

RESEARCH ARTICLE

# Inferring Protective CD8<sup>+</sup> T-Cell Epitopes for NS5 Protein of Four Serotypes of Dengue Virus Chinese Isolates Based on HLA-A, -B and -C Allelic Distribution: Implications for Epitope-Based Universal Vaccine Design

Jiandong Shi<sup>1,2</sup>, Jing Sun<sup>1,2</sup>, Meini Wu<sup>1,2</sup>, Ningzhu Hu<sup>1,2</sup>, Jianfan Li<sup>1,2</sup>, Yanhan Li<sup>1,2</sup>, Haixuan Wang<sup>1,2</sup>, Yunzhang Hu<sup>1,2\*</sup>

**1** Institute of Medical Biology, Chinese Academy of Medical Sciences and Peking Union Medical College, Kunming, 650118, China, **2** Yunnan Provincial Key Laboratory of Arbo Infectious Disease Control Research (Preparing), Institute of Medical Biology, Chinese Academy of Medical Sciences, Kunming, 650118, China

\* [huyunzhangym@126.com](mailto:huyunzhangym@126.com)



**OPEN ACCESS**

**Citation:** Shi J, Sun J, Wu M, Hu N, Li J, Li Y, et al. (2015) Inferring Protective CD8<sup>+</sup> T-Cell Epitopes for NS5 Protein of Four Serotypes of Dengue Virus Chinese Isolates Based on HLA-A, -B and -C Allelic Distribution: Implications for Epitope-Based Universal Vaccine Design. *PLoS ONE* 10(9): e0138729. doi:10.1371/journal.pone.0138729

**Editor:** Lisa F.P. Ng, Singapore Immunology Network, Agency for Science, Technology and Research (A\*STAR), SINGAPORE

**Received:** July 5, 2015

**Accepted:** September 2, 2015

**Published:** September 18, 2015

**Copyright:** © 2015 Shi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This research was supported by a grant from the Applied and Fundamental Research program of Yunnan Province (Grant No. 2013FA025), which was received by YZH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Dengue is one of the most globally serious vector-borne infectious diseases in tropical and subtropical areas for which there are currently no effective vaccines. The most highly conserved flavivirus protein, NS5, is an indispensable target of CD8<sup>+</sup> T-cells, making it an ideal vaccine design target. Using the Immune Epitope Database (IEDB), CD8<sup>+</sup> T-cell epitopes of the dengue virus (DENV) NS5 protein were predicted by genotypic frequency of the HLA-A, -B, and -C alleles in Chinese population. Antigenicity scores of all predicted epitopes were analyzed using VaxiJen v2.0. The IEDB analysis revealed that 116 antigenic epitopes for HLA-A (21), -B (53), and -C (42) had high affinity for HLA molecules. Of them, 14 had 90.97–99.35% conservancy among the four serotypes. Moreover, five candidate epitopes, including <sup>200</sup>NS5<sup>210</sup> (94.84%, A\*11:01), <sup>515</sup>NS5<sup>525</sup> (98.71%, A\*24:02), <sup>225</sup>NS5<sup>232</sup> (99.35%, A\*33:03), <sup>516</sup>NS5<sup>523</sup> (98.71%, A\*33:03), and <sup>284</sup>NS5<sup>291</sup> (98.06%, A\*33:03), were presented by HLA-A. Four candidate epitopes, including <sup>234</sup>NS5<sup>241</sup> (96.77%, B\*13:01), <sup>92</sup>NS5<sup>99</sup> (98.06%, B\*15:01, B\*15:02, and B\*46:01), <sup>262</sup>NS5<sup>269</sup> (92.90%, B\*38:02), and <sup>538</sup>NS5<sup>547</sup> (90.97%, B\*51:01), were presented by HLA-B. Another 9 candidate epitopes, including <sup>514</sup>NS5<sup>522</sup> (98.71%, C\*01:02), <sup>514</sup>NS5<sup>524</sup> (98.71%, C\*01:02 and C\*14:02), <sup>92</sup>NS5<sup>99</sup> (98.06%, C\*03:02 and C\*15:02), <sup>362</sup>NS5<sup>369</sup> (44.84%, C\*03:04 and C\*08:01), <sup>225</sup>NS5<sup>232</sup> (99.35%, C\*04:01), <sup>234</sup>NS5<sup>241</sup> (96.77%, C\*04:01), <sup>361</sup>NS5<sup>369</sup> (94.84%, C\*04:01), <sup>515</sup>NS5<sup>522</sup> (98.71%, C\*14:02), <sup>515</sup>NS5<sup>524</sup> (98.71%, C\*14:02), were presented by HLA-C. Further data showed that the four-epitope combination of <sup>92</sup>NS5<sup>99</sup> (B\*15:01, B\*15:02, B\*46:01, C\*03:02 and C\*15:02), <sup>200</sup>NS5<sup>210</sup> (A\*11:01), <sup>362</sup>NS5<sup>369</sup> (C\*03:04, C\*08:01), and <sup>514</sup>NS5<sup>524</sup> (C\*01:02, C\*14:02) could vaccinate >90% of individuals in China. Further *in vivo* study of our inferred novel epitopes will be needed for a T-cell epitope-based universal vaccine development that may prevent all four China-endemic DENV serotypes.

**Competing Interests:** The authors have declared that no competing interests exist.

## Introduction

Dengue virus (DENV) can cause dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS), globally important mosquito-borne diseases [1, 2]. These are among the most serious epidemic arbovirus diseases and endemic in tropical and subtropical regions of the world. The causative viruses are members of the genus *Flavivirus* within the family *Flaviviridae* and can be grouped into four antigenically distinct serotypes (DENV1-4) that share 67–75% sequence homology [3, 4]. DENV is transmitted to humans through the bites of infected *Aedes aegypti* and *Aedes albopictus* mosquitoes. Nearly half of the world's population is under risk of contracting dengue. It is estimated that up to 390 million infections occur annually worldwide with approximately 96 million symptomatic cases [5]. Despite more than 60 years of effort, no licensed vaccine is currently available. Thus, the search for a safe and effective vaccine is growing more imperative.

Dengue is hyperendemic and has become a serious public health concern in China. The first outbreak of dengue was reported in Guangdong Province of China in 1978 [6, 7]. Since then, annual DENV epidemics have occurred, followed by a dengue epidemic in Guangxi, Fujian, Zhejiang, and other areas of China. In 2014, the most serious dengue epidemic in history occurred in Guangdong province of China with a total of 48,162 infected individuals [8]. This outbreak is considered an imported epidemic from neighboring Southeast Asian countries [9, 10]. In recent years, the scope of the epidemic is further expanding from the coastal city of China to inland cities. In 2013, an outbreak of DENV occurred in Yunnan province of China with more than 2,000 infected individuals [11]. A safe and effective dengue vaccine is urgent need in China.

CD8<sup>+</sup> T-cell-mediated immunity plays an important role for eliminating intracellular pathogens. Thus, eliciting robust CD8<sup>+</sup> T-cell immunity is the basis for many vaccines under development. Although DENV-specific CD8<sup>+</sup> T-cell responses have been extensively studied, the vast majority of studies focused on immunopathogenic role of T-cells during DENV infection [12–14]. The viewpoint from these studies is that serotype cross-reactive CD8<sup>+</sup> T-cells may contribute to the immunopathogenesis of DHF/DSS. Thus, the vast majority of dengue vaccine candidates are designed to produce protective neutralizing antibodies with less regard for cellular immune responses. However, direct evidence linking T-cells to increased viremia or DENV-related pathology has not been demonstrated. Notably, recent extensive studies have demonstrated a protective role of CD4<sup>+</sup> and CD8<sup>+</sup> T-cells against homologous or heterotypic DENV infection in murine models [15–20]. Specifically, these studies demonstrated that CD8<sup>+</sup> T-cells can control viral replication [16], prevent antibody-dependent enhancement (ADE) of infection [19], and DENV-induced CNS disease [18].

These findings are consistent with the murine model data of a recent study supporting the concept of a protective role of T-cells against DENV infection in humans. The results of this study showed that the secondary DENV infection in humans was not significantly associated with disease severity [21]. Further, another recent study provided the first comprehensive map of the CD8<sup>+</sup> T-cell response to DENV in humans and support a HLA-linked protective but not pathogenic role for CD8<sup>+</sup> T-cells against DENV infection in humans [22]. Collectively, these findings strongly imply a protective role for CD8<sup>+</sup> T-cells against severe DENV disease in humans. Based on these studies, it is inferred that the lack of induction of a robust DENV-specific T-cell response may be a reason for the results of a recent efficacy trial of the most advanced dengue vaccine candidate, a tetravalent live-attenuated chimeric vaccine (CYD) based on the 17D-attenuated yellow fever virus backbone that showed only partial protection despite the induction of DENV-specific neutralizing antibody to each serotype in most subjects [23]. This means that the roles of T-cells in the context of DENV vaccination should not be

ignored, and it raises the possibility that T-cell responses against all DENV serotypes might be beneficial or even required for vaccine protective efficacy. The advent of a T-cell epitope-based vaccine may offer an alternative that avoids ADE. Considering the important role of serotype-specific CD8<sup>+</sup> T-cells in controlling DENV infection, a novel strategy for developing prophylactic and therapeutic CD8<sup>+</sup> T-cell epitope-based vaccines is needed. Therefore, a T-cell epitope-based universal vaccine that induces a broad dengue-specific, multifunctional, and cross-reactive CD8<sup>+</sup> T-cell responses among all four DENV serotypes may be a more promising strategy against DENV infections.

