



POTN: A Human Leukocyte Antigen-A2 Immunogenic Peptides Screening Model and Its Applications in Tumor Antigens Prediction

Qingqing Meng^{1†}, Yahong Wu^{1†}, Xinghua Sui², Jingjie Meng¹, Tingting Wang¹, Yan Lin¹, Zhiwei Wang¹, Xiuman Zhou¹, Yuanming Qi¹, Jiangfeng Du^{1*} and Yanfeng Gao^{1,2*}

OPEN ACCESS

¹ School of Life Sciences, Zhengzhou University, Zhengzhou, China, ² School of Pharmaceutical Sciences (Shenzhen), Sun Yat-sen University, Shenzhen, China

Edited by:

Yoshihiko Hirohashi,
Sapporo Medical University, Japan

Reviewed by:

Tetsuya Nakatsura,
National Cancer Centre, Japan
Eliana Ruggiero,
San Raffaele Hospital (IRCCS), Italy
Terufumi Kubo,
Sapporo Medical University, Japan

*Correspondence:

Jiangfeng Du
jiangfengdu@zzu.edu.cn
Yanfeng Gao
gaoyf29@mail.sysu.edu.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Cancer Immunity
and Immunotherapy,
a section of the journal
Frontiers in Immunology

Received: 12 March 2020

Accepted: 11 August 2020

Published: 07 October 2020

Citation:

Meng Q, Wu Y, Sui X, Meng J,
Wang T, Lin Y, Wang Z, Zhou X, Qi Y,
Du J and Gao Y (2020) POTN: A
Human Leukocyte Antigen-A2
Immunogenic Peptides Screening
Model and Its Applications in Tumor
Antigens Prediction.
Front. Immunol. 11:02193.
doi: 10.3389/fimmu.2020.02193

Whole genome/exome sequencing data for tumors are now abundant, and many tumor antigens, especially mutant antigens (neoantigens), have been identified for cancer immunotherapy. However, only a small fraction of the peptides from these antigens induce cytotoxic T cell responses. Therefore, efficient methods to identify these antigenic peptides are crucial. The current models of major histocompatibility complex (MHC) binding and antigenic prediction are still inaccurate. In this study, 360 9-mer peptides with verified immunological activity were selected to construct a prediction of tumor neoantigen (POTN) model, an immunogenic prediction model specifically for the human leukocyte antigen-A2 allele. Based on the physicochemical properties of amino acids, such as the residue propensity, hydrophobicity, and organic solvent/water, we found that the predictive capability of POTN is superior to that of the prediction programs SYPEITHI, IEDB, and NetMHCpan 4.0. We used POTN to screen peptides for the cancer-testis antigen located on the X chromosome, and we identified several peptides that may trigger immunogenicity. We synthesized and measured the binding affinity and immunogenicity of these peptides and found that the accuracy of POTN is higher than that of NetMHCpan 4.0. Identifying the properties related to the T cell response or immunogenicity paves the way to understanding the MHC/peptide/T cell receptor complex. In conclusion, POTN is an efficient prediction model for screening high-affinity immunogenic peptides from tumor antigens, and thus provides useful information for developing cancer immunotherapy.

Keywords: neoantigen prediction, peptides, immunogenicity, prediction model, cancer immunotherapy

INTRODUCTION

Cancer immunotherapy has achieved great success in several cancer types (1–3), although durable clinical responses only occur in some patients. Evidence from patients who responded to immunotherapy suggests that tumor regression is achieved by activating tumor-antigen-specific CD8⁺ cytotoxic T lymphocytes (CTLs) (4–7). Tumor antigens are generated by tumor-specific

proteins (8) and presented by the formation of peptide/major histocompatibility complex (MHC)-I complexes on cell surfaces *via* antigen presentation (9).

Generally, tumor antigens can be classified as tumor-specific antigens, including neoantigens, and as tumor-associated antigens. Neoantigens are exclusively presented on tumor cell surfaces, whereas tumor-associated antigens are highly expressed on tumor cells but are also expressed on normal cells at a low level. Using patients' specific neoantigens as tumor vaccines is a safe, feasible approach to eliciting a clinical T cell response (4). However, studies on a large-scale peptide collection found that only about 1% of the peptides can bind MHC-I molecules (10), and less than 0.3% of the peptides should be validated experimentally for immunogenicity (11). We still lack knowledge about the key features of immunogenic peptides and efficient methods to screen tumor antigen peptides from a large number of tumor mutations in personalized immunotherapy.

Tumor antigens can be identified by several approaches. Screening tumoral cDNA libraries with phage display is a powerful but labor-intensive approach to identifying tumor-associated antigens (12–14). Exome sequencing of tumor biopsy and paired normal tissues have been widely applied to screening the mutated fragments (15, 16). The fragments can be synthesized experimentally and tested further for their antigen presentation by measuring the MHC binding affinity, and for their immunogenicity *via* ELISpot, intracellular cytokine staining (ICS), and human leukocyte antigen (HLA) tetramers (15). Another approach to identifying tumor antigens is based on mass spectrometry, which identifies the sequence of peptides presented on the tumor cell surface by MHC molecules (17–19).

Reliable predictions of antigenic peptides from high-throughput sequencing data can lighten the experimental burden for identifying epitopes. *In silico* prediction programs have been developed for this purpose. For example, NetChop and ProteaSMM analyze the proteasomal cleavage pattern and the antigen processing mechanism (20–22), while NetMHCpan 4.0 and other programs predict epitopes by calculating the binding affinity of peptide/MHC allele complexes (23, 24). Other programs use a combined algorithm that integrates proteasomal cleavage prediction, the transporter associated with antigen processing (TAP) transport efficiency, and MHC binding affinity (25). These programs focus on binding capacity prediction, TAP transport prediction, and proteasomal cleavage prediction. We have used these prediction programs to identify epitopes and we found that for HLA-A2 epitopes, fewer than 20% of the predicted epitopes could induce T cell responses (26–28). Thus, the prediction accuracy of the available software packages still needs to be improved.

There are two main reasons for the limited prediction accuracy of current epitope identification programs. First, most of the programs were developed based on a pan-specific method, which does not differentiate between HLA alleles, and they are widely used to make predictions for various HLA alleles. Therefore, when they are used to identify the antigenic peptides for a particular MHC allele, the accuracy is lower because of their inherent features (29). Second, the datasets used to construct the prediction models in many programs are impure. Non-immunogenic peptides in many

datasets are randomly selected and are not experimentally validated, resulting in high false-negative rates. To avoid such shortcomings, we gathered experimental data and built a prediction model for only the most common HLA allele (30). About 5200 HLA-A alleles have been identified, among which HLA-A2 shows a high occurrence; the proportion of people with the HLA-A2 allele is 54.0% in ethnic Chinese people and 43.1% of the general population (30–32).

In this study, we selected 9-mer peptides (nonamers) with verified immunological activity and used a support vector machine (SVM) to construct the POTN prediction model for the HLA-A2 allele based on the physicochemical properties of amino acids. We validated the model by using external data. We used the POTN model to predict immunogenic peptides from the cancer-testis antigen located on the X chromosome (CT-X) and measured the binding affinity and immunological activity by ICS of the predicted peptides. We compared the prediction accuracy of POTN with that of other widely used prediction software. Our model may provide a new method to screen high-affinity immunogenic peptides from amino acid sequences or whole-exome sequencing data efficiently.

