IEDB
Immune Epitope Database & Tools

Analysis Resource Overview

tools.iedb.org
nextgen-tools.iedb.org

Presented by: Dr. Bjoern Peters, Professor

IEDB Tools

Analysis Resource & Next-Generation Tools

Epitope Prediction and Analysis Tools

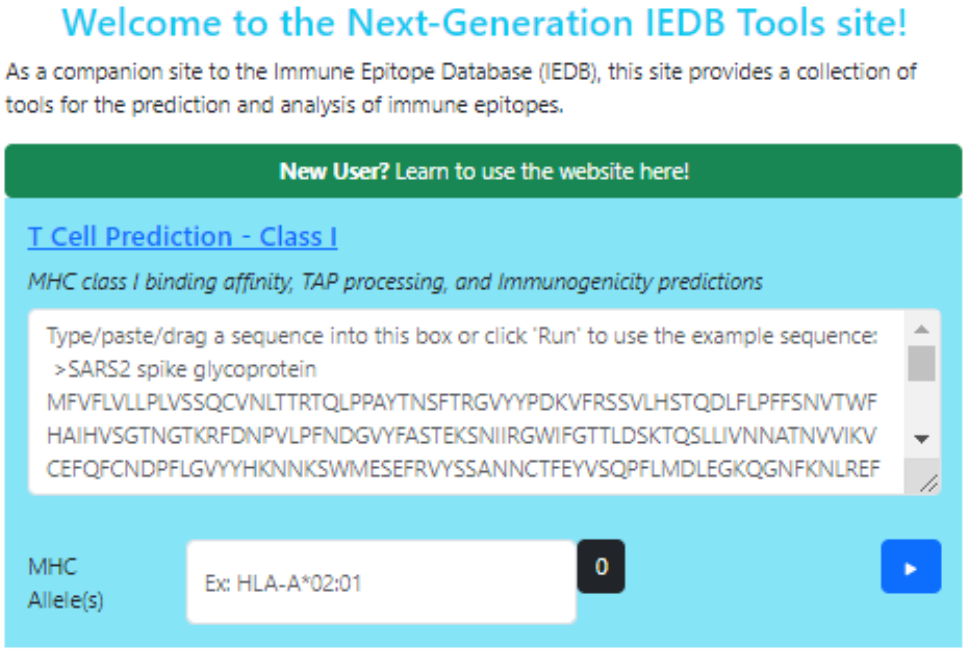
Welcome to the Immune Epitope Database Analysis Resource. This site provides a collection of tools for the prediction and analysis of immune epitopes. It serves as a companion site to the Immune Epitope Database (IEDB) , a manually curated database of experimentally characterized immune epitopes.

The tools contained fall into the following categories:

T Cell Epitope Prediction Tools
This set of tools includes MHC class I & II binding predictions, as well as peptide processing predictions and immunogenicity predictions.

B Cell Epitope Prediction Tools
The tools here are intended to predict regions of proteins that are likely to be recognized as epitopes in the context of a B cell response.

Analysis Tools
The epitope analysis tools are intended for the detailed analysis of a known epitope sequence or group of sequences.



Analysis Resource:

- T cell epitope prediction
- B cell epitope prediction
- Analysis tools

tools.iedb.org

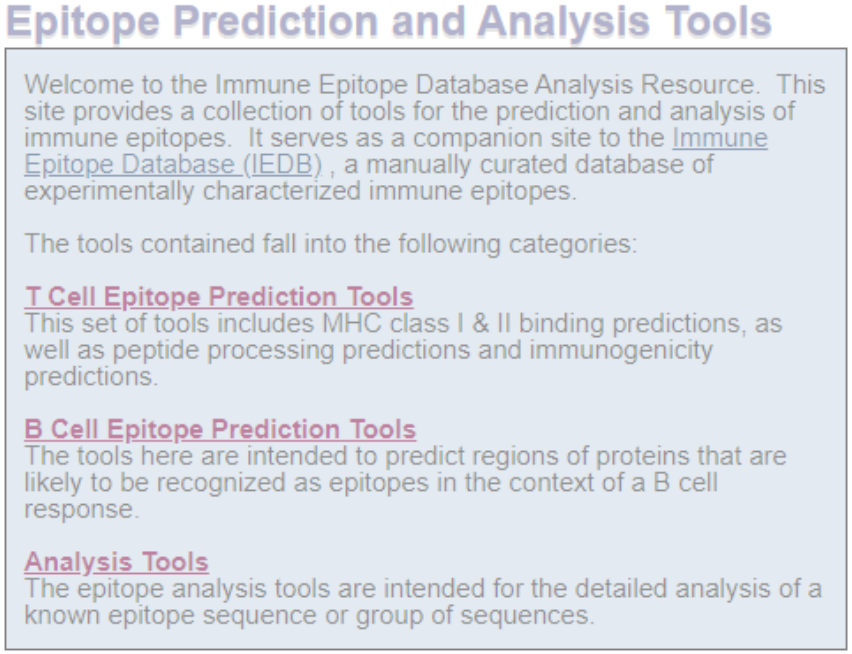
Next-Generation Tools:

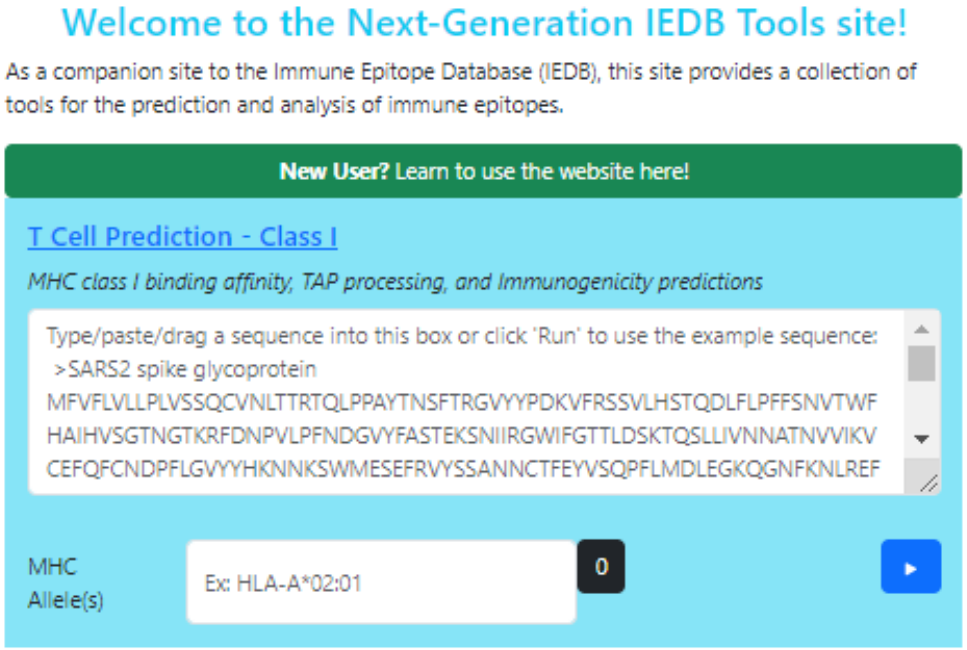
- Re-implementing existing tools with a focus on improving usability and function
- Launched in 2023
- More tools to be added

nextgen-tools.iedb.org

IEDB Tools

Analysis Resource & Next-Generation Tools





Analysis Resource:

- T cell epitope prediction
- B cell epitope prediction
- Analysis tools

tools.iedb.org

Next-Generation Tools:

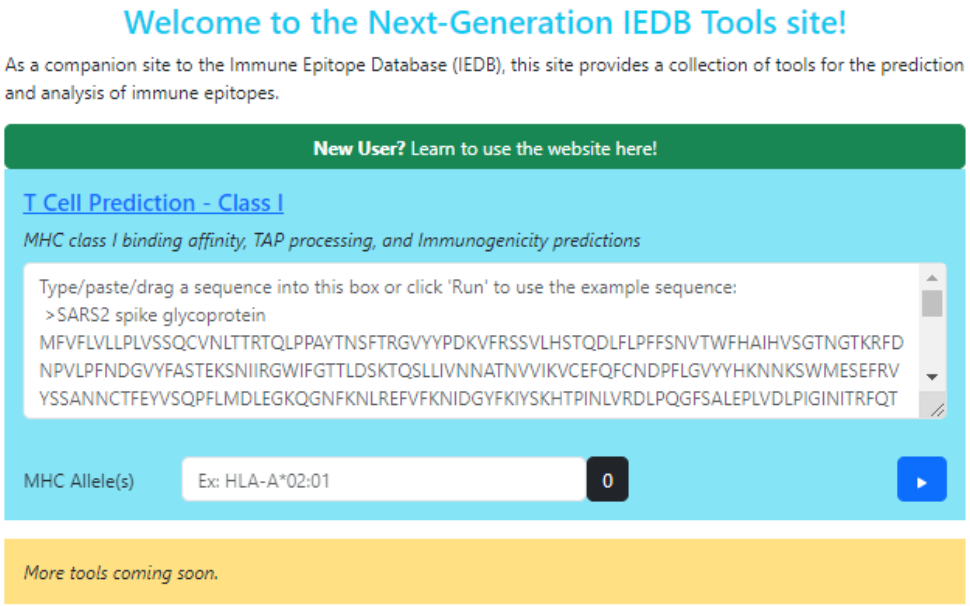
- Re-implementing existing tools with a focus on improving usability and function
- Launched in 2023
- More tools to be added

nextgen-tools.iedb.org

Next-Generation Tools

https://nextgen-tools.iedb.org/

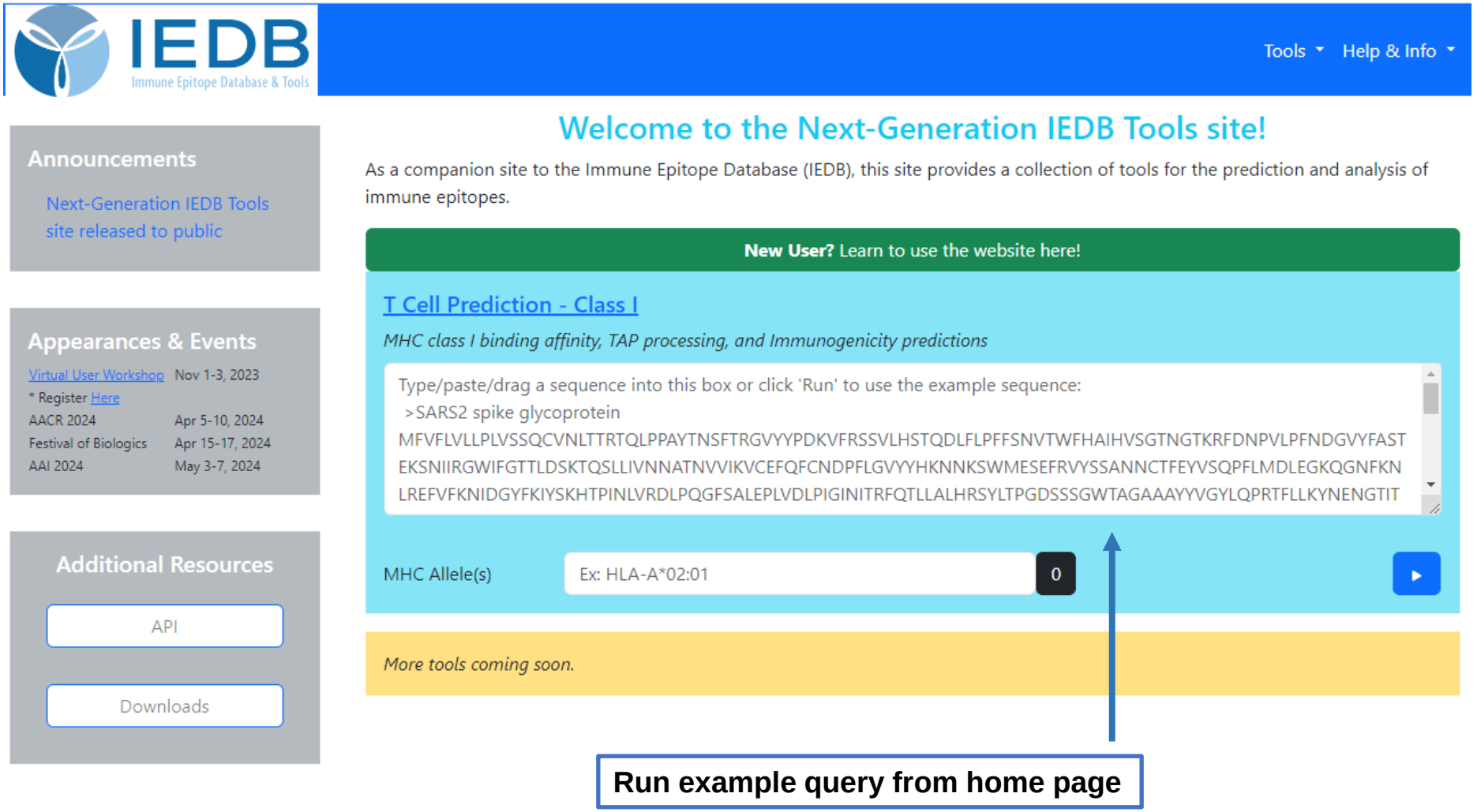
- Launched in 2023
- Multi-year and multi-release project
- The main goals of the redesign are to:
  - simplify the user experience
  - enhance existing functionality
  - add new functionality
  - ensure consistency among the tools
  - improve the layout and aesthetics
- First release focused on:
  - T cell epitope prediction class I
  - Cluster analysis
  - PEPMatch



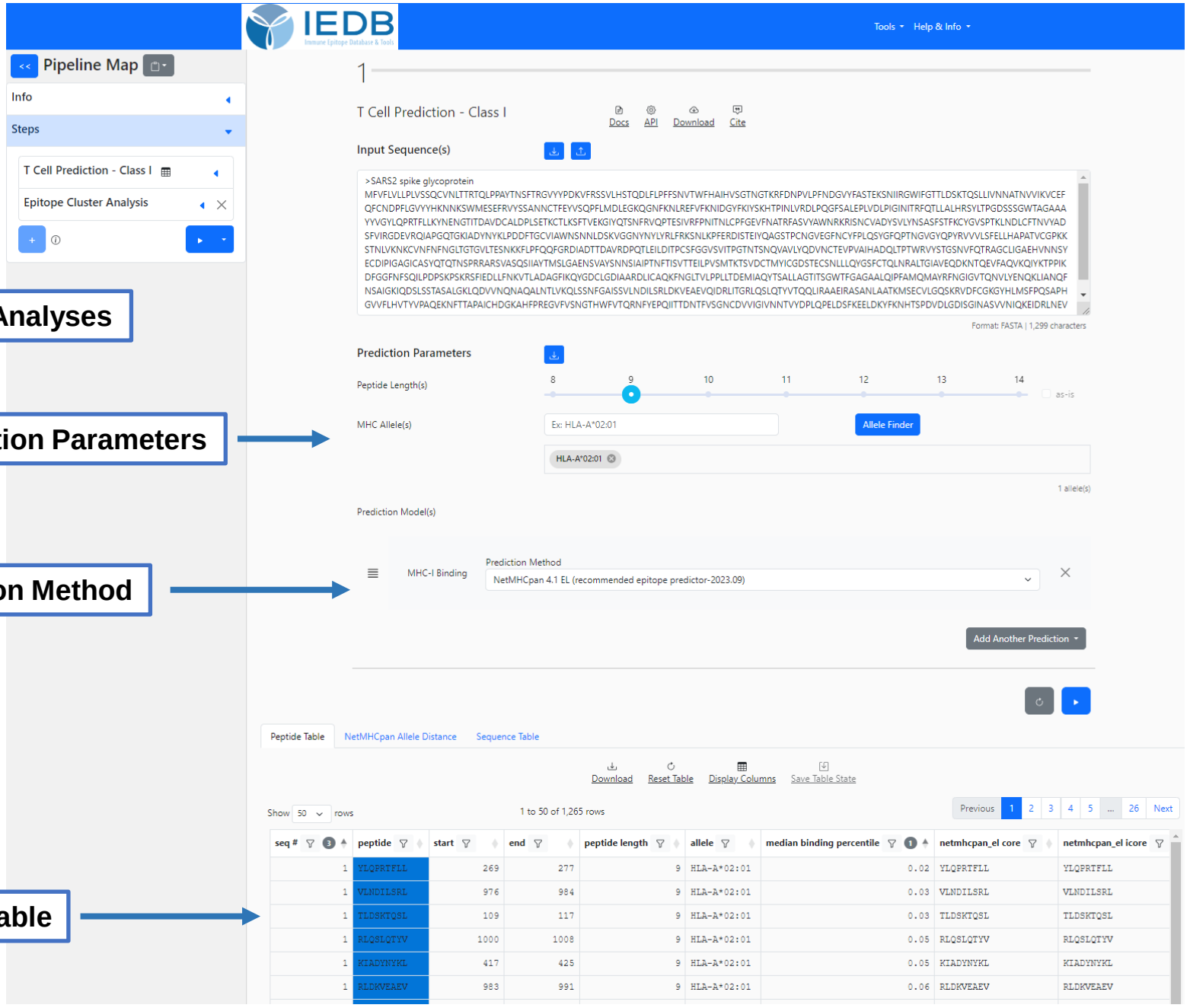
Learn more on Day 2!