The DENV genome consists of a single-stranded RNA of 10.7 kb in length. The open reading frame codes three structural proteins [capsid (C) protein, preM protein, and envelope (E) protein] and seven nonstructural proteins (NS1, NS2a, NS2b, NS3, NS4a, NS4b and NS5) [24,25]. It has been shown that CD8<sup>+</sup> T cells preferentially target the NS3 and NS5 proteins, while CD4<sup>+</sup> T cells preferentially target the E, C, and NS1 proteins [26]. Notably, NS5 is the largest and the most highly conserved protein encoded by the DENV genome, with approximately 67–82% amino acid sequence identity among the four DENV serotypes [27]. Thus, NS5 proteins could be used as a promising target in the design of a T-cell epitope-based vaccine to induce DENV-specific protective T-cell responses.

A necessary condition for a peptide to be a CD8<sup>+</sup> T-cell epitope is that it binds to human leukocyte antigen (HLA) molecules. However, HLA molecules are extremely polymorphic with several thousand variants and can bind distinct sets of peptides [28]. Each HLA variant is expressed at vastly variable frequencies in different ethnic groups and geographic regions [29]. This means that it appears that an extremely large and impractical number of peptides would have to be selected to enable the development of a broadly protective multi-epitope vaccine. A large number of studies focused on predicting epitopes from the E, prM, NS1, NS3, or NS5 proteins allowed the identification of T-cell epitopes in DENV [30–36]. However, the CD8<sup>+</sup> T-cell epitopes of the NS5 protein of DENV Chinese isolates linked with the class I HLA allele in Chinese population have been poorly revealed. Therefore, the identification of CD8<sup>+</sup> T-cell epitopes that can induce protective DENV-specific T-cell responses by a feasible immunoinformatics approach is critical and urgent for the development of a T-cell epitope-based vaccine.

In this study, based on the distribution characteristics of HLA class I alleles in Chinese population, we identified putative CD8<sup>+</sup> T-cell epitopes of NS5 protein of Chinese DENV isolates using various immunoinformatics approaches. Our results provide putative protective CD8<sup>+</sup> T-cell epitope candidates or their combination for the development of a T-cell epitope-based universal vaccine to effectively prevent all four DENV serotypes that are endemic in China.

## Materials and Methods

### Retrieving the protein sequences

The sequences of the NS5 protein from 155 Chinese isolates belong to all four serotypes of dengue virus (DENV-1, DENV-2, DENV-3 and DENV-4) were retrieved from the National Center for Biotechnology Information (NCBI) protein database (<http://www.ncbi.nlm.nih.gov/protein/>). The sequence of DENV-2 NS5 protein (accession number: KC131142.1) was used as an input for various bioinformatics tools for epitope prediction, antigenicity analysis and conservation analysis.

### HLA genotypic frequency retrieval

To improve population coverage of the CD8<sup>+</sup> T-cell epitopes, it is important to screen epitopes restricted by highly prevalent HLA alleles. Genotypic frequencies of the HLA class I alleles that include HLA-A,-B and-C loci in Chinese population were retrieved from the major

histocompatibility complex database (dbMHC) (<http://www.ncbi.nlm.nih.gov/projects/gv/mhc/main.fcgi?cmd=init>). For a broad coverage, HLA class I alleles with genotypic frequency >3% in Chinese population were selected for CD8<sup>+</sup> T-cell epitope prediction. This parameter setting covers the highly prevalent HLA-A, -B, and -C alleles found in Chinese population.

## Epitope prediction

CD8<sup>+</sup> T-cell epitope is the minimal amino acid sequence required for CD8<sup>+</sup> T-cells activation and recognition by immune system receptors. Since the affinity of an epitope binds to the HLA molecule plays a vital role in determining its immunogenicity. Hence, high affinity between epitopes and HLA molecules tends to be associated with higher immune responsiveness. The prediction of CD8<sup>+</sup> T-cell epitopes that interact with different HLA class I alleles were performed using the IEDB analysis resource (<http://tools.immuneepitope.org/mhci/>) Consensus tool [37], which combines predictions from ANN, aka NetMHC 3.4 [38][39], SMM [40] and Comblib [41], if any corresponding predictor is available for the molecule. Otherwise, NetMHCpan is used. This choice was motivated by the expected predictive performance of the methods in decreasing order: Consensus > ANN > SMM > NetMHCpan > CombLib. For the IEDB-recommended method, a low percentile indicated strong binding affinity to HLA molecules. The threshold of percentile rank was set at 1 in this study [42]. Based on the representative length of peptides that bind to HLA molecules, peptide lengths of 8–11 amino acids were selected for the prediction of epitope-based peptides in this study.

## Antigenicity analysis

Antigenicity is a key characteristic of the epitope that is recognized by immune system cells and/or antibodies. Thus, the antigenicity of the predicted epitopes is one of the most important criteria for epitope-based vaccine assessment. Since some of the predicted epitopes may lose antigenicity when analyzed, to ensure that the predicted epitopes could serve as a good CD8<sup>+</sup> T-cell epitope, all of the epitopes were screened to assess their antigenicity. VaxiJen v2.0 (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen>), an online web server used to predict the effective antigens and subunit vaccines, was used to identify and reevaluate T-cell epitope antigenicity. The predicted epitopes were uploaded in a plain sequence format and the virus was chosen as the target organism. The threshold level of an antigen was set at 0.5. The Vaxijen server performed well with 87% accuracy at a threshold of 0.5 antigenic score for viruses [43]. VaxiJen v2.0 allows antigen classification based on the physicochemical properties of proteins without recourse to sequence alignment. Finally, the epitopes with an antigenic score > 0.5 were selected as antigenic for the conservancy analysis.

## Conservancy and population coverage analysis

To obtain the universal T-cell epitopes of four DENV serotype variants, the conservancy of candidate epitopes should be considered prior to other criteria, even population coverage rate. As a universal T-cell epitope, it should be highly conserved in all viral variants. Hence, to determine the conservation level of the predicted epitopes among the NS5 protein sequences of the different DENV strains, the predicted epitopes were analyzed for their conservancy using the IEDB epitope conservancy tool ([http://tools.immuneepitope.org/tools/conservancy/iedb\\_input](http://tools.immuneepitope.org/tools/conservancy/iedb_input)) with a sequence identity threshold of 100%. The conservancy level of each potential epitope was calculated by seeking identities in all NS5 protein sequences of the four DENV serotype variants retrieved from the NCBI protein database. The epitopes that were 100% conserved in >90% of the sequences analyzed in four serotypes were selected as candidate epitopes. These highly conserved epitopes were selected and used to determine the population

coverage by the IEDB population coverage calculation tool ([http://tools.immuneepitope.org/tools/population/iedb\\_input](http://tools.immuneepitope.org/tools/population/iedb_input)). Finally, all of the selected epitopes were analyzed for similarity with human proteome using the BLAST program (<http://www.ncbi.nlm.nih.gov/BLAST/>) to verify that they would not trigger autoimmunity.

## Results and Discussion

### Retrieval of NS5 protein sequences of four DENV serotypes

The NS5 proteins sequence of all four DENV serotypes circulating in China were retrieved in FASTA format from the NCBI protein database. A total of 155 sequences of the NS5 protein of Chinese isolates of the four DENV serotypes was obtained ([S1 File](#), [S2 File](#), [S3 File](#) and [S4 File](#)) and used for the further epitope analysis.

### HLA class I alleles analysis

Since specific HLA alleles are expressed at variable frequencies in different ethnic groups and different geographic regions. Therefore, HLA allele frequencies prevalent in dengue hyperendemic areas must be considered in vaccine design. Here we focused on DENV-specific T-cell epitopes that are associated with the highly prevalent HLA alleles in a Chinese population. To this end, the genotypic frequency of the highly prevalent HLA class I alleles found in this Chinese population (>3%) that include HLA-A, -B, and -C loci were obtained from the dbMHC database. As a result, seven HLA-A alleles (A\*02:01, A\*02:03, A\*02:06, A\*02:07, A\*11:01, A\*24:02, and A\*33:03), eight HLA-B alleles (B\*13:01, B\*15:01, B\*15:02, B\*38:02, B\*40:01, B\*46:01, B\*51:01, and B\*58:01), and eight HLA-C alleles (C\*01:02, C\*03:02, C\*03:03, C\*03:04, C\*04:01, C\*08:01, C\*14:02, and C\*15:02) were obtained ([Table 1](#)).