MATERIALS AND METHODS

Peptide Data Collection

The immunogenic peptides were retrieved from the databases IEDB (33, 34), SYFPEITHI (35), and Peptide Database (36). To ensure that the dataset was not biased, peptides matching our selection criteria were randomly selected from the databases. From the IEDB database, we obtained 41 HLA-A2 cancer-associated immunogenic peptides using our initial screening criteria for the MHC-I linear epitope. From the SYFPEITHI database, 41 T cell epitopes were obtained by searching for HLA-A2 cancer-associated peptides that did not overlap with the peptides obtained from IEDB. The Peptide Database contains human tumor antigen peptides categorized as mutation, tumor-specific, differentiation, and overexpressed. We selected 64 unique peptides by excluding peptides that overlapped with the peptides from the other two databases. The peptides used as a negative dataset were screened from the IEDB database and the literature, and 214 peptides that were experimentally validated as non-immunogenic peptides were obtained (Table S1).

The final dataset consisted of a total of 360 HLA-A2 peptides, including 146 immunogenic peptides and 214 non-immunogenic peptides (Table 1). For the total dataset, 60% of the immunogenic peptides and 60% of the non-immunogenic peptides were selected as the training set, and the remaining 40% of the peptides were used as the test set (Table S1), where approximately 6% of the dataset were eluted peptides.

Selection and Calculation of Potential Immunogenic Properties

To obtain the most useful properties, we searched the literature to find features that may be relevant to immunogenicity. The accessible surface area (ASA) has been used to understand

TABLE 1 | Data collection for the model construction and evaluation.

Resources	T cell response		Total (n = 360)
	Yes (n = 146)	No (n = 214)	
IEDB	41	16	57
Peptide database	64	0	64
SYFPEITHI	41	0	41
Literatures	0	198	198

various biological problems, such as protein-protein interactions (37, 38), structural epitopes (39), and active sites (40), and it was used as a feature to build the model. The polarity and charge of amino acids in a peptide are highly correlated with binding affinity (41, 42), and thus these features were used in model construction. In addition, physicochemical properties, including isotropic surface area (ISA), electronic charge index (ECI), hydrophobicity, entropy, molecular weight (Mw), aromatic residues, organic solvent/water, and isoelectric point (PI), have been studied (7, 43–50). The physicochemical properties of 20 amino acids were obtained from the amino acid index database (51).

The properties for binding, protein cleavage, and TAP transport efficiency of each peptide were calculated by online server NetCTL 1.2 with default parameters (52). The T cell recognition score and the stability of the peptide/MHC complexes were considered (48, 53).

Because some residues tend to be in specific positions in the immunogenic peptides (54), we calculated the residue propensity, which is defined as the probability of an amino acid being at an individual position of a peptide, as

$$\text{Residue propensity (R}_i\text{P}_j) = \frac{P_j \cdot 100 - N_j \cdot 100}{P_j \cdot 100 + N_j \cdot 100}$$

where P_j is the frequency of residue i at position j for immunogenic peptides and N_j is the frequency of residue i at position j for non-immunogenic peptides.

To understand the discriminative power of predictors better, we calculated the statistical significance (p -values) of each predictor for immunogenic peptides versus non-immunogenic peptides in the training set using Student's t -test. Only predictors with significant differences ($p < 0.05$) between immunogenic and non-immunogenic peptides were included in the final model (Table 2).

Construction of the Immunogenic Prediction Model

SVM is a supervised learning model based on the principles of structure risk minimization and the kernel method (55), and it has been widely used to predict T cell epitopes (56). Here, SVM with a radial basis (Gaussian) kernel was used to construct the POTN model based on the selected immunogenicity predictors. The regularization parameter (C), which controls the trade-off between the margin and the training error, was tested for model construction and optimization. In optimizing the model construction, several C values ($C \in \{0.25, 0.50, 1, 2, 4\}$) were used to construct the model, and the values were validated by the leave-one-out approach in R (version 3.5.2).

TABLE 2 | Selected features for model construction. The selected features were highly correlated with immunogenicity (indicated by p -value).

Features	References	Position	Description	p -value
ASA	(39)	P3	Accessible surface area	0.026
Charged value	(42)	P3	Net charge	0.039
ECI	(44)	P3	Electronic charge index	0.009
Entropy	(46)	P3	Entropy of formation	0.001
Hydrophobicity	(45)	P3	Modified Kyte-Doolittle hydrophobicity scale, more hydrophobic residues are preferable to be at P4, P7, and P8.	2.35E-05
ISA	(44)	P3	Isotropic surface area	1.62E-06
Mw	(47)	P3	Molecular weight	0.042
Organic solvent/water	(50)	P3	Transfer energy, organic solvent/water	4.58E-05
Organic solvent/water	(50)	P4	Transfer energy, organic solvent/water	0.019
PI	(49)	P5	Isoelectric point	0.028
Polarity	(42)	P3	Polarity	0.002
Residue propensity	(54)	P1	Score based on frequency assigned of each amino acid (see Figure 1)	0.000
Residue propensity	(54)	P2		0.003
Residue propensity	(54)	P3		2.11E-14
Residue propensity	(54)	P4		4.73E-06
Residue propensity	(54)	P5		4.10E-07
Residue propensity	(54)	P6		8.73E-05
Residue propensity	(54)	P7		2.57E-06
Residue propensity	(54)	P8		1.41E-06
Residue propensity	(54)	P9		0.000
Residue propensity	(54)	sum		2.84E-41
Aff	(7, 52)		binding affinity	7.64E-08
Aff_rescale	(52)		Rescale binding affinity	7.63E-08
Cle	(7)		C terminal cleavage affinity	0.003
Combined score	(7)		Combined prediction score	3.04E-08
Pred	(53)		pMHC stability score	2.42E-08
Thalf	(53)		pMHC stability score	0.001
NB	(53)		pMHC stability score	1.46E-09

Peptide Prediction and Synthesis

Candidate peptides from CT-X were predicted using the POTN model and 34 peptides with the highest scores were selected, of which 22 peptides with satisfactory solubility were synthesized by the standard solid-phase Fmoc strategy (57) and purified by reverse phase high-performance liquid chromatography (58). All synthesized peptides had a purity of >95%, as measured by electrospray ionization mass spectrometry.

Binding Affinity Measurement

The T2 binding assay was used to determine the binding affinity of the candidate peptides and HLA-A2 molecule by using a previously described protocol (27). The T2 cell line (HLA-A2) was supplied by Professor Yuzhang Wu (Third Military Medical University, Chongqing, China). In brief, T2 cells (500 μ L, 1×10^6 cells/mL) were incubated with the peptide (25 μ g, 50 μ g/mL; dissolved in DMSO at a concentration of 10 mg/mL) in serum-free IMDM medium, supplemented with human β 2-microglobulin (3 μ g/mL, Merck, USA) at 37°C for 18 h. The T2 cells were washed twice and incubated with the anti-human HLA-A2-PE-cy7 antibody (BB7.2, eBioscience, USA) at 4°C for 30 min. The mean fluorescence intensity (FI) of each group was analyzed by flow cytometry (FACSCalibur, Becton-Dickinson, USA). Based on the FI, the binding affinity of the candidate peptides toward HLA-A2 molecule was calculated by

$$FI = \frac{a - b}{b}$$

where *a* is the mean PE-cy7 FI with the peptide and *b* is the mean PE-cy7 FI without the peptide.

ICS Assay for Immunogenicity

We determined whether the high-binding affinity peptides elicited a T cell response in peripheral blood samples from five HLA-A2⁺ healthy donors. The blood samples were obtained from Henan Red Cross Blood Center (Zhengzhou, China) with the approval of the Institutional Ethics Review Board. All research was performed under the approval of the Ethics Committee of Zhengzhou University. An ICS assay was used to quantify IFN- γ production of CD3⁺CD8⁺ T cells. Peripheral blood mononuclear cells (PBMCs) were stimulated by each peptide (10 μ g/mL) once-weekly for 3 weeks according to our previous work (59). On day 21, the induced T cells from the PBMCs were used as effector cells, and T2 cells were incubated with the synthesized peptides (50 μ g/mL) for 4 h as the stimulator cells. The effector cells (1×10^6) and stimulator cells (1×10^6) were co-incubated for 3 h, and brefeldin A (2 μ g/mL, Sigma-Aldrich, USA) was added to block the release of produced cytokines for another 5 h at 37°C and 5% CO₂. The cells were washed and stained with eFlour 710 labeled anti-human CD3 antibody and APC-labeled anti-human CD8 antibody (eBioscience) for 30 min at 4°C before fixation and permeabilization. Permeabilized cells were intracellularly stained with the PE-labeled anti-human IFN- γ antibody (BioLegend, Inc., USA) for 30 min on ice in the dark. Cells were resuspended in buffer for acquisition and analysis using a flow cytometer (FACSCalibur, Becton Dickinson).

