

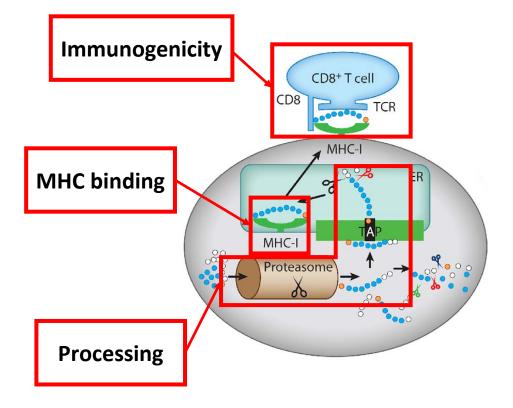
# T Cell Class I Tools Binding, Processing, Immunogenicity

Introduction

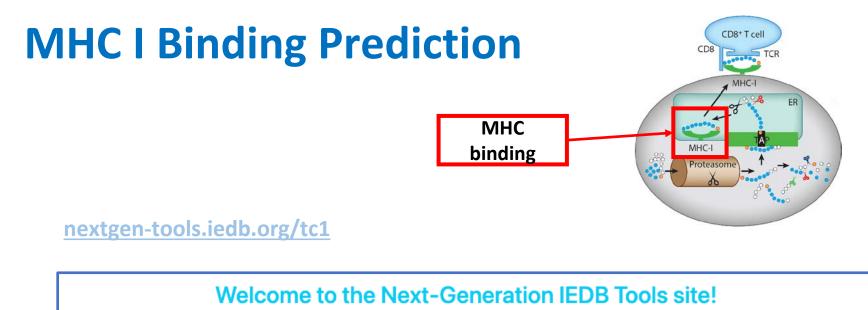
Presented by: J. Greenbaum, Bioinformatics Core Director

2023 IEDB User Workshop

### Endogonous Antigen Processing Pathway (Class I)



- Antigens generated within the cell
  - Viral particles
  - Self proteins
  - DRiPs (Defective Ribosomal Particles
- Different factors influence peptide being "epitope"



As a companion site to the Immune Epitope Database (IEDB), this site provides a collection of tools for the prediction and analysis of immune epitopes.

#### New User? Learn to use the website here!

### T Cell Prediction - Class I

MHC class I binding affinity, TAP processing, and Immunogenicity predictions

Type/paste/drag a sequence into this box or click 'Run' to use the example sequence:

>SARS2 spike glycoprotein

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFND GVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFL MDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYY

MHC Allele(s)

Ex: HLA-A\*02:01



# **Class I MHC Molecule**

- Expressed by almost all nucleated cells
- Presents antigen to CD8+ T cells (Cytotoxic T cells)
- One MHC encoded polymorphic chain (α) (2nd chain – β2microglobulin, aka B2M)
- The binding groove is closed at both ends and can accommodate peptides of 8-11 AA
- Only  $\alpha$  chain impacts binding

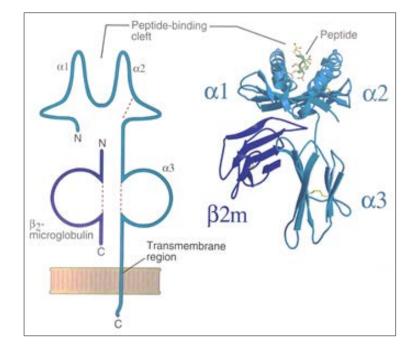


Figure Source Cellular & Molecular Immunology, 5th Ed by Abbas and Lichtman

# **MHC Binding Predictions**

- MHC molecules are highly polymorphic thousands of different variants exist
- MHC-peptide binding is **promiscuous** in nature
- Experimental characterization of peptide–MHC interactions is highly cost-intensive
- Prediction methods facilitate selection of potential epitopes from a pool of peptides

Peptide binding data HLA-A\*01:01

Peptide	IC <sub>50</sub> (nM)
ASFCGSPY	51.4
LTDFGLSK	739.3
FTSFFYRY	1285.0
KSVFNSLY	1466.0
RDWAHNSL	1804.6
FSSCPVAY	1939.4
RNWAHSSL	2201.7
LSCAASGF	2830.1
LASIDLKY	3464.0

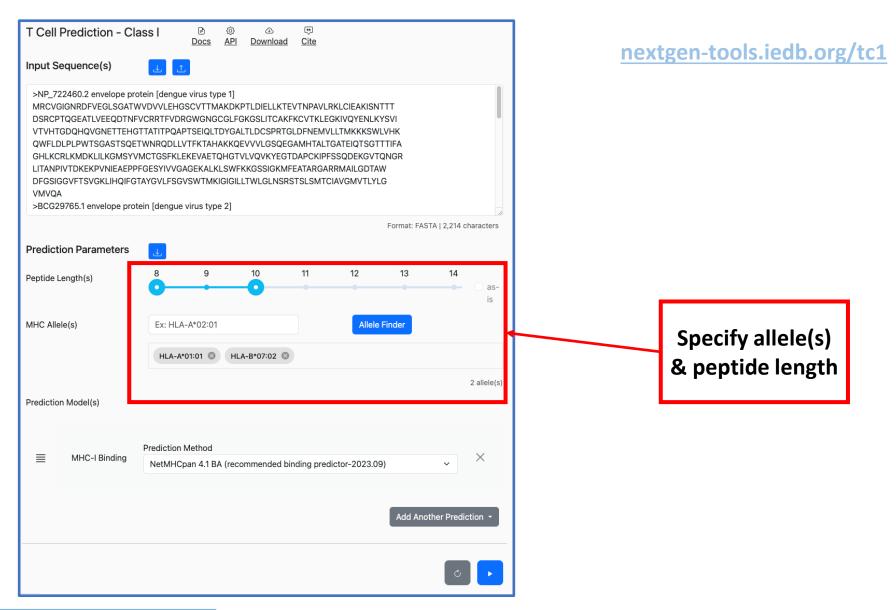
### Machine learning algorithms



# **MHC-I Binding Prediction – Example**



### **MHC-I Binding Prediction – Example**



### **Allele Selection – Reference Set for Global Coverage**

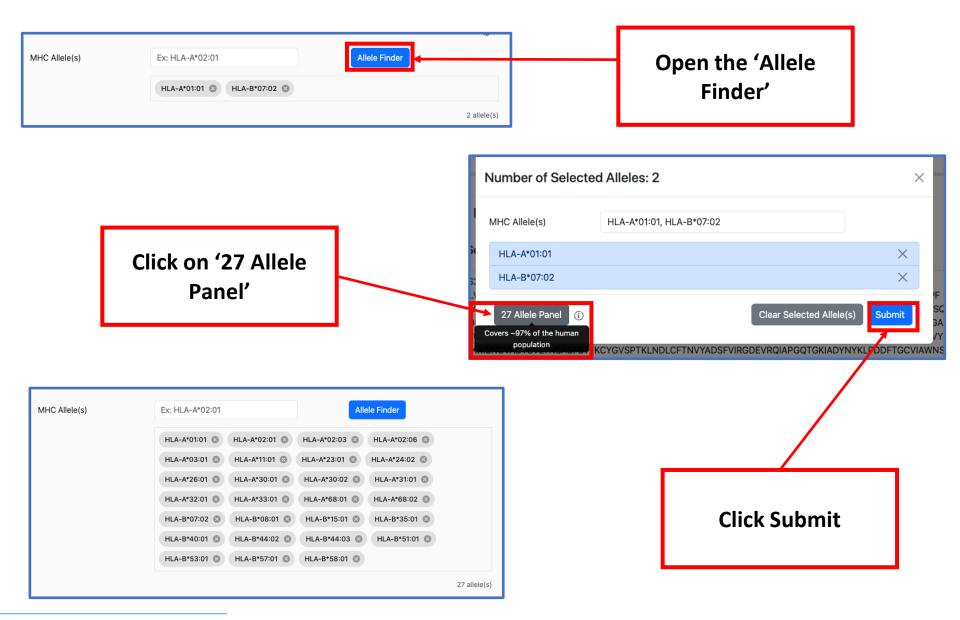
- Reference set of
   27 alleles
- Covers > 97% of population

