



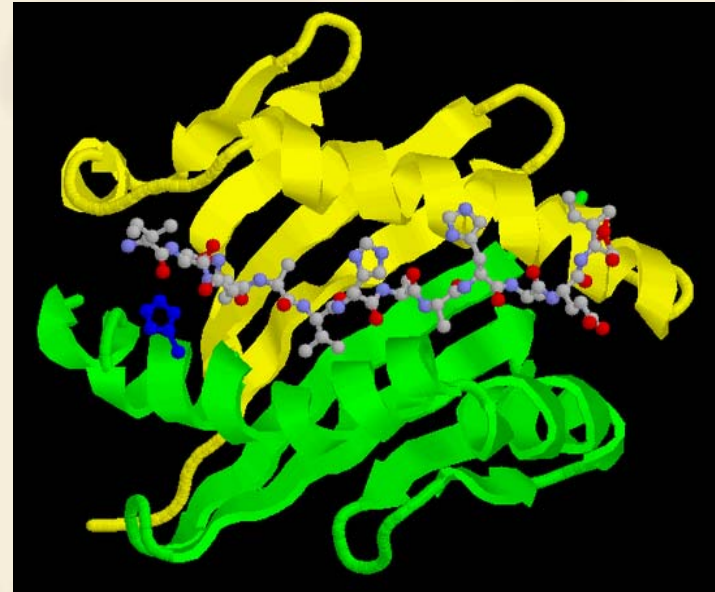
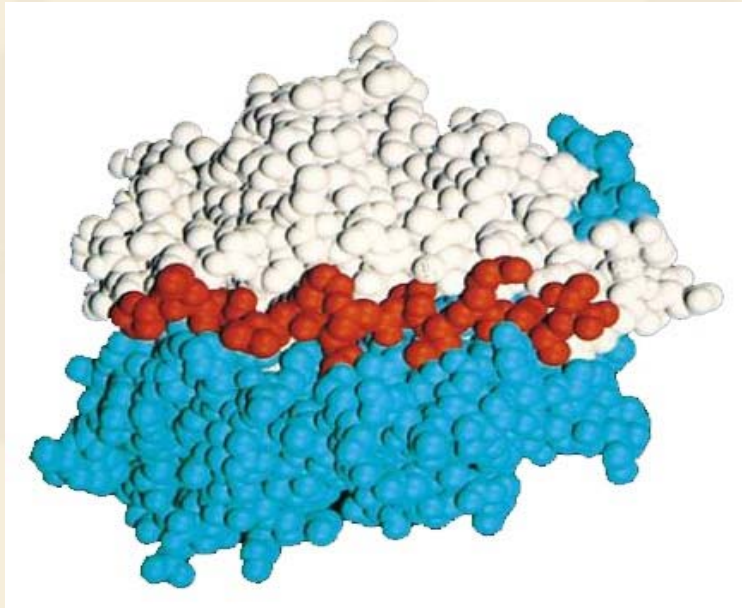
# **MetaMHCIpan: a consensus approach for pan-specific HLA-DR binding predictions**

**Shanfeng Zhu**

School of Computer Science and  
Shanghai Key Lab of  
Intelligent Information Processing,  
Fudan University, China