**Table 1. Frequency of HLA class I alleles (>3%) in Chinese population.**

Allele	Frequency
HLA-A*02:01	0.053
HLA-A*02:03	0.108
HLA-A*02:06	0.035
HLA-A*02:07	0.094
HLA-A*11:01	0.277
HLA-A*24:02	0.172
HLA-A*33:03	0.115
HLA-B*13:01	0.082
HLA-B*15:01	0.044
HLA-B*15:02	0.071
HLA-B*38:02	0.071
HLA-B*40:01	0.149
HLA-B*46:01	0.115
HLA-B*51:01	0.050
HLA-B*58:01	0.089
HLA-C*01:02	0.169
HLA-C*03:02	0.087
HLA-C*03:03	0.041
HLA-C*03:04	0.128
HLA-C*04:01	0.044
HLA-C*08:01	0.126
HLA-C*14:02	0.036
HLA-C*15:02	0.037

doi:10.1371/journal.pone.0138729.t001



## Prediction and antigenic analysis of CD8<sup>+</sup> T-cell epitopes

CD8<sup>+</sup> T-cell responses play a substantial role in eliminating DENV infected cells that cannot be managed by antibody responses. An effective vaccine that can provide protection against dengue virus infection requires robust, broad, and multi-functional CD8<sup>+</sup> T-cell responses. Therefore, the CD8<sup>+</sup> T-cell epitopes that bind to different HLA class I alleles with varying affinities must first be identified. Here a large number of the antigenic epitopes with a high binding affinity score of <1 percentile and an antigenicity score > 0.5 were obtained from NS5 proteins of DENV Chinese isolates against HLA-A, -B, and -C alleles. Most epitopes bind with high affinity to single HLA-A, -B, or -C molecules. As a consequence, a total of 21 antigenic epitopes were obtained against the seven alleles of the HLA-A loci (Table 2). A total of 53 antigenic epitopes were obtained against the eight alleles of the HLA-B loci (Table 3), while 42 antigenic epitopes were obtained against the eight alleles of the HLA-C loci (Table 4). Surprisingly, none of the epitopes bind to HLA-A\*02:07. Notably, some epitopes, like <sup>92</sup>AMTDTTPF<sup>99</sup> (B\*15:01, B\*15:02, B\*46:01, C\*03:02, and C\*15:02), <sup>515</sup>MYFHRRDLRL<sup>524</sup> (A\*24:02 and C\*14:02), <sup>225</sup>WYMWLGAR<sup>232</sup> (A\*33:03 and C\*04:01), <sup>234</sup>LEFEALGF<sup>241</sup> (B\*13:01 and C\*04:01), <sup>514</sup>LMYFHRRDLRL<sup>524</sup> (C\*01:02 and C\*14:02), and <sup>362</sup>FTNMEAQL<sup>369</sup> (C\*03:04 and C\*08:01) can be presented by multiple HLA molecules, suggesting that they can cover a broader population and may be better epitope vaccine candidates. Further, the potential of the epitope and HLA binding is essential in the assessment of the immunogenic potential of epitopes. Hence, as a ligand of the HLA-A molecule, epitopes <sup>225</sup>WYMWLGAR<sup>233</sup> and <sup>224</sup>IWYMWLGARF<sup>233</sup> presented by HLA-A\*24:02 were the best binders based on their 0.2 percentile. As ligands of HLA-B molecules, epitopes <sup>89</sup>TQMAMTDTTPF<sup>99</sup> and <sup>530</sup>CSAVPSHW<sup>537</sup>, which were presented by HLA-B\*15:01 and HLA-B\*58:01, respectively, were the best binders based on their 0.1 percentile. Likewise, as a ligand of HLA-C molecule, epitope <sup>91</sup>MAMTDTTPF<sup>99</sup> presented by HLA-C\*03:02 was the best binder based on its 0.1 percentile. Additionally, HLA-A\*24:02 has the highest number of binding epitopes (9/21), followed by A\*33:03 (6/21) in the HLA-A allele. HLA-B\*13:01 and HLA-B\*58:01 presented the same most frequent epitopes (13/53), followed by B\*15:01 (12/53), B\*46:01 (12/53), B\*38:02 (11/53), and B\*15:02 (9/53) in the HLA-B alleles. Moreover, HLA-C\*04:01 presented the most frequent epitopes (17/42), followed by C\*14:02 (13/42), C\*01:02 (8/42), C\*03:02 (8/42), and C\*15:02 (7/42) in the HLA-C alleles. These calculations were made on the basis of HLA genotypic frequencies assuming non-linkage disequilibrium between HLA loci. Overall, these results provide *in silico* insight for class I HLA allele-restricted CD8<sup>+</sup> T-cell epitopes against the NS5 protein of DENV in Chinese population.

## Conservancy and population coverage of CD8<sup>+</sup> T-cell epitopes

As an effective vaccine formulation, the epitope-based universal vaccine must include highly conserved CD8<sup>+</sup> T-cell epitopes among all DENV serotypes to induce cross-reactive T-cell responses based on the fact that the conserved epitope candidates are more likely to confer cross-protection between pathogen variants. Here, conservancy analysis revealed a total of 14 highly conserved epitopes with ≥90% protein sequence matching in a total of 155 NS5 protein sequences from four DENV serotypes of Chinese isolates. For the HLA-A allele, five of the 21 epitopes were conserved with ≥90% conservancy (Table 2), including <sup>200</sup>CVYNMMGKREK<sup>210</sup> (94.84%, A\*11:01), <sup>515</sup>MYFHRRDLRL<sup>524</sup> (98.71%, A\*24:02), <sup>225</sup>WYMWLGAR<sup>232</sup> (99.35%, A\*33:03), <sup>516</sup>YFHRRDLR<sup>523</sup> (98.71%, A\*33:03), and <sup>284</sup>DTAGWDTR<sup>291</sup> (98.06%, A\*33:03). For the HLA-B allele, four of the 53 epitopes were conserved with ≥90% conservancy (Table 3), including <sup>234</sup>LEFEALGF<sup>241</sup> (96.77%, B\*13:01), <sup>92</sup>AMTDTTPF<sup>99</sup> (98.06%, B\*15:01, B\*15:02 and B\*46:01), <sup>262</sup>LHKLGYL<sup>269</sup> (92.90%, B\*38:02), and <sup>538</sup>VPTSRTTWSI<sup>547</sup> (90.97%,

**Table 2. HLA-A restricted epitopes of the NS5 protein and their binding affinity, antigenicity and conservation (in percentages) in different serotypes.**

No.	Allele	Start <sup>a</sup>	End <sup>a</sup>	Peptide length	Sequence	Method used	Percentile rank	Antigenicity score	Percent of protein sequence matches at identity ≥ 100%
001	HLA-A*02:01	553	562	10	WMTTEDMLTV	Consensus (ann/smm)	1.0	0.6668	73.55% (114/155)
002	HLA-A*02:03	599	607	9	SLIGLTSRA	Consensus (ann/smm)	0.85	2.0423	68.39% (106/155)
003	HLA-A*02:03	553	562	10	WMTTEDMLTV	Consensus (ann/smm)	0.65	0.5568	73.55% (114/155)
004	HLA-A*02:06	553	562	10	WMTTEDMLTV	Consensus (ann/smm)	0.7	0.5568	73.55% (114/155)
005	HLA-A*11:01	338	346	9	TVMDIISRR	Consensus (ann/smm)	0.55	1.0053	10.97% (17/155)
006	HLA-A*11:01	125	133	9	KITAEWLWK	Consensus (ann/smm)	0.95	0.9815	10.97% (17/155)
007	HLA-A*11:01	317	326	10	AIFRLTYQNK	Consensus (ann/smm)	0.5	1.0707	1.94% (3/155)
<b>008</b>	<b>HLA-A*11:01</b>	<b>200</b>	<b>210</b>	<b>11</b>	<b>CVYNNMGKREK</b>	<b>Consensus (ann/smm)</b>	<b>0.65</b>	<b>0.6848</b>	<b>94.84% (147/155)</b>
009	HLA-A*24:02	225	233	9	WYMWLGARF	Consensus (ann/smm)	0.2	0.9423	32.26% (50/155)
010	HLA-A*24:02	224	233	10	IWYMWLGARF	Consensus (ann/smm)	0.2	0.7264	32.26% (50/155)
011	HLA-A*24:02	227	236	10	MWLGARFLEF	Consensus (ann/smm)	0.25	1.0434	32.26% (50/155)
012	HLA-A*24:02	225	234	10	WYMWLGARFL	Consensus (ann/smm)	0.3	0.6068	32.26% (50/155)
013	HLA-A*24:02	544	553	10	TWSIHATHEW	Consensus (ann/smm)	0.5	1.1355	9.03% (14/155)
<b>014</b>	<b>HLA-A*24:02</b>	<b>515</b>	<b>524</b>	<b>10</b>	<b>MYFHRRDLRL</b>	<b>Consensus (ann/smm)</b>	<b>0.55</b>	<b>1.6075</b>	<b>98.71% (153/155)</b>
015	HLA-A*24:02	449	458	10	GWNDWTQVPF	Consensus (ann/smm)	0.65	0.9735	9.68% (15/155)
016	HLA-A*24:02	232	241	10	RFLEFEALGF	Consensus (ann/smm)	0.95	1.7726	32.26% (50/155)
017	HLA-A*24:02	226	236	11	YMWLGARFLEF	Consensus (ann/smm)	0.95	1.0316	32.26% (50/155)
018	HLA-A*33:03	513	520	8	SLMYFHRR	netmhcpan	0.2	1.5676	80.00% (124/155)
<b>019</b>	<b>HLA-A*33:03</b>	<b>225</b>	<b>232</b>	<b>8</b>	<b>WYMWLGAR</b>	<b>netmhcpan</b>	<b>0.4</b>	<b>1.2415</b>	<b>99.35% (154/155)</b>
020	HLA-A*33:03	512	519	8	WSLMYFHR	netmhcpan	0.4	0.7229	80.00% (124/155)
<b>021</b>	<b>HLA-A*33:03</b>	<b>516</b>	<b>523</b>	<b>8</b>	<b>YFHRRDLR</b>	<b>netmhcpan</b>	<b>0.4</b>	<b>1.6748</b>	<b>98.71% (153/155)</b>
022	HLA-A*33:03	395	402	8	NWLVRVGR	netmhcpan	0.6	1.2717	3.87% (6/155)
<b>023</b>	<b>HLA-A*33:03</b>	<b>284</b>	<b>291</b>	<b>8</b>	<b>DTAGWDTR</b>	<b>netmhcpan</b>	<b>0.7</b>	<b>1.7765</b>	<b>98.06% (152/155)</b>

<sup>a</sup>The epitopes location in NS5 protein are from accession number: KC131142.1.