https://iedb.zendesk.com/entrie
s/25054538-HLA-allele-
<u>frequencies</u>

HLA-A	Frequency	HLA-B	Frequency
A*01:01	16.2	B*07:02	13.3
A*02:01	25.2	B*08:01	11.5
A*02:03	3.3	B*15:01	5.2
A*02:06	4.9	B*35:01	6.5
A*03:01	15.4	B*40:01	10.3
A*11:01	12.9	B*44:02	9.2
A*23:01	6.4	B*44:03	7.6
A*24:02	16.8	B*51:01	5.5
A*26:01	4.7	B*53:01	5.4
A*30:01	5.1	B*57:01	3.2
A*30:02	5.0	B*58:01	3.6
A*31:01	4.7		
A*32:01	5.7		
A*33:01	3.2		
A*68:01	4.6		
A*68:02	3.3		

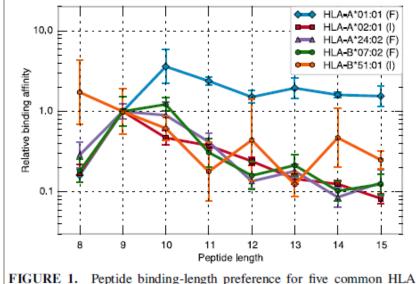
2023 IEDB User Workshop

### **Using the Reference Allele Panel**



### **Natural Length Distribution in Epitope Prediction**

• Alleles differ in their preference for lengths on binding and presentation of peptides



alleles. The length preference for each HLA was determined by measuring

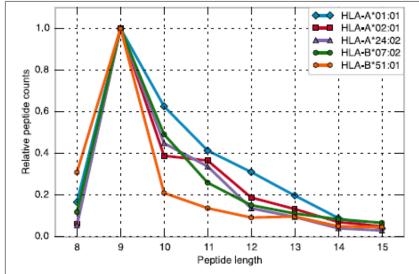


FIGURE 2. Length profiles of naturally presented peptides for five HLA molecules. Large datasets of HLA-I ligands were determined by the elu-

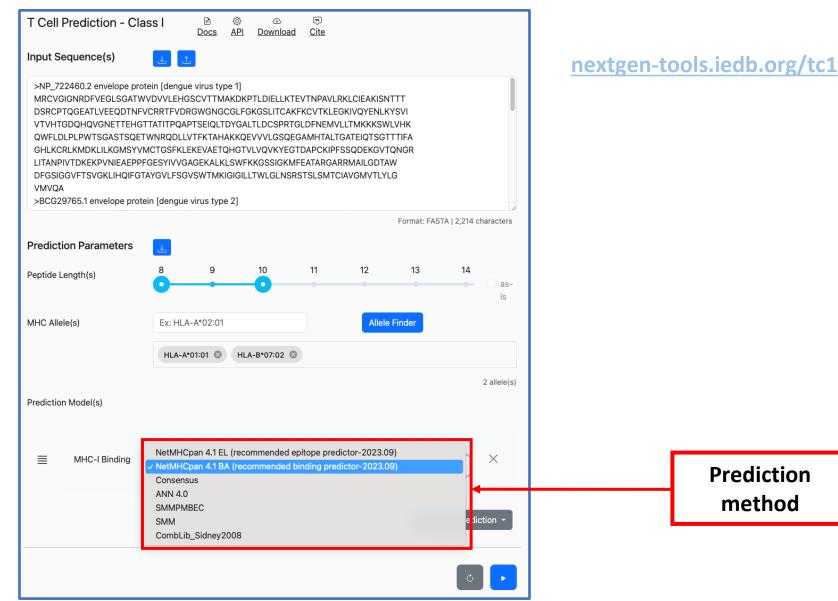
J Immunol. 2016 Feb 15;196(4):1480-7. doi: 10.4049/jimmunol.1501721. Epub 2018 Jan 18.

The Length Distribution of Class I-Restricted T Cell Epitopes Is Determined by Both Peptide Supply and MHC Allele-Specific Binding Preference.

Trolle T<sup>1</sup>, McMurtrey CP<sup>2</sup>, Sidney J<sup>3</sup>, Bardet W<sup>2</sup>, Osborn SC<sup>2</sup>, Kaever T<sup>3</sup>, Sette A<sup>3</sup>, Hildebrand WH<sup>2</sup>, Nielsen M<sup>4</sup>, Peters B<sup>5</sup>.

PMID: 26783342 PMCID: PMC4744552 DOI: 10.4049/jimmunol.1501721

### **MHC-I Binding Prediction – Example**



method

# **Guidelines: Choosing the Prediction Method**

IEDB Tools Version	Recommended Method
2023.09 (current)	NetMHCPan 4.1 EL (epitope) NetMHCPan 4.1 BA (binding)
2023.05	NetMHCPan 4.1 EL
2020.04	NetMHCPan 4.0 EL
2.22 and earlier	Consensus, if available; otherwise, NetMHCpan

- IEDB-recommended **epitope** predictor
  - Employs NetMHCpan **EL** 4.1 across all alleles
- IEDB-recommended **binding** predictor
  - Employs NetMHCpan **BA** 4.1 across all alleles
- Recommendation will change with the new benchmark studies

<sup>2023</sup> IEDB User Workshop

### **MHC Class I binding prediction benchmarks**

tools.iedb.org/auto\_bench/mhci/weekly/

### **MHC I Automated Server Benchmarks**

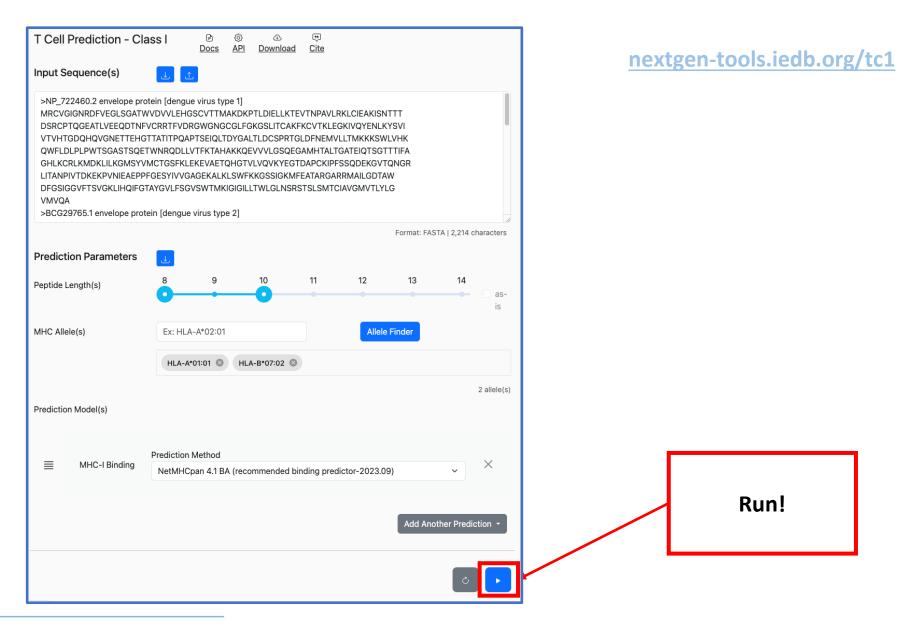
This is a <u>live</u> ranking of MHC I servers based on performance, which continues to be reevaluated over time. The weekly IEDB releases are automatically checked for datasets large enough to add to the benchmarks. The benchmark metrics in the table below will only be updated on releases where such new data is becoming available.

