Brief Report

Identification of Potential SARS-CoV-2 CD8+ T Cell Escape Mutants

Syed Faraz Ahmed [1,2,†] , Muhammad Saqib Sohail [2,†] , Ahmed Abdul Quadeer [2,*] and Matthew R. McKay [1,2,3,4,*]

1 Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, VIC 3010, Australia; faraz.ahmed@unimelb.edu.au
2 Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China; mssohail@connect.ust.hk
3 Department of Microbiology and Immunology, The Peter Doherty Institute for Infection and Immunity, University of Melbourne, Melbourne, VIC 3000, Australia
4 Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China
* Correspondence: eeaaquadeer@ust.hk (A.A.Q.); matthew.mckay@unimelb.edu.au (M.R.M.)
† These authors contributed equally to the work.

Abstract: Memory SARS-CoV-2-specific CD8+ T cell responses induced upon infection or COVID-19 vaccination have been important for protecting against severe COVID-19 disease while being largely robust against variants of concern (VOCs) observed so far. However, T cell immunity may be weakened by genetic mutations in future SARS-CoV-2 variants that lead to widespread T cell escape. The capacity for SARS-CoV-2 mutations to escape memory T cell responses requires comprehensive experimental investigation, though this is prohibited by the large number of SARS-CoV-2 mutations that have been observed. To guide targeted experimental studies, here we provide a screened list of potential SARS-CoV-2 T cell escape mutants. These mutants are identified as candidates for T cell escape as they lie within CD8+ T cell epitopes that are commonly targeted in individuals and are predicted to abrogate HLA–peptide binding.

Keywords: SARS-CoV-2; COVID-19; CD8+ T cell epitopes; immunoprevalent; immune escape; mutations; variants of concern

1. Introduction

SARS-CoV-2-specific neutralizing antibodies and T cells are generated in response to COVID-19 infection and vaccines. The level of immunity that they provide is challenged by the evolution of SARS-CoV-2 and the rise of new genetic variants such as Omicron (see [1] for details on SARS-CoV-2 variants). This presents a key issue for neutralizing antibodies for which genetic mutations enabling escape have been widely reported (e.g., [2–4]). For SARS-CoV-2-specific T cell responses, multiple studies have suggested that they are largely robust to mutations observed in the variants of concern (VOCs) so far [5–8]. However, in a small fraction of recovered and vaccinated individuals, loss of CD8+ T cell responses has been reported against variants such as Omicron [8,9]. Moreover, for some specific CD8+ T cell epitopes, it has been reported that mutations have weakened or abrogated T cell recognition in some individuals [10,11]. These recent immunological studies [8–11] indicate at least some capacity for SARS-CoV-2 to escape T cell responses.

CD8+ T cell responses present an important weapon against COVID-19, having been associated with limiting disease severity in infected [12] as well as vaccinated individuals [13]. Mutations that facilitate T cell escape could potentially compromise the protection provided by COVID-19 vaccines and infection-induced immunity against severe disease [14,15]. However, the degree to which circulating SARS-CoV-2 viruses may evade CD8+ T cell

responses is not yet well understood and requires detailed experimental investigation. Here, we report results to aid such studies by identifying and screening mutations observed among sequenced SARS-CoV-2 viruses that may facilitate CD8[+] T cell escape.

2. Materials and Methods

2.1. Acquisition of SARS-CoV-2 Epitope Data

We obtained SARS-CoV-2 CD8[+] T cell epitope data from the dashboard reported in [5] (https://www.mckayspcb.com/SARS2TcellEpitopes/; accessed on 1 February 2022). This dashboard provides up-to-date information on SARS-CoV-2 epitopes targeted by T cells in COVID-19 recovered individuals. The downloaded data comprised a total of 753 distinct HLA-specific CD8[+] T cell epitopes.

2.2. Acquisition of SARS-CoV-2 Sequence Data

SARS-CoV-2 sequence data were obtained from the GISAID database (https://www.gisaid.org/) accessed on 6 September 2021. Only the complete (full genome) sequences derived from human hosts with high coverage were downloaded. All (2,436,276) downloaded sequences were aligned to the SARS-CoV-2 reference genome (GenBank: NC_045512.2) using MAFFT [16]. The aligned genomic sequences were then translated using an in-house code to obtain the protein sequences. The positions of the open reading frames provided with the reference sequence were used to identify the respective protein regions of the full genome.

2.3. Identification of Epitope Mutants

For each SARS-CoV-2 epitope, we screened the processed sequence data at the epitope location and defined an *epitope mutant* as any epitope sequence which differed from the reported sequence of that epitope. Epitope mutants comprising any ambiguous amino acid were discarded. The count of each identified epitope mutant, i.e., the number of sequences comprising the epitope mutant, was also recorded.