RESULTS

Identification of Features and Key Residues for Immunogenic Peptides

Feature selection is a crucial step in model construction. To avoid overlaid features and decrease the less-valuable features in the model, we selected properties that have been linked to

immunogenicity. We found that 28 features were significantly different between the immunogenic and non-immunogenic groups of peptides (Table 2). Aromatic amino acids were not significantly different at either a single position or a sum of points, and TAP also made no significant difference in our dataset.

By statistically analyzing the differences in the residual properties for each position, we found that many physicochemical properties are significantly different at position 3 (P3) between the immunogenic peptides and the non-immunogenic peptides, which has not been reported before (Table 2) (60). Thus, we hypothesized that the residues at P3 should be small and flexible, which may contribute to the binding of P4–P7 to the MHC/peptide/T cell receptor complex (61). To test our hypothesis, we screened for pairs of peptides with only one amino acid different at P3, where one peptide was immunogenic and the other was non-immunogenic. We found the peptides QLCDVMFYI (immunogenic)/QLRDVMFYI (non-immunogenic), EVKEKHEFL (immunogenic)/EVREKHEFL (non-immunogenic), and GLCTLVAML (immunogenic)/GLLTLVAML (non-immunogenic) in the literature (62–66). Compared with non-immunogenic peptides, the third amino acid of the immunogenic peptide is smaller than that in non-immunogenic peptides. The evidence of the peptide pairs appeared to support our hypothesis, and we proposed that the physicochemical properties at P3 could also determine the immunogenicity of a peptide.

To investigate the amino acid preferences of the individual position between immunogenic and non-immunogenic peptides further, we compared the frequency of the amino acid at each position (Figure 1). Both immunogenic and non-immunogenic peptides had conserved residues, with leucine conserved at P2 and leucine and valine conserved at P9. P3, P4, and P6 had slight differences between immunogenic and non-immunogenic peptides. Based on this finding, the residue propensity value for each amino acid at a specific position was calculated and used as a feature for model construction.

POTN Construction and Immunogenicity Prediction

The overall workflow for model construction is shown in Figure 2. cDNA, RNA, and amino acids can be processed by POTN, which can split the sequences into nonamers. The model analyzes the properties and calculates the predicted scores, which are used to predict the immunogenicity of peptides. The R implementation of POTN is available in supplementary materials.

In order to construct a high-quality model, the cost parameter *C* (*C* value) was continually adjusted until the optimal output was reached by leave-one-out cross-validation experiment, where the *C* value was set to 1 and the optimal model was called POTN. POTN showed a high prediction power in both the training set and the test set. For the training set, the area under the curve (AUC) was 0.773 and the accuracy (ACC) was 0.653 (Figure 3A). For the test set, the AUC was 0.748 and ACC was 0.701.

To illustrate the predictive power of the POTN model further, we compared the predictive power with the prediction programs SYFPEITHI, IEDB, and NetMHCpan 4.0 (Figure 3B). The performance of POTN was better than that of the other models

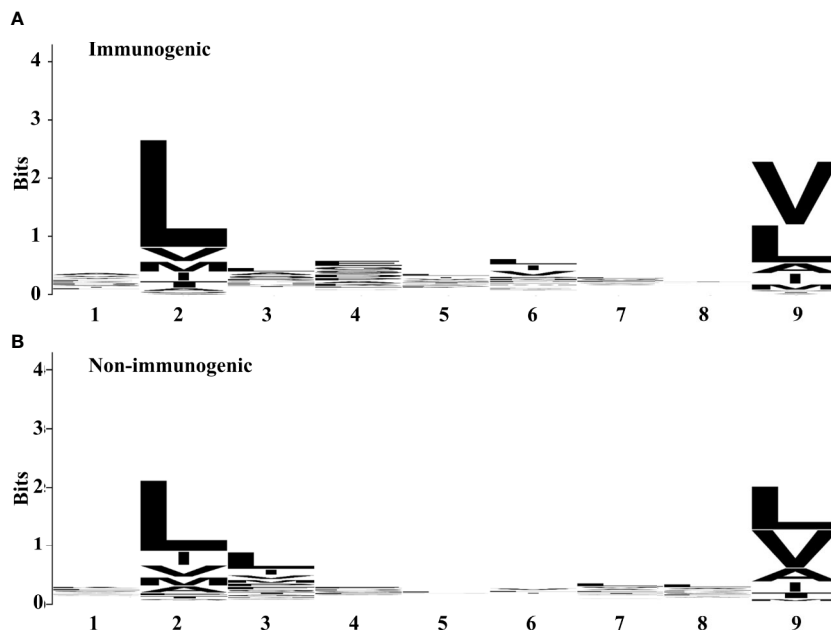


FIGURE 1 | Residue propensity between (A) immunogenic and (B) non-immunogenic peptides from the training set. The height of amino acid letters within a column indicates the relative frequency of each amino acid at the given position. The overall height of the column indicates the residual conservation at the position. (A) and (B) were generated by using WebLogo (67).

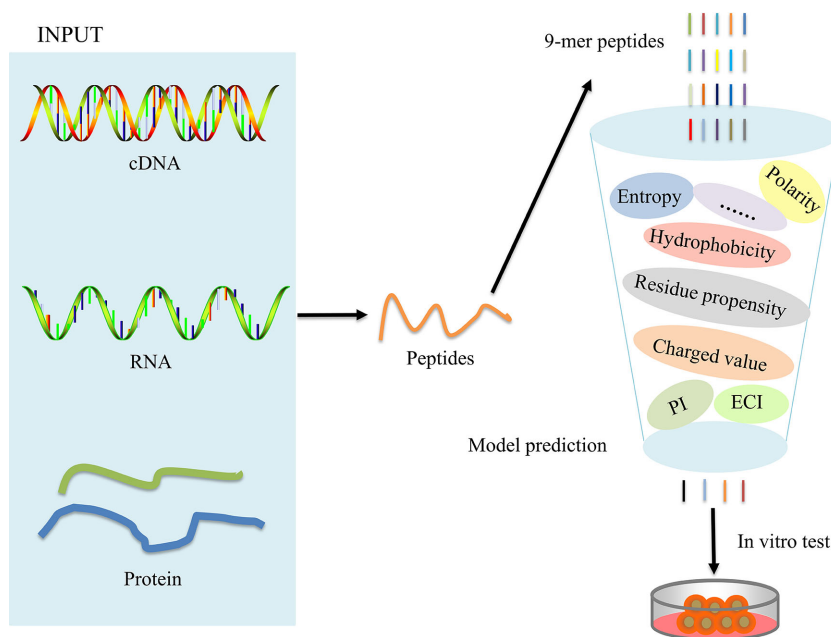


FIGURE 2 | Overall working flowchart for the POTN model construction.

with the test set (Figures 3B, C). Receiver operating characteristic curves (ROC) based on the four models were plotted. The AUC in the whole test set were 0.748, 0.635, 0.689, and 0.720 for POTN, SYFPEITHI, IEDB, and NetMHCpan 4.0, respectively. The ACC in

the whole test set were 0.701 and 0.653 for POTN and NetMHCpan 4.0, respectively. The AUC were also analyzed at different false-positive rates (FPR) (Figure 3D). The AUC was 0.01 at an FPR of 0.05 for POTN, which showed the best performance of the