Next-Generation Tools

https://nextgen-tools.iedb.org/



Results



seq #	peptide	start	end	peptide length	allele	median binding percentile	netmhcpan_el core	netmhcpan_el icore
1	YLQPRTFLL	269	277	9	HLA-A*02:01	0.02	YLQPRTFLL	YLQPRTFLL
1	VLNDILSRL	976	984	9	HLA-A*02:01	0.03	VLNDILSRL	VLNDILSRL
1	TLDSKTQSL	109	117	9	HLA-A*02:01	0.03	TLDSKTQSL	TLDSKTQSL
1	RLQSLQTYV	1000	1008	9	HLA-A*02:01	0.05	RLQSLQTYV	RLQSLQTYV
1	KIADYNYKL	417	425	9	HLA-A*02:01	0.05	KIADYNYKL	KIADYNYKL
1	RLDKVEAEV	983	991	9	HLA-A*02:01	0.06	RLDKVEAEV	RLDKVEAEV

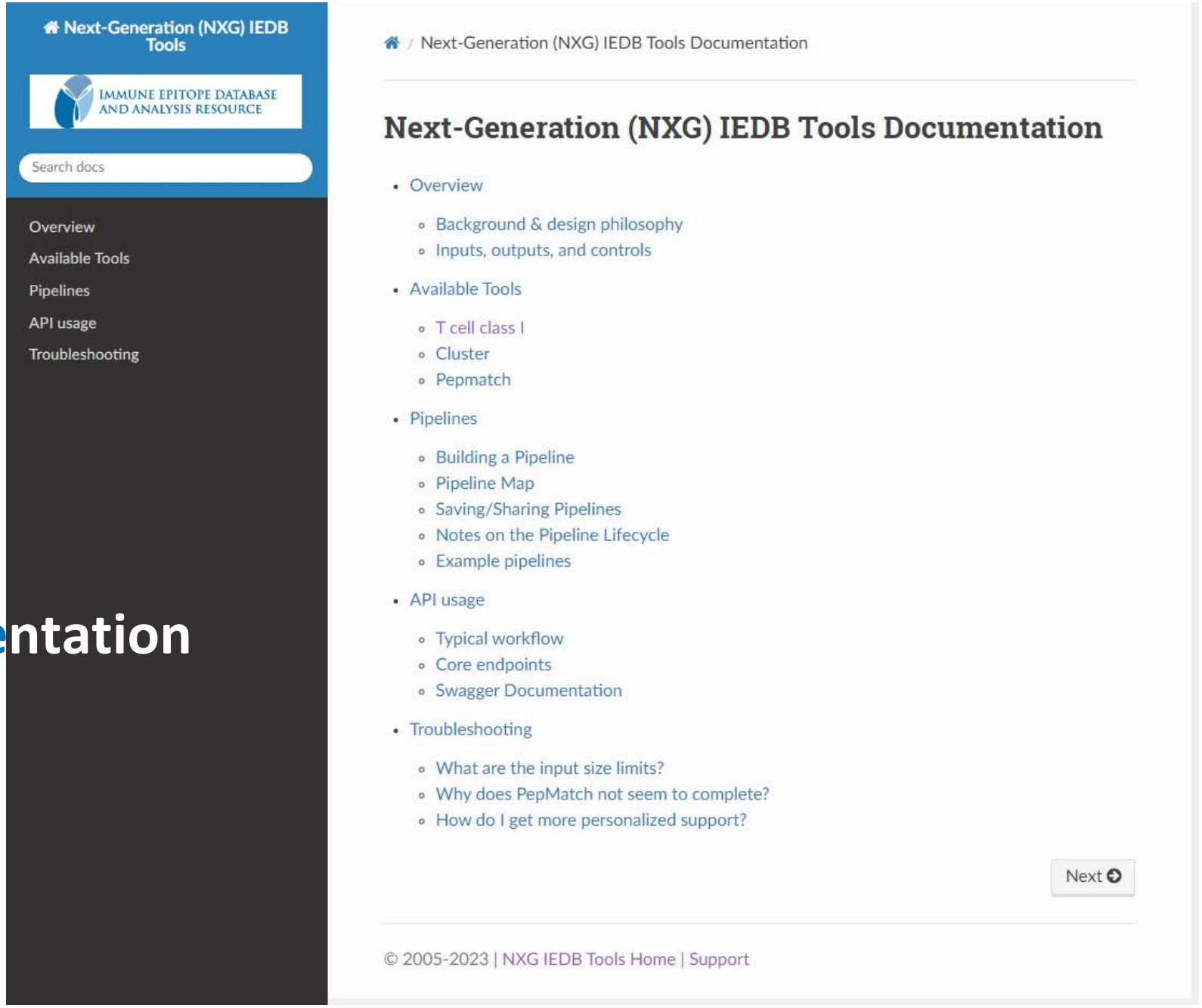
Build Pipeline Analyses

Change Prediction Parameters

Select Prediction Method

Results Table

Help Documentation



IEDB Tools

Analysis Resource & Next-Generation Tools

Epitope Prediction and Analysis Tools

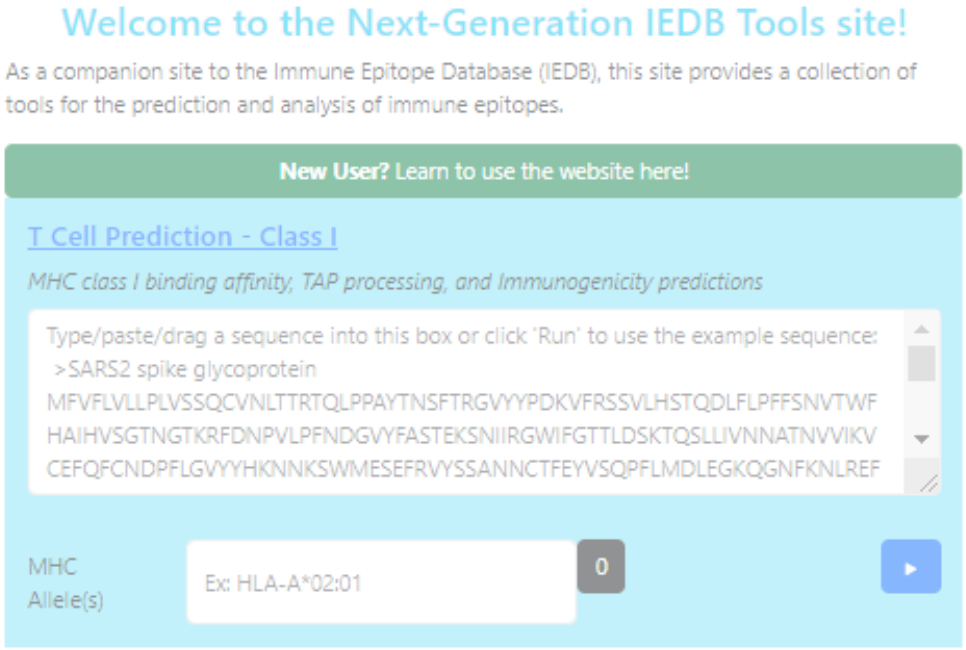
Welcome to the Immune Epitope Database Analysis Resource. This site provides a collection of tools for the prediction and analysis of immune epitopes. It serves as a companion site to the Immune Epitope Database (IEDB) , a manually curated database of experimentally characterized immune epitopes.

The tools contained fall into the following categories:

T Cell Epitope Prediction Tools
This set of tools includes MHC class I & II binding predictions, as well as peptide processing predictions and immunogenicity predictions.

B Cell Epitope Prediction Tools
The tools here are intended to predict regions of proteins that are likely to be recognized as epitopes in the context of a B cell response.

Analysis Tools
The epitope analysis tools are intended for the detailed analysis of a known epitope sequence or group of sequences.



Analysis Resource:

- T cell epitope prediction
- B cell epitope prediction
- Analysis tools

tools.iedb.org

Next-Generation Tools:

- Re-implementing existing tools with a focus on improving usability and function
- Launched in 2023
- More tools to be added

nextgen-tools.iedb.org

IEDB Analysis Resource

tools.iedb.org

Day 2 will go into much more detail!

T Cell Tools

MHC binding prediction
- MHC I binding prediction
- MHC II binding prediction
- TepiTool

MHC ligand prediction
- NetChop/NetCTL/NetCTLpan
- MHC-NP, MHCII-NP
- AXEL-F

Immunogenicity prediction
- CD4 & CD8 T cell immunogenicity prediction tools

Other
- Deimmunization tool
- TCRMatch
- Docktope

B Cell Tools

Linear epitope prediction
- BepiPred
- Other methods

Discontinuous epitope prediction
- DiscoTope
- ElliPro

Antibody and TCR structure prediction
- LYRA
- SCEptRe

Analysis Tools

Population coverage of epitope set
- Population coverage tool

Degree of conservation
- Conservancy analysis tool

Group peptides based on sequence identity
- Epitope cluster analysis

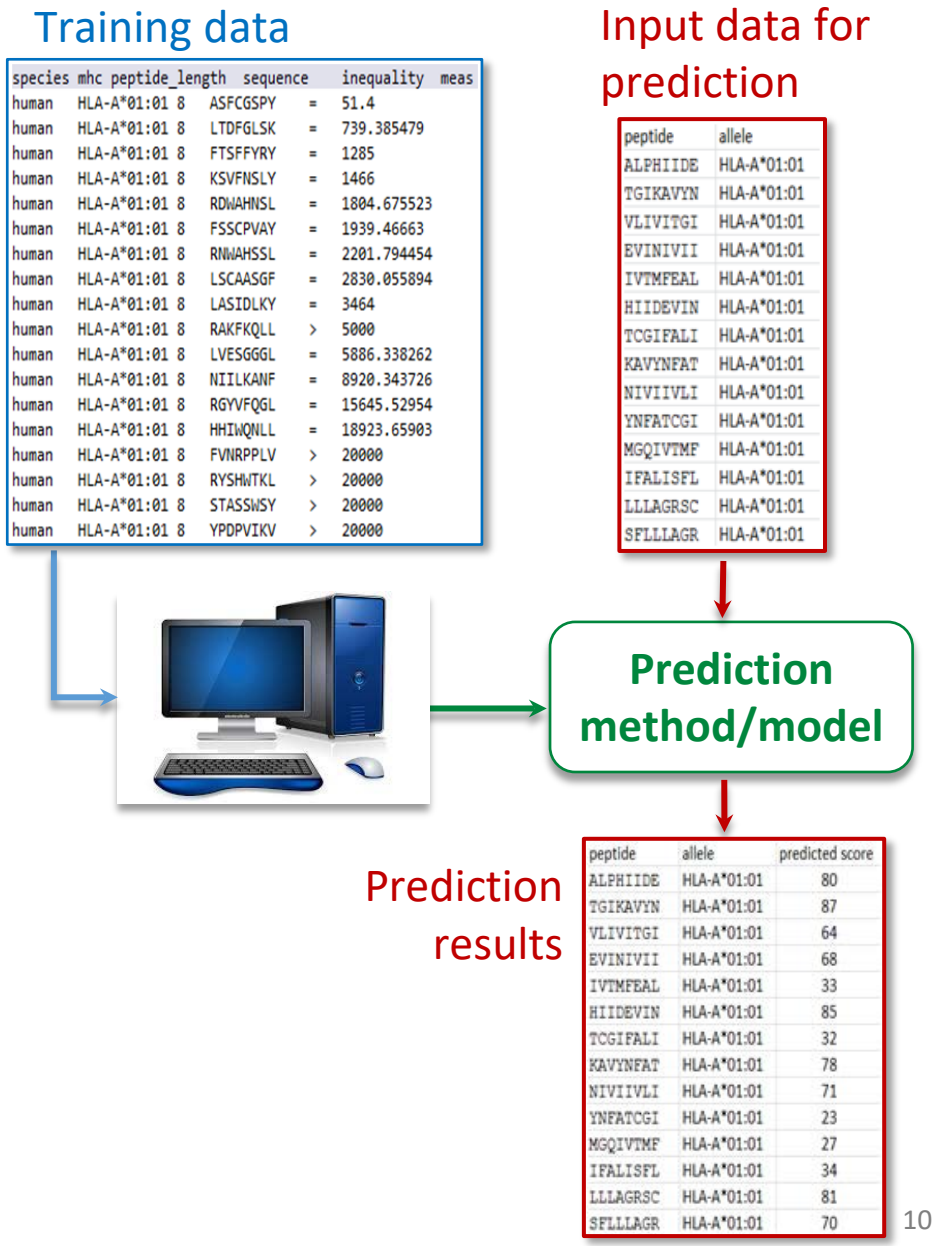
Infer restriction in HLA typed subjects
- RATE

Aggregate heterogeneous immune response
- ImmunomeBrowser

Other
- PepSySco
- PepX

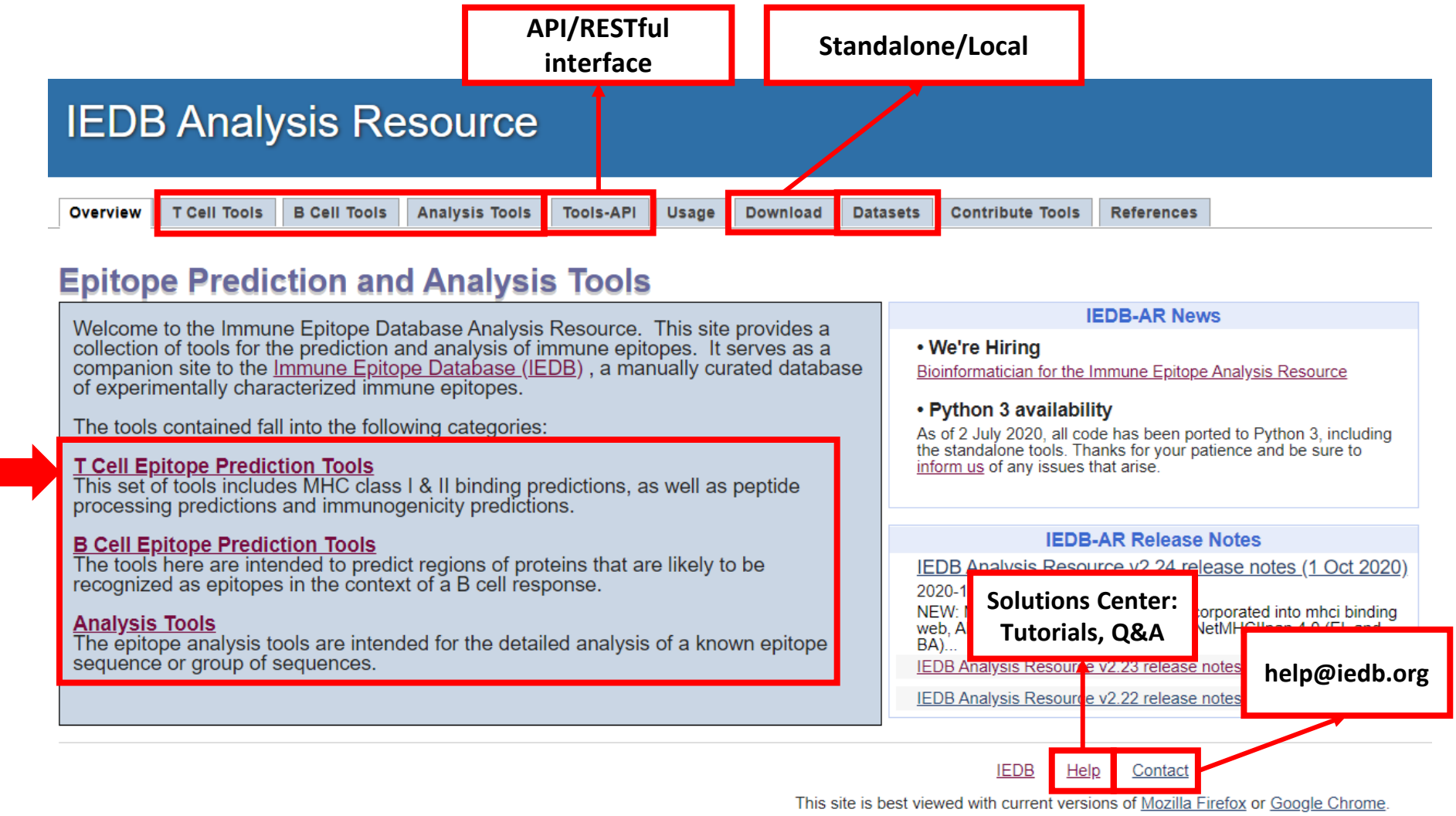
Development of Prediction Tools

- Based on machine learning techniques
  - Experimentally derived data as training input (E.g. MHC binding data, 3D crystal structures, T cell assay data) → collected in the IEDB
  - Various prediction models (e.g. Neural Networks, linear weighted models, etc.)
  - Extrapolate identified patterns to new examples