2.4. Prediction of Peptide–HLA Binding Prediction

To understand if an identified epitope mutant impacts the binding of the associated epitope with its cognate HLA allele, we used the well-known peptide–HLA binding prediction method, NetMHCpan4.1 [17]. All predictions were made using the default parameters of the prediction method.

3. Results and Discussion

To identify potential CD8[+] T cell escape mutants, we analyzed mutations lying within the 753 experimentally determined SARS-CoV-2 epitopes based on ~2.5 million SARS-CoV-2 sequences (Materials and Methods). This analysis revealed over 35,000 epitope mutants across proteins. Comprehensive experimental testing of these mutants for their potential to evade CD8[+] T cell responses is prohibitive, necessitating further screening. A more refined set was obtained by considering those mutants that are associated with epitopes that are commonly targeted globally and which thereby have the potential to result in widespread T cell escape. A list of 20 such immunoprevalent SARS-CoV-2 CD8[+] T cell epitopes was reported in [5], which are collectively estimated to be targeted by at least 85% of the global population based on their HLA alleles association [18]. Individually, the global coverage of the HLA alleles associated with the immunoprevalent epitopes ranged from 13% to 39%. An important feature of these epitopes is that T cell responses against them were recorded in multiple studies across different geographical locations and in more than half of the COVID-19 recovered individuals tested [5]. These 20 immunoprevalent T cell epitopes have been remarkably conserved so far and do not comprise any mutation specific to VOCs observed to date [7]. Restricted to this immunoprevalent epitope set, the number of HLA-specific epitope mutants was reduced to 914.

To determine the likely escape mutations within immunoprevalent epitopes, a further screening step can be performed using tools to predict the impact of mutations on peptide–HLA binding (a prerequisite for T cell recognition [19]). For this purpose, we employed NetMHCpan4.1 [17], a computational method whose predictions have lined up well with SARS-CoV-2 experimental studies describing CD8[+] T cell epitopes [20], as well as those reporting HLA binding of specific T cell epitopes [10,11].

Examination of each mutant of the immunoprevalent epitopes revealed that the majority (744/914) were predicted to have no effect on HLA binding, while very few increased binding (4/914). However, almost one-fifth (166/914) of the epitope mutants were predicted to decrease binding, and hence, these presented candidates for enabling escape from T cell responses against commonly targeted SARS-CoV-2 CD8[+] T cell epitopes. We recommend that these 166 epitope mutants (listed in Supplementary Material Table S1) be considered for experimental investigation. This may involve conducting assays such as targeted HLA binding or HLA ligand elution assays, in combination with standard T cell assays such as multimer qualitative binding, enzyme-linked immunospot (ELISPOT), or intracellular cytokine staining (ICS) interferon-γ (IFN-γ) release, and activation-induced markers (AIM) assays.

Among the identified 166 epitope mutants, most have been observed so far at only low frequencies within the global SARS-CoV-2 sequence data. Only 83 epitope mutants have been observed five times or more (listed in Table 1). Among these, about half (42/83) are associated with HLA-A*02:01, the most prevalent HLA allele globally. Of the 83 epitope mutants, 16 are derived from the spike (S) protein, which could be particularly important given that most COVID-19 vaccines target the S antigen. These also include a mutant of a notable immunoprevalent epitope, YLQPRTFLL, which has been shown to be targeted during COVID-19 infection in more than 10 different cohorts so far [5]. The remaining 67/83 epitope mutants are derived from SARS-CoV-2 proteins besides S, which are of interest for T cell responses induced by inactivated COVID-19 vaccines or prior infection.

Table 1. List of SARS-CoV-2 immunoprevalent HLA-specific CD8[+] T cell epitope mutants recommended for experimental investigation.

