

Article

Identification of Novel Candidate CD8⁺ T Cell Epitopes of the SARS-CoV2 with Homology to Other Seasonal Coronaviruses

Pradeep Darshana Pushpakumara ¹, Deshan Madhusanka ¹, Saubhagya Dhanasekara ¹, Chandima Jeewandara ¹, Graham S. Ogg ² and Gathsaurie Neelika Malavige ^{1,2,*}

¹ Centre for Dengue Research, Department of Immunology and Molecular Medicine, University of Sri Jayewardenepura, Nugegoda 10250, Sri Lanka; mail2medarshana@gmail.com (P.D.P.); deshans993@gmail.com (D.M.); saubhagyadanasekara@gmail.com (S.D.); jeewandara@sjp.ac.lk (C.J.)

² MRC Human Immunology Unit, MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford OX1 4BH, UK; graham.ogg@ndm.ox.ac.uk

* Correspondence: gathsaurie.malavige@ndm.ox.ac.uk; Tel.: +94-(0)-772443193; Fax: +94-(0)-112802026

Abstract: Cross-reactive T cell immunity to seasonal coronaviruses (HCoVs) may lead to immunopathology or protection during SARS-CoV2 infection. To understand the influence of cross-reactive T cell responses, we used IEDB (Immune epitope database) and NetMHCpan (ver. 4.1) to identify candidate CD8⁺ T cell epitopes, restricted through HLA-A and B alleles. Conservation analysis was carried out for these epitopes with HCoVs, OC43, HKU1, and NL63. 12/18 the candidate CD8⁺ T cell epitopes (binding score of ≥ 0.90), which had a high degree of homology ($>75\%$) with the other three HCoVs were within the NSP12 and NSP13 proteins. They were predicted to be restricted through HLA-A*2402, HLA-A*201, HLA-A*206, and HLA-B alleles B*3501. Thirty-one candidate CD8⁺ T cell epitopes that were specific to SARS-CoV2 virus ($<25\%$ homology with other HCoVs) were predominantly identified within the structural proteins (spike, envelop, membrane, and nucleocapsid) and the NSP1, NSP2, and NSP3. They were predominantly restricted through HLA-B*3501 (6/31), HLA-B*4001 (6/31), HLA-B*4403 (7/31), and HLA-A*2402 (8/31). It would be crucial to understand T cell responses that associate with protection, and the differences in the functionality and phenotype of epitope specific T cell responses, presented through different HLA alleles common in different geographical groups, to understand disease pathogenesis.

Keywords: CD8⁺ T cell epitopes; SARS-CoV2; seasonal coronaviruses; HLA alleles; OC-43; HKU-1; NL-63



Citation: Pushpakumara, P.D.; Madhusanka, D.; Dhanasekara, S.; Jeewandara, C.; Ogg, G.S.; Malavige, G.N. Identification of Novel Candidate CD8⁺ T Cell Epitopes of the SARS-CoV2 with Homology to Other Seasonal Coronaviruses. *Viruses* **2021**, *13*, 972. <https://doi.org/10.3390/v13060972>

Academic Editors: Concetta Castilletti, Luisa Barzon and Francesca Colavita

Received: 29 March 2021

Accepted: 21 May 2021

Published: 24 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Infection due to the SARS-CoV2 virus is currently the leading cause of mortality among elderly and vulnerable groups in many countries. Outbreaks of COVID-19 are now reported in 220 countries in the world [1], with some countries experiencing a massive second wave. The majority of those who are infected with SARS-CoV2 experience asymptomatic or mild illness, whereas severe illness and fatalities are seen in older individuals and in those with comorbidities such as diabetes, cancer, and cardiovascular disease [2–4]. However, the observed case fatality ratios (CFRs) vary widely between countries. For instance, countries such as Mexico, Italy, and Iran report CFRs of over 3.5%, while countries such as India, Turkey, South Korea report CFRs of $<2\%$ [5,6]. The deaths per 100,000 population also varies widely among countries, with many European countries and the US reporting rates of over 100 deaths/100,000 population, while the majority of South Asian and South East Asian countries report death rates <25 deaths/100,000 population, despite the ongoing large outbreaks in the South East Asian region [5].