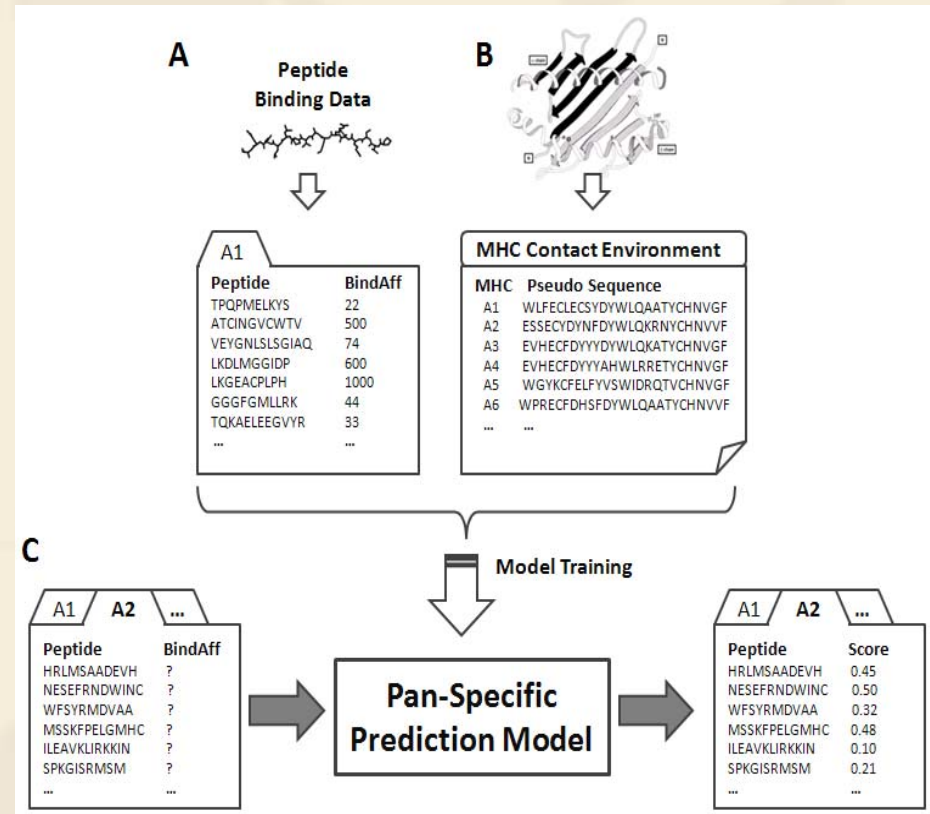
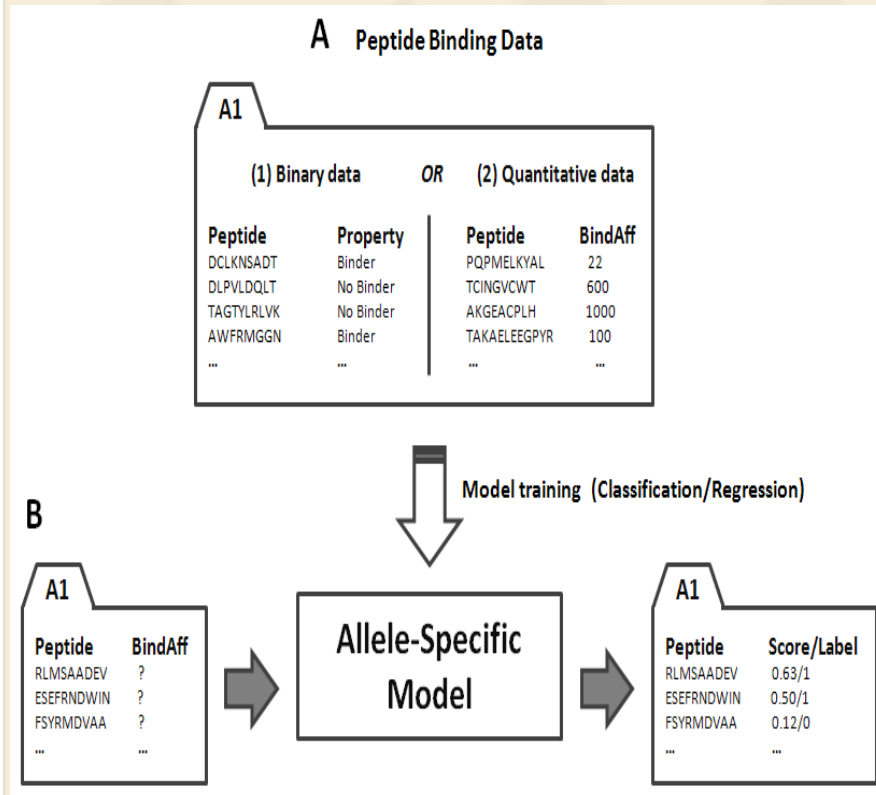
# MHC molecular and binding peptide