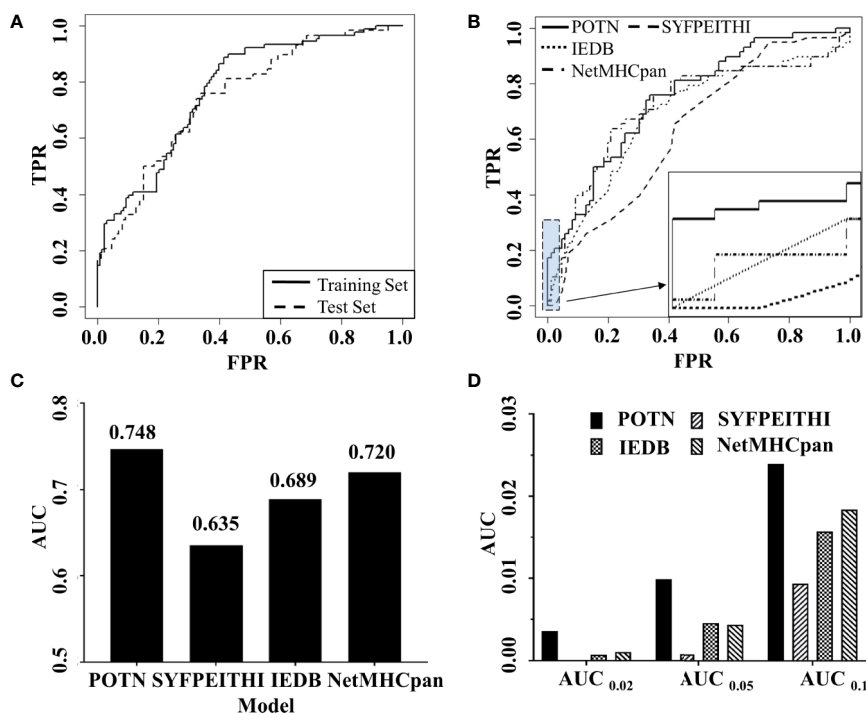


FIGURE 3 | Comparison of the performance of POTN with other programs. **(A)** ROC curves generated by the POTN model with the training set ($n = 216$) and test set ($n = 144$). The black solid line shows the ROC curve for the training set. The short-dashed line shows the ROC curve for the test set. **(B)** ROC curves generated by the POTN (black solid line), SYFPEITHI (short-dashed line), IEDB (dotted line), and NetMHCpan 4.0 (dashed-dotted lines) models with the test set. **(C)** AUC generated by the four models with the test set. **(D)** AUC at different FPR.

prediction models. In addition, we also compared the precision indicator, which was calculated by the ratio of the true positive to the predicted positive peptides. In the test set, the precision indicator of NetMHCpan 4.0 was 54.55%, while the precision indicator of POTN was 67.44%, with 23.63% improvement [the method for improvement rate calculation was referred to (68)].

Application to CT-X Antigen Dataset

We applied the POTN model to a dataset of CT-X antigens, which are tumor antigens overexpressed in the testis and other malignancies, as an antigen resource to screen epitope candidates. The amino acid sequences of these antigens were cleaved into nonamers, and POTN obtained a total of 17,310 nonamers from more than 50 antigens after excluding duplicates (Table S2) (Figure 4). The immunogenic value of each peptide was predicted by POTN, and the top 0.2%, consisting of 34 peptides, was selected based on the predicted values. The solubilities of these 34 peptides were predicted using the MOE package, and 22 of 34 peptides were selected as being sufficiently soluble (Table 3).

The 22 peptides were synthesized to test the activity. The binding affinity of the synthesized peptides to HLA-A2 was measured *via* a binding assay (FI) with the T2 cell line (27). Based on the FI, the peptides were clustered into three groups: weak binding affinity ($FI < 0.5$), moderate binding affinity ($FI \geq 0.5$ to < 1.5), and high binding affinity ($FI \geq 1.5$) (Figure 5A).

Most of the synthesized peptides (59.09%+36.36%) had a moderate or high FI value ($FI \geq 0.5$), of which a large proportion (61.9%) had a high binding affinity and a smaller proportion (38.1%) had moderate binding affinity. Of the 22 synthesized peptides (Table 3), eight peptides had moderate binding affinity ($FI \geq 0.5$ to < 1.5) and 13 peptides had a high binding affinity ($FI \geq 1.5$), which showed that the POTN model had an accuracy rate of 95.45% (21 of 22 synthesized peptides) in predicting the HLA-A2 binding peptides.

Next, we examined the T cell responses of the 13 synthesized peptides with high binding affinity by detecting the percentages of IFN- γ^+ CD8 $^+$ T cells from five HLA-A2 $^+$ healthy donors. A higher percentage of IFN- γ^+ CD8 $^+$ T cells in the total CD8 $^+$ T cell population than that of the negative control indicated an immunogenic peptide. In donor 1, 12 were immunogenic (Figure 5B); in donor 2, eight peptides were immunogenic (Figure 5C); in donor 3, 10 peptides were immunogenic (Figure 5D); in donor 4, six peptides were immunogenic (Figure 5E); and in donor 5, 12 peptides were immunogenic (Figure 5F). These results showed that more than half of the peptides elicited immune responses in at least three donors, whereas peptide KLSSIIPSA only elicited a response in donor 5 (Figures 5F, G). In other words, any of the 13 high-affinity peptides could stimulate a T lymphocyte response in at least one donor (Figures 5G, H).

In addition, we compared the virtual screening performance of the POTN model with that of NetMHCpan 4.0. The