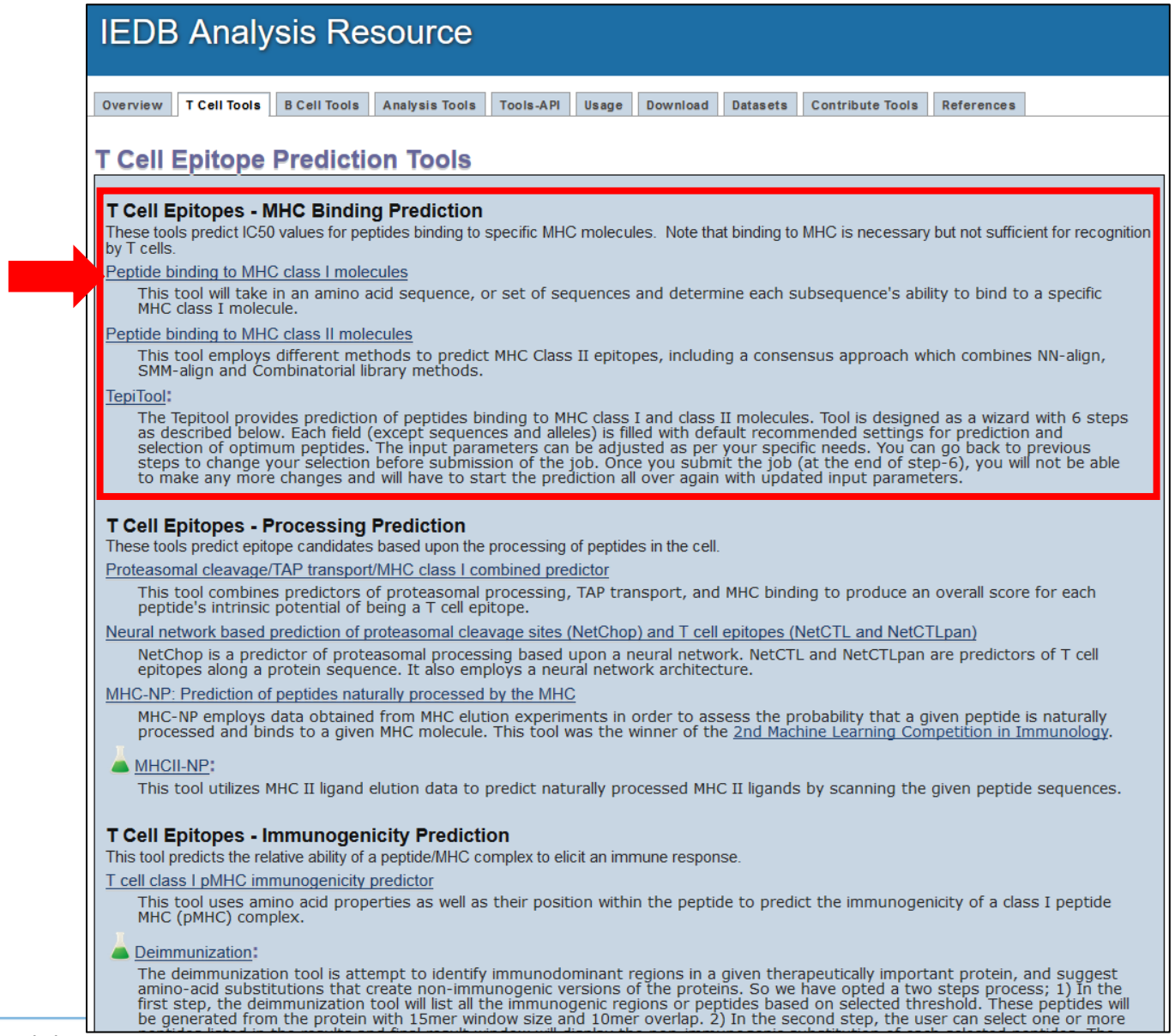
Accessing the Analysis Resource

Available Resources



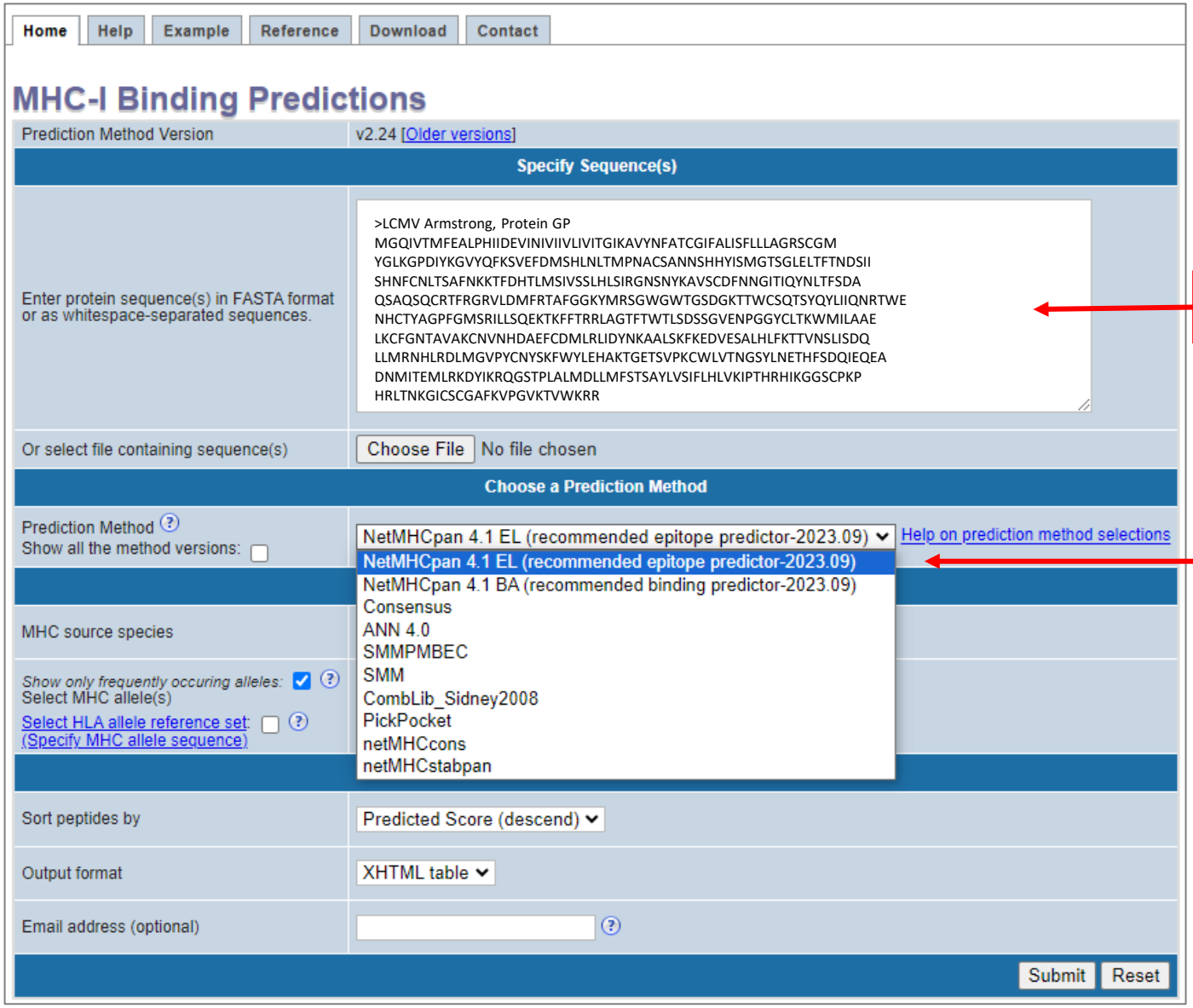
T Cell Tools

tools.iedb.org/main/tcell/



MHC I Binding Prediction

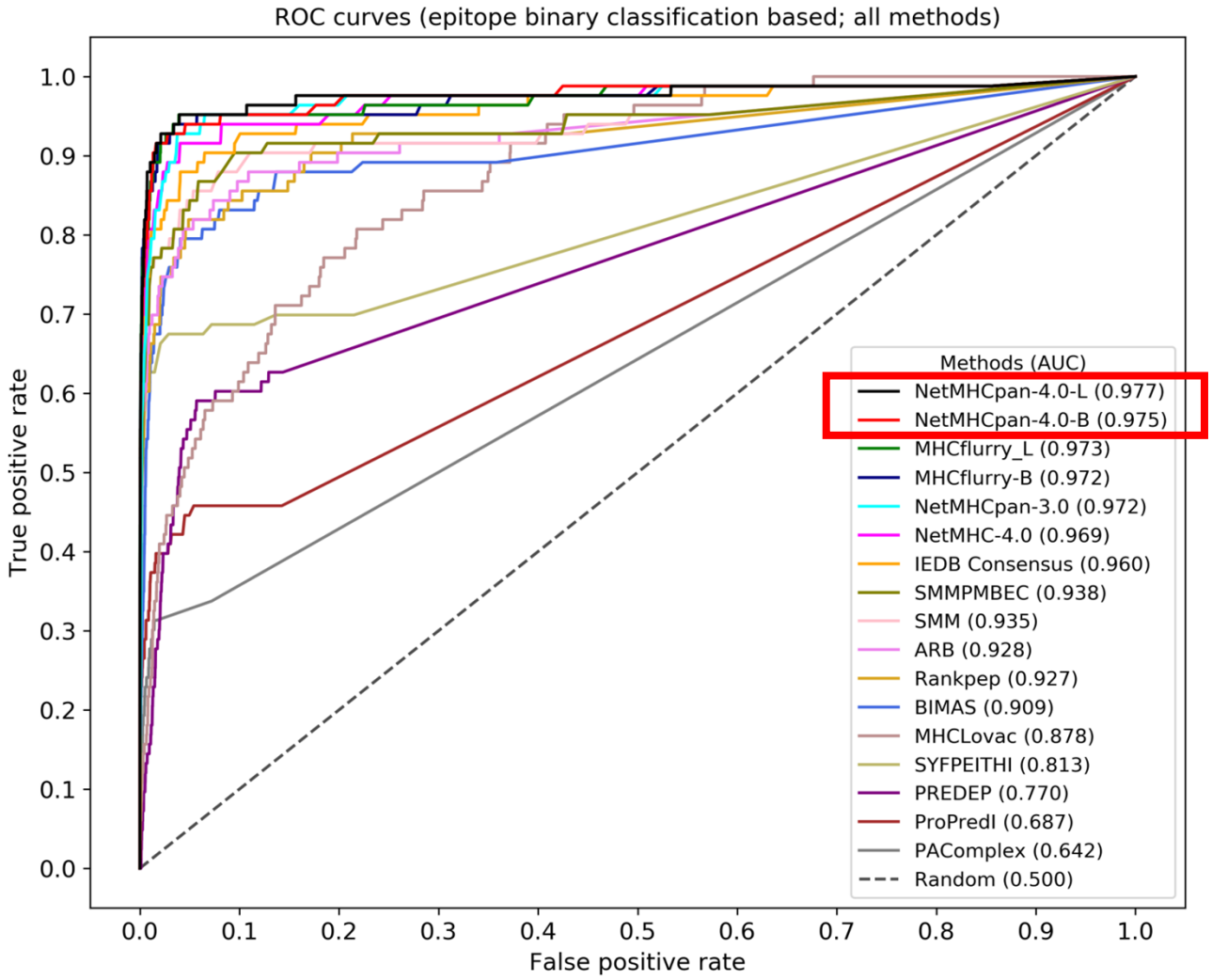
tools.iedb.org/mhci/



Sequence

Prediction method

MHC I Binding Prediction Methods - Benchmarking



- Comprehensive evaluation of different MHC binding prediction methods to identify T cell epitopes

Paul et al, PLoS Comp Biol, 2020

MHC I Binding Prediction Methods - Benchmarking

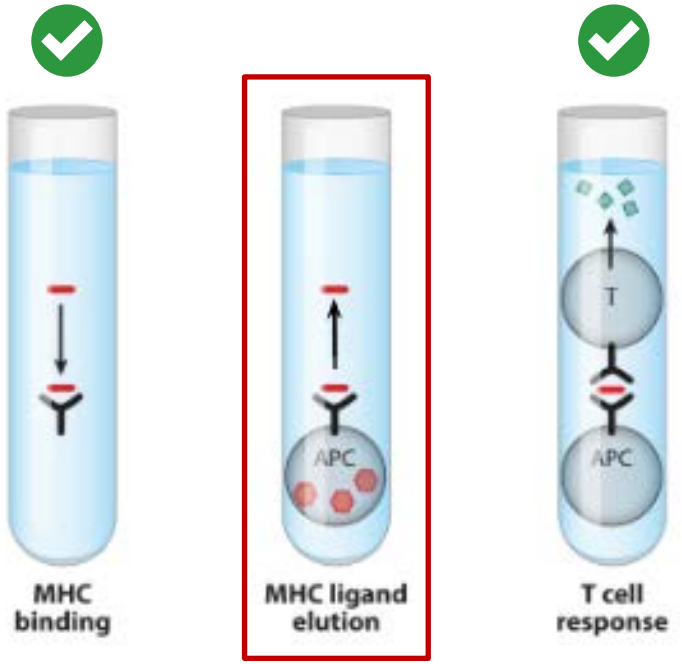
- Automated evaluation of MHC binding predictions on newly released datasets in the IEDB
- http://tools.iedb.org/auto_bench/mhci/weekly/
- Prediction of absolute and ranked binding affinities
- Consistent scoring of the NetMHCPan 4.0+ BA methods being on top for binding predictions