Bold and italic- indicates the percentage of epitope that is 100% conserved in more than 90% of the sequences analysed in four serotypes.

doi:10.1371/journal.pone.0138729.t002

B\*51:01). For the HLA-C allele, nine of the 42 epitopes were conserved with ≥90% conservancy (Table 4), including <sup>514</sup>LMYFHRRDL<sup>522</sup> (98.71%, C\*01:02), <sup>514</sup>LMYFHRRDLRL<sup>524</sup> (98.71%, C\*14:02 and C\*01:02), <sup>92</sup>AMTDTPF<sup>99</sup> (98.06%, C\*03:02 and C\*15:02), <sup>362</sup>FTNMEAQL<sup>369</sup> (94.84%, C\*03:04 and C\*08:01), <sup>225</sup>WYMWLGAR<sup>232</sup> (99.35%, C\*04:01),

**Table 3. HLA-B restricted epitopes of the NS5 protein and their binding affinity, antigenicity and conservation (in percentages) in different serotypes.**

No.	Allele	Start <sup>a</sup>	End <sup>a</sup>	Peptide length	Sequence	Method used	Percentile rank	Antigenicity score	Percent of protein sequence matches at identity $\geq$ 100%
<b>001</b>	<b>HLA-B*13:01</b>	<b>234</b>	<b>241</b>	<b>8</b>	<b>LEFEALGF</b>	<b>netmhcpan</b>	<b>0.2</b>	<b>2.1459</b>	<b>96.77% (150/155)</b>
002	HLA-B*13:01	128	135	8	AEWLWKEL	netmhcpan	0.3	0.8381	10.97% (17/155)
003	HLA-B*13:01	510	517	8	QMWSLMYF	netmhcpan	0.4	0.5350	80.65% (125/155)
004	HLA-B*13:01	226	233	8	YMWLGARF	netmhcpan	0.6	0.8453	32.26% (50/155)
005	HLA-B*13:01	371	378	8	RQMEGEGV	netmhcpan	0.8	0.5155	77.42% (120/155)
006	HLA-B*13:01	19	26	8	SETPNLDI	netmhcpan	0.9	0.5435	5.16% (8/155)
007	HLA-B*13:01	234	242	9	LEFEALGFL	netmhcpan	0.5	1.7009	78.71% (122/155)
008	HLA-B*13:01	19	27	9	SETPNLDI	netmhcpan	0.7	0.5209	5.16% (8/155)
009	HLA-B*13:01	520	529	10	RDLRLAANAI	netmhcpan	0.7	0.8520	29.03% (45/155)
010	HLA-B*13:01	382	391	10	IQHLTVTEEI	netmhcpan	0.9	0.7084	9.68% (15/155)
011	HLA-B*13:01	580	589	10	VESWEEIPLY	netmhcpan	1	0.5190	10.97% (17/155)
012	HLA-B*13:01	89	99	11	TQMAMDTTPF	netmhcpan	0.2	0.9856	77.42% (120/155)
013	HLA-B*13:01	226	236	11	YMWLGARFLEF	netmhcpan	0.3	1.0316	32.26% (50/155)
<b>014</b>	<b>HLA-B*15:01</b>	<b>92</b>	<b>99</b>	<b>8</b>	<b>AMTDTTPF</b>	<b>ann</b>	<b>0.2</b>	<b>0.6985</b>	<b>98.06% (152/155)</b>
015	HLA-B*15:01	226	233	8	YMWLGARF	ann	0.5	0.8453	32.26% (50/155)
016	HLA-B*15:01	142	149	8	RMCTREEF	ann	0.6	1.0232	10.97% (17/155)
017	HLA-B*15:01	510	517	8	QMWSLMYF	ann	0.6	0.5350	80.65% (125/155)
018	HLA-B*15:01	91	99	9	MAMTDTTPF	Consensus (ann/ comblib_sidney2008/ smm)	0.2	0.6735	77.42% (120/155)
019	HLA-B*15:01	228	236	9	WLGARFLEF	Consensus (ann/ comblib_sidney2008/ smm)	0.6	1.5566	32.26% (50/155)
020	HLA-B*15:01	90	99	10	QAMAMDTTPF	Consensus (ann/smm)	0.3	0.9860	77.42% (120/155)
021	HLA-B*15:01	89	99	11	TQMAMDTTPF	ann	0.1	0.9856	77.42% (120/155)
022	HLA-B*15:01	226	236	11	YMWLGARFLEF	ann	0.3	1.0316	32.26% (50/155)
023	HLA-B*15:01	330	340	11	VQRPTPRGTVM	ann	0.3	0.8415	9.68% (15/155)
024	HLA-B*15:02	226	233	8	YMWLGARF	ann	0.2	0.8453	32.26% (50/155)
025	HLA-B*15:02	622	629	8	SLIGNEEY	ann	0.3	0.8841	10.32% (16/155)
<b>026</b>	<b>HLA-B*15:02</b>	<b>92</b>	<b>99</b>	<b>8</b>	<b>AMTDTTPF</b>	<b>ann</b>	<b>0.7</b>	<b>0.6985</b>	<b>98.06% (152/155)</b>
027	HLA-B*15:02	510	517	8	QMWSLMYF	ann	1	0.5350	80.65% (125/155)
028	HLA-B*15:02	90	99	10	QAMAMDTTPF	ann	0.3	0.9860	77.42% (120/155)
029	HLA-B*15:02	89	99	11	TQMAMDTTPF	ann	0.3	0.9856	77.42% (120/155)
030	HLA-B*15:02	226	239	11	YMWLGARFLEF	ann	0.3	1.0316	32.26% (50/155)
031	HLA-B*15:02	622	632	11	SLIGNEEYTDY	ann	0.3	0.9393	10.32% (16/155)
032	HLA-B*15:02	330	340	11	VQRPTPRGTVM	ann	0.9	0.8415	9.68% (15/155)
033	HLA-B*38:02	124	131	8	MKITAEWL	netmhcpan	1	0.5377	10.97% (17/155)
<b>034</b>	<b>HLA-B*38:02</b>	<b>262</b>	<b>269</b>	<b>8</b>	<b>LHKLGYL</b>	<b>netmhcpan</b>	<b>1</b>	<b>0.7094</b>	<b>92.90% (144/155)</b>
035	HLA-B*38:02	383	391	9	QHLTVTEEI	netmhcpan	0.2	0.9107	9.68% (15/155)
036	HLA-B*38:02	42	50	9	WHYDQDHPY	netmhcpan	0.4	0.5520	9.03% (14/155)
037	HLA-B*38:02	91	99	9	MAMTDTTPF	netmhcpan	0.8	0.6735	77.42% (120/155)
038	HLA-B*38:02	89	99	11	TQMAMDTTPF	netmhcpan	0.3	0.9856	77.42% (120/155)
039	HLA-B*38:02	383	393	11	QHLTVTEEIAV	netmhcpan	0.4	0.8018	9.68% (15/155)
040	HLA-B*38:02	535	545	11	SHWVPTSRTTW	netmhcpan	0.5	1.4792	10.97% (17/155)
041	HLA-B*38:02	626	636	22	NEEYTDYMPSM	netmhcpan	0.8	0.6913	10.32% (16/155)
042	HLA-B*38:02	480	490	11	NQDELIGRARI	netmhcpan	0.9	0.6578	14.19% (22/155)

(Continued)



Table 3. (Continued)