Accurate identification of MHC-restricted peptides is of great importance for elucidating the underlying mechanism of immune recognition, as well as for developing effective epitope-based vaccines and promising immunotherapies for many severe diseases.

# Allele-specific vs. Pan-specific methods

[Zhang et al., 2012]



The advantage of pan-specific method.



# Current methods

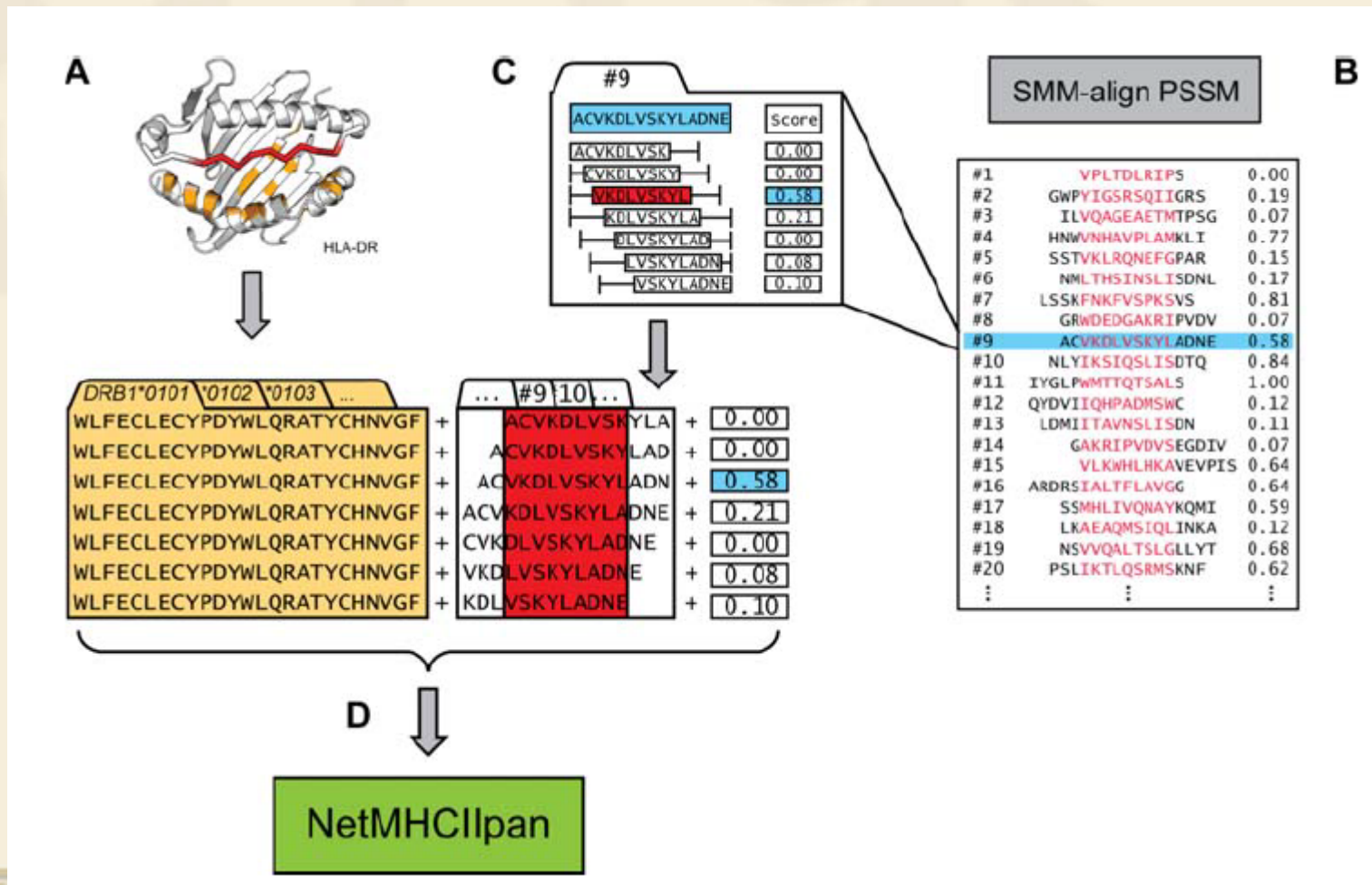
## ❖ Pan-specific methods:

- ❧ TEPITOPE
- ❧ NetMHCIIpan-1.0
- ❧ NetMHCIIpan-2.0
- ❧ MultiRTA



# NetMHCIIpan-1.0 and 2.0

(Nielsen et al., 2008;2010)



[Zhang et al., 2012]

**Table 8:** Comparison of MHC-II pan-specific methods on independent binding data

Allele	No.	No. of bind	TEPITOPE	MultiRTA	NetMHCpan-1.0	NetMHCpan-2.0
DRBI*01:01	732	270	0.763	0.767	<b>0.778</b>	0.77
DRBI*03:01	557	247	0.608	0.623	0.652	<b>0.665</b>
DRBI*04:01	779	386	0.737	0.744	<b>0.763</b>	0.76
DRBI*04:02	113	31	<b>0.782</b>	0.753	0.78	0.766
DRBI*04:04	204	65	0.784	0.764	<b>0.802</b>	0.779
DRBI*04:05	95	36	0.724	<b>0.728</b>	0.719	0.692
DRBI*07:01	507	222	0.65	0.646	0.639	<b>0.655</b>
DRBI*08:02	68	32	0.694	<b>0.747</b>	0.739	0.681
DRBI*11:01	575	219	0.743	0.734	0.729	<b>0.762</b>
DRBI*12:01	693	444	–	0.7	0.753	<b>0.78</b>
DRBI*13:01	168	40	<b>0.722</b>	0.71	0.693	0.642
DRBI*13:02	88	40	0.556	0.591	<b>0.645</b>	0.619
DRBI*15:01	251	149	0.717	0.639	0.713	<b>0.751</b>
DRB3*01:01	146	23	–	0.548	<b>0.627</b>	0.586
DRB3*02:02	648	315	–	0.685	0.735	<b>0.812</b>
DRB4*01:01	156	67	–	0.542	0.628	<b>0.637</b>
DRB5*01:01	217	151	0.888	0.851	0.874	<b>0.892</b>
Avg (17)	5997	2737		0.692	<b>0.722</b>	<b>0.721</b>
In Tepitope(13)	4354	1888	<b>0.721</b>	0.715	<b>0.733</b>	0.726
!In Tepitope(4)	1643	849		0.619	0.686	<b>0.704</b>



# Comparison of different methods

- ① Different method, based on different principles, have different performance.
- ② The prediction accuracy for MHC II should be improved.
- ③ NetMHCIIpan is the best predictor among the four methods. Although TEPITOPE is proposed early, the predicted result is satisfying and especially good for ligand and binding core prediction.



# TEPITOPEpan

[Zhang et al., 2012]

- ❖ TEPITOPE has a library of 11 PSSMs. One PSSM is a  $20 \times 9$  matrix where nine binding specificity vectors correspond to nine pockets. Each of the 11 PSSMs corresponds to one of 11 known DRB alleles.
- ❖ TEPITOPEpan uses this library to generate a PSSM for an arbitrary HLA-DRB allele. In a generated PSSM, each vector is a weighted average of binding specificity vectors of the corresponding pocket over 11 DRB alleles.