Server	2023-10-19	2023-08-10	2023-08-03	2023-06-15	2023-05-11	2023-04-20	2022-11-24	2022-10-28	2022-10-06	2022-07-22	2022-05-20	2022-04-29	2021-12-24	2021-12-10	2021-11-05	2021-10-08	2021-10-01	2021-09-03	2021-08-20	2021-08-06
NetMHCpan 4.1 BA	73	74	74	65	66	67	66	72	83	76	80	75	72	64	64	63	63	70	71	76
NetMHCcons	69	70	65	64	60	57	60	56	49	61	60	69	68	57	56	49	50	52	61	71
NetMHCpan 4.0 BA	68	66	61	48	60	60	60	68	76	76	80	80	80	71	70	62	61	66	58	69
NetMHCpan 3.0	65	65	64	58	68	65	70	73	72	74	74	73	72	68	64	58	57	56	49	64
ANN 3.4	64	66	62	73	64	63	61	50	44	48	45	47	45	46	50	53	53	51	63	69
mhcflurry 1.2.0	63	56	51	56	58	61	56	59	42	33	49	54	60	61	59	68	70	74	72	50
NetMHCpan 2.8	60	65	59	41	60	57	63	69	46	54	48	65	65	63	60	48	49	50	51	64
ANN 4.0	56	55	60	53	64	65	67	73	76	64	57	56	51	54	52	50	53	46	54	67
SMMPMBEC	50	56	73	75	57	60	61	61	73	62	50	43	46	53	53	64	61	67	67	54
IEDB Consensus	48	52	58	65	71	72	69	71	69	58	60	52	54	58	57	62	59	60	63	62
SMM	46	51	54	55	56	57	56	61	58	52	54	39	43	50	49	63	62	65	68	60
NetMHCpan 4.0 EL	45	42	37	51	46	43	46	37	25	36	34	44	42	50	47	42	44	28	40	60
NetMHCpan 4.1 EL	34	32	32	56	51	48	42	36	25	26	39	50	49	57	52	44	46	39	26	38
PickPocket	34	40	46	54	57	57	68	68	55	41	40	34	38	52	54	56	55	59	49	60
ARB	31	34	48	52	48	49	53	41	37	42	34	27	24	26	27	34	37	44	28	29

Trolle et al, Bioinformatics, 2015
Trevizani et al, Briefings in Bioinformatics, 2022

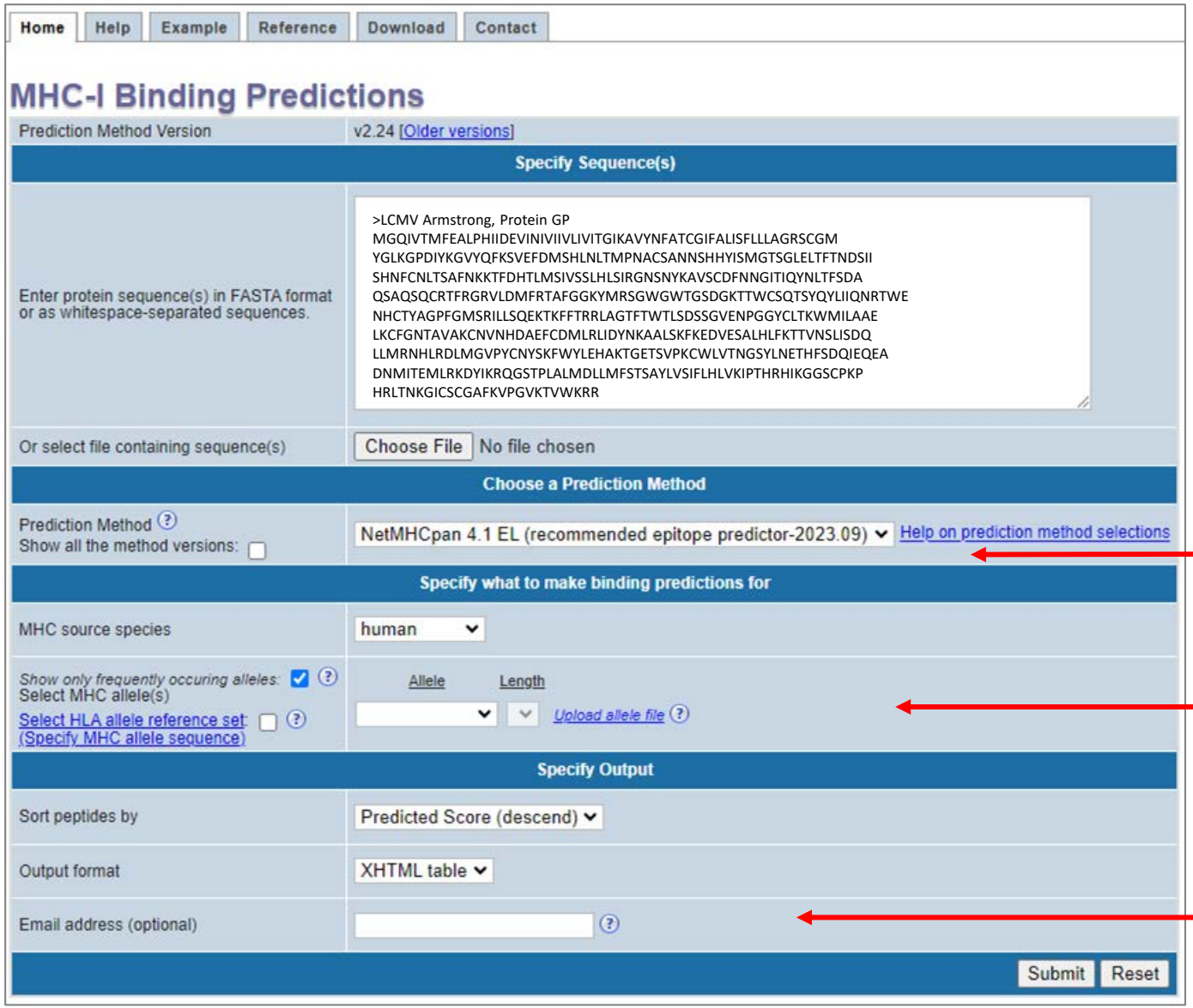
New Tools Benchmarks *with Ligand Elution Data*

- MHC ligand experiments generate massive amounts of data
- MHC ligands incorporate features of T cell epitopes beyond MHC binding
- Generation of negative datasets is not trivial
- Best results so far using peptides from UniProt (significantly differs from random)



MHC I Binding Prediction

tools.iedb.org/mhci/



Prediction method

Allele & length

Email ID

MHC I Binding Prediction - Result

Home | Help | Example | Reference | Download | Contact

tools.iedb.org/mhci/

MHC-I Binding Prediction Results

Input Sequences

#	Name	Sequence
1	LCMV Armstrong, Protein GP	MGQIVTMFEALPHIIDEVINIVIIVLIVITGIKAVYNFATCGIFALISFLLLAGRSCGMYGLKGPDIYKG VYQFKSVEFDMSHLNLTMPNACSANNSHHYISMGTSGLELTFTNDSIISHNFCNLTSAFNKKTF DHTLMSIVSSLHLSIRGNSNYKAVSCDFNNGITIQYNLTFSDAQSAQSQCRTFRGRVLDMFRTA FGGKYMRSGWGWTGSDGKTTWCSQTSYQYLIIQNRTWENHCTYAGPFGMSRILLSQEKTKF FTRRLAGTFTWTLSDSSGVENPGGYCLTKWMILAAELKCFGNTAVAKCNVNHDAEFCDMLRLI DYNKAALSKFKEDVESALHLFKTTVNSLISDQLLMRNHLRDLMGVPYCNYSKFWYLEHAKTGE TSVPKCWLVTNGSYLNETHFSDQIEQEADNMITEMLRKDYIKRQGSTPLALMDLLMFSTSAYLV SIFLHLVKIPTHRHIKGGSCPKPHRLTNKGICSCGAFKVPGVKTVWKRR

NetMHCpan allele distance

Input Allele	Closest Allele	Distance
HLA-A02:01	HLA-A02:01	0.000

Prediction method: NetMHCpan EL 4.1 High Score = good binder

* The 'IEDB recommended' method was updated in September 2020 to NetMHCpan EL 4.1 More information is available on the help page.

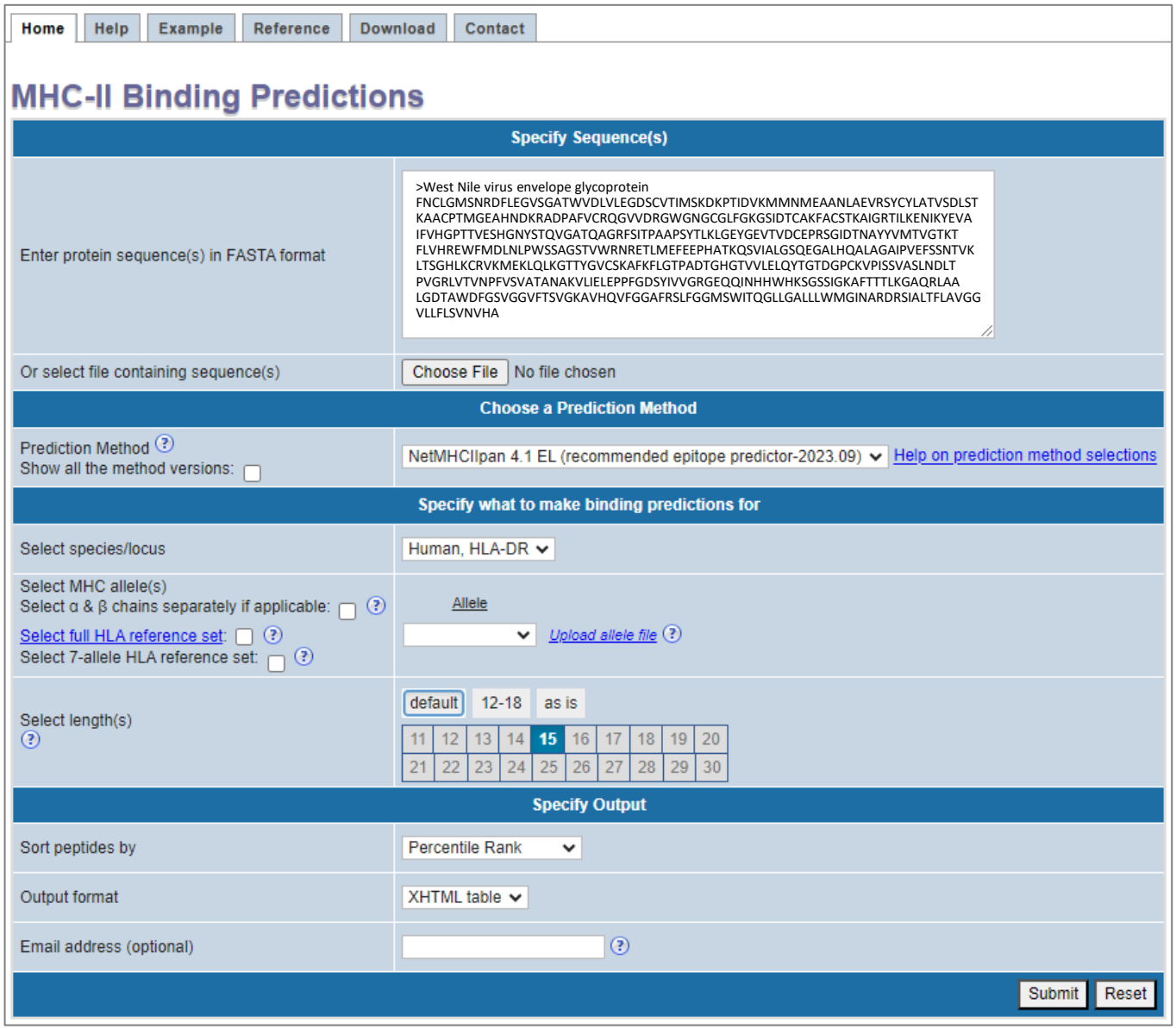
Download result

Citations

Allele	#	Start	End	Length	Peptide	Core	Icore	Score	Percentile Rank
HLA-A*02:01	1	6	14	9	TMFEALPHI	TMFEALPHI	TMFEALPHI	0.942547	0.03
HLA-A*02:01	1	10	18	9	ALPHIIDEV	ALPHIIDEV	ALPHIIDEV	0.920331	0.03
HLA-A*02:01	1	137	145	9	TLMSIVSSL	TLMSIVSSL	TLMSIVSSL	0.882391	0.04
HLA-A*02:01	1	447	455	9	YLVSIFLHL	YLVSIFLHL	YLVSIFLHL	0.855633	0.06
HLA-A*02:01	1	14	22	9	IIDEVINIV	IIDEVINIV	IIDEVINIV	0.807736	0.08
HLA-A*02:01	1	13	21	9	HIIDEVINI	HIIDEVINI	HIIDEVINI	0.779939	0.09
HLA-A*02:01	1	339	347	9	ALHLFKTTV	ALHLFKTTV	ALHLFKTTV	0.565798	0.22
HLA-A*02:01	1	450	458	9	SIFLHLVKI	SIFLHLVKI	SIFLHLVKI	0.530163	0.24
HLA-A*02:01	1	45	53	9	ALISFLLLA	ALISFLLLA	ALISFLLLA	0.525738	0.25
HLA-A*02:01	1	349	357	9	SLISDQLLM	SLISDQLLM	SLISDQLLM	0.510305	0.26
HLA-A*02:01	1	440	448	9	LMFSTSAYL	LMFSTSAYL	LMFSTSAYL	0.483579	0.27
HLA-A*02:01	1	320	328	9	RLIDYNKAA	RLIDYNKAA	RLIDYNKAA	0.438929	0.32
HLA-A*02:01	1	435	443	9	ALMDLLMFS	ALMDLLMFS	ALMDLLMFS	0.386032	0.39
HLA-A*02:01	1	27	35	9	IVITGIKAV	IVITGIKAV	IVITGIKAV	0.36266	0.42
HLA-A*02:01	1	42	50	9	GIFALISFL	GIFALISFL	GIFALISFL	0.285315	0.57

MHC II Binding Prediction

tools.iedb.org/mhcii/



MHC II Binding Prediction Methods - Benchmarking

http://tools.iedb.org/auto_bench/mhcii/weekly/

MHC II Automated Server Benchmarks

This is a live ranking of MHC II 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.

