

Preplanned Studies

SARS-CoV-2 Omicron Variant is Expected to Retain Most of the Spike Protein Specific Dominant T-Cell Epitopes Presented by COVID-19 Vaccines — Worldwide, 2021

Jiajing Jiang¹; Yingying Du^{1,*}; Tao Peng^{2,*}

Summary

What is already known about this topic?

The newly emerged variant of Omicron, which carries many of the mutations found in other variants of concern (VOCs), as well as a great number of new mutations that may enhance its immune escape, has spread rapidly around the world. This has raised public concern about the effectiveness of the current coronavirus disease 2019 (COVID-19) vaccine.

What is added by this report?

In this study, different bioinformatic softwares were applied to predict the dominant Omicron spike (S) protein cytotoxic T lymphocyte (CTL) and T helper (Th) epitopes in representative world population and Chinese population. Compared to the original severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) S protein, limited mutations were identified within the dominant CTL and Th epitopes in Omicron variant.

What are the implications for public health practice?

The results of this study suggested that the current COVID-19 vaccine-induced T-cell immunity may still provide significant protection against Omicron variant infection in fully vaccinated individuals.

The World Health Organization (WHO) categorized the new B.1.1.529 severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variant (Omicron) as a variant of concern (VOC) on November 26, 2021. Since then, this newly emerged variant has placed the world on high alert. Compared to other VOCs, Omicron has an unusual constellation of mutations. It contains over 50 mutations in various locations of its genome. In the spike protein gene, Omicron has over 30 mutations, doubling the number associated with the Delta variant. In the receptor binding domain (RBD) alone, Omicron has over 10 mutations while Delta only has 2. Many of these

mutations have been shown to enhance the interaction between the viral spike and the cellular receptor angiotensin-converting enzyme 2 (ACE2). Based on this observation, it is predicted that the Omicron variant could be highly contagious.

In general, a vaccine provides two arms of immune protection. On one hand, the vaccine results in the production of neutralizing antibodies (NAbs) by B cells. These NAbs bind to the spike protein of the virus and inhibit its ability to infect the host cells. For previously immunized people, the memory humoral immune response is the first line of defense against Omicron infection. Unfortunately, Omicron variant carries many mutations on the spike protein that neutralizing antibodies recognize, reducing vaccinated individuals' immunity to this variant. The human body's second line of defense is heavily reliant on the human leukocyte antigen (HLA)-restricted T-cell response mechanism, in which viral epitopes are presented by dendritic cells to CD8+ T lymphocytes through interactions with HLA class I alleles and CD4+ T lymphocytes through HLA class II alleles. Viral epitope presentation by HLA class I leads to clonal expansion of HLA-restricted CD8+ cytotoxic T lymphocytes (CTLs), which are primed to perform antiviral defense during acute infection. Subsequently, reinfection of the virus is controlled by memory CTLs. The recognition of viral epitope-HLA class II complexes by CD4+ T cell enhances cell-mediated immune response by inducing cytokines and facilitates antibody production by activating B cells. The urgent question right now is how many of the memory CTLs and T helper (Th) epitopes remain in the heavily mutated Omicron variant.

T-cell epitopes were predicted in this paper using IEDB recommended 2020.09 and SYFPEITHI for major histocompatibility complex Class I (MHC Class I), and IEDB recommended 2.22 and Propred for MHC Class II, before real-world data were available to answer this question.

Wuhan-Hu-1 (NCBI Reference Sequence: YP_009724390.1) was obtained from the GenBank database of the National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/>). EPI_ISL:EPI_ISL_6640916 sequence (Omicron) was obtained from the GISAID database (<https://www.gisaid.org/>).