#### Accumulated overall ranking scores

Ranking scores based on data sets submitted to the IEDB for the last at least 5 references.

Server 🖕	<u>2023-</u> <u>10-19</u> ≑	<u>2023-</u> <u>08-10</u> ≑	<u>2023-</u> <u>08-03</u> ≑	<u>2023-</u> <u>06-15</u> ≑	<u>2023-</u> <u>05-11</u> ≑	<u>2023-</u> <u>04-20</u> ♦	<u>2022-</u> <u>11-24</u> ≑	<u>2022-</u> <u>10-28</u> ♦	<u>2022-</u> <u>10-06</u> ≑	<u>2022-</u> 07-22 \$	<u>2022-</u> 05-20 ♦	<u>2022-</u> 04-29 ♦	<u>2021-</u> <u>12-24</u> ≑	<u>2021-</u> <u>12-10</u> ≑	<u>2021-</u> <u>11-05</u> ♦	<u>2021-</u> <u>10-08</u> ♦
<u>NetMHCpan</u> <u>4.1 BA</u>	73	74	74	65	66	67	66	72	83	76	80	75	72	64	64	63
NetMHCcons	69	70	65	64	60	57	60	56	49	61	60	69	68	57	56	49
<u>NetMHCpan</u> <u>4.0 BA</u>	68	66	61	48	60	60	60	68	76	76	80	80	80	71	70	62
NetMHCpan <u>3.0</u>	65	65	64	58	68	65	70	73	72	74	74	73	72	68	64	58
<u>ANN 3.4</u>	64	66	62	73	64	63	61	50	44	48	45	47	45	46	50	53
<u>mhcflurry</u> <u>1.2.0</u>	63	56	51	56	58	61	56	59	42	33	49	54	60	61	59	68
NetMHCpan 2.8	60	65	59	41	60	57	63	69	46	54	48	65	65	63	60	48
<u>ANN 4.0</u>	56	55	60	53	64	65	67	73	76	64	57	56	51	54	52	50
<b>SMMPMBEC</b>	50	56	73	75	57	60	61	61	73	62	50	43	46	53	53	64
<u>IEDB</u> <u>Consensus</u>	48	52	58	65	71	72	69	71	69	58	60	52	54	58	57	62
SMM	46	51	54	55	56	57	56	61	58	52	54	39	43	50	49	63
<u>NetMHCpan</u> <u>4.0 EL</u>	45	42	37	51	46	43	46	37	25	36	34	44	42	50	47	42
NetMHCpan 4.1 EL	34	32	32	56	51	48	42	36	25	26	39	50	49	57	52	44
<b>PickPocket</b>	34	40	46	54	57	57	68	68	55	41	40	34	38	52	54	56
ARB	31	34	48	52	48	49	53	41	37	42	34	27	24	26	27	34

### **MHC-I Binding Prediction – Example**



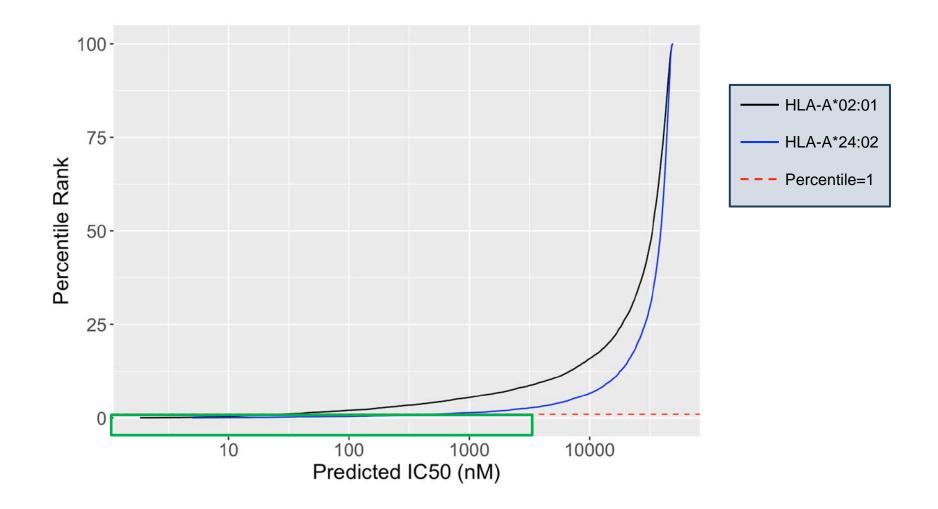
### **How the Tool Works**

- Breaks the sequence into all possible peptides of chosen length(s), unless 'as-is' option is selected
- Predicts the binding affinity / elution score for each peptide based on the method
- Compares the predicted affinity / elution score to that of a large set of randomly selected peptides
- Assigns a percentile rank depending on individual predicted affinity / elution score
- Calculates the median percentile rank

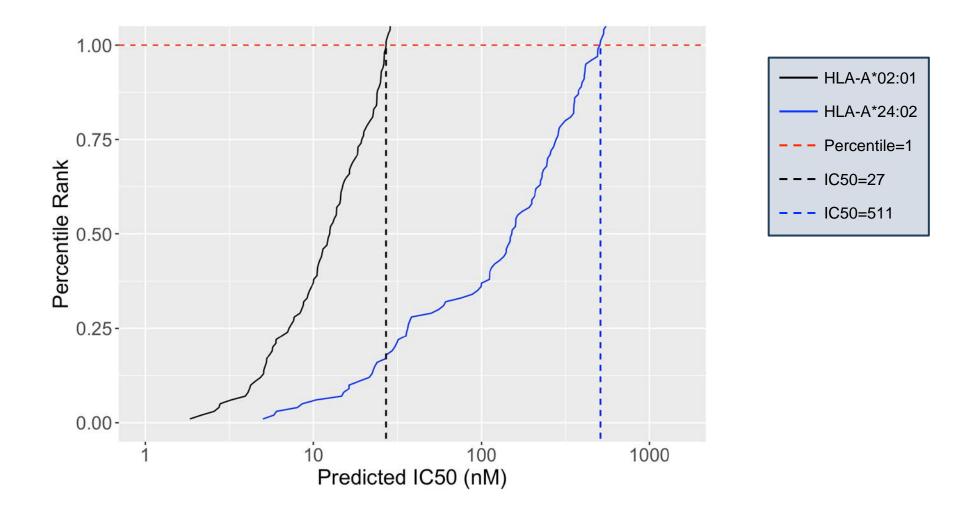
### **MHC-I Binding Prediction – Example**

	Ç	Calculating re	sults, please wa	it Click <u>here</u> to receive em	Cancel Run	<u>nextgen-tools.iedb.org/to</u>
Peptide Table Net	MHCpan Allele Dis	stance Sequen	ce Table			]
Show 10 V rot	NS	ی <u>Download</u> <u>Re</u> 1 to 10 of 11,6		Image: marked state state       ay Columns       Save Table State       Previous     1	2 3 4 5 1168 Next	<b>Output</b> (sorted low-to-high by percentile rank)
seq # 🖓 🍦	peptide	start 🖓 🍦	allele 🖓 🍦		netmhcpan_ba percentile 🖓 🗍	
3	TPRSPSVEV		HLA-B*07:02	8.93	0.03	A percentile rank
2	ITEAELTGY	170	HLA-A*01:01	33.71	0.04	for a peptide is
1	TSEIQLTDY	170	HLA-A*01:01	54.13	0.04	the percentage of
3	TPTWNRKEL	226	HLA-B*07:02	16.40	0.05	randomly
3	STTEAILPEY	167	HLA-A*01:01	121.67	0.08	sampled peptides
2	SITEAELTGY	169	HLA-A*01:01	133.29	0.09	scoring better
3	RCPTQGEAAL	73	HLA-B*07:02	33.24	0.1	than the peptide.
1	TPQAPTSEI		HLA-B*07:02	37.66	0.11	
4	ITPRSPSVEV	164	HLA-B*07:02	40.58	0.11	