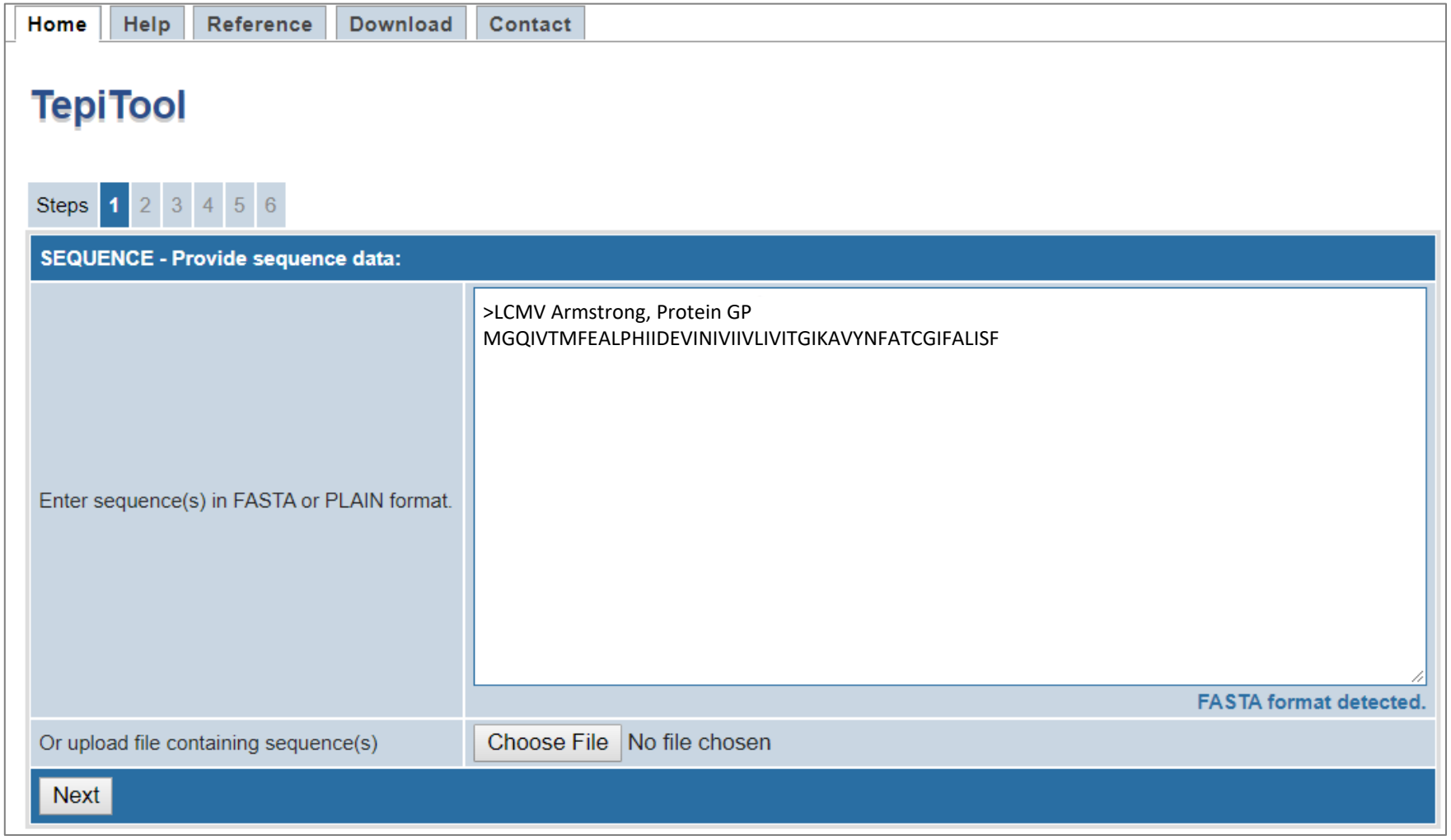
Server	2023-09-01	2023-07-28	2023-04-21	2023-03-03	2023-01-27	2022-10-28	2022-09-16	2022-09-02	2022-06-17	2022-06-10	2022-04-08	2022-04-01	2021-12-10	2021-11-05	2021-10-29	2021-07-23	2021-04-23	2021-04-16	2021-02-19	2021-01-22	2020-12-18	2020-10-30	2020-10-23	2020-08-07	2020-06-26	2020-05-01	2020-03-27	2020-03-16	2020-01-03	2019-07-02	2019-05-24	2019-03-22	2018-11-23
NetMHCIIpan-4.1 BA	85	82	70	71	70	70	76	65	65	65	65	65	67	66	65	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
NetMHCIIpan-4.0 BA	76	78	68	70	68	68	70	65	66	66	63	63	63	62	61	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
NetMHCIIpan-3.2	67	67	64	65	64	63	52	51	51	50	61	63	69	70	77	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
NN-align 2.3	56	59	51	49	49	48	45	63	64	67	67	67	65	67	67	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
NetMHCIIpan-3.1	53	54	66	56	54	55	41	47	47	41	46	46	48	50	54	58	57	58	55	58	46	49	60	61	60	64	55	64	65	65	71	84	79
Consensus IEDB method	50	58	58	59	59	61	57	66	66	68	65	66	63	62	63	66	66	71	73	69	73	76	62	60	60	58	58	61	64	66	64	64	64
NetMHCIIpan-4.1 EL	45	39	51	54	57	57	54	32	32	33	36	36	38	36	38	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SMM-align	43	42	50	50	47	48	42	55	56	52	50	47	50	51	53	49	51	43	43	43	57	61	57	57	54	56	38	41	43	39	43	49	35
NN-align	38	42	44	43	43	42	43	58	58	57	57	59	55	56	57	68	64	73	73	75	59	60	73	74	78	79	88	75	71	71	66	54	63
NetMHCIIpan-4.0 EL	27	24	47	47	50	51	48	31	31	29	34	33	37	34	38	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Tepitope (Sturniolo)	26	30	39	39	41	41	37	31	27	41	36	37	41	39	29	21	24	17	18	14	21	29	27	25	27	26	33	28	26	26	29	21	47
Comblib matrices	20	14	7	4	4	4	13	17	17	15	18	16	16	16	17	10	11	8	8	7	30	38	25	25	21	10	0	0	5	5	5	10	4

- Similar to the MHC Class I benchmark set-up

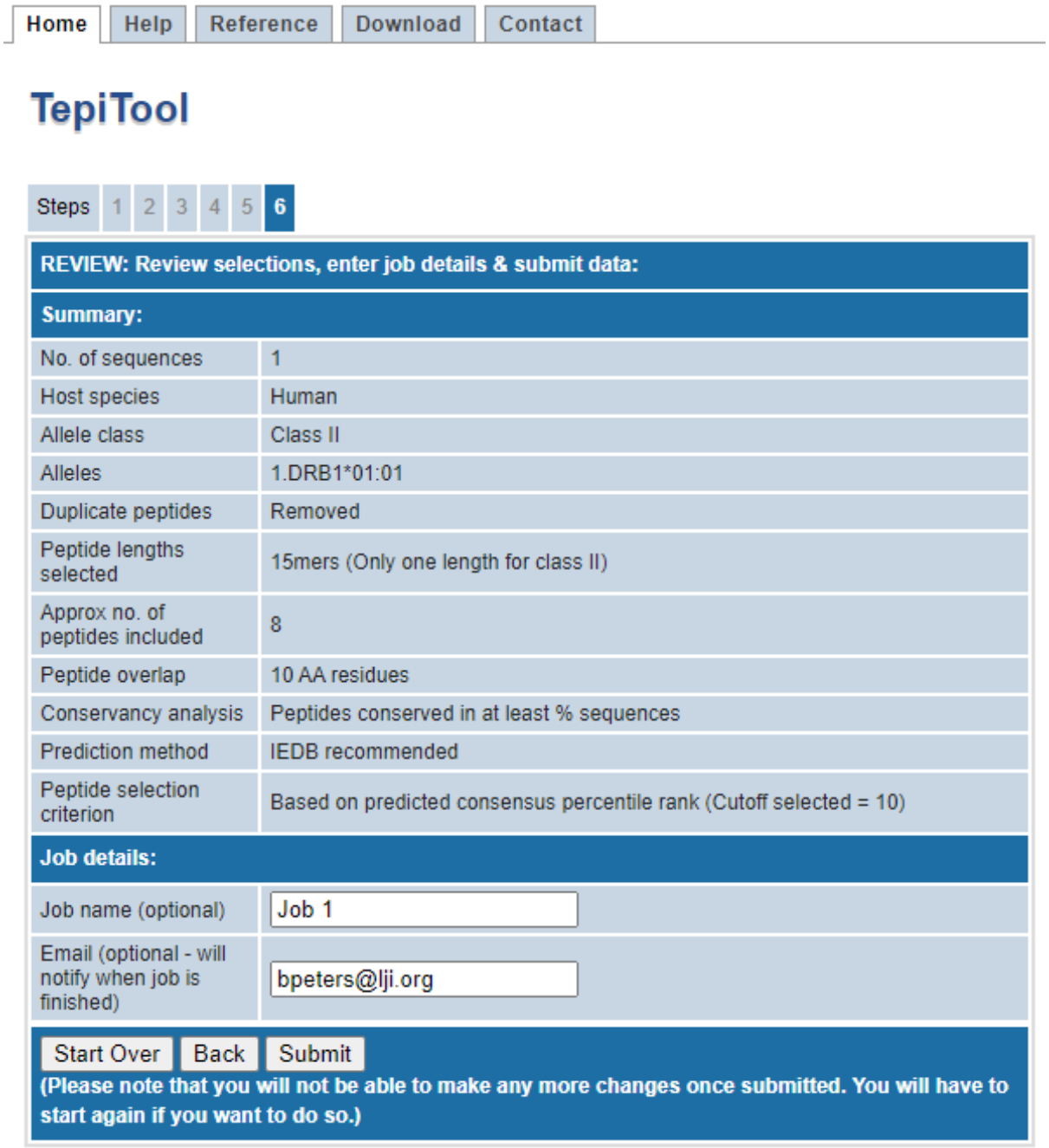
Andreatta et al, Bioinformatics, 2017

TepiTool – MHC I and II Binding Prediction

tools.iedb.org/tepitool/



TepiTool – MHC I and II Binding Prediction



tools.iedb.org/tepitool/

TepiTool – MHC I and II Binding Prediction

TepiTool

Prediction results - concise (Download table):

Seq #	Peptide start	Peptide end	Peptide sequence	Consensus percentile rank	Allele
1	23	37	IIVLIVITGIKAVYN	0.2644	HLA-DRB1*01:01
1	4	18	IVTMFEALPHIIDEV	0.1227	HLA-DRB1*01:01
1	28	42	VITGIKAVYNFATCG	0.0354	HLA-DRB1*01:01
1	11	25	LPHIIDEVINIVIIV	0.0053	HLA-DRB1*01:01
1	33	47	KAVYNFATCGIFALI	0.0024	HLA-DRB1*01:01
1	16	30	DEVINIVIIVLIVIT	0.0016	HLA-DRB1*01:01

Download results details:

Non-redundant results	Prediction results with redundant peptides within each sequence removed - Includes positives and negatives
Complete results	Prediction results of all peptides

Citation information:

If you use these predictions in a manuscript, please include the following in the method section:

For complete list of references please click here: References

Input sequences:

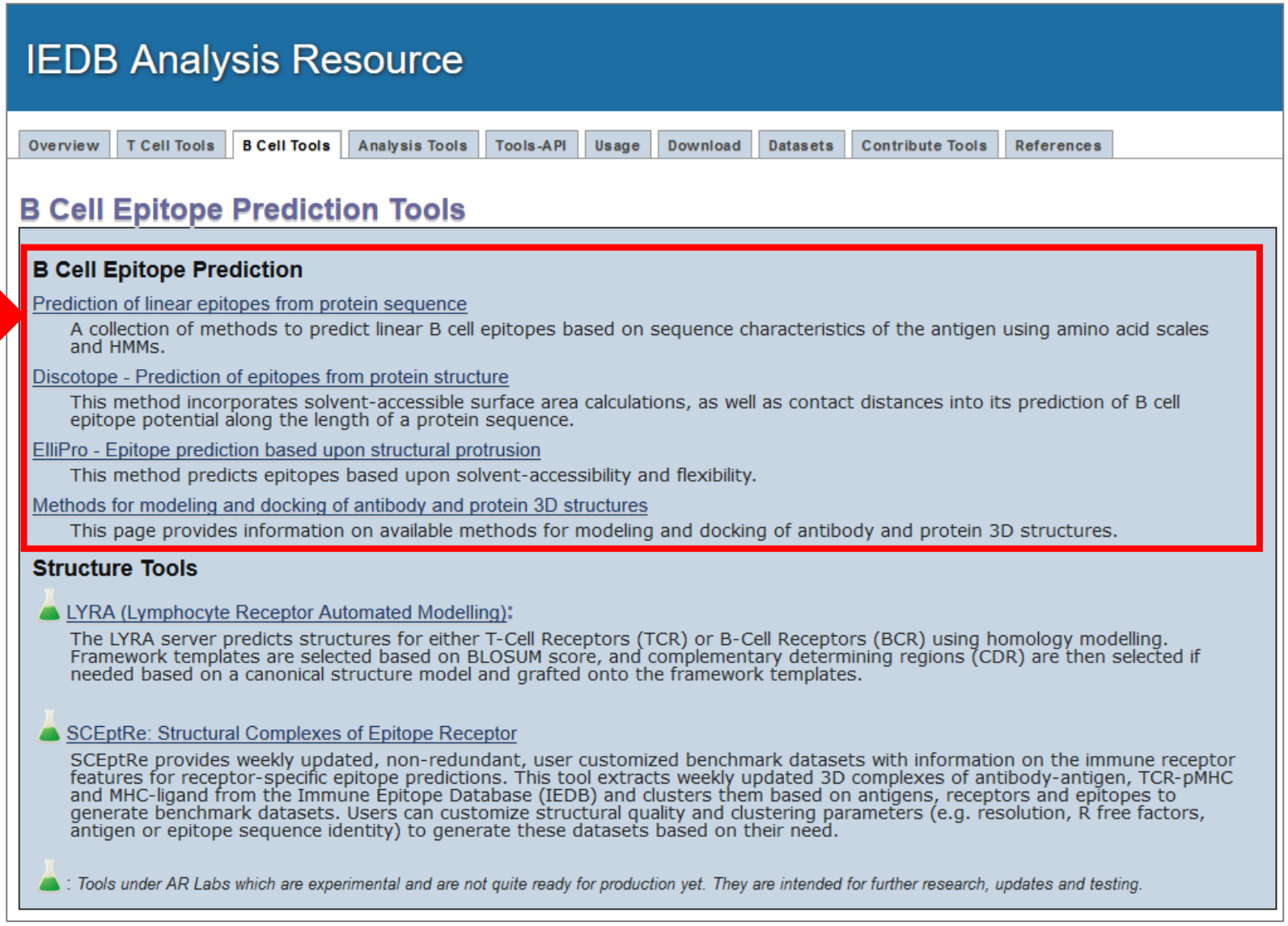
Seq #	Seq title	Sequence
1	LCMV Armstrong, Protein GP	MGQIVTMFEALPHIIDEVINIVIIVLIVITGIKAVYNFATCGIFALISF

Other input parameters:

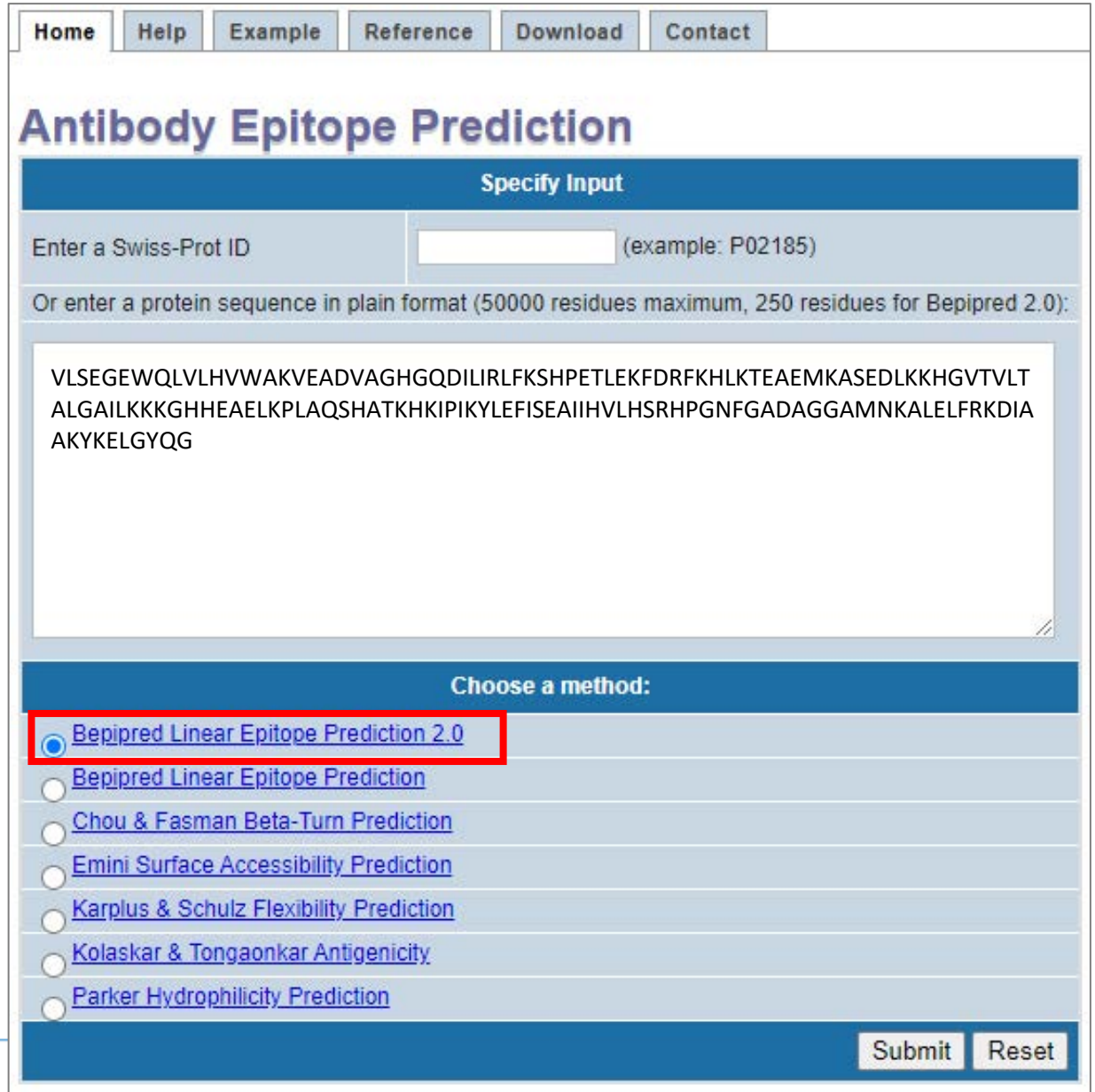
Input summary:	
No. of sequences	1
Host species	Human
Allele class	Class II
Alleles	DRB1*01:01
Duplicate peptides	Removed
Peptide lengths selected	15mers (Only one length for class II)
Peptide overlap	10 AA residues
Conservancy analysis	No
Prediction method	IEDB recommended
Peptide selection criterion	Predicted percentile rank
Cutoff for peptide selection criterion	10
Job name	Job 1
Email	nblazeska@lji.org