No.	Allele	Start <sup>a</sup>	End <sup>a</sup>	Peptide length	Sequence	Method used	Percentile rank	Antigenecity score	Percent of protein sequence matches at identity $\geq 100\%$
043	HLA-B*38:02	226	236	11	YMWLGARFLEF	netmhcpan	1	1.0316	32.26% (50/155)
044	HLA-B*40:01	250	257	8	RENSLSGV	Consensus (ann/smm)	0.75	0.8542	28.39% (44/155)
045	HLA-B*40:01	234	242	9	LEFEALGFL	Consensus (ann/smm)	0.2	1.7009	78.71% (122/155)
046	HLA-B*40:01	19	27	9	SETPNLDII	Consensus (ann/smm)	0.85	0.5209	5.16% (8/155)
047	HLA-B*40:01	580	589	10	VESWEEIPLY	Consensus (ann/smm)	0.5	0.5109	10.97% (17/155)
048	HLA-B*40:01	182	191	10	WELVDKERNL	Consensus (ann/smm)	0.85	1.4932	10.97% (17/155)
<b>049</b>	<b>HLA-B*46:01</b>	<b>92</b>	<b>99</b>	<b>8</b>	<b>AMTDTTPF</b>	<b>ann</b>	<b>0.3</b>	<b>0.6985</b>	<b>98.06% (152/155)</b>
050	HLA-B*46:01	226	233	8	YMWLGARF	ann	0.4	0.8453	32.26% (50/155)
051	HLA-B*46:01	510	517	8	QMWSLMYF	ann	0.7	0.5350	80.65% (125/155)
052	HLA-B*46:01	467	474	8	IMKDGRVL	ann	1	0.6354	10.97% (17/155)
053	HLA-B*46:01	629	636	8	YTDYMPSM	ann	1	0.8868	10.97% (17/155)
054	HLA-B*46:01	91	99	9	MAMTDTTPF	Consensus (ann/smm)	0.2	0.6735	77.42% (120/155)
055	HLA-B*46:01	272	280	9	VSKKEGGAM	Consensus (ann/smm)	0.65	0.5314	10.97% (17/155)
056	HLA-B*46:01	248	257	10	FSRENSLSGV	ann	0.7	1.2193	28.39% (44/155)
057	HLA-B*46:01	90	99	10	QMAMTDTTPF	ann	0.8	0.9860	77.42% (120/155)
058	HLA-B*46:01	78	87	10	LTKPWDVIPM	ann	0.9	0.9964	32.26% (50/155)
059	HLA-B*46:01	226	236	11	YMWLGARFLEF	ann	0.2	1.0316	32.26% (50/155)
060	HLA-B*46:01	89	99	11	TQMAMTDTTPF	ann	0.8	0.9856	77.42% (120/155)
061	HLA-B*51:01	85	93	9	IPMVTQMAM	Consensus (ann/ comblib_sidney2008/ smm)	0.5	0.5370	10.97% (17/155)
062	HLA-B*51:01	80	88	9	KPWDVIPMV	Consensus (ann/ comblib_sidney2008/ smm)	0.8	1.2766	32.26% (50/155)
<b>063</b>	<b>HLA-B*51:01</b>	<b>538</b>	<b>547</b>	<b>10</b>	<b>VPTSRTTWSI</b>	<b>Consensus (ann/smm)</b>	<b>0.3</b>	<b>1.0517</b>	<b>90.97% (141/155)</b>
064	HLA-B*58:01	530	537	8	CSAVPSHW	ann	0.1	0.8475	10.97% (17/155)
065	HLA-B*58:01	125	132	8	KITAEWLW	ann	0.3	1.0187	10.97% (17/155)
066	HLA-B*58:01	123	130	8	LMKITAEW	ann	1	1.0313	10.97% (17/155)
067	HLA-B*58:01	545	553	9	KLMKITAEW	Consensus (ann/ comblib_sidney2008/ smm)	0.2	1.0764	10.97% (17/155)
068	HLA-B*58:01	91	99	9	MAMTDTTPF	Consensus (ann/ comblib_sidney2008/ smm)	0.3	0.6735	77.42% (120/155)
069	HLA-B*58:01	124	132	9	MKITAEWLW	Consensus (ann/ comblib_sidney2008/ smm)	0.3	0.7346	10.97% (17/155)
070	HLA-B*58:01	601	609	9	IGLTSRATW	Consensus (ann/ comblib_sidney2008/ smm)	0.4	1.5513	68.39% (106/155)
071	HLA-B*58:01	621	629	9	RSLIGNEEY	Consensus (ann/ comblib_sidney2008/ smm)	0.9	0.9167	10.32% (16/155)
072	HLA-B*58:01	544	553	10	TWSIHATHEW	Consensus (ann/smm)	0.6	1.1355	9.03% (14/155)
073	HLA-B*58:01	33	42	10	KIKQEHETSW	Consensus (ann/smm)	0.6	0.9160	10.97% (17/155)
074	HLA-B*58:01	123	132	10	LMKITAEWLW	Consensus (ann/smm)	0.65	0.9159	10.97% (17/155)
075	HLA-B*58:01	237	247	11	EALGFLNEDHW	Consensus (ann/smm)	0.75	1.0740	77.42% (120/155)

(Continued)

**Table 3.** (Continued)

No.	Allele	Start <sup>a</sup>	End <sup>a</sup>	Peptide length	Sequence	Method used	Percentile rank	Antigenicity score	Percent of protein sequence matches at identity $\geq 100\%$
076	HLA-B*58:01	599	609	11	SLIGLTSRATW	Consensus (ann/smm)	0.95	1.7930	68.39% (106/155)

<sup>a</sup>The epitopes location in NS5 protein are from accession number: KC131142.1.

Bold and italic- indicates the percentage of epitope that is 100% conserved in more than 90% of the sequences analysed in four serotypes.

doi:10.1371/journal.pone.0138729.t003

<sup>234</sup>LEFEALGF<sup>241</sup> (96.77%, C\*04:01), <sup>361</sup>TFTNMEAQL<sup>369</sup> (94.84%, C\*04:01), <sup>515</sup>MYFHRRDL<sup>522</sup> (98.71%, C\*14:02), and <sup>515</sup>MYFHRRDLRL<sup>524</sup> (98.71%, C\*14:02).

Since the HLA allele frequencies vary among populations due to different genetic backgrounds, to design an effective vaccine, we should consider the candidate epitopes that specifically bind with the prevalent HLA molecules in the target population where the vaccine will be employed. Therefore, here we examined the population coverage of the proposed epitope vaccine candidate in a Chinese population. The results showed that the epitope <sup>92</sup>AMTDTTPF<sup>99</sup> (B\*15:01, B\*15:02, B\*46:01, C\*03:02, and C\*15:02) has the highest percentage of population coverage (47.16%), followed by <sup>200</sup>CVYNMMGKREK<sup>210</sup> (A\*11:01, 43.48%), <sup>362</sup>FTNMEAQL<sup>369</sup> (C\*03:04 and C\*08:01, 36.60%), <sup>514</sup>LMYFHRRDLRL<sup>524</sup> (C\*01:02 and C\*14:02, 33.53%), <sup>515</sup>MYFHRRDLRL<sup>524</sup> (A\*24:02 and C\*14:02, 28.69%), and <sup>514</sup>LMYFHRRDL<sup>522</sup> (C\*01:02, 27.68%) in China (Table 5). It is worth noting that the combination of the epitopes <sup>92</sup>AMTDTTPF<sup>99</sup> (B\*15:01, B\*15:02, B\*46:01, C\*03:02 and C\*15:02), <sup>200</sup>CVYNMMGKREK<sup>210</sup> (A\*11:01), <sup>362</sup>FTNMEAQL<sup>369</sup> (C\*03:04 and C\*08:01), and <sup>514</sup>LMYFHRRDLRL<sup>524</sup> (C\*01:02 and C\*14:02) could vaccinate >90% of the Chinese population, suggesting that the four epitopes are better candidates for a multiple T-cell epitope-based vaccine. These highly conserved HLA restricted epitopes with acceptable population coverage could be putative epitope vaccine candidates in their combinations to elicit DENV-specific T-cell responses. Finally, to avoid the autoimmune response, all of the predicted class I HLA-binding antigenic epitopes were analyzed for their homology with human proteome, but no epitope was homologous with human proteome. Based on these results, we proposed that the combination of these highly conserved epitopes could be as universal CD8<sup>+</sup> T-cell epitope vaccine candidates to induce DENV-specific T-cell responses against four DENV serotypes that are endemic in China.

### Conclusion

HLA-restricted epitopes for prophylactic or therapeutic vaccines against infectious diseases to induce a T-cell response that eliminates infected cells is a promising vaccine strategy. In this study, we identified 14 universal CD8<sup>+</sup> T-cell epitope candidates using immunoinformatic approach, and they are highly conserved among all four DENV serotypes that are endemic in China. The combination of four epitopes, including <sup>92</sup>AMTDTTPF<sup>99</sup> (B\*15:01, B\*15:02, B\*46:01, C\*03:02 and C\*15:02), <sup>200</sup>CVYNMMGKREK<sup>210</sup> (A\*11:01), <sup>362</sup>FTNMEAQL<sup>369</sup> (C\*03:04 and C\*08:01), and <sup>514</sup>LMYFHRRDLRL<sup>524</sup> (C\*01:02 and C\*14:02), could vaccinate >90% of individuals in China. These epitopes are valuable T-cell epitope-based vaccine candidates for the development of a universal dengue vaccine that is capable of eliciting specific and robust protective T-cell responses against four DENV serotype variants. In conclusion, our study highlights that it is possible to design an epitope-based universal vaccine against all four DENV serotypes based on protective CD8<sup>+</sup> T-cell-mediated cellular immune responses.