IEDB recommended 2020.09 (NetMHCpan 4.1EL) (<http://tools.immuneepitope.org/mhci/>, National Institute of Allergy and Infectious Diseases, USA) and SYFPEITHI (<http://www.syfpeithi.de/>, BMI Biomedical Informatics, Heidelberg, Germany) were used to predict HLA-A*02:01, HLA-A*11:01 restricted epitopes. NetMHCpan4.1 based on artificial neural network (ANN) applied binding affinity (BA) and mass spectrometry (MS) eluted ligands (EL) data as a model to analyze the affinity between target peptides and their ligands (MHC). This algorithm not only integrates affinity data and mass spectrometry eluted ligand data, but also covers information in the process of antigen processing and presentation (1–2). The smaller the value, the higher the prediction score of the corresponding random natural peptide. SYFPEITHI is based on the motif matrix algorithm (3), according to the natural ligands, T-cell epitopes, or the frequency of amino acids in the binding peptides to predict the target peptides. The anchors and their auxiliary anchors that appear frequently get higher scores in the prediction results. The higher the score, the greater the possibility that the peptide will become an antigenic peptide.

IEDB recommended 2.22 (<http://tools.immuneepitope.org/mhcii/>, National Institute of Allergy and Infectious Diseases, USA) and Propred (<http://www.imtech.res.in/raghava/propred/>, Department of Computational Biology, Indraprastha Institute of Information Technology, New Delhi, India) were used to predict HLA-DRB1*01:01 and HLA-DRB1*15:01 restricted epitopes. IEDB recommended 2.22 apply Consensus combined with Combinatorial library (4), SMM-align (5), NN-align (6) to analyze the target peptides and combine their predicted values to obtain more accurate results. Propred based on an algorithm called quantitative affinity matrix (QAM) predicted the score of the sequence by comparing the degree of match between the target sequence and the binding pocket of HLA (7).

In this study, we analyzed 4 representative HLA alleles (including HLA-A*02:01, HLA-A*11:01, HLA-DRB1*01:01, and HLA-DRB1*15:01) restricted peptides among which HLA-A*02:01 and HLA-

DRB1*01:01 represented the dominant HLA alleles in the world population, HLA-A*11:01 and HLA-DRB1*15:01 represent the dominant HLA alleles in the Chinese population (8–9). Various bio-information software was used to predict the HLA restricted CTL or Th epitopes, and the top ten dominant epitopes were finally screened. The comparison between Wuhan-Hu-1 and Omicron was performed.

There were 17 dominant HLA-A*02:01 restricted CTL epitopes (ranking in the top 10) derived from spike (S) protein of Wuhan-Hu-1, among which 2 epitopes mutated in Omicron variant (Table 1). These mutations in Omicron variant exhibited minor decline in epitope rank and may slightly impair corresponding CTL response in viral clearance. However, 15 dominant HLA-DRB1*01:01 restricted Th epitopes in S protein of Wuhan-Hu-1 (ranking in the top ten) remain consistent in Omicron (Table 1). These results indicate that dominant protective CTL and Th epitopes derived from original Wuhan-Hu-1 strain or vaccines constructed based on Wuhan-Hu-1 still provided good T cell protection for convalescents or vaccinated individuals, although limited mutations were found within dominant CTL epitopes in a large proportion of the world's population.

There were 12 dominant HLA-A*11:01-restricted CTL epitopes (ranking in the top 10) in the S protein of Wuhan-Hu-1, of which 3 epitopes mutated in Omicron variant (Table 2). Similarly, three of the top-ten dominant HLA-DRB1*15:01 restricted Th epitopes of Wuhan-Hu-1 strain S protein mutated in Omicron variant (Table 2). The mutations in the CTL epitopes may lead to a decrease in CTL against Omicron variant, while the mutated Th epitopes, may alter its capacity to promote T cell mediated immune response via cytokines and antibody generation via B cell activation, thus may have some impact on protective effect of vaccine against Omicron variant.

DISCUSSION

Since the WHO designated Omicron as a VOC, Omicron has caused concern for the world for its significant transmissibility and infectivity. Previously, Delta variant, also classified as VOC, had been spreading rapidly worldwide and causing serious outbreaks due to its high transmissibility, short incubation period, and high viral load (10). Reduced levels of neutralizing antibody in serum against the Delta variant in vaccinated people and patients who recovered from coronavirus disease 2019 (COVID-19)

TABLE 1. Comparison of HLA-A*02:01-restricted CTL epitopes and HLA-DRB1*01:01-restricted Th epitopes in Wuhan-Hu-1 S protein and Omicron S protein.

