Class I Epitope Prediction Example

MHC Class I & II Epitope Prediction by Vaxitope:

P Value Cutoff: [help](#)

MHC Host Species:

MHC Allele:
 Supertype of MHC Class I alleles
 Supertype of MHC Class II alleles
 HLA-A*01:01
HLA-A*02:01
 HLA-A*02:02

Epitope Length:

Epitope Location (Alpha Helix):



(A)

MHC I Binding Prediction Order by allele name

Run MHC I epitope prediction using IEDB consensus method and compare with Vaxitope

Index	Epitope	Epitope Length	MHC Allele	P value	Matching from	Matching to	Location
1	GLSQHYPPHY	10	HLA-A*02:01	0.0212	63	72	outside
2	HQYPGVLFSG	10	HLA-A*02:01	0.0267	74	83	outside
3	DLFVSCMMGA	10	HLA-A*02:01	0.0495	319	328	outside

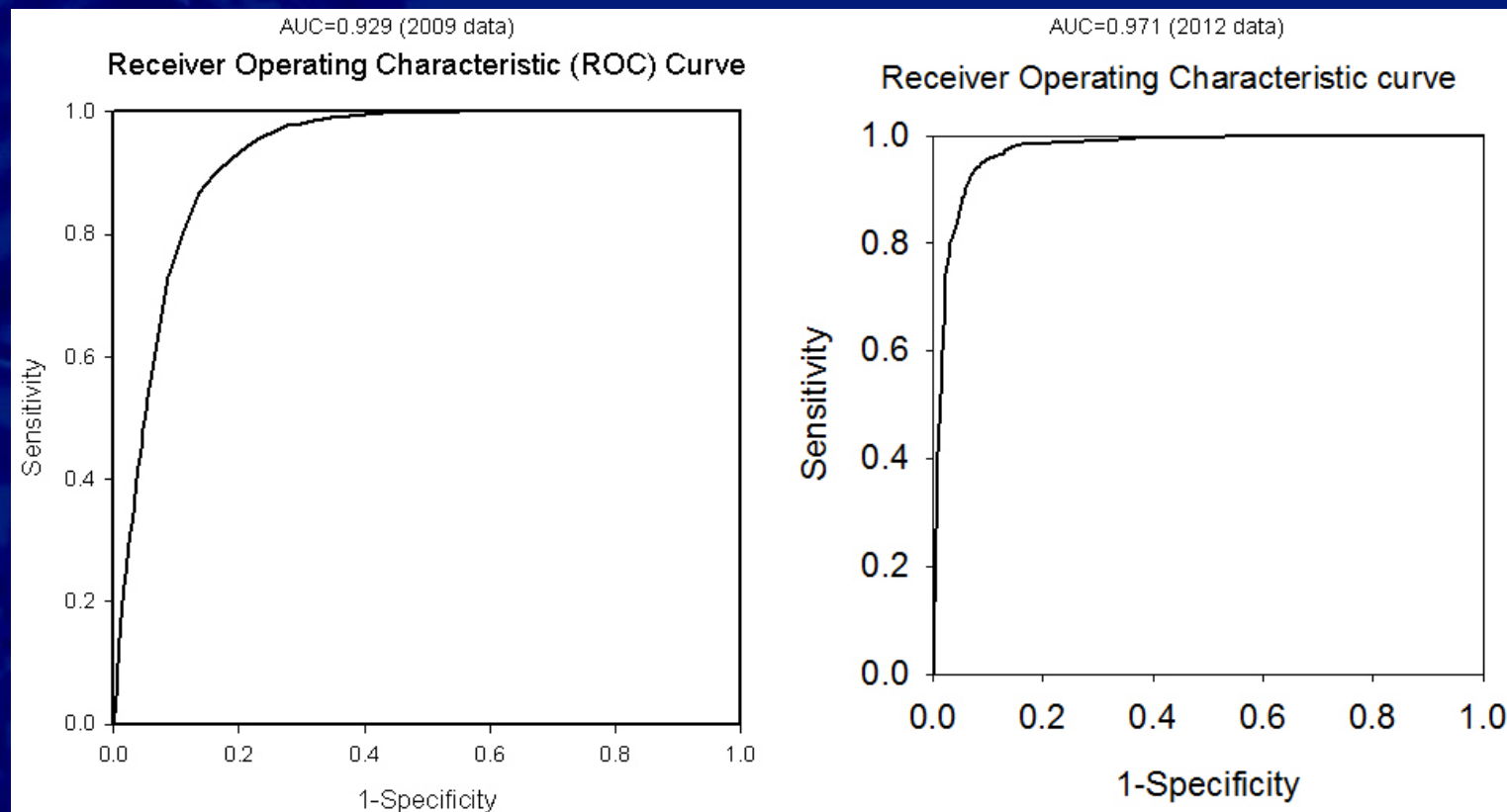
1 unique MHC I alleles.