**Table 4. HLA-C restricted epitopes of the NS5 protein and their binding affinity, antigenicity and conservation (in percentages) in different serotypes.**

No.	Allele	Start <sup>a</sup>	End <sup>a</sup>	Peptide length	Sequence	Method used	Percentile rank	Antigenicity score	Percent of protein sequence matches at identity ≥100%
001	HLA-C*01:02	629	636	8	YTDYMPSM	netmhcpan	0.6	0.8868	10.97% (17/155)
002	HLA-C*01:02	91	99	9	MAMTDTTPF	netmhcpan	0.2	0.6735	77.42% (120/155)
003	HLA-C*01:02	85	93	9	IPMVTQMAM	netmhcpan	0.7	0.5370	10.97% (17/155)
<b>004</b>	<b>HLA-C*01:02</b>	<b>514</b>	<b>522</b>	<b>9</b>	<b>LMYFHRRDL</b>	<b>netmhcpan</b>	<b>0.8</b>	<b>1.5116</b>	<b>98.71% (153/155)</b>
005	HLA-C*01:02	84	93	10	VIPMVTQMAM	netmhcpan	0.6	0.7337	10.97% (17/155)
006	HLA-C*01:02	513	522	10	SLMYFHRRDL	netmhcpan	0.7	1.4737	80.00% (124/155)
007	HLA-C*01:02	226	236	11	YMWLGARFLEF	netmhcpan	0.8	1.0316	32.26% (50/155)
<b>008</b>	<b>HLA-C*01:02</b>	<b>514</b>	<b>524</b>	<b>11</b>	<b>LMYFHRRDLRL</b>	<b>netmhcpan</b>	<b>0.9</b>	<b>1.5932</b>	<b>98.71% (153/155)</b>
009	HLA-C*03:02	629	636	8	YTDYMPSM	netmhcpan	0.4	0.8868	10.97% (17/155)
<b>010</b>	<b>HLA-C*03:02</b>	<b>92</b>	<b>99</b>	<b>8</b>	<b>AMTDTTPF</b>	<b>netmhcpan</b>	<b>0.9</b>	<b>0.6985</b>	<b>98.06% (152/155)</b>
011	HLA-C*03:02	226	233	8	YMWLGARF	netmhcpan	1	0.8453	32.26% (50/155)
012	HLA-C*03:02	91	99	9	MAMTDTTPF	netmhcpan	0.1	0.6735	77.42% (120/155)
013	HLA-C*03:02	615	623	9	TAINQVRSL	netmhcpan	0.7	0.6391	10.97% (17/155)
014	HLA-C*03:02	545	554	10	WSIHATHEWM	netmhcpan	0.8	0.8628	9.03% (14/155)
015	HLA-C*03:02	226	236	11	YMWLGARFLEF	netmhcpan	0.5	1.0316	32.26% (50/155)
016	HLA-C*03:02	89	99	11	TQMAMTDTTPF	netmhcpan	1	0.9856	77.42% (120/155)
017	HLA-C*03:03	615	623	9	TAINQVRSL	Consensus (ann/smm)	0.6	0.6391	10.97% (17/155)
018	HLA-C*03:03	501	511	11	TACLGKSYAQM	ann	0.9	0.6954	29.03% (45/155)
019	HLA-C*03:03	474	484	11	LVVPCRNDDEL	ann	1	1.0079	14.19% (22/155)
020	HLA-C*03:04	629	636	8	YTDYMPSM	netmhcpan	0.5	0.8868	10.97% (17/155)
<b>021</b>	<b>HLA-C*03:04</b>	<b>362</b>	<b>369</b>	<b>8</b>	<b>FTNMEAQL</b>	<b>netmhcpan</b>	<b>1</b>	<b>0.7649</b>	<b>94.84% (147/155)</b>
022	HLA-C*03:04	91	99	9	MAMTDTTPF	netmhcpan	0.2	0.6735	77.42% (120/155)
023	HLA-C*03:04	615	623	9	TAINQVRSL	netmhcpan	0.4	0.6391	10.97% (17/155)
024	HLA-C*04:01	287	294	8	GDWTRITL	ann	0.2	1.6194	10.97% (17/155)
025	HLA-C*04:01	582	589	8	SWEEIPLY	ann	0.4	0.9783	10.97% (17/155)
<b>026</b>	<b>HLA-C*04:01</b>	<b>225</b>	<b>232</b>	<b>8</b>	<b>WYMWLGAR</b>	<b>ann</b>	<b>0.5</b>	<b>1.2415</b>	<b>99.35% (154/155)</b>
027	HLA-C*04:01	226	233	8	YMWLGARF	ann	0.5	0.8453	32.26% (50/155)
<b>028</b>	<b>HLA-C*04:01</b>	<b>234</b>	<b>241</b>	<b>8</b>	<b>LEFEALGF</b>	<b>ann</b>	<b>1</b>	<b>2.1459</b>	<b>96.77% (150/155)</b>
029	HLA-C*04:01	287	295	9	GDWTRITLE	Consensus (ann/smm)	0.25	1.4743	10.97% (17/155)
030	HLA-C*04:01	225	233	9	WYMWLGARF	Consensus (ann/smm)	0.75	0.9423	32.26% (50/155)
<b>031</b>	<b>HLA-C*04:01</b>	<b>361</b>	<b>369</b>	<b>9</b>	<b>TFTNMEAQL</b>	<b>Consensus (ann/smm)</b>	<b>1</b>	<b>0.8131</b>	<b>94.84% (147/155)</b>
032	HLA-C*04:01	450	458	9	WNDWTQVPF	Consensus (ann/smm)	0.6	0.9685	9.68% (15/155)
033	HLA-C*04:01	225	234	10	WYMWLGARFL	ann	0.2	0.6068	32.26% (50/155)
034	HLA-C*04:01	449	458	10	GWNDWTQVPF	ann	0.3	0.9735	9.68% (15/155)
035	HLA-C*04:01	224	233	10	IWYMWLGARF	ann	0.6	0.7264	32.26% (50/155)
036	HLA-C*04:01	223	233	11	AIWYMWLGARF	ann	0.7	0.5658	32.26% (50/155)
037	HLA-C*04:01	89	99	11	TQMAMTDTTPF	ann	0.9	0.9856	77.42% (120/155)
038	HLA-C*04:01	225	235	11	WYMWLGARFLE	ann	0.9	0.6782	32.26% (50/155)
039	HLA-C*04:01	226	236	11	YMWLGARFLEF	ann	0.9	1.0316	32.26% (50/155)
040	HLA-C*04:01	582	592	11	SWEEIPLYGKR	ann	0.9	1.2892	10.97% (17/155)
041	HLA-C*08:01	629	636	8	YTDYMPSM	netmhcpan	0.2	0.8868	10.97% (17/155)

(Continued)

Table 4. (Continued)

No.	Allele	Start <sup>a</sup>	End <sup>a</sup>	Peptide length	Sequence	Method used	Percentile rank	Antigenicity score	Percent of protein sequence matches at identity ≥100%
<b>042</b>	<b><i>HLA-C*08:01</i></b>	<b>362</b>	<b>369</b>	<b>8</b>	<b><i>FTNMEAQL</i></b>	<b><i>netmhcpn</i></b>	<b>0.9</b>	<b>0.7649</b>	<b>94.84% (147/155)</b>
043	HLA-C*08:01	91	99	9	MAMTDTTPF	netmhcpn	0.3	0.6735	77.42% (120/155)
<b>044</b>	<b><i>HLA-C*14:02</i></b>	<b>515</b>	<b>522</b>	<b>8</b>	<b><i>MYFHRRDL</i></b>	<b><i>ann</i></b>	<b>0.2</b>	<b>1.5109</b>	<b>98.71% (153/155)</b>
045	HLA-C*14:02	247	254	8	WFSRENSL	ann	0.8	0.5342	28.39% (44/155)
046	HLA-C*14:02	232	239	8	RFLEFEAL	ann	0.9	1.2524	32.26% (50/155)
047	HLA-C*14:02	43	50	8	HYDQDHPY	ann	1	0.6739	9.03% (14/155)
048	HLA-C*14:02	226	233	8	YMWLGARF	ann	1	0.8453	32.26% (50/155)
049	HLA-C*14:02	628	636	9	EYTDYMPSM	Consensus (ann/smm)	0.35	0.8846	10.97% (17/155)
050	HLA-C*14:02	225	233	9	WYMWLGARF	Consensus (ann/smm)	0.65	0.9423	32.26% (50/155)
<b>051</b>	<b><i>HLA-C*14:02</i></b>	<b>515</b>	<b>524</b>	<b>10</b>	<b><i>MYFHRRDLRL</i></b>	<b><i>ann</i></b>	<b>0.2</b>	<b>1.6075</b>	<b>98.71% (153/155)</b>
052	HLA-C*14:02	225	234	10	WYMWLGARFL	ann	0.3	0.6068	32.26% (50/155)
053	HLA-C*14:02	84	93	10	VIPMVTQMAM	ann	1	0.7337	10.97% (17/155)
054	HLA-C*14:02	226	236	11	YMWLGARFLEF	ann		1.0316	32.26% (50/155)
<b>055</b>	<b><i>HLA-C*14:02</i></b>	<b>514</b>	<b>524</b>	<b>11</b>	<b><i>LMYFHRRDLRL</i></b>	<b><i>ann</i></b>	<b>0.9</b>	<b>1.5932</b>	<b>98.71% (153/155)</b>
056	HLA-C*14:02	89	99	11	TQMAMTDTTPF	ann	1	0.9856	77.42% (120/155)
057	HLA-C*15:02	226	233	8	YMWLGARF	ann	0.2	0.8453	32.26% (50/155)
058	HLA-C*15:02	622	629	8	SLIGNEEY	ann	0.3	0.8841	10.32% (16/155)
<b>059</b>	<b><i>HLA-C*15:02</i></b>	<b>92</b>	<b>99</b>	<b>8</b>	<b><i>AMTDTTPF</i></b>	<b><i>ann</i></b>	<b>0.7</b>	<b>0.6985</b>	<b>98.06% (152/155)</b>
060	HLA-C*15:02	510	517	8	QMWSLMYF	ann	1	0.5350	80.65% (125/155)
061	HLA-C*15:02	362	370	9	FTNMEAQLI	Consensus (ann/smm)	0.25	0.7034	29.68% (46/155)
062	HLA-C*15:02	154	162	9	RSNAALGAI	Consensus (ann/smm)	0.6	1.0339	9.03% (14/155)
063	HLA-C*15:02	360	370	11	NTFTNMEAQLI	ann	0.9	0.7529	29.68% (46/155)

<sup>a</sup>The epitopes location in NS5 protein are from accession number: KC131142.1.