# Illustration of TEPITOPEpan

Binding specificity vector of allele DRB1\*0813 at P4

Source MHC allele

DRB1\*0101-**P4** weight 0.1

DRB1\*0301-**P4** weight 0.3

⋮

DRB5\*0101-**P4** weight 0.2

Linear weighted  
combination

DRB1\*0813-**P4**

{ ..., **P4**, ... }

Same for P1~P9

# Performance of TEPITOPEpan

AUC of different methods on 10 new alleles without training data in  
Nielsen-Set2 (Nielsen et al., 2010)

[Zhang et al., 2012]

MHC	Count	Binder	NetMHCIIpan 2.0	NetMHCIIpan- 1.0	MultiRTA	TEPITOPE	TEPITOPEpan
DRB1*03:02	148	44	<b>0.759</b>	0.688	0.549		0.602
DRB1*08:06	118	91	<b>0.902</b>	0.703	0.652	0.87	0.886
DRB1*08:13	1370	455	0.666	0.763	0.712	0.746	<b>0.768</b>
DRB1*08:19	116	54	<b>0.813</b>	0.677	0.63		0.714
DRB1*12:01	117	81	0.798	0.587	0.62		<b>0.832</b>
DRB1*12:02	117	79	<b>0.879</b>	0.66	0.663		0.842
DRB1*14:02	118	78	<b>0.846</b>	0.713	0.672		0.725
DRB1*14:04	30	16	0.679	0.571	0.563		<b>0.683</b>
DRB1*14:12	116	63	<b>0.897</b>	0.797	0.688		0.804
DRB3*03:01	160	70	0.765	0.739	0.729		0.771
Ave			<b>0.8</b>	0.69	0.683		<b>0.763</b>



# Performance of TEPITOPEpan

[Zhang et al., 2012]

The prediction accuracy on 9-11 length ligand and epitope dataset.

Data	Len	Count	NetMHCIIpan	TEPITOPEpan
SYF Ligand	9	10	0.809	<b>0.874</b>
	10	10	<b>0.840</b>	0.714
	11	20	0.730	<b>0.733</b>
IEDB Epitope	9	12	0.644	<b>0.758</b>
	10	54	<b>0.776</b>	0.722
	11	26	<b>0.889</b>	0.849

# Performance of TEPITOPEpan

[Zhang et al., 2012]

TEPITOPEpan has the best performance on binding core prediction.

The number of errors on predicting binding cores of 20 complexes in EpanCore-Set8.

PDB	#complexes	#alleles	NetMHCIIpan-2.0	NetMHCIIpan-1.0	MultiRTA	TEPITOPEpan
Count	20	7	5 errors	3 errors	3 errors	2 error