### **IC50 value distributions vary by allele**

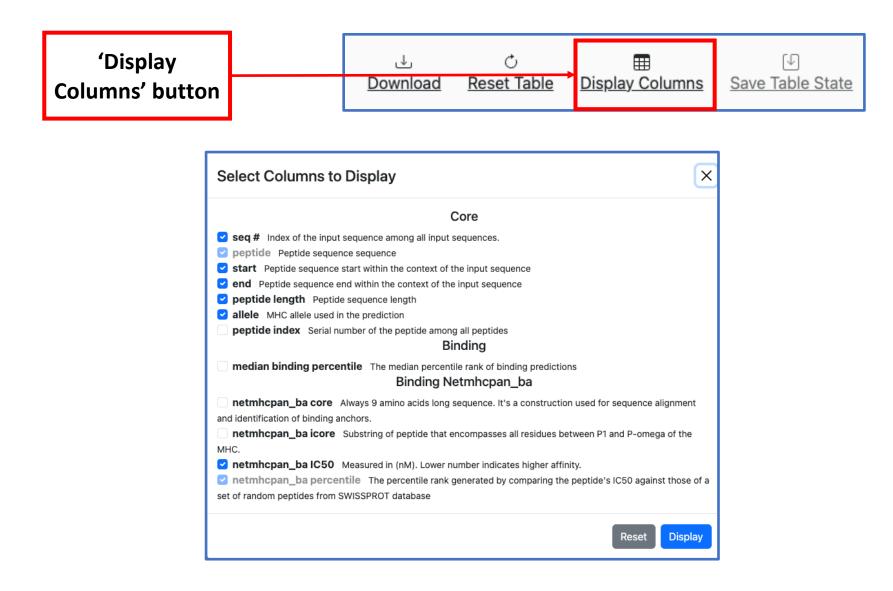


### Mapping IC50 values to percentile ranks enables comparisons across alleles



#### 2023 IEDB User Workshop

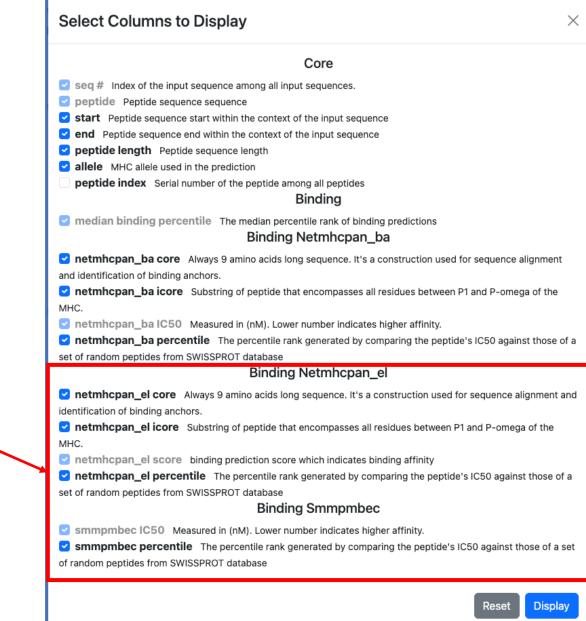
### **Show / Hide Output Columns**



# With more methods, more fields are visible

(Mo methods, Mo columns)

> Additional selected methods



### **Downloaded Prediction Results**

لي Download Res	ਂ et Table Dis	⊞ lay Columns	<u>s Sav</u>	L	IJ ble S	State	
All rows 🕨	TSV	Pre	evious	1	2	3	4
Displayed rows →	csv	<b>₽</b>	peptic	de le	ngth	ı V	k
IIAQY	JSON	873					9

	A	В	С	D	E	F	G	н	I	J	К	L
1	seq #	peptide	start	end	peptide length	allele	peptide index	median binding percentile	netmhcpan_ba core	netmhcpan_ba icore	netmhcpan _ba IC50	netmhcpan _ba percentile
2	3	TTEAILPEY	168	176	9	HLA-A*01:01	3576	0.03	TTEAILPEY	TTEAILPEY	30.97	0.03
3	4	TPRSPSVEV	165	173	9	HLA-B*07:02	5030	0.03	TPRSPSVEV	TPRSPSVEV	8.93	0.03
4	1	TSEIQLTDY	170	178	9	HLA-A*01:01	658	0.04	TSEIQLTDY	TSEIQLTDY	54.13	0.04
5	2	ITEAELTGY	170	178	9	HLA-A*01:01	2119	0.04	ITEAELTGY	ITEAELTGY	33.71	0.04
6	3	TPTWNRKEL	226	234	9	HLA-B*07:02	3634	0.05	TPTWNRKEL	TPTWNRKEL	16.4	0.05
7	3	STTEAILPEY	167	176	10	HLA-A*01:01	4060	0.08	STEAILPEY	STTEAILPEY	121.67	0.08
8	2	SITEAELTGY	169	178	10	HLA-A*01:01	2605	0.09	STEAELTGY	SITEAELTGY	133.29	0.09
9	3	RCPTQGEAAL	73	82	10	HLA-B*07:02	3966	0.1	RPTQGEAAL	RCPTQGEAAL	33.24	0.1
10	1	LTDYGALTL	175	183	9	HLA-A*01:01	663	0.11	LTDYGALTL	LTDYGALTL	202.56	0.11
11	1	TPQAPTSEI	165	173	9	HLA-B*07:02	653	0.11	TPQAPTSEI	TPQAPTSEI	37.66	0.11
12	4	ITPRSPSVEV	164	173	10	HLA-B*07:02	5516	0.11	IPRSPSVEV	ITPRSPSVEV	40.58	0.11
13	1	PTSEIQLTDY	169	178	10	HLA-A*01:01	1144	0.13	PTSEIQLTY	PTSEIQLTDY	236.77	0.13
14	3	CPTQGEAAL	74	82	9	HLA-B*07:02	3482	0.14	CPTQGEAAL	CPTQGEAAL	60.04	0.14
15	4	VPHAKRQDV	242	250	9	HLA-B*07:02	5107	0.16	VPHAKRQDV	VPHAKRQDV	66.2	0.16

### **Emailed Prediction Results and Link**

Pipeline Nan	ne and Email	×	
Add or change the finished.	e pipeline title and email address and we will send you an email when your job has		Fill in details and click Submit
Pipeline Name	Workshop demo		and click Submit
Email	my-email@iedb.org		
	Sut	mit	

IEDB Tools prediction Result for job "Workshop Demo"       Image: Constraint of the second seco										
<b>IEDB Tools</b> <noreply+nextgen-tools@iedb.org> to me ▼</noreply+nextgen-tools@iedb.org>	@ 4:06 PM (0 minutes ago) 🖌 🛣	← Reply :								
IEDB Tools prediction Result for job "Workshop Demo" below. Please go back to the browser or visit <u>https://nextgen-to 0fc6e23a28e5</u> for details.										
One attachment • Scanned by Gmail ()										