Epitope [1]	HLA	Epitope Mutant [2]	Count
S			
$_{691}$SIIAYTMSL$_{699}$	A*02:01	SII**V**YTMSL	720
		SIIAYTM**LL**	655
		PIIAYTMSL	205
		SIIAYTM**AL**	181
		S**F**IAYTMSL	38
		SI**V**AYTMSL	24
		SIIAY**A**MSL	7
		SIIAYTMS**F**	7
		TIIAYTMSL	5
		SII**F**YTMSL	5
$_{1208}$QYIKWPWYI$_{1216}$	A*24:02	QYIKWPWY**T**	314
		Q**H**IKWPWYI	15
		QYIKWPWY**S**	13
$_{1000}$RLQSLQTYV$_{1008}$	A*02:01	R**F**QSLQTYV	20
		RLQSLQTY**A**	10
$_{269}$YLQPRTFLL$_{277}$	A*02:01	**C**LQPRTFLL	6

Table 1. *Cont.*

Epitope [1]	HLA	Epitope Mutant [2]	Count
M			
$_{26}$FLFLTWICL$_{34}$	A*02:01	FLFLTWIC**F**	1478
		FLFL**I**WICL	1394
		FLFLT**C**ICL	154
		LLFLTWICL	85
		F**I**FLTWICL	33
		CLFLTWICL	17
		F**V**FLTWICL	14
		FLFLTWIC**I**	13
		VLFLTWICL	8
		FL**L**LIWICL	7
N			
$_{338}$KLDDKDPNF$_{346}$	A*02:01	KLDDKD**S**NF	633
		KLD**N**KDPNF	226
		KL**N**DKDPNF	177
		K**F**DDKDPNF	118
		KLD**Y**KDPNF	78
		KLD**H**KDPNF	62
		KL**G**DKDPNF	38
		KLDDKD**Q**NF	35
		KLD**V**KDPNF	13
		KL**E**DKDPNF	11
		KLD**G**KDPNF	10
		KL**A**DKDPNF	6
		KL**V**DKDPNF	6
		KLD**A**KDPNF	6
		KLDDKDPN**S**	5
$_{361}$KTFPPTEPK$_{369}$	A*03:01	K**K**FPPTEPK	300
		KTFPPTEP**N**	107
		KTFPPTEP**I**	57
		K**R**FPPTEPK	30
		KTFPPTEP**E**	16
		KTFPPTEP**T**	14
		KTFPPTEP**L**	6
$_{361}$KTFPPTEPK$_{369}$	A*11:01	K**K**FPPTEPK	300
		KTFPPTEP**N**	107
		KTFPPTEP**I**	57
		K**R**FPPTEPK	30
		KTFPPTEP**E**	16
		KTFPPTEP**T**	14
		KTFPPTEP**L**	6
$_{134}$ATEGALNTPK$_{143}$	A*11:01	A**I**EGALNTPK	9685
		VTEGALNTPK	1162
		A**A**EGALNTPK	196
		A**N**EGALNTPK	101
		A**P**EGALNTPK	36
		TTEGALNTPK	27
$_{361}$KTFPPTEPKK$_{370}$	A*03:01	K**K**FPPTEPKK	300
		K**R**FPPTEPKK	30
		KTFP**S**TEPK**N**	28
$_{105}$SPRWYFYYL$_{113}$	B*07:02	S**S**RWYFYYL	23

Table 1. *Cont.*

Epitope [1]	HLA	Epitope Mutant [2]	Count
ORF3a			
$_{112}$VYFLQSINF$_{120}$	A*24:02	V**H**FLQSINF	339
		VYFLQSIN**C**	112
		VYFLQSIN**S**	50
$_{139}$LLYDANYFL$_{147}$	A*02:01	L**F**YDANYFL	2582
		LLYDANYF**F**	1276
$_{207}$FTSDYYQLY$_{215}$	A*01:01	FTSDYYQL**C**	121
		FTSDYYQL**H**	64
ORF1a			
$_{1637}$TTDPSFLGRYM$_{1647}$	A*01:01	TT**N**PSFLGRYM	2106
		T**I**DPSFLGRYM	1368
		T**N**DPSFLGRYM	102
		ITD**L**SFLGRYM	55
		T**I**D**L**SFLGRYM	5
$_{1321}$PTDNYITTY$_{1329}$	A*01:01	P**P**DNYITTY	141
		P**K**DNYITTY	13
		PTDNYITT**H**	10
$_{2332}$ILFTRFFYV$_{2340}$	A*02:01	I**F**FTRFFYV	696
		IL**C**TRFFYV	10
$_{1636}$HTTDPSFLGRY$_{1646}$	A*01:01	HTTDPSFLGR**H**	5

[1] Reported as immunoprevalent in [5]. The numbers in the subscript before and after the epitope sequence are the start and end indices of the epitope within the corresponding protein. [2] An epitope mutant represents an experimentally determined CD8^{+} T cell epitope with at least one mutation (shown in bold) observed in its sequence. The listed epitope mutants are predicted (using the default parameters of NetMHCpan4.1) to decrease HLA binding and occur at least five times in SARS-CoV-2 global sequence data (~2.5 million). The epitope mutants belonging to a specific protein are grouped together. Within each group, the epitope mutants are presented in descending order according to the number of times they are observed in the global sequence data.