Bold and italic- indicates the percentage of epitope that is 100% conserved in more than 90% of the sequences analysed in four serotypes.

doi:10.1371/journal.pone.0138729.t004

Table 5. Population coverage rate (%) for the highly conserved epitopes that could be as multiple epitope-based universal vaccine candidates.

Epitope candidates	Position (aa) <sup>a</sup>	HLA class I alleles	Population coverage (%)
AMTDTTPF	92–99	HLA-B*15:01, HLA-B*15:02, HLA-B*46:01, HLA-C*03:02, HLA-C*15:02	47.16
CVYNMMGKREK	200–210	HLA-A*11:01	43.48
FTNMEAQL	362–369	HLA-C*03:04, HLA-C*08:01	36.60
LMYFHRRDLRL	514–524	HLA-C*01:02, HLA-C*14:02	33.53
MYFHRRDLRL	515–524	HLA-A*24:02, HLA-C*14:02	28.69
LMYFHRRDL	514–522	HLA-C*01:02	27.68
WYMWLGAR	225–232	HLA-A*33:03, HLA-C*04:01	18.46
LEFEALGF	234–241	HLA-B*13:01, HLA-C*04:01	17.48
YFHRRDLR	516–523	HLA-A*33:03	9.78
DTAGWDTR	284–291	HLA-A*33:03	9.78
TFTNMEAQL	361–369	HLA-C*04:01	9.62
VPTSRTTWSI	538–547	HLA-B*51:01	7.39
MYFHRRDL	515–522	HLA-C*14:02	6.91
LHKLGYIL	262–269	HLA-B*38:02	5.22

<sup>a</sup>The epitopes location in NS5 protein are from accession number: KC131142.1.

doi:10.1371/journal.pone.0138729.t005

## Supporting Information

**S1 File. All available NS5 protein sequences of dengue virus serotype 1 Chinese isolates.** (TXT)

**S2 File. All available NS5 protein sequences of dengue virus serotype 2 Chinese isolates.** (TXT)

**S3 File. All available NS5 protein sequences of dengue virus serotype 3 Chinese isolates.** (TXT)

**S4 File. All available NS5 protein sequences of dengue virus serotype 4 Chinese isolates.** (TXT)

## Author Contributions

Conceived and designed the experiments: JDS YZH. Performed the experiments: JDS JS MNW NZH JFL YHL HXW. Analyzed the data: JDS. Contributed reagents/materials/analysis tools: JS NZH. Wrote the paper: JDS.

## References

1. Kraemer MU, Sinka ME, Duda KA, Mylne AQ, Shearer FM, Barker CM, et al. The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. *Elife*. 2015; 4:e08347. doi: [10.7554/eLife.08347](https://doi.org/10.7554/eLife.08347) PMID: [26126267](https://pubmed.ncbi.nlm.nih.gov/26126267/).
2. Russell PK, Nisalak A. Dengue virus identification by the plaque reduction neutralization test. *J Immunol*. 1967; 99(2):291–6. PMID: [4961907](https://pubmed.ncbi.nlm.nih.gov/4961907/).
3. Ghosh A, Dar L. Dengue vaccines: challenges, development, current status and prospects. *Indian J Med Microbiol*. 2015; 33(1):3–15. PMID: [25559995](https://pubmed.ncbi.nlm.nih.gov/25559995/).
4. Kuno G, Chang GJ, Tsuchiya KR, Karabatsos N, Cropp CB. Phylogeny of the genus *Flavivirus*. *J Virol*. 1998; 72(1):73–83. PMID: [9420202](https://pubmed.ncbi.nlm.nih.gov/9420202/).
5. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, et al. The global distribution and burden of dengue. *Nature*. 2013; 496(7446):504–7. Epub 2013/04/07. doi: [10.1038/nature12060](https://doi.org/10.1038/nature12060) PMID: [23563266](https://pubmed.ncbi.nlm.nih.gov/23563266/).
6. Qiu FX, Gubler DJ, Liu JC, Chen QQ. Dengue in China: a clinical review. *Bull World Health Organ*. 1993; 71(3–4):349–59. PMID: [8324854](https://pubmed.ncbi.nlm.nih.gov/8324854/).
7. Qiu FX, Chen QQ, Ho QY, Chen WZ, Zhao ZG, Zhao BW. The first epidemic of dengue hemorrhagic fever in the People's Republic of China. *Am J Trop Med Hyg*. 1991; 44(4):364–70. PMID: [2042704](https://pubmed.ncbi.nlm.nih.gov/2042704/).
8. National Health and Family Planning Commission of the People's Republic of China. The report of the progress of dengue fever control and prevention. Available: <http://www.nhfpc.gov.cn/>. Accessed 2 August 2015.
9. Luo L, Liang HY, Hu YS, Liu WJ, Wang YL, Jing QL, et al. Epidemiological, virological, and entomological characteristics of dengue from 1978 to 2009 in Guangzhou, China. *J Vector Ecol*. 2012; 37(1):230–40. doi: [10.1111/j.1948-7134.2012.00221.x](https://doi.org/10.1111/j.1948-7134.2012.00221.x) PMID: [22548558](https://pubmed.ncbi.nlm.nih.gov/22548558/).
10. Xu G, Dong H, Shi N, Liu S, Zhou A, Cheng Z, et al. An outbreak of dengue virus serotype 1 infection in Cixi, Ningbo, People's Republic of China, 2004, associated with a traveler from Thailand and high density of *Aedes albopictus*. *Am J Trop Med Hyg*. 2007; 76(6):1182–8. PMID: [17556633](https://pubmed.ncbi.nlm.nih.gov/17556633/).
11. Wang B, Li Y, Feng Y, Zhou H, Liang Y, Dai J, et al. Phylogenetic analysis of dengue virus reveals the high relatedness between imported and local strains during the 2013 dengue outbreak in Yunnan, China: a retrospective analysis. *BMC Infect Dis*. 2015; 15:142. doi: [10.1186/s12879-015-0908-x](https://doi.org/10.1186/s12879-015-0908-x) PMID: [25885572](https://pubmed.ncbi.nlm.nih.gov/25885572/).
12. Mongkolsapaya J, Dejnirattisai W, Xu XN, Vasanawathana S, Tangthawornchaikul N, Chairunsri A, et al. Original antigenic sin and apoptosis in the pathogenesis of dengue hemorrhagic fever. *Nat Med*. 2003; 9(7):921–7. PMID: [12808447](https://pubmed.ncbi.nlm.nih.gov/12808447/).
13. Duangchinda T, Dejnirattisai W, Vasanawathana S, Limpitikul W, Tangthawornchaikul N, Malasit P, et al. Immunodominant T-cell responses to dengue virus NS3 are associated with DHF. *Proc Natl Acad Sci U S A*. 2010; 107(39):16922–7. Epub 2010/09/13. doi: [10.1073/pnas.1010867107](https://doi.org/10.1073/pnas.1010867107) PMID: [20837518](https://pubmed.ncbi.nlm.nih.gov/20837518/).