# **Selection of "Binders"**

- Pick peptides below percentile rank 1.0
- Pick peptides below predicted binding affinity of 500 nM
  - IC50 < 50 nM high affinity
  - IC50 < 500 nM intermediate affinity
  - IC50 < 5000 nM low affinity
  - Sette et al. 1994, J. Immunology (PMID: 7527444)
  - Ensures that all peptides have reasonable affinity
- Pick top 1% of peptides for each allele/length combination to cover most of immune responses
  - Moutaftsi et al. 2006 (PMID: 16767078)
  - Kotturi et al. 2007 (PMID: 17329346)
  - Ensures equal number of peptides per allele
- Select based on allele specific binding affinity threshold

<sup>2023</sup> IEDB User Workshop

## **Allele-specific Thresholds**

Tools -Help & Info 🝷

Support

Cite Us

Licensing

Contribute a Tool

**Tools Documentation** 

### Next-Generation (NXG) IEDB Tools IMMUNE EPITOPE DATABASE AND ANALYSIS RESOURCE Search docs Overview Available Tools

#### Thresholds and interpreting scores

peptide 🖓

1 YLQPRTFLL

1 FIAGLIAIV

The IEDB currently recommends using the percentile rank as the metric for ranking binding predictions. A percentile rank of <= 1% has been demonstrated to cover 80% of the immune response for many alleles. For more information on selecting thresholds, please consult these guidelines.

Once the prediction is completed, an output table will be displayed similar to the one shown below. end 🖓

269

1220

1000

417

2

q.

allele 🖓

277 HLA-A\*02:01

1228 HLA-A\*02:01

1008 HLA-A\*02:01

425 HLA-A\*02:01

10 HLA-A\*02:01

smm score 🖓 🙏

11.9980

20.1888

25.4162

36.7375

38,8249

□ T cell class I

#### Selecting thresholds (cut-offs) for MHC class I and II binding predictions

Ward Fleri

8 Ward Hien posted this on May 21, 2013 04:33 PM

#### MHC class I

For MHC class I T cell epitope predictions, selection of predicted binders can be done based on the percentile rank or MHC binding affinity. The IEDB currently recommends making selections based on a percentile rank of <= 1% for each (MHC allele, length) combination to cover most of the immune responses.1, 2 Alternatively, a binding affinity (IC50) threshold of 500 nM identifies peptide binders recognized by T cells and this threshold can be used to select peptides.<sup>3</sup> Recently, a paper from our group showed that absolute binding affinity threshold correlates better with immunogenicity and also that, for even better correlation, MHC-specific thresholds should be used.<sup>4</sup> The tables below show the allele-specific thresholds for the 38 most common HLA-A and HLA-B alleles, representative of the nine major supertypes. The tables can also be downloaded as an RTF file (see attached file).

Alleles s	orted by popula	ation frequency	All	Alleles sorted by name			
Allele	Population frequency of allele	Allele specific affinity cutoff (IC50 nM)	Allele	Population frequency of allele	Allele specific affinity cutoff (IC50 nM)		
A*0201	25.2	255	A*0101	16.2	884		
A*2402	16.8	849	A*0201	25.2	255		
A*0101	16.2	884	A*0203	3.3	92		
A*0301	15.4	602	A*0206	4.9	60		
B*0702	13.3	687	A*0301	15.4	602		
A*1101	12.9	382	A*1101	12.9	382		
B*0801	11.5	663	A*2301	6.4	740		
B*4001	10.3	639	A*2402	16.8	849		
B*4402	9.2	904	A*2501	2.5	795		
B*4403	7.6	780	A*2601	4.7	815		
B*3501	6.5	348	A*2902	2.9	641		
A*2301	6.4	740	A*3001	5.1	109		
A*3201	5.7	131	A*3002	5	674		
B*5101	5.5	939	A*3101	4.7	329		
B*5301	5.4	538	A*3201	5.7	131		
B*1501	5.2	528	A*3301	3.2	606		
A*3001	5.1	109	A*6801	4.6	197		
A*3002	5	674	A*6802	3.3	259		

J Immunol. 2013 Dec 15;191(12):5831-9. doi: 10.4049/jimmunol.1302101. Epub 2013 Nov 4.

 $\nabla$ 

0.3

0.4

0.5

0.7

0.7

#### HLA class I alleles are associated with peptide-binding repertoires of different size, affinity, and immunogenicity.

Paul S<sup>#1</sup>, Weiskopf D<sup>#1</sup>, Angelo MA<sup>1</sup>, Sidney J<sup>1</sup>, Peters B<sup>1</sup>, Sette A<sup>1</sup>.

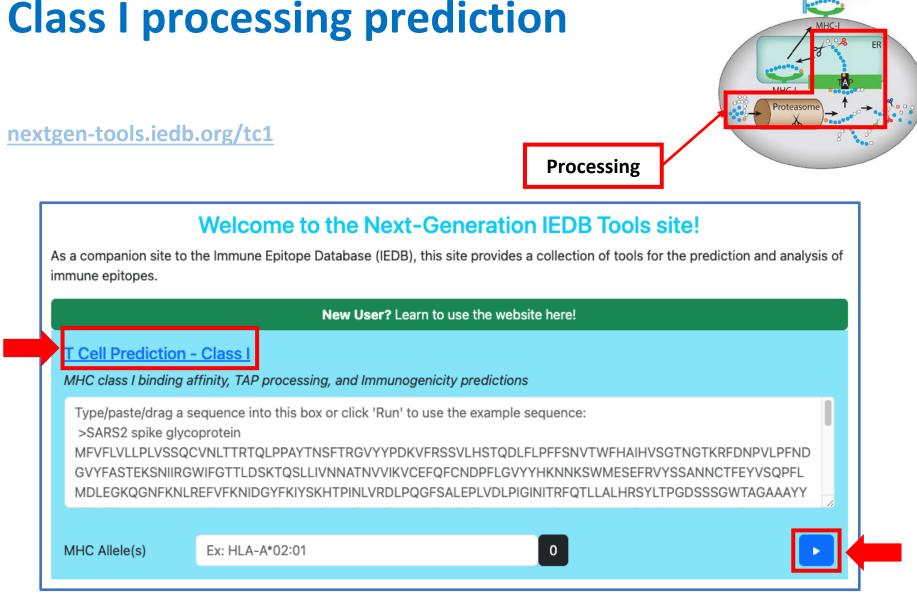
PMID: 24190657 PMCID: PMC3872965 DOI: 10.4049/jimmunol.1302101

#### https://doi.org/10.4049/jimmunol.1302101

2023 IEDB User Workshop

### Recommendations

- All approaches (affinity and ranking) are reasonable, and have been applied in numerous studies
- Thresholds can be combined (peptides in top 1% and IC50 <500nM)</li>
- Current studies suggest that allele specific thresholds can be derived