tools.iedb.org/tepitool/

B Cell Tools


IEDB Analysis Resource
Overview
T Cell Tools
B Cell Tools
Analysis Tools
Tools-API
Usage
Download
Datasets
Contribute Tools
References
B Cell Epitope Prediction Tools
B Cell Epitope Prediction
Prediction of linear epitopes from protein sequence
A collection of methods to predict linear B cell epitopes based on sequence characteristics of the antigen using amino acid scales and HMMs.
Discotope - Prediction of epitopes from protein structure
This method incorporates solvent-accessible surface area calculations, as well as contact distances into its prediction of B cell epitope potential along the length of a protein sequence.
ElliPro - Epitope prediction based upon structural protrusion
This method predicts epitopes based upon solvent-accessibility and flexibility.
Methods for modeling and docking of antibody and protein 3D structures
This page provides information on available methods for modeling and docking of antibody and protein 3D structures.
Structure Tools
LYRA (Lymphocyte Receptor Automated Modelling):
The LYRA server predicts structures for either T-Cell Receptors (TCR) or B-Cell Receptors (BCR) using homology modelling. Framework templates are selected based on BLOSUM score, and complementary determining regions (CDR) are then selected if needed based on a canonical structure model and grafted onto the framework templates.
SCEptRe: Structural Complexes of Epitope Receptor
SCEptRe provides weekly updated, non-redundant, user customized benchmark datasets with information on the immune receptor features for receptor-specific epitope predictions. This tool extracts weekly updated 3D complexes of antibody-antigen, TCR-pMHC and MHC-ligand from the Immune Epitope Database (IEDB) and clusters them based on antigens, receptors and epitopes to generate benchmark datasets. Users can customize structural quality and clustering parameters (e.g. resolution, R free factors, antigen or epitope sequence identity) to generate these datasets based on their need.
: Tools under AR Labs which are experimental and are not quite ready for production yet. They are intended for further research, updates and testing.

B Cell Epitope Prediction – Sequence Based



tools.iedb.org/bcell/

B Cell Epitope Prediction – Sequence Based

tools.iedb.org/bcell/

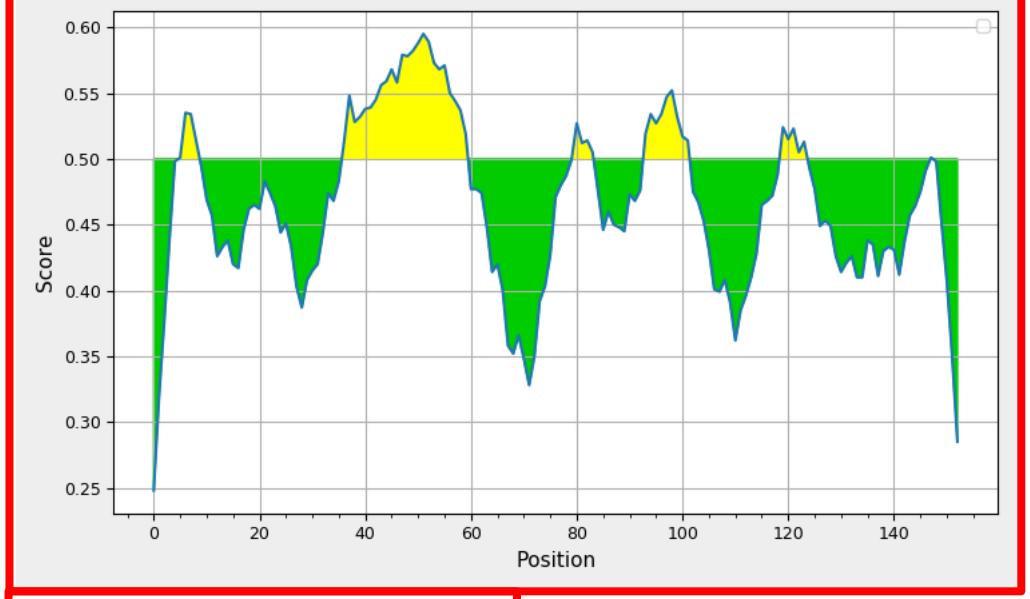
Home | Help | Example | Reference | Download | Contact

Bepipred Linear Epitope Prediction 2.0 Results

Input Sequences

```
  1 VLSEGEWQLV LHVWAKVEAD VAGHGQDILI RLFKSHPETL EKFDRFKHLK TEAEMKASED
 61 LKKHGVTVLT ALGAILKKKG HHEAELKPLA QSHATKHKIP IKYLEFISEA IIHVLHSRHP
121 GNFGADAGGA MNKALELFRK DIAAKYKELG YQG
```

Center position: 4 Threshold: 0.500 Recalculate



Average: 0.466 **Minimum:** 0.248 **Maximum:** 0.595

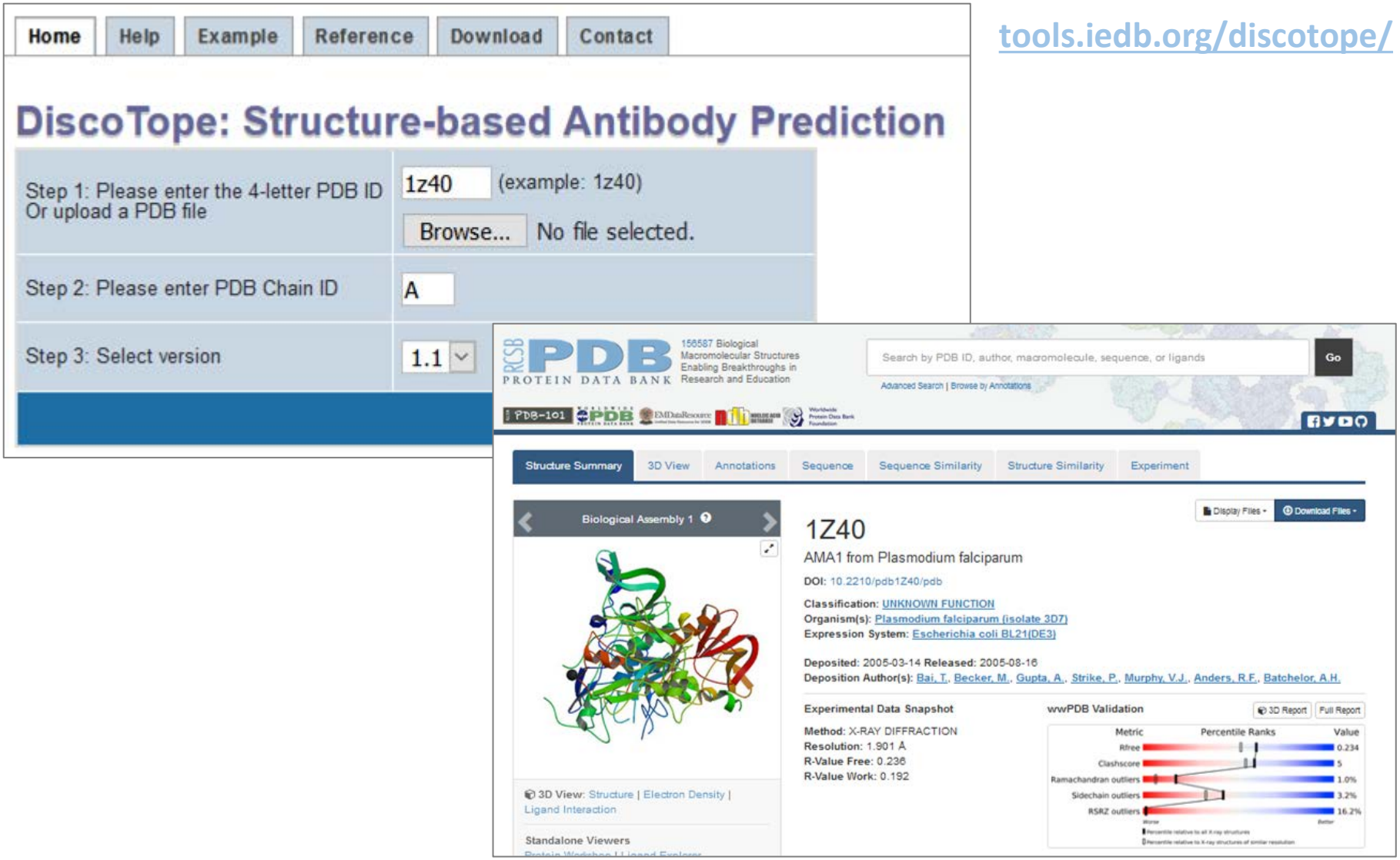
Predicted peptides:

No.	Start	End	Peptide	Length
1	6	9	EWQL	4
2	37	60	PETLEKFDRFKHLKTEAEMKASED	24
3	81	84	HHEA	4
4	94	102	ATKHKIPIK	9
5	120	124	PGNFG	5
6	148	148	E	1

Predicted residue scores:

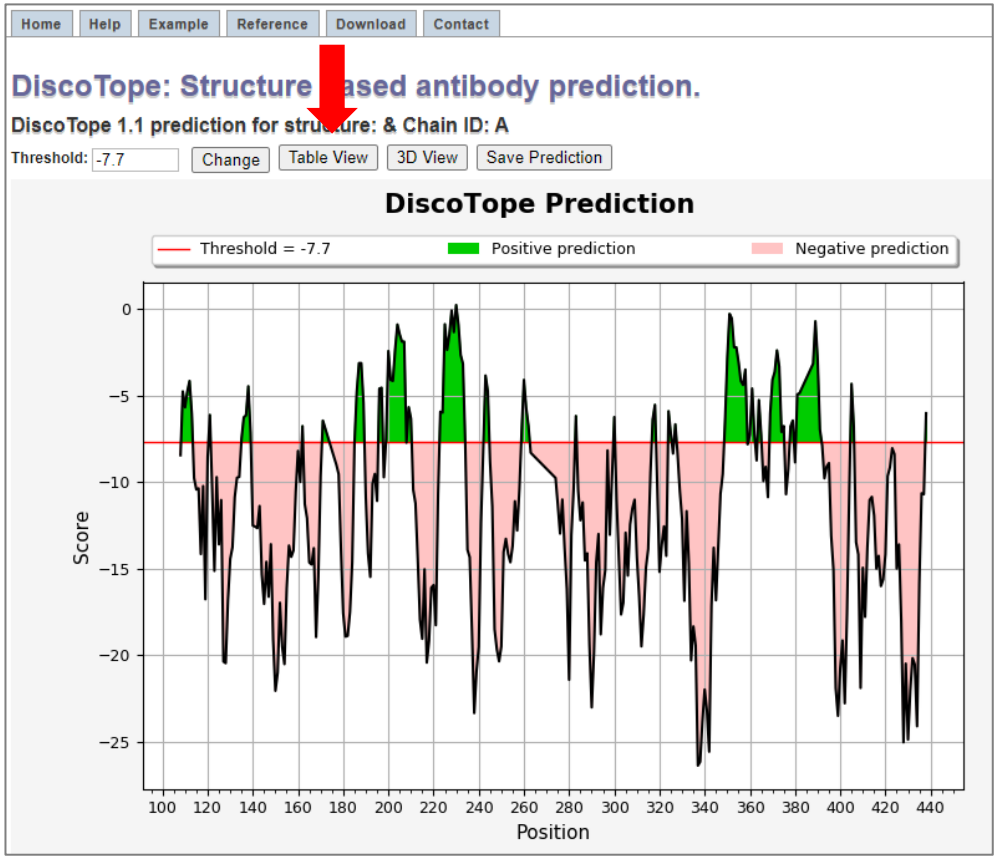
Position	Residue	Score	Assignment
0	V	0.248	.
1	L	0.319	.
2	S	0.377	.
3	E	0.440	.
4	G	0.498	.
5	E	0.501	E

B Cell Epitope Prediction – Structure Based



tools.iedb.org/discotope/

B Cell Epitope Prediction – Structure Based



DiscoTope - Result

DiscoTope 1.1 ediction for structure: & Chain ID: A

The positive predi ons are displayed in green.

Chart View | 3D View | Save Prediction

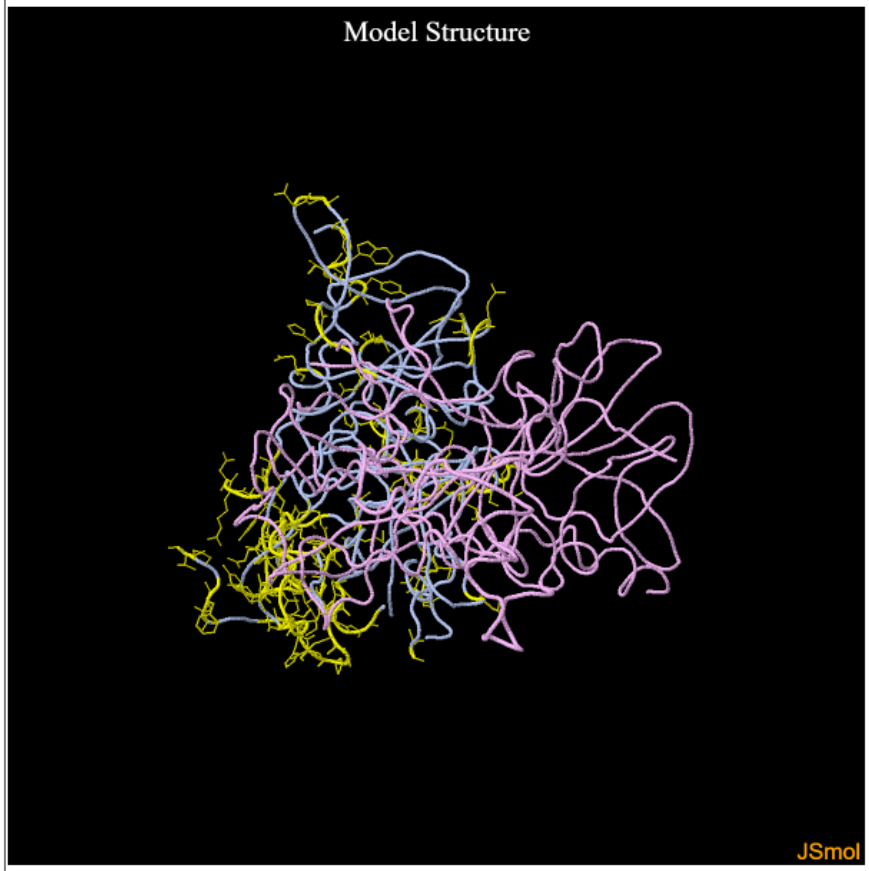
Chain ID	Residue ID	Residue Name	Contact Number	Propensity Score	Discotope Score
A	108	ASN	14	-1.459	-8.459
A	**109**	**PRO**	**11**	**0.724**	**-4.776**
A	**110**	**TRP**	**13**	**0.804**	**-5.696**
A	**111**	**THR**	**12**	**1.211**	**-4.789**
A	**112**	**GLU**	**11**	**1.331**	**-4.169**
A	**113**	**TYR**	**14**	**0.929**	**-6.071**
A	114	MET	18	-0.779	-9.779
A	115	ALA	20	-0.444	-10.444
A	116	LYS	21	0.122	-10.378
A	117	TYR	24	-2.172	-14.172
A	118	ASP	21	0.257	-10.243
A	119	ILE	32	-0.783	-16.783
A	120	GLU	21	1.954	-8.546
A	**121**	**GLU**	**15**	**1.366**	**-6.134**
A	122	VAL	20	-0.374	-10.374
A	123	HIS	28	-1.144	-15.144
A	124	GLY	22	1.274	-9.726
A	125	SER	29	0.887	-13.613
A	126	GLY	28	2.951	-11.049
A	127	ILE	35	-2.881	-20.381
A	128	ARG	29	-5.973	-20.473
A	129	VAL	30	-1.817	-16.817
A	130	ASP	31	1.048	-14.452
A	131	LEU	31	1.727	-13.773
A	132	GLY	25	1.617	-10.883
A	133	GLU	19	-0.26	-9.76
A	134	ASP	18	-0.714	-9.714
A	**135**	**ALA**	**15**	**-0.017**	**-7.517**
A	**136**	**GLU**	**15**	**1.225**	**-6.275**
A	**137**	**VAL**	**22**	**4.846**	**-6.154**
A	**138**	**ALA**	**15**	**3.024**	**-4.476**
A	**139**	**GLY**	**12**	**-1.166**	**-7.166**
A	140	THR	18	-3.524	-12.524
A	141	GLN	20	-2.583	-12.583
A	142	TYR	24	-0.673	-12.673
A	143	ARG	24	0.606	-11.394