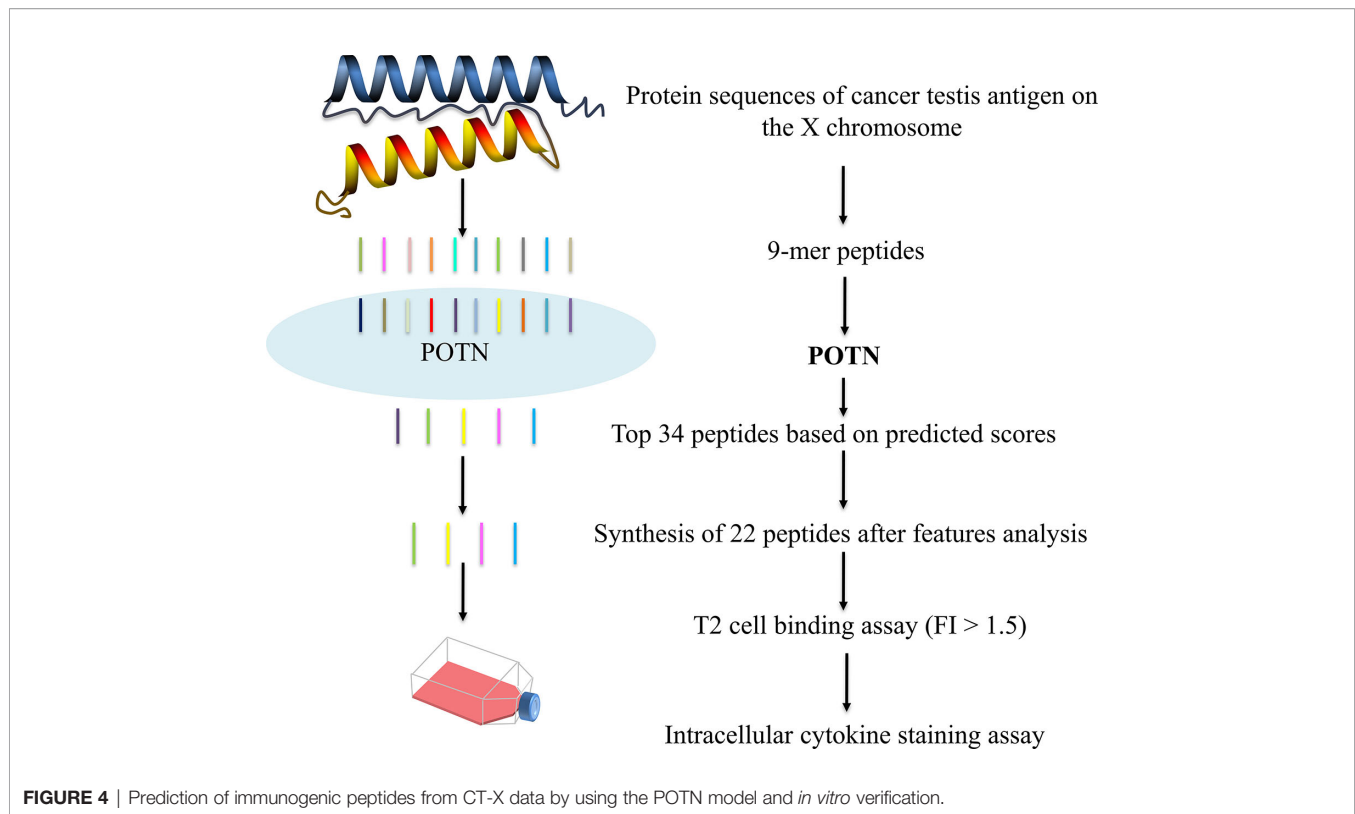


TABLE 3 | Overview of the immunogenicity and HLA-A2 binding affinity of candidate peptides predicted by the POTN model.

Peptide	Prediction score	Binding	Immunogenicity*
KLSSIIPSA	1.1299	++	1/5
FLAKLNNTV	1.1257	+	\
FLSKLSSII	1.1157	-	\
VLSAVTPEL	1.1020	+	\
VLSNVLSGL	1.1010	+	\
SIDDLSFYV	1.0988	++	4/5
ILDRANQSV	1.0906	++	3/5
YLATADMPA	1.0898	++	3/5
ALDEKVAEL	1.0847	++	4/5
ALSTVLPGL	1.0832	++	2/5
TLDEKVAEL	1.0777	++	5/5
TLDQVLDEV	1.0680	++	5/5
AMASASPSV	1.0663	++	3/5
VLSTAPPQL	1.0654	++	4/5
KVADLIHFL	1.0653	+	\
KVAELVHFL	1.0627	+	\
LMDVQIPTA	1.0546	++	5/5
ALSVMGVYV	1.0541	+	\
FLAMLKNTV	1.0498	+	\
KVAKLVHFL	1.0493	+	\
KMAGELIKI	1.0386	++	4/5
FIDKLVESV	1.0359	++	5/5

Prediction scores were ranked and retained with four decimal digits. -, $FI < 0.5$; +, $FI \geq 0.5$ to < 1.5 ; ++, $FI \geq 1.5$; \, no experimental data. *The response ratio of each peptide in five donors detected by the percentages of IFN- γ -secreting CD8⁺ T cells.

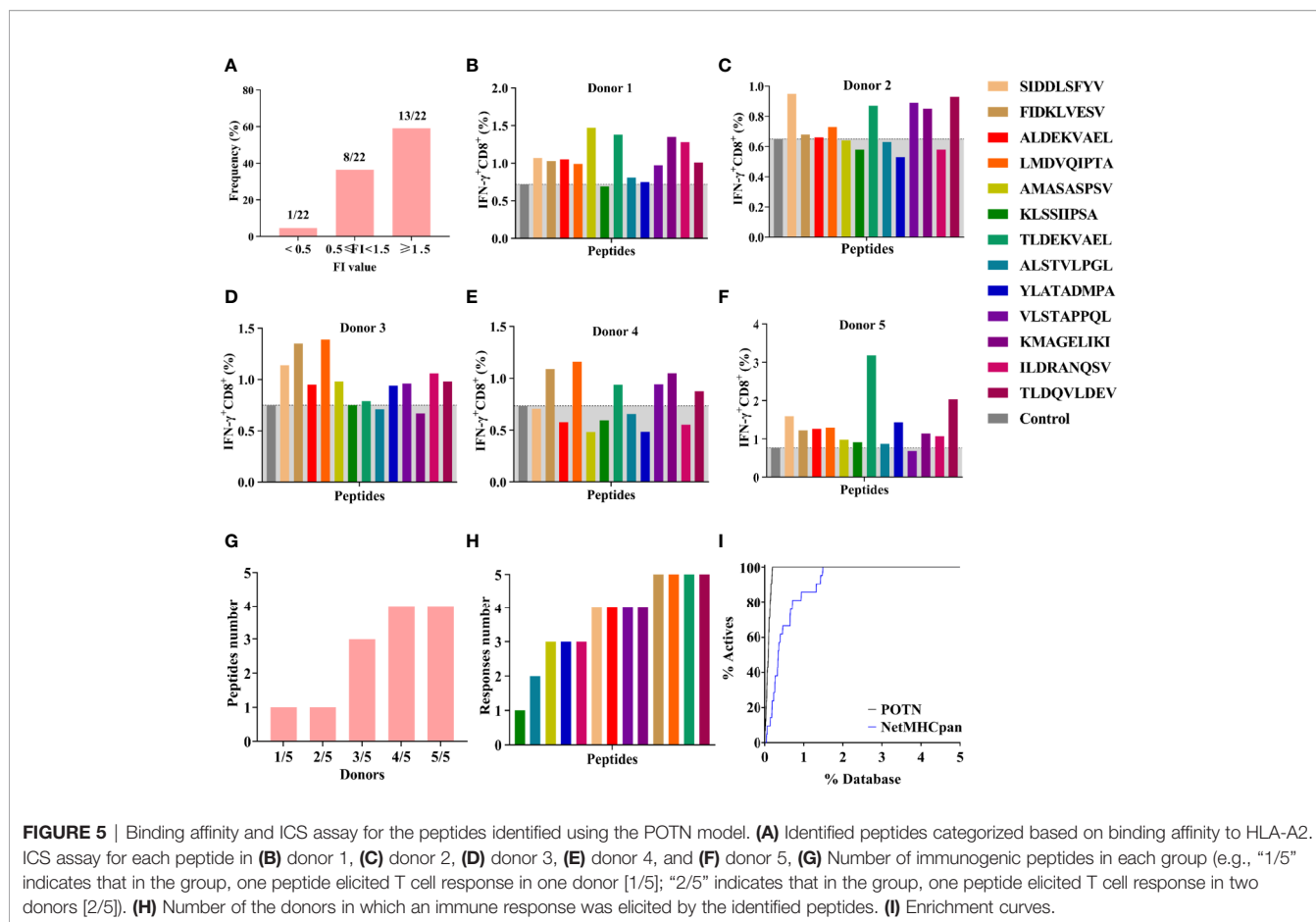
enrichment curves of the two models showed that both programs efficiently distinguished the immunogenic peptides from the database (**Figure 5I**). All immunogenic peptides were

identified in the top 1% of the database by using POTN, and they were identified in the top 2% of the database by using NetMHCpan 4.0. The results indicated that the screening performance of POTN was two-fold better than that of NetMHCpan 4.0.