2023 IEDB User Workshop

CD8<sup>+</sup> T cel

				1						
T Cell I	Prediction - Clas	sl	₽ <u>Docs</u>	③ <u>API</u>	④ Download	•• <u>Cite</u>				
Input Se	equence(s)	1								
MRCVG DSRCP VTVHTQ QWFLD GHLKCI LITANP DFGSIG VMVQA	22460.2 envelope prote 900NRDFVEGLSGATWV TQGEATLVEEQDTNFVC GDQHQVGNETTEHGTT 12PLPWTSGASTSQETW RLKMDKLILKGMSYVM IVTDKEKPVNIEAEPPFG 30VFTSVGKLIHQIFGTA 4 9765.1 envelope proteir	DVVLEHO CRTFVDI TATITPQA /NRQDLL CTGSFKL ESYIVVG YGVLFSG	GSCVTTM/ RGWGNGC PTSEIQLTI VTFKTAH/ EKEVAETC AGEKALKI	AKDKF CGLFG DYGAI AKKQE QHGT\ LSWFF IGIGILI	KGSLITCAK TLDCSPRT( VVVLGSQE (LVQVKYEG KGSSIGKM	FKCVTKLI GLDFNEM GAMHTAL TDAPCKIP FEATARGA	EGKIVQYENLK VLLTMKKKSW TGATEIQTSG1 FSSQDEKGV1 ARRMAILGDTA	(YSVI /LVHK ITTIFA IQNGR W		
								Format: FAS	TA   2,214 (	characters
Predicti	ion Parameters	ل								
Peptide L	ength(s)	8	9		10	11	12	13	14	as-
										is
MHC Alle	le(s)	Ex: HLA	A-A*02:01				Allele	Finder		
		HLA-A	•01:01 🕲	HL	A-B*07:02 🔇					
Prediction	n Model(s)									2 allele(s
	MHC-I Processing		tion Metho c Processir		dictions				~	$\times$
≡	MHC-I Binding Me	thods	NetMH	Cpan	4.1 BA (reco	mmended	binding predi	ctor-2023.09	)	~
	Proteasome			Tran	sporter asso	ociated wi	th antigen pro	ocessing		
	Cleavage immuno ~		Max Precu Extension	irsor	1	A	lpha Factor	0.2		
								Add Anoth	ner Predict	tion -
									Ċ	

# **Class I Processing 'Combined Predictor'**

nextgen-tools.iedb.org/tc1

- Combines predictions for:
  - MHC binding
  - proteasomal cleavage
  - TAP transport
- Trained on specific *in vitro* datasets

### **Proteasomal Cleavage & TAP Transport Parameters**

MHC-I Binding Methods
NetMHCpan 4.1 BA (recommended binding predictor-2023.09)

Proteasome
Transporter associated with antigen processing

Cleavage
Max Precursor

immuno
1

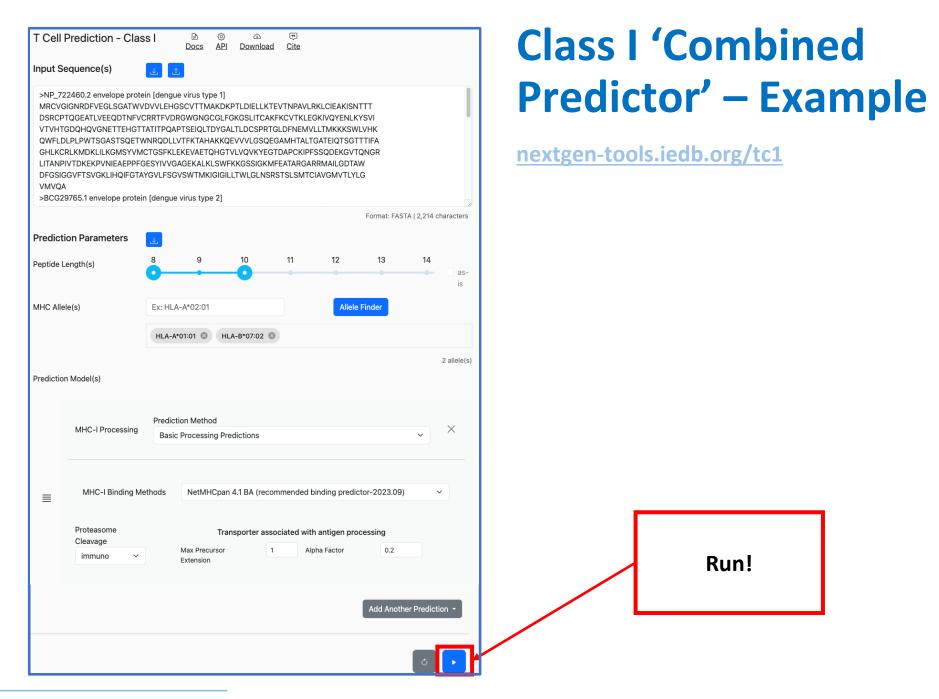
Alpha Factor
0.2

### **Proteasomal Cleavage**

- Proteasomes create the C-terminal end of peptides
- Prediction looks for sequence motif up and downstream of potential cleavage site
- Cells may switch between immuno and constitutive proteasome machinery depending upon state

### **TAP Transport**

- TAP transport efficiency of peptides is sequence dependent; motif derived based on *in vitro* assays
- Overall TAP transport efficiency of a presented MHC ligand can be result of a collection of precursors
- Unless paper thoroughly read and details about the precursor length distribution are known, keep parameters unchanged



# **Class I 'Combined Predictor' – Example**

### nextgen-tools.iedb.org/tc1

			ر⊾ Download	් End Table Display Co				
how 10 🗸 ro	ows		1 to 10 of 11,676 rows				Previous 1 2 3 4 5 116	68 Nex
seq # 🖓 💧	peptide 🖓 🕴	allele 🖓 🕴	netmhcpan_ba IC50 💡 🕴	proteasome score 🛛 🕴	tap score 💡 🍦	mhc score 💡 🍦	processing score 🖓 🕴 processing total sco	re 🖓
3	TTEAILPEY	HLA-A*01:01	30.97	1.50	1.21	-1.49	2.71	1.2
2	ITEAELTGY	HLA-A*01:01	33.71	1.36	1.23	-1.53	2.59	1.0
3	STTEAILPEY	HLA-A*01:01	121.67	1.50	1.33	-2.09	2.84	0.7
1	TSEIQLTDY	HLA-A*01:01	54.13	1.20	1.15	-1.73	2.35	0.6
2	SITEAELTGY	HLA-A*01:01	133.29	1.36	1.36	-2.12	2.72	0.5
3	TPTWNRKEL	HLA-B*07:02	16.40	1.43	0.19	-1.21	1.62	0.4
3	RCPTQGEAAL	HLA-B*07:02	33.24	1.26	0.55	-1.52	1.81	0.2
4	TPRSPSVEV	HLA-B*07:02	8.93	1.14	0.02	-0.95	1.16	0.2
2	VVQPENLEY	HLA-A*01:01	677.26	1.50	1.39	-2.83	2.89	0.0
4	LVQIENLEY	HLA-A*01:01	469.87	1.33	1.38	-2.67	2.71	0.0

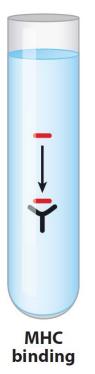
- Higher scores = higher efficiency for MHC-I presentation
- MHC binding score = -log10(IC50) (sign change)
- Combined scores are additive
  - Processing = proteasome + TAP
  - Total = proteasome + TAP + MHC

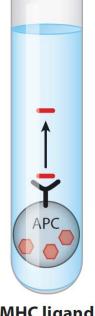
<sup>2023</sup> IEDB User Workshop

### **Caveats / Performance of Processing Predictions**

- Predictions help understand why a given peptide may not be a good MHC ligand, despite strong predicted binding
- Most high-affinity binders are also efficiently processed due to co-evolution of MHC molecules with the proteosome and TAP
- Total processing score can be used as an additional filter for peptide selection, but should not be used without also considering binding independently

### **<u>Recommended Alternative</u>: Use Predictors Directly Trained on Eluted Ligand Data**





MHC ligand elution

- Mass spectrometry of eluted ligands allows for the identification of a very large number of ligands in a single experiments
- Ligand sequences contain signals from both binding and processing
- NetMHCPan EL predictions (trained on eluted ligands) performs excellent, and can be used just like a regular MHC binding prediction