B Cell Epitope Prediction – Structure Based

Home | Help | Example | Reference | Download | Contact

JSmol-Rendered PDB Structure

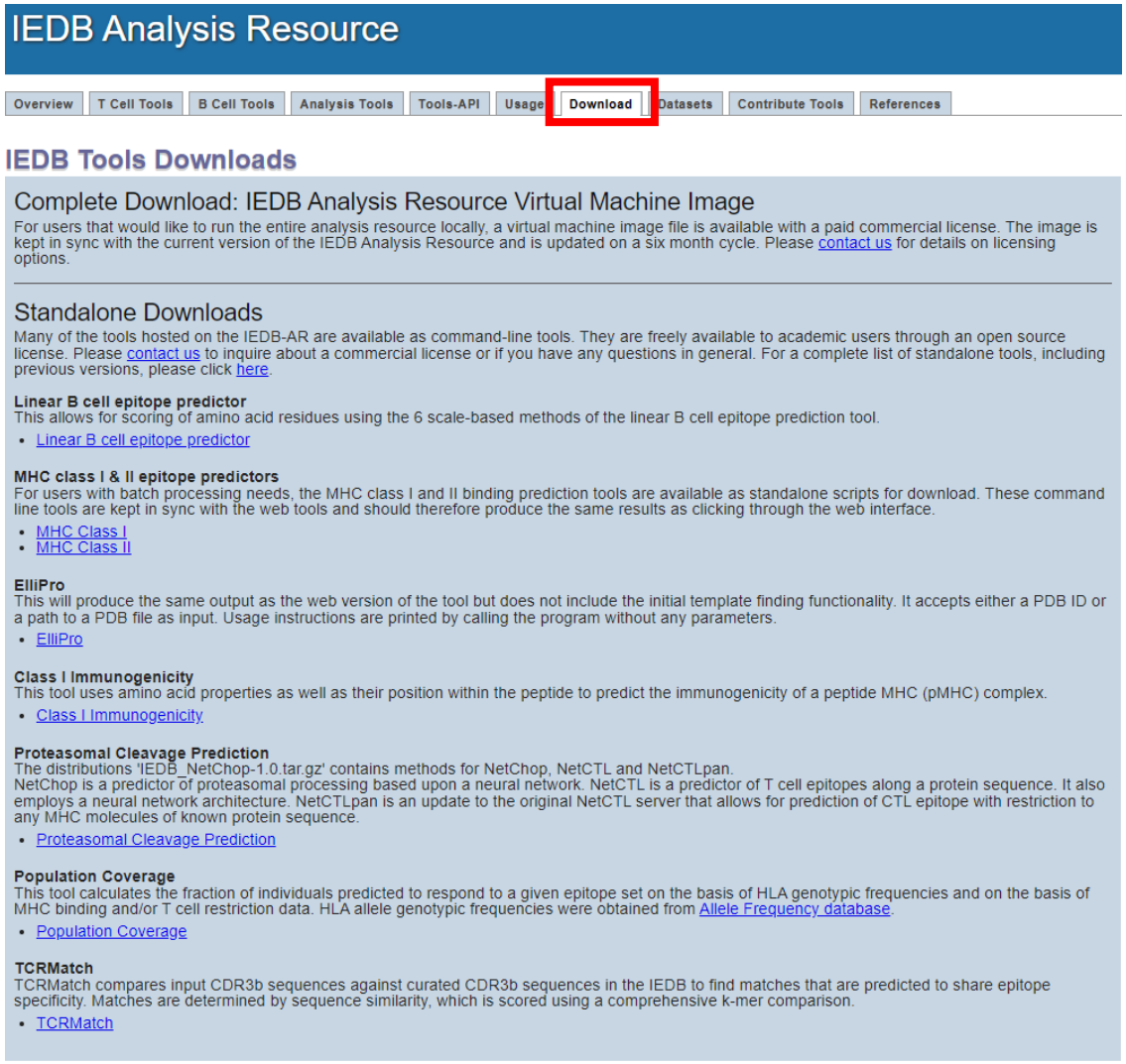
Chart View | Table View | Save Prediction



Chain ID	Residue ID	Residue Name	Contact Number	Propensity Score	Discotope Score	View
A	109	PRO	11	0.724	-4.776	CPK
A	110	TRP	13	0.804	-5.696	CPK
A	111	THR	12	1.211	-4.789	CPK
A	112	GLU	11	1.331	-4.169	CPK
A	113	TYR	14	0.929	-6.071	CPK
A	121	GLU	15	1.366	-6.134	CPK
A	135	ALA	15	-0.017	-7.517	CPK
A	136	GLU	15	1.225	-6.275	CPK
A	137	VAL	22	4.846	-6.154	CPK
A	138	ALA	15	3.024	-4.476	CPK
A	139	GLY	12	-1.166	-7.166	CPK
A	162	ASN	11	-1.279	-6.779	CPK
A	171	THR	9	-1.963	-6.463	CPK
A	186	THR	10	0.194	-4.806	CPK
A	187	GLU	7	0.361	-3.139	CPK
A	188	PRO	7	0.361	-3.139	CPK
A	189	LEU	11	0.049	-5.451	CPK
A	196	ASP	12	1.384	-4.616	CPK
A	197	GLU	15	2.938	-4.562	CPK
A	199	ARG	20	2.492	-7.508	CPK
A	200	HIS	13	4.057	-2.443	CPK
A	201	PHE	16	3.888	-4.112	CPK

Standalone (Local) Version

tools.iedb.org/main/download/



- Run programs on your local machine
- Advantages:
  - No internet needed
  - Very helpful in case of large data sets
  - Free for non-profit & academia
  - Available for industry at a nominal fee
- license@iedb.org

Standalone (Local) Version

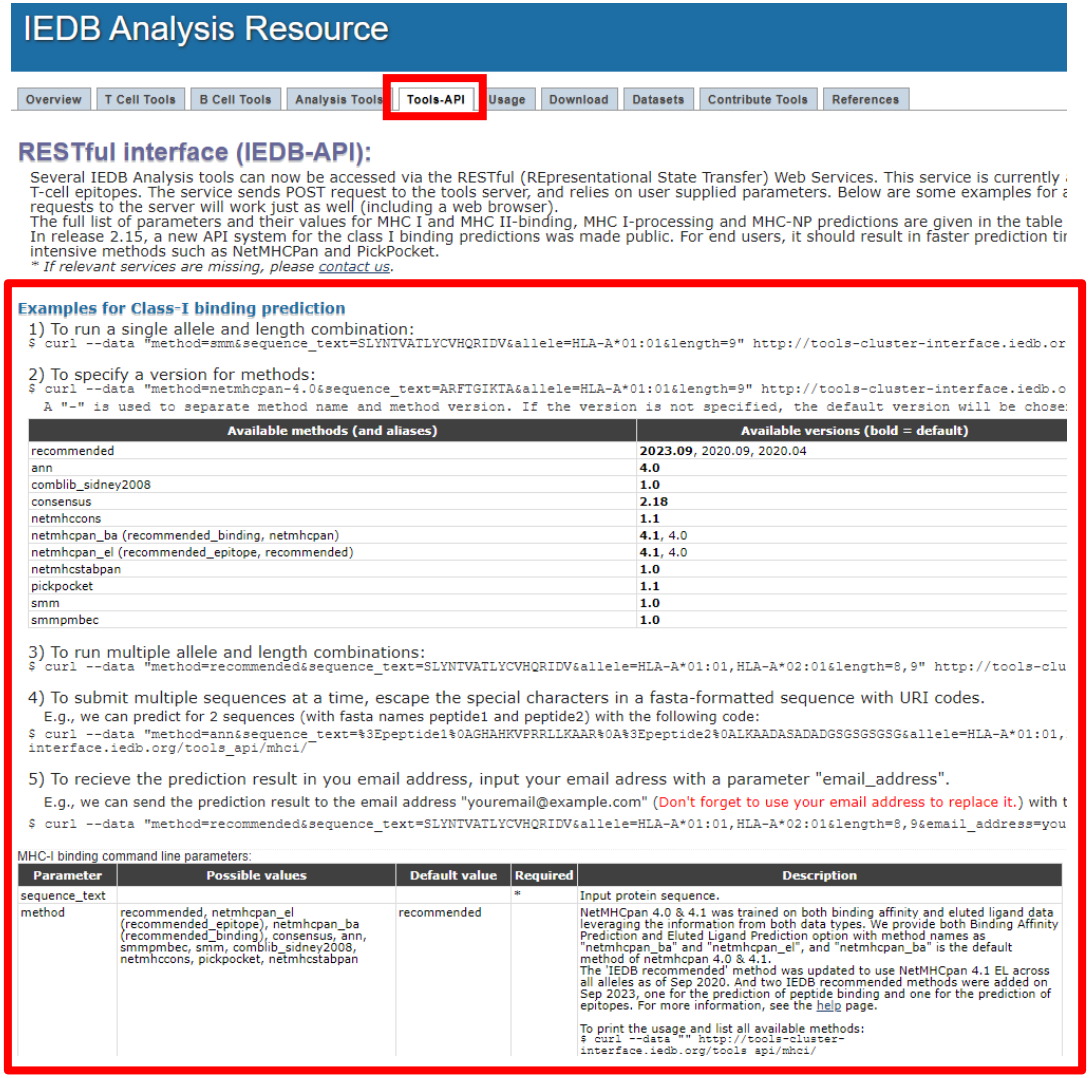
```
spaul@ubuntu:~/tools/mhc_i$ ./src/predict_binding.py netmhcpan "HLA-A*02:01" 9 test_sequence.fasta
allele  seq_num start   end     length  peptide ic50    rank
HLA-A*02:01     1       6       14      9       TMFEALPHI       4.3     0.03
HLA-A*02:01     1       10      18      9       ALPHIIDEV       12.6    0.12
HLA-A*02:01     1       14      22      9       IIDEVINIV       43.3    0.47
HLA-A*02:01     1       38      46      9       FATCGIFAL       65.2    0.64
HLA-A*02:01     1       13      21      9       HIIDEVINI       97.3    0.87
HLA-A*02:01     1       2       10      9       GQIVTMFEA       245.5   1.8
HLA-A*02:01     1       27      35      9       IVITGIKAV       324.4   2.1
HLA-A*02:01     1       24      32      9       IVLIVITGI       651.0   3.0
HLA-A*02:01     1       17      25      9       EVINIVIIV       3263.2  7.2
HLA-A*02:01     1       18      26      9       VINIVIIVL       3491.7  7.5
HLA-A*02:01     1       3       11      9       QIVTMFEAL       3523.2  7.5
HLA-A*02:01     1       20      28      9       NIVIIVLIV       4861.5  9.0
HLA-A*02:01     1       21      29      9       IVIIVLIVI       5959.9  12
HLA-A*02:01     1       31      39      9       GIKAVYNFA       6530.5  12
HLA-A*02:01     1       7       15      9       MFEALPHII       11337.5 17
HLA-A*02:01     1       37      45      9       NFATCGIFA       11579.5 17
HLA-A*02:01     1       26      34      9       LIVITGIKA       12185.2 18
HLA-A*02:01     1       33      41      9       KAVYNFATC       12713.3 18
HLA-A*02:01     1       22      30      9       VIIVLIVIT       13298.3 19
HLA-A*02:01     1       19      27      9       INIVIIVLI       14913.1 21
HLA-A*02:01     1       4       12      9       IVTMFEALP       15618.8 22
HLA-A*02:01     1       34      42      9       AVYNFATCG       18955.1 25
HLA-A*02:01     1       35      43      9       VYNFATCGI       19064.3 26
HLA-A*02:01     1       25      33      9       VLIVITGIK       19910.8 27
HLA-A*02:01     1       36      44      9       YNFATCGIF       21579.6 29
HLA-A*02:01     1       23      31      9       IIVLIVITG       22420.2 30
HLA-A*02:01     1       28      36      9       VITGIKAVY       28040.2 39
HLA-A*02:01     1       1       9       9       MGQIVTMFE       28618.9 40
HLA-A*02:01     1       32      40      9       IKAVYNFAT       28756.4 41
HLA-A*02:01     1       5       13      9       VTMFEALPH       28919.3 41
HLA-A*02:01     1       30      38      9       TGIKAVYNF       29272.8 42
HLA-A*02:01     1       11      19      9       LPHIIDEVI       29396.6 42
HLA-A*02:01     1       15      23      9       IDEVINIVI       32112.8 48
HLA-A*02:01     1       16      24      9       DEVINIVII       32893.9 50
HLA-A*02:01     1       9       17      9       EALPHIIDE       37902.4 66
HLA-A*02:01     1       29      37      9       ITGIKAVYN       38010.0 66
HLA-A*02:01     1       8       16      9       FEALPHIID       40600.3 76
HLA-A*02:01     1       12      20      9       PHIIDEVIN       48177.9 100
spaul@ubuntu:~/tools/mhc_i$ ./src/predict_binding.py netmhcpan "HLA-A*02:01" 9 test_s
spaul@ubuntu:~/tools/mhc_i$
```

prediction_results.txt

	allele	seq_num	start	end	length	peptide	ic50	rank
2	HLA-A*02:01	1	6	14	9	TMFEALPHI	4.3	0.03
3	HLA-A*02:01	1	10	18	9	ALPHIIDEV	12.6	0.12
4	HLA-A*02:01	1	14	22	9	IIDEVINIV	43.3	0.47
5	HLA-A*02:01	1	38	46	9	FATCGIFAL	65.2	0.64
6	HLA-A*02:01	1	13	21	9	HIIDEVINI	97.3	0.87
7	HLA-A*02:01	1	2	10	9	GQIVTMFEA	245.5	1.8
8	HLA-A*02:01	1	27	35	9	IVITGIKAV	324.4	2.1
9	HLA-A*02:01	1	24	32	9	IVLIVITGI	651.0	3.0
10	HLA-A*02:01	1	17	25	9	EVINIVIIV	3263.2	7.2
11	HLA-A*02:01	1	18	26	9	VINIVIIVL	3491.7	7.5
12	HLA-A*02:01	1	3	11	9	QIVTMFEAL	3523.2	7.5
13	HLA-A*02:01	1	20	28	9	NIVIIVLIV	4861.5	9.0
14	HLA-A*02:01	1	21	29	9	IVIIVLIVI	5959.9	12
15	HLA-A*02:01	1	31	39	9	GIKAVYNFA	6530.5	12
16	HLA-A*02:01	1	7	15	9	MFEALPHII	11337.5	17
17	HLA-A*02:01	1	37	45	9	NFATCGIFA	11579.5	17
18	HLA-A*02:01	1	26	34	9	LIVITGIKA	12185.2	18
19	HLA-A*02:01	1	33	41	9	KAVYNFATC	12713.3	18
20	HLA-A*02:01	1	22	30	9	VIIVLIVIT	13298.3	19
21	HLA-A*02:01	1	19	27	9	INIVIIVLI	14913.1	21
22	HLA-A*02:01	1	4	12	9	IVTMFEALP	15618.8	22
23	HLA-A*02:01	1	34	42	9	AVYNFATCG	18955.1	25
24	HLA-A*02:01	1	35	43	9	VYNFATCGI	19064.3	26
25	HLA-A*02:01	1	25	33	9	VLIVITGIK	19910.8	27
26	HLA-A*02:01	1	36	44	9	YNFATCGIF	21579.6	29
27	HLA-A*02:01	1	23	31	9	IIVLIVITG	22420.2	30
28	HLA-A*02:01	1	28	36	9	VITGIKAVY	28040.2	39
29	HLA-A*02:01	1	1	9	9	MGQIVTMFE	28618.9	40
30	HLA-A*02:01	1	32	40	9	IKAVYNFAT	28756.4	41
31	HLA-A*02:01	1	5	13	9	VTMFEALPH	28919.3	41
32	HLA-A*02:01	1	30	38	9	TGIKAVYNF	29272.8	42
33	HLA-A*02:01	1	11	19	9	LPHIIDEVI	29396.6	42
34	HLA-A*02:01	1	15	23	9	IDEVINIVI	32112.8	48
35	HLA-A*02:01	1	16	24	9	DEVINIVII	32893.9	50
36	HLA-A*02:01	1	9	17	9	EALPHIIDE	37902.4	66
37	HLA-A*02:01	1	29	37	9	ITGIKAVYN	38010.0	66
38	HLA-A*02:01	1	8	16	9	FEALPHIID	40600.3	76
39	HLA-A*02:01	1	12	20	9	PHIIDEVIN	48177.9	100