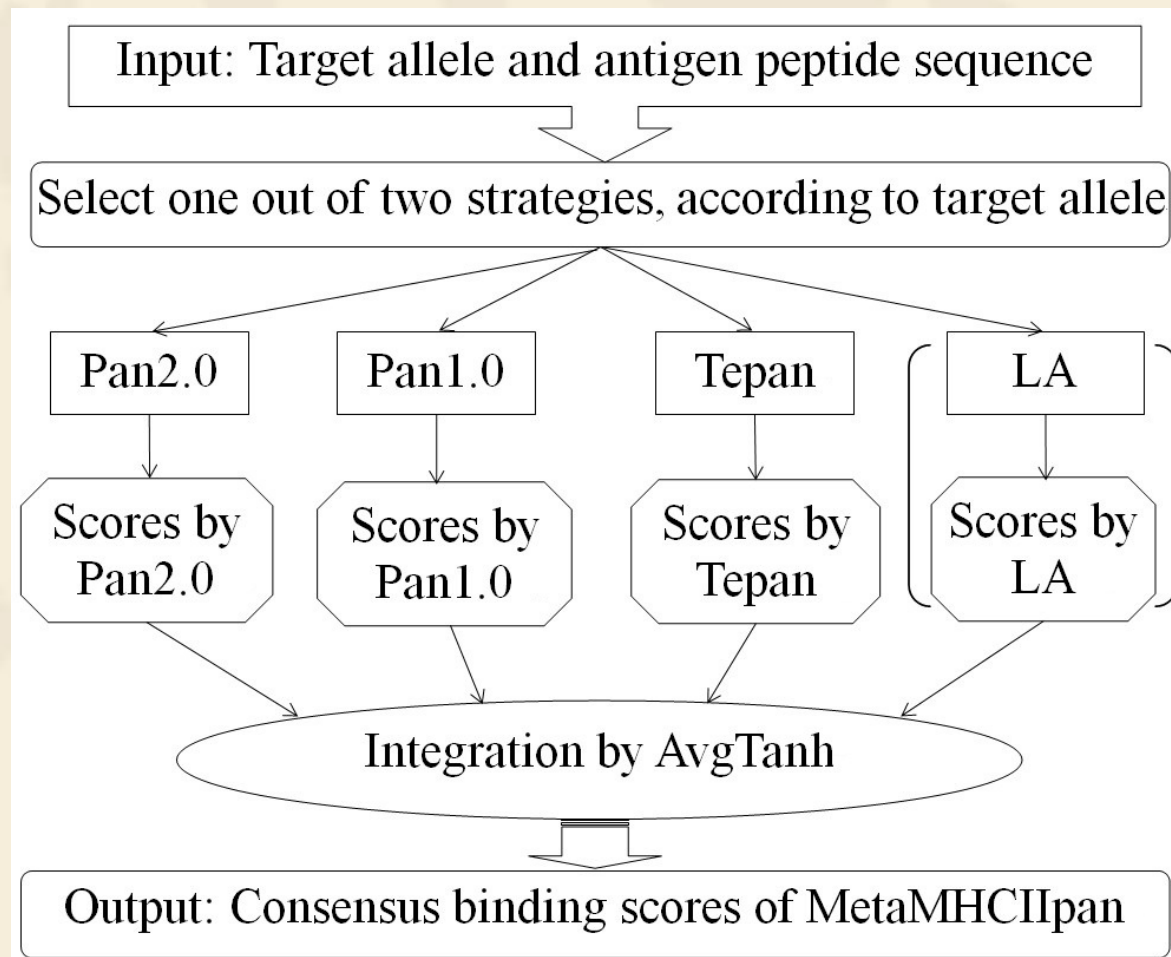
# MetaMHCIIpan

- ❖ MetaMHCIIpan is a consensus method which integrates by averaging the tanh-transformed outputs of multiple pan-specific methods for predicting binding peptides.
- ❖ After checking the performance of various combinations, an optimal one is selected as the default setting of MetaMHCIIpan.
- ❖ Formula used by MetaMHCIIpan:

$$Score_{\tanh} = \frac{1}{2} \left\{ \tanh\left(0.1 \left(\frac{Score - \mu}{\sigma}\right)\right) \right\}$$