# **Incorporating Antigen Expression: Axel-F**

### tools.iedb.org/axelf

- Increased expression of an antigen in a cell increases the likelihood that peptides derived from it are processed and presented
- Axel-F tool integrates expression data into MHC ligand predictions

### iScience

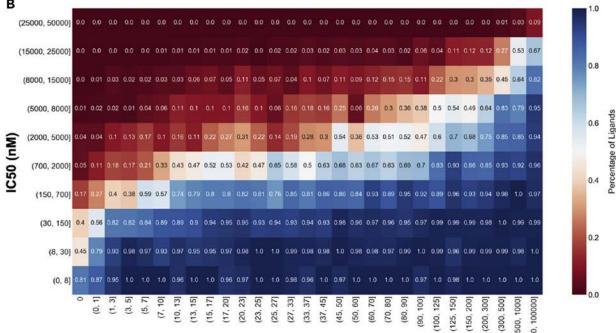


Volume 25, Issue 2, 18 February 2022, 103850

Article

Combined assessment of MHC binding and antigen abundance improves T cell epitope predictions

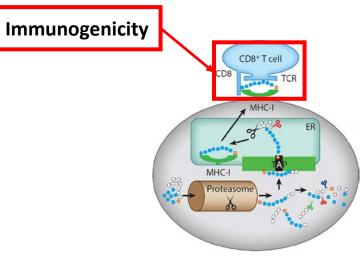
Zeynep Koşaloğlu-Yalçın <sup>1</sup>, Jenny Lee <sup>1</sup>, Jason Greenbaum <sup>1</sup>, Stephen P. Schoenberger <sup>2, 3</sup>, Aaron Miller <sup>2, 3</sup>, Young J. Kim <sup>4</sup>, Alessandro Sette <sup>1, 5</sup>, Morten Nielsen <sup>6, 7</sup>, Bjoern Peters <sup>1, 5, 8</sup>  $\Re$  📾



#### **Expression (TPM)**

# **Class I immunogenicity** prediction

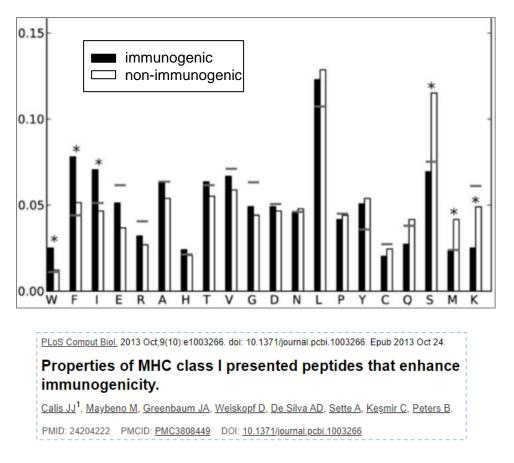
nextgen-tools.iedb.org/tc1



	New User? Learn to use the website here!	
T Cell Predict	ion - Class I ing affinity, TAP processing, and Immunogenicity predictions	
	g a sequence into this box or click 'Run' to use the example sequence:	n
>SARS2 spike		
	SSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFND	
GVYFASTEKS	IIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFL	

### **Class I Immunogenicity Prediction**

- Approach: Assemble two datasets of peptides with similar MHC binding affinity, that are (i) recognized or (ii) not recognized by T cells
- Enrichment of W,F,I and depletion of S,M,K in immunogenic peptides
- Use enrichments to calculate propensity scores



#### T Cell Prediction - Class I

② ④ API Download ••

Cite

Input Sequence(s)

>NP\_722460.2 envelope protein [dengue virus type 1]

MRCVGIGNRDFVEGLSGATWVDVVLEHGSCVTTMAKDKPTLDIELLKTEVTNPAVLRKLCIEAKISNTTT DSRCPTQGEATLVEEQDTNFVCRRTFVDRGWGNGCGLFGKGSLITCAKFKCVTKLEGKIVQYENLKYSVI VTVHTGDQHQVGNETTEHGTTATITPQAPTSEIQLTDYGALTLDCSPRTGLDFNEMVLLTMKKKSWLVHK QWFLDLPLPWTSGASTSQETWNRQDLLVTFKTAHAKKQEVVVLGSQEGAMHTALTGATEIQTSGTTTIFA GHLKCRLKMDKLILKGMSYVMCTGSFKLEKEVAETQHGTVLVQVKYEGTDAPCKIPFSSQDEKGVTQNGR LITANPIVTDKEKPVNIEAEPPFGESYIVVGAGEKALKLSWFKKGSSIGKMFEATARGARRMAILGDTAW DFGSIGGVFTSVGKLIHQIFGTAYGVLFSGVSWTMKIGIGILLTWLGLNSRSTSLSMTCIAVGMVTLYLG VMVQA

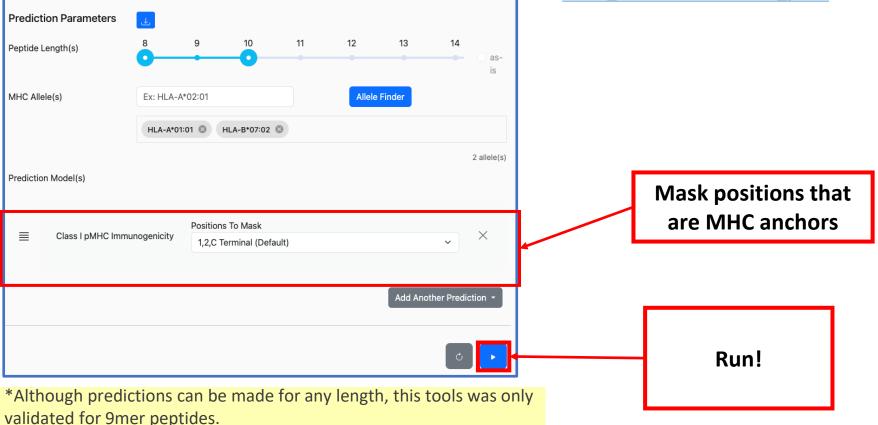
Ð

Docs

>BCG29765.1 envelope protein [dengue virus type 2]

# Class I Immunogenicity Prediction – Example

### nextgen-tools.iedb.org/tc1



Format: FASTA | 2,214 characters

# **Class I Immunogenicity Prediction – Example**

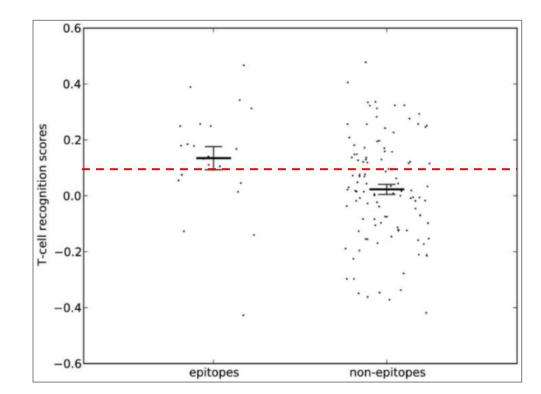
### nextgen-tools.iedb.org/tc1

		لي <u>Download</u>	ै <u>Reset Table</u> <u>D</u>		↓ able State	
now 10 🗸 re	ows	1 to 10 of 11	,676 rows	Pre	evious 1 2	3 4 5 1168 Nex
seq # 🖓 💧	peptide 🎖 🕴	start 🖓	end $\bigtriangledown$	peptide length $\nabla$	allele 🖓 🕴	immunogenicity score 💡
2	IGVVITWIGM	459	468	10	HLA-A*01:01	0.56297
2	IGVVITWIGM	459	468	10	HLA-B*07:02	0.56297
2	GVVITWIGM	460	468	9	HLA-A*01:01	0.52575
2	GVVITWIGM	460	468	9	HLA-B*07:02	0.52575
2	LIGVVITWIG	458	467	10	HLA-A*01:01	0.52358
2	LIGVVITWIG	458	467	10	HLA-B*07:02	0.52358
3	FSCIAIGIIT	475	484	10	HLA-A*01:01	0.50610
3	FSCIAIGIIT	475	484	10	HLA-B*07:02	0.50610
2	GVVITWIG	460	467	8	HLA-A*01:01	0.50595
2	GVVITWIG	460	467	8	HLA-B*07:02	0.50595

- Scores are sums of propensity scores at all unmasked positions
- High scores = peptide is more likely to be immunogenic
- If 'allele-specific' is unselected, the same score will apply over all alleles

### Class I Immunogenicity Prediction Caveats / Performance

- Experimentally, many MHC binding peptides can be immunogenic (~50%)
- Cross validation gave AUC values ~ 0.65. Test on independent blind set gave AUC = 0.69
- Recommendation: Use as filter (cutoff 0) if high specificity is desired.