There have been many factors that have been shown to influence the CFRs in different countries including the proportion of the population aged >70 years of age, the GDP per capita, BCG vaccination status [7,8], climate, population density, and social distancing

measures, which could account for the differences in the CFRs. However, there could be many immune factors that also lead to these differences, including immunity to human seasonal coronaviruses (HCoV), providing cross protection against SARS-CoV2. Children are more frequently exposed to HCoVs compared to adults and recently it was shown that a large proportion of children, who were sampled between the years of 2011 to 2018, between the ages of 1 to 16, had IgG antibodies that cross-react with the spike protein of SARS-CoV-2 [9]. Sera from SARS-CoV2 uninfected individuals who had such cross-reactive antibodies were able to neutralize the pseudotypes of SARS-CoV-2 without causing antibody dependent enhancement of infection [9].

Higher lymphocyte counts and specifically CD8⁺ T cell counts have been associated with early viral clearance and reduced disease severity [10]. Robust SARS-CoV2 specific CD4⁺ and CD8⁺ T cell responses have been detected following natural infection [11,12]. SARS-CoV2 cross reactive T cells have been detected following SARS-CoV1 infection in 2003, which were predominantly directed against the N and structural proteins (NSP7 and NSP13) [13]. Interestingly, SARS-CoV2 cross reactive CD4⁺ T cells have also been detected in a large proportion of unexposed individuals in some studies, which are thought to be due the presence of T cells specific to HCoVs [12,14]. However, differences in T cell methodology have resulted in differential findings of cross-reactivity related to altered sensitivity and specificity. Ultimately, it will be important to know if T cells can respond to virally infected target cells where naturally processed epitopes are presented. Nevertheless, the presence of such cross-reactive T cell responses has been speculated to protect against infection with the SARS-CoV2 virus and have also speculated to cause disease pathogenesis or response to vaccines [13,14].

As there is a difference in the CFRs among different countries, which are due to multiple factors, it would be important to understand if prior immunity to HCoVs could influence disease outcome when infected with the SARS-CoV2 virus. The predominant CD8⁺ and CD4⁺ T cell epitopes recognized by a population are also influenced by the frequency of different HLA types of a population. Therefore, in this study, we analyzed the cross reactivity of all the proteins of the SARS-CoV2 virus with three HCoVs (OC43, NL63, and HKU1) and those that are specific to the SARS-CoV2 virus and then used immune epitope database (IEDB) to predict the CD8⁺ T cell epitopes predicted to be restricted through the common HLA-Class I alleles in the Sri Lankan population.

2. Materials and Methods

2.1. Data Collection for Seasonal Human Corona Virus (HCoV) Sequence Analysis

The full-length protein sequences of the SARS-CoV2 (NC_045512), OC43 (NC_006213), HKU1 (NC_006577), and NL63 (NC_005831) were obtained from the National Center for Biotechnology Information (NCBI) GenBank (<https://www.ncbi.nlm.nih.gov/>, 5 October 2020). Four structural and 16 non-structural proteins (NSPs) were individually used for T cell epitope prediction and conservation analysis.

2.2. T-Cell Epitope Prediction for MHC Class I HLA Alleles

IEDB (Immune epitope database) (www.iedb.org, 10 October 2020) [15] along with NetMHCpan (ver. 4.1) [16] was used for finding putative peptide sequences that were restricted through the MHC Class-I HLA alleles, as these methods use algorithms (generated by artificial neural networks), which have accuracy of epitope prediction. HLA-A*02, HLA-A*24, HLA-A*33, HLA-B*35, HLA-B*40, HLA-B*44, HLA-C*07, HLA-C*06, HLA-C*04, and HLA-C*03 alleles were used for this epitope prediction, as these five alleles are the predominant alleles (each allele seen in >10% in the population) in the Sri Lankan population [17]. The epitope predictions were carried out for 8mer, 9mer, and 10mer epitopes. According to the epitope prediction NetMHCpan EL 4, which gives a score ranging from 0.0 to 1.0. Epitopes with a higher score are considered as stronger binders.