DISCUSSION

Cancer immunotherapy has achieved great clinical success, and many studies have shown that the clinical effect depends on the presence of tumor-specific T lymphocytes in patients (69). The tumor-specific T lymphocytes kill tumor cells by secreting cytokines, releasing granzymes, and producing perforin when the MHC-bound peptide is recognized by CTLs. With the development of next-generation sequencing technologies, tumor antigens from cancer patients can be identified easily by sequencing the cancer biopsy. These proteins can be fragmented into numerous peptide sequences, some of which can be presented by the MHC molecule and trigger a specific T cell response targeting the peptide-expressing tumor cells. However, efficiently identifying the MHC binding and immunogenic peptides from the huge amount of sequencing data remains a challenge.

Current programs used for either MHC binding or antigenic prediction are still inaccurate. Possible reasons include the lack of experimental data for many HLA alleles, the non-immunogenic peptides selected for model building include false negatives, and the use of pan-specific methods. To overcome these problems,



we designed the POTN model to predict T cell response of peptides to HLA-A2, a common allele of MHC-I.

For current programs, the negative data sets selected for many predictive models are random peptides, which allow some potentially immunogenic peptides to be classified as non-immunogenic. To construct a model with a better predictive effect, 360 nonamers verified by *in vitro* immunological activity experiments were used to construct the POTN model. We selected non-immunogenic peptides with experimental data as our negative data set. These peptides have binding affinity but are not immunogenic, and they have properties that are more similar to the immunogenic peptides. Thus, we chose these peptides as our dataset to identify properties that are directly related to immunogenicity and build a better model.

We collected 216 peptides as the training set and 144 peptides as the test set for the model. To effectively distinguish the MHC binding nonamers from the sequence database, we used all of the peptide features to construct a predictive model. Statistically significant features were selected for model construction. Because the peptides had nine amino acids, these features were further decomposed into 28 descriptors for each peptide (**Table 2**). The relationships between the peptide features and immunogenicity indicated that many features were statistically different at P3, and that P3 may be an important position for distinguishing immunogenicity (**Figure S1**). This result was

unsurprising, because the amino acids that came into contact with the MHC/peptide/T cell receptor complex in the nonamers were typically at P4–P7. The features of the amino acid at P3, which is adjacent to sites P4–P7, may indeed be a factor affecting immunogenicity.

The performance of the POTN model was superior to that of the other widely used prediction programs, IEDB, SYFPEITHI, and NetMHCpan 4.0 (**Figure 3B**). The high true positive rate and low false negative rate of the POTN model indicated that it could accurately predict epitopes from a peptide sequence database, which may facilitate the development of personalized cancer immunotherapy based on exome sequencing. The performance of the POTN model proved that the properties of peptides, such as polarity, charges, and entropy, give useful information about how likely it is that a peptide is an epitope, which indicates a new direction for software development.

Antigen presentation is crucial to the function of the adaptive immune response, where the HLA molecule presents the antigenic peptides (epitopes) to T cells and stimulates their proliferation and activation. HLA-A2 is a common allele in humans. Therefore, a prediction model that can specifically identify HLA-A2 epitopes is useful for cancer vaccine development. Our model is designed for this purpose and only predicts epitopes for HLA-A2 (30, 70). Therefore, the

current version of the POTN model is restricted to predicting HLA-A2-bound peptides. However, other MHC allele-specific prediction models could be built with the same approach if the experimental binding data for the allele are provided. In addition, only nonamers were evaluated using the model, so the prediction power of the model for peptides with other lengths is not clear. In addition, we wonder about the performance of the POTN system for the peptides from thymic selection. The mechanism of central immune tolerance allows immature T cells of the central immune organ to develop immune tolerance when exposed to self-antigens, and therefore the tolerated self-peptide after thymic selection should not have the characteristics as that from immunogenic peptides, and they can theoretically be excluded by the POTN system. To test the performance of the POTN system for self-peptides, we deliberately selected two self-proteins for study, and the prediction results showed that POTN predicted several self-peptides as immunogenic peptides, although the false positive rates were extremely low (0.36% and 1.7%, separately). The results indicated that the POTN system cannot absolutely exclude self-peptides from immunogenic peptides and the input data for POTN system is suggested to the mutated sequencing data.

Finally, we selected peripheral blood samples from five healthy donors to test the high-affinity HLA-A2 binding peptides, and at least half of the peptides elicited a T cell response in three or more donors. The results showed that anti-tumor immunity could be activated by these peptides in cancer patients, which should be investigated further in an *in vivo* study of tumor treatment with the identified peptides.

CONCLUSION

The easy acquisition of personalized exome sequencing data from cancer patients requires a tool for identifying epitopes with high prediction power. In this study, we developed the POTN model to predict the immunogenicity of HLA-A2 peptides, and our model showed superior performance compared with the most commonly used programs, SYPEITHI, IEDB, and NetMHCpan 4.0. POTN may help to identify tumor neoepitopes efficiently from sequencing data, and the approach behind the model may provide a method for constructing prediction models for other MHC alleles. We used the POTN model to identify several epitopes from the CT-X database and four of the peptides elicited a T cell response in all five healthy donors. These peptides could serve as starting points for developing new cancer treatments.

REFERENCES

- Doran SL, Stevanovic S, Adhikary S, Gartner JJ, Jia L, Kwong MLM, et al. T-Cell Receptor Gene Therapy for Human Papillomavirus-Associated Epithelial Cancers: A First-in-Human, Phase I/II Study. *J Clin Oncol* (2019) 37(30):2759–68. doi: 10.1200/JCO.18.02424
- Mehta GU, Malekzadeh P, Shelton T, White DE, Butman JA, Yang JC, et al. Outcomes of Adoptive Cell Transfer With Tumor-infiltrating Lymphocytes

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by: Ethics Committee of Zhengzhou University and Henan Red Cross Blood Center with the approval of the Institutional Ethics Review Board. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

YG, YQ, and YW designed the experiments for peptide synthesis, binding assay, and T cell response. JD and QM designed the *in silico* experiments for model construction and data analysis. QM, YW, JM, TW, and YL performed the experiments with critical support from ZW and XZ. XS and QM analyzed the data. QM, JD, and XS wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by National Natural Science Foundation of China (Project No. 31500620, U1604286, 81601448), the Henan Province and the Key Scientific Research Projects of Henan Higher Education Institutions (No. 19A180007, 19A180009).

ACKNOWLEDGMENTS

We thank Professor Yuzhang Wu for providing T2 cell lines.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2020.02193/full#supplementary-material>

for Metastatic Melanoma Patients With and Without Brain Metastases. *J Immunother* (2018) 41(5):241–7. doi: 10.1097/CJI.0000000000000223

- Tran E, Robbins PF, Lu YC, Prickett TD, Gartner JJ, Jia L, et al. T-Cell Transfer Therapy Targeting Mutant KRAS. *N Engl J Med* (2017) 376(7):e11. doi: 10.1056/NEJMc1616637
- Ott PA, Hu Z, Keskin DB, Shukla SA, Sun J, Bozym DJ, et al. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* (2017) 547(7662):217–21. doi: 10.1038/nature22991