The approach that we have adopted, which screens potential T cell epitope mutants based on their ability to evade peptide–HLA binding, is a conservative strategy. This is because a loss of peptide–HLA binding is irreparable since such a peptide would not be expressed on the cell surface for recognition by T cells. It is also possible that epitope mutants that do not decrease HLA binding may still result in T cell escape, mediated by other mechanisms such as abrogation of T cell receptor (TCR) binding with the peptide–HLA complex [21]. However, T cell escape via this means may be recovered by TCR recognition by other T cell clonotypes present in an individual [22]. Thus, while T cell escape through abrogation of TCR binding should not be ignored, the epitope mutants that we have identified to decrease peptide–HLA binding appear to be clear candidates for conferring T cell escape.

4. Conclusions

SARS-CoV-2 may evolve mutations to broadly escape T cell responses. While memory T cell responses induced by prior infection or vaccination have so far mainly remained robust [5–8], some evidence of T cell escape has been reported [8–11]. Continued monitoring of the capacity of SARS-CoV-2 to broadly evade T cell responses is required, as widespread T cell escape could compromise population-level immunity and have significant clinical and public health implications. To aid such monitoring efforts, our study considered SARS-CoV-2-specific CD8^{+} T cell responses and sought to identify potential SARS-CoV-2 CD8^{+} T cell escape mutants.

A comprehensive experimental evaluation of all potential SARS-CoV-2 CD8^{+} T cell escape mutants would involve testing all relevant mutations observed in available genetic sequence data, which is prohibitive as it would involve evaluating tens of thousands

of mutations within known SARS-CoV-2 epitopes. To help address this challenge, we performed a computational analysis to facilitate targeted experimental screening of T cell escape mutants. The main outcome of our study, given in Table 1, is a selective list of 83 observed SARS-CoV-2 mutations that lie within commonly targeted CD8+ T cell epitopes and have been observed in the SARS-CoV-2 global sequence data. An important feature of these epitope mutants is that they are expected to decrease peptide–HLA binding, which is a prerequisite for T cell recognition. Escape from these epitopes could be problematic, as it could affect vaccine and infection-induced T cell immunity in a large fraction of the population.

Overall, our computational study leverages available SARS-CoV-2 sequence and immunological data to direct attention towards commonly targeted T cell epitope mutants that emerge as clear candidates to confer escape from the vaccine or infection-induced responses. These epitope mutants are recommended for immediate experimental investigation.

References

1. Rotondo, J.C.; Martini, F.; Maritati, M.; Mazziotta, C.; Di Mauro, G.; Lanzillotti, C.; Barp, N.; Gallerani, A.; Tognon, M.; Contini, C. SARS-CoV-2 infection: New molecular, phylogenetic, and pathogenetic Insights. Efficacy of current vaccines and the potential risk of variants. *Viruses* **2021**, *13*, 1687. [CrossRef] [PubMed]
2. Abdool Karim, S.S.; de Oliveira, T. New SARS-CoV-2 variants—clinical, public health, and vaccine implications. *N. Engl. J. Med.* **2021**, *384*, 1866–1868. [CrossRef] [PubMed]
3. Planas, D.; Saunders, N.; Maes, P.; Guivel-Benhassine, F.; Planchais, C.; Buchrieser, J.; Bolland, W.-H.; Porrot, F.; Staropoli, I.; Lemoine, F.; et al. Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature* **2022**, *602*, 671–675. [CrossRef]
4. Planas, D.; Bruel, T.; Grzelak, L.; Guivel-Benhassine, F.; Staropoli, I.; Porrot, F.; Planchais, C.; Buchrieser, J.; Rajah, M.M.; Bishop, E.; et al. Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. *Nat. Med.* **2021**, *27*, 917–924. [CrossRef]
5. Quadeer, A.A.; Ahmed, S.F.; McKay, M.R. Landscape of epitopes targeted by T cells in 852 individuals recovered from COVID-19: Meta-analysis, immunoprevalence, and web platform. *Cell Rep. Med.* **2021**, *2*, 100312. [CrossRef] [PubMed]
6. Tarke, A.; Sidney, J.; Methot, N.; Yu, E.D.; Zhang, Y.; Dan, J.M.; Goodwin, B.; Rubiro, P.; Sutherland, A.; Wang, E.; et al. Impact of SARS-CoV-2 variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals. *Cell Rep. Med.* **2021**, *2*, 100355. [CrossRef]
7. Ahmed, S.F.; Quadeer, A.A.; McKay, M.R. SARS-CoV-2 T cell responses elicited by COVID-19 vaccines or infection are expected to remain robust against Omicron. *Viruses* **2022**, *14*, 79. [CrossRef]