API Version (RESTful Interface)

tools.iedb.org/main/tools-api/



IEDB Analysis Resource

Overview | T Cell Tools | B Cell Tools | Analysis Tools | Tools-API | Usage | Download | Datasets | Contribute Tools | References

RESTful interface (IEDB-API):

Several IEDB Analysis tools can now be accessed via the RESTful (REpresentational State Transfer) Web Services. This service is currently
T-cell epitopes. The service sends POST request to the tools server, and relies on user supplied parameters. Below are some examples for
requests to the server will work just as well (including a web browser).
The full list of parameters and their values for MHC I and MHC II-binding, MHC I-processing and MHC-NP predictions are given in the table
In release 2.15, a new API system for the class I binding predictions was made public. For end users, it should result in faster prediction
intensive methods such as NetMHCPan and PickPocket.
** If relevant services are missing, please contact us.*

Examples for Class-I binding prediction

1) To run a single allele and length combination:

```
$ curl --data "method=smm&sequence_text=SLYNTVATLYCVHQRIDV&allele=HLA-A*01:01&length=9" http://tools-cluster-interface.iedb.or
```

2) To specify a version for methods:

```
$ curl --data "method=netmhcpan-4.0&sequence_text=ARFTGIKTA&allele=HLA-A*01:01&length=9" http://tools-cluster-interface.iedb.o
```

A "-" is used to separate method name and method version. If the version is not specified, the default version will be chose

Available methods (and aliases)	Available versions (bold = default)
recommended	**2023.09**, 2020.09, 2020.04
ann	**4.0**
comblib_sidney2008	**1.0**
consensus	**2.18**
netmhccons	**1.1**
netmhcpan_ba (recommended_binding, netmhcpan)	**4.1**, 4.0
netmhcpan_el (recommended_epitope, recommended)	**4.1**, 4.0
netmhcstabpan	**1.0**
pickpocket	**1.1**
smm	**1.0**
smmpmbec	**1.0**

3) To run multiple allele and length combinations:

```
$ curl --data "method=recommended&sequence_text=SLYNTVATLYCVHQRIDV&allele=HLA-A*01:01,HLA-A*02:01&length=8,9" http://tools-clu
```

4) To submit multiple sequences at a time, escape the special characters in a fasta-formatted sequence with URI codes.
E.g., we can predict for 2 sequences (with fasta names peptide1 and peptide2) with the following code:

```
$ curl --data "method=ann&sequence_text=%3Epeptide1%0AGHAHKVPRRLLKAAR%0A%3Epeptide2%0ALKAADASADADGSGSGSGSG&allele=HLA-A*01:01,
interface.iedb.org/tools_api/mhci/
```

5) To recieve the prediction result in you email address, input your email adress with a parameter "email_address".
E.g., we can send the prediction result to the email address "youremail@example.com" (Don't forget to use your email address to replace it.) with t

```
$ curl --data "method=recommended&sequence_text=SLYNTVATLYCVHQRIDV&allele=HLA-A*01:01,HLA-A*02:01&length=8,9&email_address=you
```

MHC-I binding command line parameters:

Parameter	Possible values	Default value	Required	Description
sequence_text			*	Input protein sequence.
method	recommended, netmhcpan_el (recommended_epitope), netmhcpan_ba (recommended_binding), consensus, ann, smmpmbec, smm, comblib_sidney2008, netmhccons, pickpocket, netmhcstabpan	recommended		NetMHCpan 4.0 & 4.1 was trained on both binding affinity and eluted ligand data leveraging the information from both data types. We provide both Binding Affinity Prediction and Eluted Ligand Prediction option with method names as "netmhcpan_ba" and "netmhcpan_el", and "netmhcpan_ba" is the default method of netmhcpan 4.0 & 4.1. The 'IEDB recommended' method was updated to use NetMHCpan 4.1 EL across all alleles as of Sep 2020. And two IEDB recommended methods were added on Sep 2023, one for the prediction of peptide binding and one for the prediction of epitopes. For more information, see the help page. To print the usage and list all available methods: `$ curl --data "" http://tools-cluster-interface.iedb.org/tools_api/mhci/`

- Sends prediction request to the tools server at LJI
- No need to install tools on your machine
- Freely available to all users
- Can be incorporated in prediction pipelines
- Automatic update without reinstalling

API Version (RESTful Interface) – Example

	A	B
1	peptide	allele
2	EALPHIIDEVINI	HLA-B*58:01
3	AVAKCNVNHDAEFC	HLA-A*68:01
4	SKFKEDVESA	HLA-A*68:02
5	SHLNLTMPNA	HLA-A*01:01
6	LMRNHLRDLMGV	HLA-A*32:01
7	NPGGYCLTKWMILA	HLA-A*26:01
8	AQSAQSQCRT	HLA-A*01:01
9	LSIRGNSNYKAVSC	HLA-A*03:01
10	QCRTFRGRVLDMF	HLA-B*53:01
11	GTSGLELTFTND	HLA-A*11:01
12	NLTSAFNKK	HLA-A*23:01
13	CDMLRLIDYNKAA	HLA-B*53:01
14	YIKRQGSTPL	HLA-A*26:01
15	YMRSGWGWTG	HLA-A*23:01
16	LVTNGSYLNETHF	HLA-B*58:01
17	TKFFTRRL	HLA-B*57:01
18	NVNHDAEFCDMLRL	HLA-B*08:01
19	HIKGGSCPKPH	HLA-A*30:01
20	DGKTTWCSQTS	HLA-A*32:01
21	HFSDQIEQEADNM	HLA-A*32:01
22	FSDQIEQEADNMI	HLA-B*57:01
23	CNYSKFWY	HLA-B*58:01
24	MSHLNLTMPNAC	HLA-A*02:01
25	SGVENPGGYC	HLA-B*44:03
26	VIIVLIVITGIK	H-2-Kb
27	FRGRVLDMFR	HLA-B*51:01
28	HIKGGSCPKPHR	HLA-B*44:02
29	SIRGNSNYKAVS	H-2-Kb
30	IQYNLTFSDA	HLA-A*02:06
31	RTFRGRVL	HLA-B*15:01
32	DAQSAQSQCRTFRG	HLA-B*44:02
33	QNRTWENHCTYAGP	HLA-B*15:01
34	AFGGKYMRSGWGWT	HLA-B*07:02
35	SRILLSQEKTKFFT	HLA-A*31:01
36	SALHLFKTTVNSLI	H-2-Kb
37	RKDYIKRQGSTP	HLA-B*58:01
38	FKSVEFDMSHLNLT	HLA-B*58:01

Sheet1 Sheet2 Sheet3

api_predictor.py

```
import pandas as pd
import shlex, subprocess

peptide_data = pd.read_csv('peptides_to_predict.txt', sep='\t')
alleles = peptide_data['allele'].tolist()
peptides = peptide_data['peptide'].tolist()


def prediction(peptide, allele, length):
    command = 'curl --data "method=recommended&sequence_text=' + peptide + \
              '&allele=' + allele + \
              '&length=' + str(length) + \
              '" http://tools-cluster-interface.iedb.org/tools_api/mhci/'
    args = shlex.split(command)
    process = subprocess.Popen(args,stdout=subprocess.PIPE,stderr=subprocess.PIPE)
    output = process.communicate()
    consensus_percentile = output[0].decode('utf8').split('\t')[22]
    return consensus_percentile


output_file_name = 'prediction_results.txt'
output_file = open(output_file_name,'w')
with open(output_file_name,'a') as file_to_write:
    file_to_write.write('peptide\tallele\tlength\tpercentile_rank\n')
    for i in range(len(alleles)):
        peptide = peptides[i]
        allele = alleles[i]
        length = len(peptide)
        consensus_percentile = prediction(peptide, allele,
        write_line = peptide + '\t' + allele + '\t' + str(
        file_to_write.write(write_line)
output_file.close()
```

prediction_results.txt

	peptide	allele	length	percentile_rank
2	EALPHIIDEVINI	HLA-B*58:01	13	53.0
3	AVAKCNVNHDAEFC	HLA-A*68:01	14	56.0
4	SKFKEDVESA	HLA-A*68:02	10	36.5
5	SHLNLTMPNA	HLA-A*01:01	10	46.5
6	LMRNHLRDLMGV	HLA-A*32:01	12	45.0
7	NPGGYCLTKWMILA	HLA-A*26:01	14	57.0
8	AQSAQSQCRT	HLA-A*01:01	10	28.5
9	LSIRGNSNYKAVSC	HLA-A*03:01	14	13.0
10	QCRTFRGRVLDMF	HLA-B*53:01	13	41.0
11	GTSGLELTFTND	HLA-A*11:01	12	46.0
12	NLTSAFNKK	HLA-A*23:01	9	28.5
13	CDMLRLIDYNKAA	HLA-B*53:01	13	53.0
14	YIKRQGSTPL	HLA-A*26:01	10	10.25
15	YMRSGWGWTG	HLA-A*23:01	10	13.55

Versions of IEDB Analysis Resource tools

Web

http://tools.iedb.org

- Client uses browsers to submit data
- Predictions run on IEDB tools server
- Can be run on Windows/Mac/Linux
- Internet is needed
- May not be suitable for very large data sets
- Automatically updated

Standalone

http://tools.iedb.org/main/download

- Uses command line interface
- Downloaded from IEDB website
- Installed and run on local machine
- Runs on Linux only (can use virtual machines to run Linux on other OS)
- Internet not needed once installed
- Better for very large data sets
- Need to update for every release
- Free to academia/non-profit; License fee for industry

API

http://tools.iedb.org/main/tools-api

- Uses command line interface
- Predictions run on IEDB server
- Clients send parameters to IEDB server using commands or scripts
- Internet is needed
- Can be used to make custom scripts for use with large data sets
- Automatically updated
- Free to all

Points to Remember

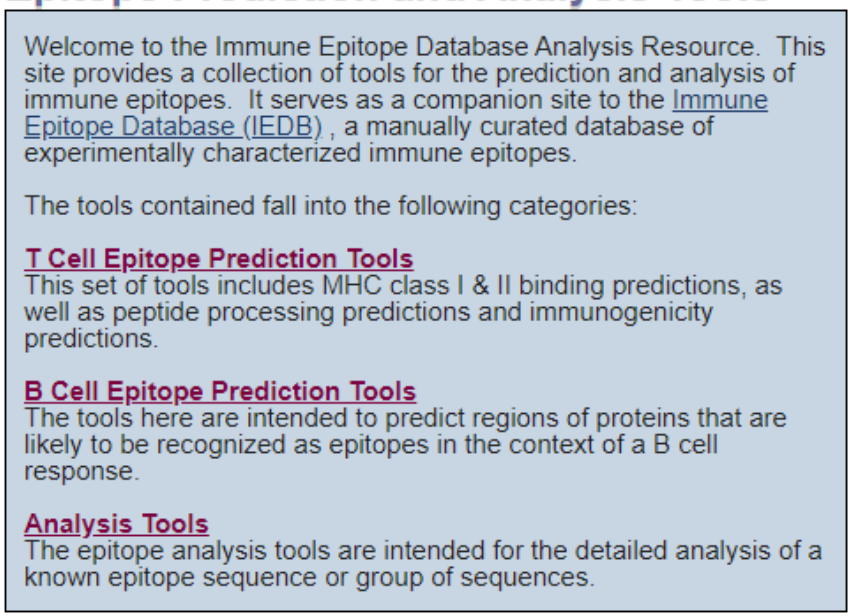
- First stop is IEDB database

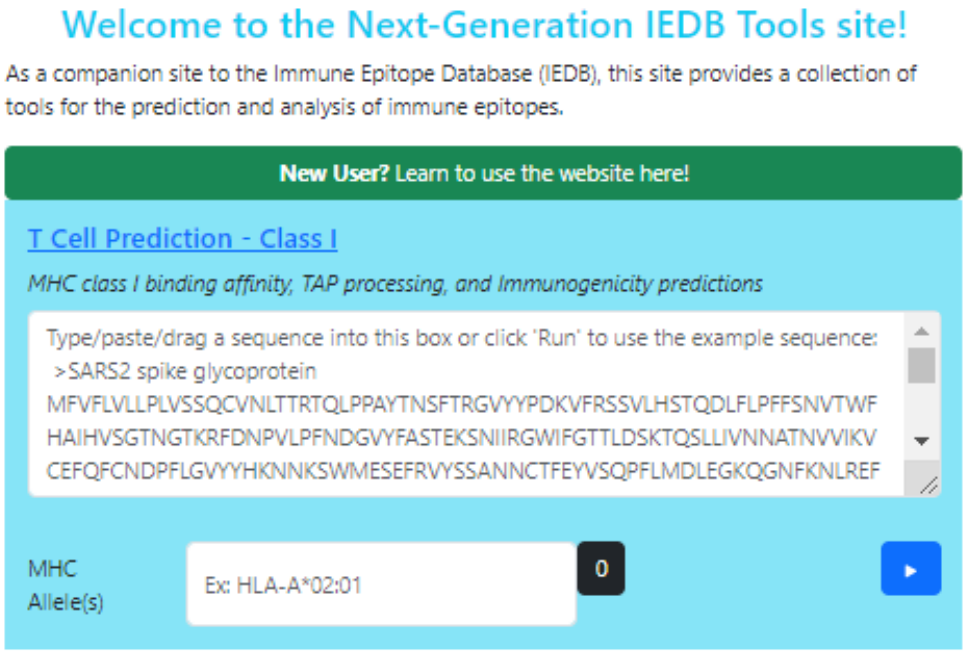
- Epitope prediction tools extrapolate from existing data to identify new candidate epitopes
  - 'Machine learning' approaches identify patterns
  - ROC curves / AUC values as preferred performance metrics
  - Prediction is a screening step, not confirmatory
  - Predicted peptides should be experimentally tested for verification

- Analysis tools help to examine existing sets of epitopes and gain new knowledge
  - No single metric of performance
  - Broad array of applications

IEDB Tools

Analysis Resource & Next-Generation Tools





Analysis Resource:

- T cell epitope prediction
- B cell epitope prediction
- Analysis tools

tools.iedb.org

Next-Generation Tools:

- Re-implementing existing tools with a focus on improving usability and function
- Launched in 2023
- More tools to be added

nextgen-tools.iedb.org

Test them out!