5. Sahin U, Derhovanessian E, Miller M, Kloke BP, Simon P, Lower M, et al. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature* (2017) 547(7662):222–6. doi: 10.1038/nature23003
6. van der Burg SH, Arens R, Ossendorp F, van Hall T, Melief CJ. Vaccines for established cancer: overcoming the challenges posed by immune evasion. *Nat Rev Cancer* (2016) 16(4):219–33. doi: 10.1038/nrc.2016.16
7. Capietto AH, Jhunjhunwala S, Delamarre L. Characterizing neoantigens for personalized cancer immunotherapy. *Curr Opin Immunol* (2017) 46:58–65. doi: 10.1016/j.coi.2017.04.007
8. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013
9. Coulie PG, Van den Eynde BJ, van der Bruggen P, Boon T. Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. *Nat Rev Cancer* (2014) 14(2):135–46. doi: 10.1038/nrc3670
10. Paul S, Weiskopf D, Angelo MA, Sidney J, Peters B, Sette A. HLA class I alleles are associated with peptide-binding repertoires of different size, affinity, and immunogenicity. *J Immunol* (2013) 191(12):5831–9. doi: 10.4049/jimmunol.1302101
11. Vitiello A, Zanetti M. Neoantigen prediction and the need for validation. *Nat Biotechnol* (2017) 35(9):815–7. doi: 10.1038/nbt.3932
12. Minenkova O, Pucci A, Pavoni E, De Tomassi A, Fortugno P, Gargano N, et al. Identification of tumor-associated antigens by screening phage-displayed human cDNA libraries with sera from tumor patients. *Int J Cancer* (2003) 106(4):534–44. doi: 10.1002/ijc.11269
13. van der Bruggen P, Traversari C, Chomez P, Lurquin C, De Plaen E, Van den Eynde BJ, et al. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science* (1991) 254(5038):1643–7. doi: 10.1126/science.1840703
14. Ma W, Germeau C, Vigneron N, Maernoudt AS, Morel S, Boon T, et al. Two new tumor-specific antigenic peptides encoded by gene *MAGE-C2* and presented to cytolytic T lymphocytes by HLA-A2. *Int J Cancer* (2004) 109(5):698–702. doi: 10.1002/ijc.20038
15. Gubin MM, Artyomov MN, Mardis ER, Schreiber RD. Tumor neoantigens: building a framework for personalized cancer immunotherapy. *J Clin Invest* (2015) 125(9):3413–21. doi: 10.1172/JCI80008
16. Martin SD, Wick DA, Nielsen JS, Little N, Holt RA, Nelson BH. A library-based screening method identifies neoantigen-reactive T cells in peripheral blood prior to relapse of ovarian cancer. *Oncimmunology* (2017) 7(1):e1371895. doi: 10.1080/2162402X.2017.1371895
17. Schirle M, Keilholz W, Weber B, Gouttefangeas C, Dumrese T, Becker HD, et al. Identification of tumor-associated MHC class I ligands by a novel T cell-independent approach. *Eur J Immunol* (2000) 30(8):2216–25. doi: 10.1002/1521-4141(2000)30:8<2216::AID-IMMU2216>3.0.CO;2-7
18. Freudenmann LK, Marcu A, Stevanovic S. Mapping the tumour human leukocyte antigen (HLA) ligandome by mass spectrometry. *Immunology* (2018) 154(3):331–45. doi: 10.1111/imm.12936
19. Abelin JG, Keskin DB, Sarkizova S, Hartigan CR, Zhang W, Sidney J, et al. Mass Spectrometry Profiling of HLA-Associated Peptidomes in Mono-allelic Cells Enables More Accurate Epitope Prediction. *Immunity* (2017) 46(2):315–26. doi: 10.1016/j.immuni.2017.02.007
20. Calis JJ, Reinink P, Keller C, Kloetzel PM, Kesmir C. Role of peptide processing predictions in T cell epitope identification: contribution of different prediction programs. *Immunogenetics* (2015) 67(2):85–93. doi: 10.1007/s00251-014-0815-0
21. Kesmir C, Nussbaum AK, Schild H, Detours V, Brunak S. Prediction of proteasome cleavage motifs by neural networks. *Protein Eng* (2002) 15(4):287–96. doi: 10.1093/protein/15.4.287
22. Tenzer S, Stoltz L, Schonfisch B, Dengiel J, Muller M, Stevanovic S, et al. Quantitative analysis of prion-protein degradation by constitutive and immunosuppressive proteasomes indicates differences correlated with disease susceptibility. *J Immunol* (2004) 172(2):1083–91. doi: 10.4049/jimmunol.172.2.1083
23. Hoof I, Peters B, Sidney J, Pedersen LE, Sette A, Lund O, et al. NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics* (2009) 61(1):1–13. doi: 10.1007/s00251-008-0341-z
24. Karosiene E, Lundegaard C, Lund O, Nielsen M. NetMHCcons: a consensus method for the major histocompatibility complex class I predictions. *Immunogenetics* (2012) 64(3):177–86. doi: 10.1007/s00251-011-0579-8
25. Stranzl T, Larsen MV, Lundegaard C, Nielsen M. NetCTLpan: pan-specific MHC class I pathway epitope predictions. *Immunogenetics* (2010) 62(6):357–68. doi: 10.1007/s00251-010-0441-4
26. Wu Y, Zhai W, Zhou X, Wang Z, Lin Y, Ran L, et al. HLA-A2-Restricted Epitopes Identified from MTA1 Could Elicit Antigen-Specific Cytotoxic T Lymphocyte Response. *J Immunol Res* (2018) 2018:2942679. doi: 10.1155/2018/2942679
27. Liu W, Zhai M, Wu Z, Qi Y, Wu Y, Dai C, et al. Identification of a novel HLA-A2-restricted cytotoxic T lymphocyte epitope from cancer-testis antigen *PLAC1* in breast cancer. *Amino Acids* (2012) 42(6):2257–65. doi: 10.1007/s00726-011-0966-3
28. Lv H, Gao Y, Wu Y, Zhai M, Li L, Zhu Y, et al. Identification of a novel cytotoxic T lymphocyte epitope from CFP21, a secreted protein of *Mycobacterium tuberculosis*. *Immunol Lett* (2010) 133(2):94–8. doi: 10.1016/j.imlet.2010.07.007
29. The editorial. The problem with neoantigen prediction. *Nat Biotechnol* (2017) 35(2):97. doi: 10.1038/nbt.3800
30. Chen KY, Liu J, Ren EC. Structural and functional distinctiveness of HLA-A2 allelic variants. *Immunol Res* (2012) 53(1-3):182–90. doi: 10.1007/s12026-012-8295-5
31. Robinson J, Soormally AR, Hayhurst JD, Marsh SGE. The IPD-IMGT/HLA Database - New developments in reporting HLA variation. *Hum Immunol* (2016) 77(3):233–7. doi: 10.1016/j.humimm.2016.01.020
32. Sidney J, Grey HM, Kubo RT, Sette A. Practical, biochemical and evolutionary implications of the discovery of HLA class I supermotifs. *Immunol Today* (1996) 17(6):261–6. doi: 10.1016/0167-5699(96)80542-1
33. Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, Greenbaum J, et al. Immune epitope database analysis resource. *Nucleic Acids Res* (2012) 40(Web Server issue):W525–30. doi: 10.1093/nar/gks438
34. Fleri W, Vaughan K, Salimi N, Vita R, Peters B, Sette A. The Immune Epitope Database: How Data Are Entered and Retrieved. *J Immunol Res* (2017) 2017:5974574. doi: 10.1155/2017/5974574
35. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* (1999) 50(3-4):213–9. doi: 10.1007/s002510050595
36. Vigneron N, Stroobant V, Van den Eynde BJ, van der Bruggen P. Database of T cell-defined human tumor antigens: the 2013 update. *Cancer Immun* (2013) 13:15.
37. Jones S, Thornton JM. Analysis of protein-protein interaction sites using surface patches. *J Mol Biol* (1997) 272(1):121–32. doi: 10.1006/jmbi.1997.1234
38. Jones S, Thornton JM. Prediction of protein-protein interaction sites using patch analysis. *J Mol Biol* (1997) 272(1):133–43. doi: 10.1006/jmbi.1997.1233
39. Haste Andersen P, Nielsen M, Lund O. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein Sci* (2006) 15(11):2558–67. doi: 10.1110/ps.062405906
40. Panchenko AR, Kondrashov F, Bryant S. Prediction of functional sites by analysis of sequence and structure conservation. *Protein Sci* (2004) 13(4):884–92. doi: 10.1110/ps.03465504
41. Patronov A, Doytchinova I. T-cell epitope vaccine design by immunoinformatics. *Open Biol* (2013) 3(1):120139. doi: 10.1098/rsob.120139
42. Zen J, Treutlein HR, Rudy GB. Predicting sequences and structures of MHC-binding peptides: a computational combinatorial approach. *J Comput Aided Mol Des* (2001) 15(6):573–86. doi: 10.1023/A:1011145123635
43. Dunn WJ, 3rd, Koehler MG, Grigoras S. The role of solvent-accessible surface area in determining partition coefficients. *J Med Chem* (1987) 30:1121–6. doi: 10.1021/jm00390a002
44. Collantes ER, Dunn WJ, 3rd. Amino acid side chain descriptors for quantitative structure-activity relationship studies of peptide analogues. *J Med Chem* (1995) 38(14):2705–13. doi: 10.1021/jm00014a022
45. Chowell D, Krishna S, Becker PD, Cocita C, Shu J, Tan X, et al. TCR contact residue hydrophobicity is a hallmark of immunogenic CD8+ T cell epitopes. *Proc Natl Acad Sci USA* (2015) 112(14):E1754–62. doi: 10.1073/pnas.1500973112
46. Liu MK, Hawkins N, Ritchie AJ, Ganusov VV, Whale V, Brackenridge S, et al. Vertical T cell immunodominance and epitope entropy determine HIV-1 escape. *J Clin Invest* (2013) 123(1):380–93. doi: 10.1172/JCI65330
47. Dintzis HM, Dintzis RZ, Vogelstein B. Molecular determinants of immunogenicity: the immunon model of immune response. *Proc Natl Acad Sci USA* (1976) 73(10):3671–5. doi: 10.1073/pnas.73.10.3671