8. Keeton, R.; Tincho, M.B.; Ngomti, A.; Baguma, R.; Benede, N.; Suzuki, A.; Khan, K.; Cele, S.; Bernstein, M.; Karim, F.; et al. T cell responses to SARS-CoV-2 spike cross-recognize Omicron. *Nature* **2022**, *603*, 488–492. [CrossRef]

9. Naranbhai, V.; Nathan, A.; Kaseke, C.; Berrios, C.; Khatri, A.; Choi, S.; Getz, M.A.; Tano-Menka, R.; Ofoman, O.; Gayton, A.; et al. T-cell reactivity to the SARS-CoV-2 Omicron variant is preserved in most but not all individuals. *Cell* **2022**, *185*, 1041–1051.E6. [CrossRef]

10. Dolton, G.; Rius, C.; Hasan, M.S.; Wall, A.; Szomolay, B.; Behiry, E.; Whalley, T.; Southgate, J.; Fuller, A.; The COVID-19 Genomics UK (COG-UK) consortium; et al. Emergence of immune escape at dominant SARS-CoV-2 killer T-cell epitope. *medRxiv* **2021**. preprint. [CrossRef]

11. Agerer, B.; Koblischke, M.; Gudipati, V.; Montaño-Gutierrez, L.F.; Smyth, M.; Popa, A.; Genger, J.-W.; Endler, L.; Florian, D.M.; Mühlgrabner, V.; et al. SARS-CoV-2 mutations in MHC-I-restricted epitopes evade CD8+ T cell responses. *Sci. Immunol.* **2021**, *6*, eabg6461. [CrossRef] [PubMed]

12. Rydyznski Moderbacher, C.; Ramirez, S.I.; Dan, J.M.; Grifoni, A.; Hastie, K.M.; Weiskopf, D.; Belanger, S.; Abbott, R.K.; Kim, C.; Choi, J.; et al. Antigen-specific adaptive immunity to SARS-CoV-2 in acute COVID-19 and associations with age and disease severity. *Cell* **2020**, *183*, 996–1012.e19. [CrossRef] [PubMed]

13. Kalimuddin, S.; Tham, C.Y.L.; Qui, M.; de Alwis, R.; Sim, J.X.Y.; Lim, J.M.E.; Tan, H.-C.; Syenina, A.; Zhang, S.L.; Le Bert, N.; et al. Early T cell and binding antibody responses are associated with COVID-19 RNA vaccine efficacy onset. *Med* **2021**, *2*, 682–688.e4. [CrossRef] [PubMed]

14. Moss, P. The T cell immune response against SARS-CoV-2. *Nat. Immunol.* **2022**, *23*, 186–193. [CrossRef] [PubMed]

15. Bushman, M.; Kahn, R.; Taylor, B.P.; Lipsitch, M.; Hanage, W.P. Population impact of SARS-CoV-2 variants with enhanced transmissibility and/or partial immune escape. *Cell* **2021**, *184*, 6229–6242.e18. [CrossRef]

16. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software Version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef]

17. 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*, W449–W454. [CrossRef]

18. Bui, H.-H.; Sidney, J.; Dinh, K.; Southwood, S.; Newman, M.J.; Sette, A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinform.* **2006**, *7*, 153. [CrossRef]

19. Wucherpfennig, K.W.; Sethi, D. T cell receptor recognition of self and foreign antigens in the induction of autoimmunity. *Semin. Immunol.* **2011**, *23*, 84–91. [CrossRef]

20. Sohail, M.S.; Ahmed, S.F.; Quadeer, A.A.; McKay, M.R. In silico T cell epitope identification for SARS-CoV-2: Progress and perspectives. *Adv. Drug Deliv. Rev.* **2021**, *171*, 29–47. [CrossRef]

21. Peters, B.; Nielsen, M.; Sette, A. T cell epitope predictions. *Annu. Rev. Immunol.* **2020**, *38*, 123–145. [CrossRef] [PubMed]

22. Ladell, K.; Hashimoto, M.; Iglesias, M.C.; Wilmann, P.G.; McLaren, J.E.; Gras, S.; Chikata, T.; Kuse, N.; Fastenackels, S.; Gostick, E.; et al. A molecular basis for the control of preimmune escape variants by HIV-Specific CD8+ T cells. *Immunity* **2013**, *38*, 425–436. [CrossRef] [PubMed]