MHC I Binding [Show all predicted epitope bindings on one page](#)

MNEVPTSGTEAPAPEGDGSSYLWIPASHYNNQLVAGHAAPQPQPHSAFGEFPAAGSVA YGPHGA**glsqhypp**
hvAhqyppgvlfsgPS PLEAQIAALVGAIAADRQAGGQPAAGDPGVRRSGKRRRYEAGPSES YCDQDEPDA
 DYPYPGEARGAERGVDSRRAARHS PGTNETITALMGAVTSLQQELAHMRARTSAPYGMYPVAHYRPQV
 GEEPEPTTTFALCPPEAVYRPPPHSAFYGE PQGSPASHAPTEFYAPAACEPGEPEPPPCPSTQTRAPLPTEP
 AFPPAATGSQPEASNAEAGALVNASSAAHVVDVDTARAA**dlfvsqmmga**R

Vaxign-Vaxitop MHC Class I and II Epitope Predictions

- Internally developed project
- Based on position specific scoring matrices (PSSM)
- Unique: Calculate statistical P-value for each prediction



Vaxign and IEDB prediction comparison on MHC Class I Epitope Prediction

Epitope predicted by IEDB consensus method

Index	Epitope	Epitope length	MHC allele	Matching from	Matching to	IC50 (IEDB consensus)	Vaxitope P-value
1	GLSQHYPPHY	10	HLA-A*02:01	63	72	1.05	0.0212
2	DLFYSQMMGA	10	HLA-A*02:01	319	328	4.45	0.0495
3	HQYPGYLFSG	10	HLA-A*02:01	74	83	5.4	0.0267
4	ALMGAVTSLQ	10	HLA-A*02:01	174	183	5.4	>0.1
5	FGFPAAGSV	10	HLA-A*02:01	46	55	6.95	>0.1
6	TALMGAYTSL	10	HLA-A*02:01	173	182	7.05	0.087
7	YLWIPASHYN	10	HLA-A*02:01	20	29	7.2	>0.1
8	SAPYGMYPV	10	HLA-A*02:01	194	203	7.4	0.0641
9	VLFSGSPLE	10	HLA-A*02:01	79	88	8.95	>0.1
10	GMYTFYAHYR	10	HLA-A*02:01	198	207	9.45	>0.1
11	DTARAADLFV	10	HLA-A*02:01	313	322	9.45	>0.1
12	GVLFSGSPLE	10	HLA-A*02:01	78	87	9.9	>0.1

Epitope predicted by Vaxitop method

Index	Epitope	Epitope Length	MHC Allele	Matching from	Matching to	P-value	IC50 (IEDB consensus)
1	GLSQHYPPHY	10	HLA-A*02:01	63	72	0.0212	1.05
2	HQYPGYLFSG	10	HLA-A*02:01	74	83	0.0267	5.4
3	DLFYSQMMGA	10	HLA-A*02:01	319	328	0.0495	4.45
4	SAPYGMYPV	10	HLA-A*02:01	194	203	0.0641	7.4
5	GCPAAGDPGY	10	HLA-A*02:01	106	115	0.0762	>50
6	TALMGAYTSL	10	HLA-A*02:01	173	182	0.087	7.05

- Results overlap
- Vaxitop is more conservative in predicting positive results.

Conclusion

- Vaxign is a specific and sensitive predictor of vaccine targets.
- It is web-based, user-friendly, and free.
- More information in next demo and hands-on training

<http://www.violinet.org/vaxign/>

Discussion

Challenges:

How to better rank predicted vaccine targets?
What's unique about protective antigens?

Directions: integrated with microarray, proteomics, literature mining, 3D structure, and VIOLIN components.

Acknowledgements

He Laboratory / U of Michigan (UM)



**

** Zuoshuang “Allen” Xiang

Funding: NIH-NIAID Grant 1R01AI081062