HLA-A*02:01 CTL epitopes derived from S Protein	Epitope rank in Wuhan-Hu-1 [†]	Epitope mutation exist in Omicron	HLA-DRB1*01:01 Th Epitopes derived from S protein	Epitope rank in Wuhan-Hu-1 [§]	Epitope mutation exist in Omicron
YLQPRTFLL	1	No	MFVFLVLLPLVSSQC	1	No
VLNDILSR	1	L981F	FVFLVLLPLVSSQCV	1	No
TLDSKTQSL	2	No	VVLSFELLHAPATVC	2	No
KIADYNYKL	2	K417N	VLSFELLHAPATVCG	2	No
RLDKVEAEV	3	No	LSFELLHAPATVCGP	2	No
ALNTLVKQL	3	No	VVLSFELLHAPATV	3	No
FIAGLIAIV	3	No	SFELLHAPATVCGPK	3	No
RLQSLQTYV	4	No	RVVLSFELLHAPAT	4	No
NLNESLIDL	4	No	VFLVLLPLVSSQCVN	5	No
LLFNKVTLA	5	No	FELLHAPATVCGPKK	6	No
SIAYTMSL	5	No	FLVLLPLVSSQCVNL	7	No
RLNEVAKNL	6	No	ITRFQTLALHRSYL	8	No
VVFLHVTYV	7	No	TRFQTLALHRSYLT	8	No
HLMSFPQSA	8	No	GWTFGAGAALQIPFA	9	No
GLTVLPPLL	8	No	SGWTFGAGAALQIPF	10	No
VLYENQKLI	9	No			
YQDVNCTEV	10	No			

Abbreviations: S=spike; HLA=human leukocyte antigen; CTL=cytotoxic T lymphocytes; Th=T helper.

[†] This rank is derived from the combined ranking of IEDB recommended 2020.09 and SYFPEITHI.

[§] This rank is derived from the combined ranking of IEDB recommended 2.22 and ProPred.

TABLE 2. Comparison of HLA-A*11:01-restricted CTL epitopes and HLA-DRB1*15:01-restricted Th epitopes in Wuhan-Hu-1 S protein and Omicron S protein.

HLA-A*11:01 CTL epitopes derived from S protein	Epitope rank in Wuhan-Hu-1 [†]	Epitope mutation exist in Omicron	HLA-DRB1*15:01 Th Epitopes derived from S protein	Epitope rank in Wuhan-Hu-1 [§]	Epitope mutation exist in Omicron
SVLNDILSR	1	L981F	PTESIVRFPNITNLC	1	No
GVIYFASTK	2	T95I	CSNLLLQYGSFCTQL	2	No
ASANLAATK	2	No	QPTESIVRFPNITNL	2	No
VTYVPAQEK	3	No	TESIVRFPNITNLCP	2	No
SSTASALGK	4	No	SNLLLQYGSFCTQLN	3	N764K
TLKSFTVEK	5	No	ECSNLLLQYGSFCTQ	4	No
GTHWFVTQR	6	No	ESIVRFPNITNLCPF	4	No
NSASFSTFK	7	S371L/S373P/S375F	NLLLQYGSFCTQLNR	5	N764K
GVLTESNKK	8	No	TECSNLLLQYGSFCT	6	No
GVIYHKNNK	8	No	LLLQYGSFCTQLNRA	7	N764K
QIYKTPPIK	9	No	LTDEMIAQYTSALLA	8	No
EILPVSMTK	10	No	DEMIAQYTSALLAGT	9	No
			TDEMIAQYTSALLAG	10	No

Abbreviations: S=spike; HLA=human leukocyte antigen; CTL=cytotoxic T-lymphocytes; Th=T helper.

[†] This rank is derived from the combined ranking of IEDB recommended 2020.09 and SYFPEITHI.

[§] This rank is derived from the combined ranking of IEDB recommended 2.22 and ProPred.

have also been reported to contribute to severe outbreaks.

Vaccines are the safest option to achieve herd protection during the SARS-CoV-2 pandemic. Overall, 43% of the global population has been fully vaccinated with COVID-19 vaccine so far. Although some populations had changes in the dominant T-cell epitopes due to Omicron mutation, based on the results from this analysis, more than 70% of the dominant epitopes were still retained. This result predicted that vaccine-induced T-cell immunity may still provide good protection when fully vaccinated individuals are exposed to the Omicron variant.