14. Rothman AL. Immunity to dengue virus: a tale of original antigenic sin and tropical cytokine storms. *Nat Rev Immunol.* 2011; 11(8):532–43. doi: [10.1038/nri3014](https://doi.org/10.1038/nri3014) PMID: [21760609](https://pubmed.ncbi.nlm.nih.gov/21760609/).
15. Zellweger RM, Miller R, Eddy WE, White LJ, Johnston RE, Shresta S. Role of humoral versus cellular responses induced by a protective dengue vaccine candidate. *PLoS Pathog.* 2013; 9(10):e1003723. Epub 2013/10/31. doi: [10.1371/journal.ppat.1003723](https://doi.org/10.1371/journal.ppat.1003723) PMID: [24204271](https://pubmed.ncbi.nlm.nih.gov/24204271/).
16. Yauch LE, Zellweger RM, Kotturi MF, Qutubuddin A, Sidney J, Peters B, et al. A protective role for dengue virus-specific CD8<sup>+</sup> T cells. *J Immunol.* 2009; 182(8):4865–73. doi: [10.4049/jimmunol.0801974](https://doi.org/10.4049/jimmunol.0801974) PMID: [19342665](https://pubmed.ncbi.nlm.nih.gov/19342665/).
17. Yauch LE, Prestwood TR, May MM, Morar MM, Zellweger RM, Peters B, et al. CD4<sup>+</sup> T cells are not required for the induction of dengue virus-specific CD8<sup>+</sup> T cell or antibody responses but contribute to protection after vaccination. *J Immunol.* 2010; 185(9):5405–16. Epub 2010/09/24. doi: [10.4049/jimmunol.1001709](https://doi.org/10.4049/jimmunol.1001709) PMID: [20870934](https://pubmed.ncbi.nlm.nih.gov/20870934/).
18. Prestwood TR, Morar MM, Zellweger RM, Miller R, May MM, Yauch LE, et al. Gamma interferon (IFN- $\gamma$ ) receptor restricts systemic dengue virus replication and prevents paralysis in IFN- $\alpha/\beta$  receptor-deficient mice. *J Virol.* 2012; 86(23):12561–70. Epub 2012/09/12. doi: [10.1128/JVI.06743-11](https://doi.org/10.1128/JVI.06743-11) PMID: [22973027](https://pubmed.ncbi.nlm.nih.gov/22973027/).
19. Zellweger RM, Eddy WE, Tang WW, Miller R, Shresta S. CD8<sup>+</sup> T cells prevent antigen-induced antibody-dependent enhancement of dengue disease in mice. *J Immunol.* 2014; 193(8):4117–24. Epub 2014/09/12. doi: [10.4049/jimmunol.1401597](https://doi.org/10.4049/jimmunol.1401597) PMID: [25217165](https://pubmed.ncbi.nlm.nih.gov/25217165/).
20. Zompi S, Santich BH, Beatty PR, Harris E. Protection from secondary dengue virus infection in a mouse model reveals the role of serotype cross-reactive B and T cells. *J Immunol.* 2012; 188(1):404–16. Epub 2011/11/30. doi: [10.4049/jimmunol.1102124](https://doi.org/10.4049/jimmunol.1102124) PMID: [22131327](https://pubmed.ncbi.nlm.nih.gov/22131327/).
21. Simmons CP, Dong T, Chau NV, Dung NT, Chau TN, Thao le TT, et al. Early T-cell responses to dengue virus epitopes in Vietnamese adults with secondary dengue virus infections. *J Virol.* 2005; 79(9):5665–75. PMID: [15827181](https://pubmed.ncbi.nlm.nih.gov/15827181/).
22. Weiskopf D, Angelo MA, de Azeredo EL, Sidney J, Greenbaum JA, Fernando AN, et al. Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8<sup>+</sup> T cells. *Proc Natl Acad Sci U S A.* 2013; 110(22):E2046–53. Epub 2013/04/11. doi: [10.1073/pnas.1305227110](https://doi.org/10.1073/pnas.1305227110) PMID: [23580623](https://pubmed.ncbi.nlm.nih.gov/23580623/).
23. Capeding MR, Tran NH, Hadinegoro SR, Ismail HI, Chotpitayasunondh T, Chua MN, et al. Clinical efficacy and safety of a novel tetravalent dengue vaccine in healthy children in Asia: a phase 3, randomised, observer-masked, placebo-controlled trial. *Lancet.* 2014; 384(9951):1358–65. Epub 2014/07/10. doi: [10.1016/S0140-6736\(14\)61060-6](https://doi.org/10.1016/S0140-6736(14)61060-6) PMID: [25018116](https://pubmed.ncbi.nlm.nih.gov/25018116/).
24. Chambers TJ, Hahn CS, Galler R, Rice CM. Flavivirus genome organization, expression, and replication. *Annu Rev Microbiol.* 1990; 44:649–88. PMID: [2174669](https://pubmed.ncbi.nlm.nih.gov/2174669/).
25. Kuhn RJ, Zhang W, Rossmann MG, Pletnev SV, Corver J, Lenches E, et al. Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. *Cell.* 2002; 108(5):717–25. PMID: [11893341](https://pubmed.ncbi.nlm.nih.gov/11893341/).
26. Rivino L, Kumaran EA, Jovanovic V, Nadua K, Teo EW, Pang SW, et al. Differential targeting of viral components by CD4<sup>+</sup> versus CD8<sup>+</sup> T lymphocytes in dengue virus infection. *J Virol.* 2013; 87(5):2693–706. Epub 2012/11/19. doi: [10.1128/JVI.02675-12](https://doi.org/10.1128/JVI.02675-12) PMID: [23255803](https://pubmed.ncbi.nlm.nih.gov/23255803/).
27. Koonin EV. Computer-assisted identification of a putative methyltransferase domain in NS5 protein of flaviviruses and lambda 2 protein of reovirus. *J Gen Virol.* 1993; 74 (Pt 4):733–40. PMID: [8385698](https://pubmed.ncbi.nlm.nih.gov/8385698/).
28. Molero-Abraham M, Lafuente EM, Reche P. Customized predictions of peptide-MHC binding and T-cell epitopes using EPIMHC. *Methods Mol Biol.* 2014; 1184:319–32. doi: [10.1007/978-1-4939-1115-8\\_18](https://doi.org/10.1007/978-1-4939-1115-8_18) PMID: [25048133](https://pubmed.ncbi.nlm.nih.gov/25048133/).
29. Cao K, Hollenbach J, Shi X, Shi W, Chopek M, Fernández-Viña MA. Analysis of the frequencies of HLA-A, B, and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Hum Immunol.* 2001; 62(9):1009–30. PMID: [11543903](https://pubmed.ncbi.nlm.nih.gov/11543903/).
30. Duan ZL, Li Q, Wang ZB, Xia KD, Guo JL, Liu WQ, et al. HLA-A\*0201-restricted CD8<sup>+</sup> T-cell epitopes identified in dengue viruses. *Virol J.* 2012; 9:259. doi: [10.1186/1743-422X-9-259](https://doi.org/10.1186/1743-422X-9-259) PMID: [23121866](https://pubmed.ncbi.nlm.nih.gov/23121866/).
31. Tian J, Zeng G, Pang X, Liang M, Zhou J, Fang D, et al. Identification and immunogenicity of two new HLA-A\*0201-restricted CD8<sup>+</sup> T-cell epitopes on dengue NS1 protein. *Int Immunol.* 2012; 24(4):207–18. Epub 2012/01/31. doi: [10.1093/intimm/dxr115](https://doi.org/10.1093/intimm/dxr115) PMID: [22298881](https://pubmed.ncbi.nlm.nih.gov/22298881/).
32. Duan Z, Guo J, Huang X, Liu H, Chen X, Jiang M, et al. Identification of cytotoxic T lymphocyte epitopes in dengue virus serotype 1. *J Med Virol.* 2015; 87(7):1077–89. Epub 2015/03/16. doi: [10.1002/jmv.24167](https://doi.org/10.1002/jmv.24167) PMID: [25777343](https://pubmed.ncbi.nlm.nih.gov/25777343/).

33. Townsley E, Woda M, Thomas SJ, Kalayanarooj S, Gibbons RV, Nisalak A, et al. Distinct activation phenotype of a highly conserved novel HLA-B57-restricted epitope during dengue virus infection. *Immunology*. 2014; 141(1):27–38. doi: [10.1111/imm.12161](https://doi.org/10.1111/imm.12161) PMID: [23941420](https://pubmed.ncbi.nlm.nih.gov/23941420/).
34. Chang CX, Tan AT, Or MY, Toh KY, Lim PY, Chia AS, et al. Conditional ligands for Asian HLA variants facilitate the definition of CD8<sup>+</sup> T-cell responses in acute and chronic viral diseases. *Eur J Immunol*. 2013; 43(4):1109–20. Epub 2013/02/04. doi: [10.1002/eji.201243088](https://doi.org/10.1002/eji.201243088) PMID: [23280567](https://pubmed.ncbi.nlm.nih.gov/23280567/).
35. Zivna I, Green S, Vaughn DW, Kalayanarooj S, Stephens HA, Chandanayingyong D, et al. T cell responses to an HLA-B\*07-restricted epitope on the dengue NS3 protein correlate with disease severity. *J Immunol*. 2002; 168(11):5959–65. PMID: [12023403](https://pubmed.ncbi.nlm.nih.gov/12023403/).
36. Okamoto Y, Kurane I, Leporati AM, Ennis FA. Definition of the region on NS3 which contains multiple epitopes recognized by dengue virus serotype-cross-reactive and flavivirus-cross-reactive, HLA-DPw2-restricted CD4<sup>+</sup> T cell clones. *J Gen Virol*. 1998; 79 (Pt 4):697–704. PMID: [9568963](https://pubmed.ncbi.nlm.nih.gov/9568963/).
37. 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. Epub 2012/05/18. doi: [10.1093/nar/gks438](https://doi.org/10.1093/nar/gks438) PMID: [22610854](https://pubmed.ncbi.nlm.nih.gov/22610854/).
38. Nielsen M, Lundegaard C, Wornig P, Lauemøller SL, Lamberth K, Buus S, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci*. 2003; 12(5):1007–17. PMID: [12717023](https://pubmed.ncbi.nlm.nih.gov/12717023/).
39. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, Nielsen M. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8–11. *Nucleic Acids Res*. 2008; 36(Web Server issue):W509–12. Epub 2008/05/07. doi: [10.1093/nar/gkn202](https://doi.org/10.1093/nar/gkn202) PMID: [18463140](https://pubmed.ncbi.nlm.nih.gov/18463140/).
40. Peters B, Sette A. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics*. 2005; 6:132. PMID: [15927070](https://pubmed.ncbi.nlm.nih.gov/15927070/).
41. Sidney J, Assarsson E, Moore C, Ngo S, Pinilla C, Sette A, et al. Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. *Immunome Res*. 2008; 4:2. doi: [10.1186/1745-7580-4-2](https://doi.org/10.1186/1745-7580-4-2) PMID: [18221540](https://pubmed.ncbi.nlm.nih.gov/18221540/).
42. Yao Y, Huang W, Yang X, Sun W, Liu X, Cun W, et al. HPV-16 E6 and E7 protein T cell epitopes prediction analysis based on distributions of HLA-A loci across populations: an in silico approach. *Vaccine*. 2013; 31(18):2289–94. Epub 2013/03/13. doi: [10.1016/j.vaccine.2013.02.065](https://doi.org/10.1016/j.vaccine.2013.02.065) PMID: [23499609](https://pubmed.ncbi.nlm.nih.gov/23499609/).
43. Gededzha MP, Mphahlele MJ, Selabe SG. Prediction of T-cell epitopes of hepatitis C virus genotype 5a. *Virol J*. 2014; 11:187. doi: [10.1186/1743-422X-11-187](https://doi.org/10.1186/1743-422X-11-187) PMID: [25380768](https://pubmed.ncbi.nlm.nih.gov/25380768/).