# Workflow of MetaMHCIIpan



# Performance of MetaMHCIIpan

AUC comparison of individual method and ensemble method

MetaMHCIIpan

Allele	Count	Bind	Pan2.0	Pan1.0	mRta	Tepan	2.0+1.0+LA+Tepan
DRB1_0101	955	431	0.734	0.718	0.718	0.721	<b>0.738</b>
DRB1_0301	551	229	<b>0.719</b>	0.698	0.681	0.640	0.713
DRB1_0401	1017	484	0.719	0.723	0.693	0.695	<b>0.731</b>
DRB1_0404	402	215	0.825	0.814	0.791	0.754	<b>0.833</b>
DRB1_0405	337	157	0.717	0.746	0.678	0.677	<b>0.765</b>
DRB1_0701	593	285	0.697	0.689	0.692	0.695	<b>0.709</b>
DRB1_0802	163	111	0.799	0.789	0.737	0.763	<b>0.806</b>
DRB1_0901	167	98	0.779	0.715	0.736	0.747	<b>0.781</b>
DRB1_1101	504	232	0.759	0.766	0.714	0.713	<b>0.778</b>
DRB1_1302	238	153	0.754	0.761	0.686	0.691	<b>0.784</b>
DRB1_1501	510	267	0.787	0.720	0.703	0.741	<b>0.793</b>
DRB3_0101	192	48	0.716	0.709	0.714	0.653	<b>0.742</b>
DRB4_0101	217	143	0.724	0.750	0.694	0.580	<b>0.778</b>
DRB5_0101	224	158	0.842	0.829	0.754	0.762	<b>0.854</b>
DRB1_0402:	151	47	0.773	0.776	0.770	0.780	<b>0.793</b>
DRB1_0403:	115	43	0.792	0.787	0.695	0.763	<b>0.807</b>
DRB1_1201:	653	432	<b>0.832</b>	0.783	0.719	0.717	0.820
DRB1_1301:	99	28	0.597	0.657	<b>0.712</b>	0.641	0.642
DRB3_0202:	640	307	<b>0.815</b>	0.734	0.682	0.685	0.782
Ave			0.757	0.746	0.714	0.706	<b>0.771</b>

# MetaMHCIIpan Web Server

<http://www.biokdd.fudan.edu.cn/Service/MetaMHCIIpan/server.html>

## MetaMHCIIpan A pan-specific consensus server for HLA-DR binding predictions

[About](#) [Server](#) [Help](#) [Links](#) [Contact](#)

**Enter Protein Sequence(s)**

Or upload from a file   [clear inputs](#)

**Choose Input Format**

Auto detect format

**Individual Predictors**

TEPITOPEpan  MultiRTA  NetMHCIIpan-1.0  NetMHCIIpan-2.0

LA-Kernel  NN-align

**Meta-predictor**

AvgTanh

**MHC Source Species**

HLA-DR

**MHC-II Allele**

HLA-DRB1\*01:01

**Input Your MHC sequence**

Enable to input one complete MHC protein sequence in fasta format

Or upload from a file

**Maximum Peptide Length**

15-mer

**Output Interface**

Webpage



## REFERENCE

- Lianming Zhang, Yiqing Chen, Hau-san Wong, Shuigeng Zhou, Hiroshi Mamitsuka, Shanfeng Zhu. (2012) *TEPITOPEpan: Extending TEPITOPE for peptide binding prediction covering over 700 HLA-DR molecules*. **PLoS ONE** 7(2):e30483
- Lianming Zhang, Keiko Udaka, Hiroshi Mamitsuka, Shanfeng Zhu. (2012) *Toward more accurate pan-specific MHC-peptide binding prediction: a review of current methods and tools*. **Briefings in Bioinformatics** 13(3): 350-364
- Xihao Hu, Wenjian Zhou, Keiko Udaka, Hiroshi Mamitsuka, Shanfeng Zhu (2010) *MetaMHC: a meta approach to predict peptides binding to MHC molecules* **Nucleic Acids Research** 38: W474-W479
- Xihao Hu, Hiroshi Mamitsuka, Shanfeng Zhu . (2011) *Ensemble approaches for improving HLA Class I-peptide binding prediction*. **Journal of Immunological Methods**. 374: 47-52
- Shanfeng Zhu et al. (2006) *Improving MHC binding peptide prediction by incorporating binding data of auxiliary MHC molecules*. **Bioinformatics** 22(13): 1648-1655

# Acknowledgement

- ❖ Lianming Zhang @ Fudan University
- ❖ Linyuan Guo @Fudan University
- ❖ Prof. Hiroshi Mamitsuka @Kyoto University