2.3. Conservation Analysis of SARS-CoV-2 Proteins and Predicted Epitopes with OC43, HKU1, and NL63

The epitopes which had a predicted binding score of >0.8 through the above approach and were highly conserved between the SARS-CoV2 and OC43, HKU1, and NL63, were further analyzed to determine percentage of identities and similarities between SARS-CoV2 and other HCoV. The sequences were aligned using the Clustal W on MEGA X software (www.megasoftware.net, 5 October 2020). Conservation of protein sequences were analyzed by using conservation analysis tool available at European Bioinformatics Institute (EBI) (www.ebi.ac.uk, 12 October 2020), IEDB, and Jalview (www.jalview.org, 15 October 2020).

3. Results

In order to identify the possible cross-reactive regions between SARS-CoV2 and other HCoVs, we carried out conservation analysis of SARS-CoV-2 virus with OC43, HKU1, and NL63, which have been seen in Sri Lanka. We analyzed the homology and conservation of the 4 structural and the 16 NSPs of SARS-CoV2 with the 3 other HCoVs (Table S1). The NSP12, NSP13, and NSP16 of the SARS-CoV2 showed $\geq 65\%$ homology with OC43 and HKU, which are beta coronaviruses, and a $>55\%$ homology with NL63, which is an alpha coronavirus.

In contrast, the NSP1 and NSP2 proteins of the SARS-CoV2 showed $<20\%$ homology with other three viruses, suggesting that these proteins were more specific for SARS-CoV-2 compared to the other proteins. All four structural proteins (S, E, M, and N), NSP3, and NSP6 showed $<35\%$ homology with OC43, HKU1, and NL63. Therefore, immune responses directed at these proteins may also be specific to the SARS-CoV2, unlike immune responses generated against NSP12, NSP13 and NSP16. The SARS-CoV-2 proteins S, E, M, N, and the non-structural proteins showed less homology with NL63 than OC43, and HKU1, suggesting that the SARS-CoV-2 virus is genetically closer to OC43 and HKU1 than NL63.

3.1. Identification of Possible CD8⁺ Epitopes of the SARS-CoV2 Virus Restricted through HLA-A Alleles of the SARS-CoV2

The NetMHCpan EL4 epitope prediction tool gives peptide binding scores ranging from 0 to 1.0 and we considered a predicted score ≥ 0.90 for a peptide as indicative of a stronger binder. None of the 8mer peptides gave a predictive score of ≥ 0.90 . However, 39 9mer peptides gave a high binding score of ≥ 0.90 , which were restricted through different HLA-A alleles, (Table 1). Five of these epitopes were identified from the spike and NSP3 proteins and 4 epitopes each were identified from NSP4, NSP6, NSP12, NSP14, and NSP15. Peptides with high binding scores were not identified from the envelope, NSP1, NSP9, NSP10, NSP11, and NSP16 proteins. The epitopes from NSP3 ⁷²⁶YYTSNPTTF⁷³⁴ (predicted to be restricted through HLA-A*2402) and NSP6 ⁷⁰FLLPSLATV⁷⁸ (predicted to be restricted through HLA-A*0201) gave a score of 0.99, while ¹³⁴⁹NYMPYFFTL¹³⁵⁷ from NSP3, ⁴²⁰FLLNKEMYL⁴²⁸ from NSP4, and ¹⁵²ALWEIQQVV¹⁶⁰ from NSP8 also gave a score of 0.98 scores. However, they had $<45\%$ homology with the other three viruses. 23/39 of the 9mer epitopes predicted in this study were restricted through HLA-A*02 while 16/39 9mer peptides were predicted to be restricted through HLA-A*24. Although HLA-A*33 was an allele seen in over 10% of the Sri Lankan population, 9mers that had a score of ≥ 0.90 were not identified.