48. Calis JJ, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, et al. Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput Biol* (2013) 9(10):e1003266. doi: 10.1371/journal.pcbi.1003266
49. Kusov Y, Gauss-Muller V, Morace G. Immunogenic epitopes on the surface of the hepatitis A virus capsid: Impact of secondary structure and/or isoelectric point on chimeric virus assembly. *Virus Res* (2007) 130(1-2):296–302. doi: 10.1016/j.virusres.2007.06.002
50. Khatun S, Hasan M, Kurata H. Efficient computational model for identification of antitubercular peptides by integrating amino acid patterns and properties. *FEBS Lett* (2019) 593(21):3029–39. doi: 10.1002/1873-3468.13536
51. Kawashima S, Pokarowski P, Pokarowska M, Kolinski A, Katayama T, Kanehisa M. AAindex: amino acid index database, progress report 2008. *Nucleic Acids Res* (2008) 36(Database issue):D202–5. doi: 10.1093/nar/gkm998
52. Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinf* (2007) 8:424. doi: 10.1186/1471-2105-8-424
53. Jorgensen KW, Rasmussen M, Buus S, Nielsen M. NetMHCstab - predicting stability of peptide-MHC-I complexes; impacts for cytotoxic T lymphocyte epitope discovery. *Immunology* (2014) 141(1):18–26. doi: 10.1111/imm.12160
54. Tung CW, Ziehm M, Kämper A, Kohlbacher Q, Ho SY. POPISK: T-cell reactivity prediction using support vector machines and string kernels. *BMC Bioinf* (2011) 12:446. doi: 10.1186/1471-2105-12-446
55. Cortes C, Vapnik V. Support-vector networks. *Mach Learn* (1995) 20(3):273–97. doi: 10.1007/BF00994018
56. Donnes P, Kohlbacher O. SVMHC: a server for prediction of MHC-binding peptides. *Nucleic Acids Res* (2006) 34(Web Server issue):W194–7. doi: 10.1093/nar/gkl284
57. Kaspari A, Schierhorn A, Schutkowski M. Solid-phase synthesis of peptide-4-nitroanilides. *Int J Pept Protein Res* (1996) 48(5):486–94. doi: 10.1111/j.1399-3011.1996.tb00867.x
58. Mahoney WC, Hermodson MA. Separation of large denatured peptides by reverse phase high performance liquid chromatography. Trifluoroacetic acid as a peptide solvent. *J Biol Chem* (1980) 255(23):11199–203.
59. Wu YH, Gao YF, He YJ, Shi RR, Zhai MX, Wu ZY, et al. A novel cytotoxic T lymphocyte epitope analogue with enhanced activity derived from cyclooxygenase-2. *Scand J Immunol* (2012) 76(3):278–85. doi: 10.1111/j.1365-3083.2012.02738.x
60. Lee JK, Stewart-Jones G, Dong T, Harlos K, Gleria KD, Dorrell L, et al. T cell cross-reactivity and conformational changes during TCR engagement. *J Exp Med* (2004) 200(11):1455–66. doi: 10.1084/jem.20041251
61. van der Merwe PA, Davis SJ. Molecular interactions mediating T cell antigen recognition. *Annu Rev Immunol* (2003) 21:659–84. doi: 10.1146/annurev.immunol.21.120601.141036
62. Matsuda T, Leisegang M, Park JH, Ren L, Kato T, Ikeda Y, et al. Induction of Neoantigen-Specific Cytotoxic T Cells and Construction of T-cell Receptor-Engineered T Cells for Ovarian Cancer. *Clin Cancer Res* (2018) 24(21):5357–67. doi: 10.1158/1078-0432.CCR-18-0142
63. Varela-Calvino R, Skowera A, Arif S, Peakman M. Identification of a naturally processed cytotoxic CD8 T-cell epitope of coxsackievirus B4, presented by HLA-A2.1 and located in the PEVKEK region of the P2C nonstructural protein. *J Virol* (2004) 78(24):13399–408. doi: 10.1128/JVI.78.24.13399-13408.2004
64. Weinzierl AO, Rudolf D, Maurer D, Wernet D, Rammensee HG, Stevanović S, et al. Identification of HLA-A*01- and HLA-A*02-restricted CD8+ T-cell epitopes shared among group B enteroviruses. *J Gen Virol* (2008) 89(Pt 9):2090–7. doi: 10.1099/vir.0.2008/000711-0
65. Aspod C, Laurin D, Richard MJ, Vie H, Chaperot L, Plumas J. Induction of antiviral cytotoxic T cells by plasmacytoid dendritic cells for adoptive immunotherapy of posttransplant diseases. *Am J Transplant* (2011) 11(12):2613–26. doi: 10.1111/j.1600-6143.2011.03722.x
66. Benz C, Utermöhlen O, Wulf A, Villmow B, Dries V, Goeser T, et al. Activated virus-specific T cells are early indicators of anti-CMV immune reactions in liver transplant patients. *Gastroenterology* (2002) 122(5):1201–15. doi: 10.1053/gast.2002.33021
67. Crooks GE, Hon G, Chandonia GM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res* (2004) 14(6):1188–90. doi: 10.1101/gr.849004
68. Wu J, Wang W, Zhang J, Zhou B, Zhao W, Su Z, et al. DeepHLApan: A Deep Learning Approach for Neoantigen Prediction Considering Both HLA-Peptide Binding and Immunogenicity. *Front Immunol* (2019) 10:2559. doi: 10.3389/fimmu.2019.02559
69. Li L, Goedegebuure SP, Gillanders WE. Preclinical and clinical development of neoantigen vaccines. *Ann Oncol* (2017) 28(suppl_12):xii11–7. doi: 10.1093/annonc/mdx681
70. Piancatelli D, Canossi A, Aureli A, Oumhani K, Beato TD, Rocco MD, et al. Human leukocyte antigen-A, -B, and -Cw polymorphism in a Berber population from North Morocco using sequence-based typing. *Tissue Antigens* (2004) 63(2):158–72. doi: 10.1111/j.1399-0039.2004.00161.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Meng, Wu, Sui, Meng, Wang, Lin, Wang, Zhou, Qi, Du and Gao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.