#### T Cell Prediction - Class I

6 ٩ . Docs API Download Cite

Input Sequence(s)

>NP\_722460.2 envelope protein [dengue virus type 1]

MRCVGIGNRDFVEGLSGATWVDVVLEHGSCVTTMAKDKPTLDIELLKTEVTNPAVLRKLCIEAKISNTTT DSRCPTQGEATLVEEQDTNFVCRRTFVDRGWGNGCGLFGKGSLITCAKFKCVTKLEGKIVQYENLKYSVI VTVHTGDQHQVGNETTEHGTTATITPQAPTSEIQLTDYGALTLDCSPRTGLDFNEMVLLTMKKKSWLVHK QWFLDLPLPWTSGASTSQETWNRQDLLVTFKTAHAKKQEVVVLGSQEGAMHTALTGATEIQTSGTTTIFA GHLKCRLKMDKLILKGMSYVMCTGSFKLEKEVAETQHGTVLVQVKYEGTDAPCKIPFSSQDEKGVTQNGR LITANPIVTDKEKPVNIEAEPPFGESYIVVGAGEKALKLSWFKKGSSIGKMFEATARGARRMAILGDTAW DFGSIGGVFTSVGKLIHQIFGTAYGVLFSGVSWTMKIGIGILLTWLGLNSRSTSLSMTCIAVGMVTLYLG VMVQA

₿

>BCG29765.1 envelope protein [dengue virus type 2] Format: FASTA | 2,214 characters Prediction Parameters 8 11 12 13 9 10 14 Peptide Length(s) as-is MHC Allele(s) Ex: HLA-A\*02:01 Allele Finder HLA-A\*01:01 💿 HLA-A\*02:01 💿 HLA-A\*02:03 💿 HLA-A\*02:06 🔘 HLA-A\*03:01 💿 HLA-A\*11:01 💿 HLA-A\*23:01 💿 HLA-A\*24:02 💿 HLA-A\*26:01 🐼 HLA-A\*30:01 🕥 HLA-A\*30:02 🛞 HLA-A\*31:01 🙁 HLA-A\*32:01 🔘 HLA-A\*33:01 🛞 HLA-A\*68:01 🙁 HLA-A\*68:02 💿 HLA-B\*07:02 💿 HLA-B\*08:01 💿 HLA-B\*15:01 💿 HLA-B\*35:01 🛛 HLA-B\*40:01 🛞 HLA-B\*44:02 🔘 HLA-B\*44:03 🔘 HLA-B\*51:01 🔘 HLA-B\*53:01 💿 HLA-B\*57:01 🐼 HLA-B\*58:01 💿 27 allele(s) Prediction Model(s) Prediction Method X ≣ MHC-I Binding NetMHCpan 4.1 EL (recommended epitope predictor-2023.09)

Add Another Prediction -

Class I pMHC Immunogenicity MHC-I Processing

MHC Binding

### Combining **Multiple Predictors**

nextgen-tools.iedb.org/tc1

Add another prediction

### **Combining Multiple Predictors**

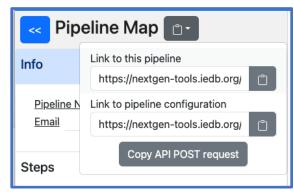
Predictio	on Model(s)						<u>nextge</u>	en-tools.iedb.o	org/tc1
≡	MHC-I Binding	Prediction Me NetMHCpar	ethod n 4.1 EL (recommended ep	itope predictor-2023.0!	9) ~	×			
	MHC-I Processing	Predictior Basic Pr	Method ocessing Predictions		~	×			
≡	MHC-I Binding M	lethods	NetMHCpan 4.1 BA (recom	nmended binding predic	ctor-2023.09)	~			
	Proteasome Cleavage immuno ~		Transporter assoc ax Precursor 1 tension	ciated with antigen pro	0.2				
≣	Class I pMHC Imm	unogenicity	Positions To Mask 1,2,C Terminal (Default	)	~	×			
					Add Another Predi	iction -			
					Ċ			Run!	

2023 IEDB User Workshop

### **Apply Filters and Save Your Work!**

Show 10 v ro	ows	1 to 10 o	f 157,626 rows		3 4 5 15763 Next		
peptide 🖓 🕴	allele 🖓 🕴	median binding percentile		percentile 🖓 🍦	netmhcpan_ba percentile 🖓 🕴	immunogenicity score 💡 🕴	processing total score $\bigtriangledown$
GESALTLHW	HLA-B*44:03		Sort A→Z	0.01	0.01	0.02095	0.69
GESALTLHW	HLA-B*44:02		Sort Z→A	0.01	0.01	0.02095	0.71
KSWLVHKQW	HLA-B*57:01	•	Min:	0.01	0.02	-0.11829	1.33
KTWLVHKQW	HLA-B*57:01		0.01	0.01	0.02	-0.11829	1.26
KAWMVHRQW	HLA-B*57:01		Vax:	0.01	0.02	-0.05815	1.17
KAWLVHRQW	HLA-B*57:01		Clear OK	0.01	0.02	0.10739	1.04
AETQHGTVL	HLA-B*40:01			0.01	0.02	0.01632	1.02

now 10 v ro	ows	1 to 10 of 276 r	ows filtered from 157,626 total rows	3)	Previous 1 2	3 4 5 28 Next
peptide 🎖 🕴	allele 🖓 💧	median binding percentile 🍞 🖣	netmhcpan_el percentile 🖓		unogenicity score 🖓 ≬	processing total score 🖓 (
GESALTLHW	HLA-B*44:03	0.01	0	So los	0.02095	0.6
GESALTLHW	HLA-B*44:02	0.01	0		0.02095	0.7
KAWLVHRQW	HLA-B*57:01	0.015	0		0.10739	1.0
AETQHGTVL	HLA-B*40:01	0.015	0	ats have 276 different	0.01632	1.0
IQKETLVTF	HLA-B*15:01	0.02	0	acial expressions, study	0.11563	1.6
KEVALLRTY	HLA-B*44:03	0.02	0	nds Source: CNN	0.09789	1.5
SETQHGTIL	HLA-B*40:01	0.02	0.01	Source. Chin	0.06996	0.5



- Pipeline link will include all selected parameters plus all data
- Pipeline 'configuration' link will only include selected parameters

### Review

- Next-generation IEDB Tools website (nextgentools.iedb.org) hosts the complete T cell, class I suite of tools
  - MHC binding & elution
  - Antigen processing
  - Immunogenicity
- Recommended methods will continue to change as algorithms are evaluated
- No uniform threshold exists for selecting peptide candidates
- Tools may be executed simultaneously, and results can be filtered and shared