Some of the mutations in Omicron S protein that cause the changes of the dominant epitopes also exist in other VOCs (for example, T95I of Delta variant and K417N of Beta variant). Reduced recognition capacity of memory T cell to viral epitope-MHC complex due to the mutations may contribute to the immune escape of these variants. Thus, for the next generation of COVID-19 vaccine, the selection of cross-variant conserved T-cell epitopes should be included in the antigen design.

Despite the advancements in bioinformatics software for epitope prediction in recent years, predictions can still be inaccurate. We chose to target HLA-A*11:01/HLA-DRB*15:01 and HLA-A*02:01/HLA-DRB1*01:01 in this study, therefore the prediction results can only be instructive for these populations. The results of this study indicated that as the SARS-CoV-2 virus mutate, variation among T cell epitopes is limited in much of the world's population (representative alleles: HLA-A*02:01 and HLA-DRB1*01:01). Since the mutation rate is up to 23%–25%, mutations in T cell epitopes in the Chinese population (representative alleles: HLA-A*11:01 and HLA-DRB*15:01) should be noticed.

Acknowledgements: Hongzhou Chen, Yanting Lan, and Leyi Li.

Funding: Supported by Guangzhou Science and Technology Innovation Committee (CYLJTD-201602); and the Guangzhou Institute of Respiratory Health Open Project (Funds Provided by China Evergrande Group:2020GIRHHMS20); and the Guangzhou Laboratory (EKPG21-26); and the

Science, Technology, and Information Bureau of Guangzhou Development District (2018-L08).

doi: 10.46234/ccdcw2022.065

Corresponding authors: Yingying Du, yingyingdu@gzhmu.edu.cn; Tao Peng, pengtao@gibh.ac.cn.

¹ State Key Laboratory of Respiratory Disease, Sino-French Hoffmann Institute, College of Basic Medical Science, Guangzhou Medical University, Guangzhou, Guangdong, China; ² Guangdong South China Vaccine, Guangzhou, Guangdong, China.

Submitted: December 08, 2021; Accepted: March 15, 2022

REFERENCES

1. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* 2020;48(W1):W449 – 54. <http://dx.doi.org/10.1093/nar/gkaa379>.
2. 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. <http://dx.doi.org/10.1007/s00251-008-0341-z>.
3. Rammensee HG, Bachmann J, Emmerich NPN, Bachor OA, Stevanović S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* 1999;50(3 – 4):213 – 9. <http://dx.doi.org/10.1007/s002510050595>.
4. 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. <http://dx.doi.org/10.1186/1745-7580-4-2>.
5. Nielsen M, Lundegaard C, Lund O. Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinformatics* 2007;8:238. <http://dx.doi.org/10.1186/1471-2105-8-238>.
6. Nielsen M, Lund O. *NN-align*. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinformatics* 2009;10:296. <http://dx.doi.org/10.1186/1471-2105-10-296>.
7. Singh H, Raghava GPS. ProPred: prediction of HLA-DR binding sites. *Bioinformatics* 2001;17(12):1236 – 7. <http://dx.doi.org/10.1093/bioinformatics/17.12.1236>.
8. Pajot A, Michel ML, Fazilleau N, Pancré V, Auriault C, Ojcius DM, et al. A mouse model of human adaptive immune functions: *HLA-A2.1/HLA-DRI*-transgenic *H-2 class I/class II*-knockout mice. *Eur J Immunol* 2004;34(11):3060 – 9. <http://dx.doi.org/10.1002/eji.200425463>.
9. He Y, Li J, Mao W, Zhang D, Liu M, Shan X, et al. HLA common and well-documented alleles in China. *HLA* 2018;92(4):199 – 205. <http://dx.doi.org/10.1111/tan.13358>.
10. Allen H, Vusirikala A, Flannagan J, Twohig KA, Zaidi A, Chudasama D, et al. Household transmission of COVID-19 cases associated with SARS-CoV-2 Delta variant (B.1.617.2): national case-control study. *Lancet Reg Health Eur* 2022;12:100252. <http://dx.doi.org/10.1016/j.lanepe.2021.100252>.