Only four 10mer peptides were predicted to have a score of ≥ 0.90 , two were from spike, and one each from NSP3 and NSP6 proteins (Table 1).

Table 1. Predicted CD8⁺ epitopes of the SARS-CoV2 virus restricted through HLA-A alleles.

9mer Peptides with a Peptide Binding Score of ≥ 0.90						
Protein	HLA-A Allele	Sequence	Score	OC43 % Identity with SARS-CoV2	HKU1 % Identity with SARS-CoV2	NL63 % Identity with SARS-CoV2
Spike	A*02:01	²⁶⁹ YLQPRIFLL ²⁷⁷	0.97	44	44	22
Spike	A*24:02	¹²⁰⁸ QYIKWPWYI ¹²¹⁶	0.95	77	77	77
Spike	A*02:01	⁹⁷⁶ VLNDILSRL ⁹⁸⁴	0.94	67	67	33
Spike	A*24:02	⁶³⁵ VYSTGSNVF ⁶⁴³	0.93	22	22	22
Spike	A*02:01	¹⁰⁹ TLDSKTQSL ¹¹⁷	0.91	22	22	33
Membrane	A*24:02	⁹⁵ YFIASFRLF ¹⁰³	0.91	77	77	77
Nucleocapsid	A*02:01	²²² LLLDRLNQL ²³⁰	0.96	33	33	33
NSP2	A*02:01	²⁶⁵ GLNDNLLEI ²⁷³	0.92	44	33	44
NSP2	A*24:02	⁴⁹⁷ TFEKLVNKF ⁵⁰⁵	0.90	33	22	22
NSP3	A*24:02	⁷²⁶ YYTSNPTTF ⁷³⁴	0.99	22	22	22
NSP3	A*24:02	¹³⁴⁹ NYMPYFFTL ¹³⁵⁷	0.98	33	22	11
NSP3	A*24:02	⁸¹⁶ YYHTTDPSPF ⁸²⁴	0.96	11	11	0
NSP3	A*24:02	³⁶⁴ LYDKLVSSF ³⁷²	0.95	33	33	11
NSP3	A*24:02	¹⁰⁸¹ YKKDNSYF ¹⁰⁸⁹	0.94	33	33	0
NSP4	A*02:01	⁴²⁰ FLLNKEMYL ⁴²⁸	0.98	33	33	33
NSP4	A*02:01	³⁵⁹ FLAHIQWMV ³⁶⁷	0.94	44	44	33
NSP4	A*24:02	³⁵¹ FYLTNDVSF ³⁵⁹	0.92	22	22	22
NSP4	A*24:02	⁴⁸⁶ LYQPPQTSI ⁴⁹⁴	0.90	66	66	66
NSP5	A*02:06	¹⁹⁴ LIQDYIQSV ²⁰²	0.95	44	33	11
NSP6	A*02:01	⁷⁰ FLLPSLATV ⁷⁸	0.99	33	33	44
NSP6	A*02:01	¹⁴¹ TLMNVLTIV ¹⁴⁹	0.92	0	0	0
NSP6	A*24:02	⁸⁴ VYMPASWVM ⁹²	0.91	0	0	11
NSP6	A*24:02	¹¹⁵ MYASAVVLL ¹²³	0.90	22	22	22
NSP7	A*02:01	¹² VLLSVLQQL ²⁰	0.95	66	66	44
NSP8	A*02:01	¹⁵² ALWEIQQVV ¹⁶⁰	0.98	44	33	22
NSP12	A*24:02	³⁷ IYNDKVAGF ⁴⁵	0.95	66	44	44
NSP12	A*02:01	¹²³ TMADLVYAL ¹³¹	0.93	77	77	77
NSP12	A*02:06	³³⁴ FVDGVPFV ³⁴²	0.93	100	100	88
NSP12	A*02:01	⁸⁵⁴ LMIERFVSL ⁸⁶²	0.91	88	88	55
NSP13	A*02:01	²³⁹ TLVPQEHYV ²⁴⁷	0.95	77	77	44
NSP13	A*24:02	³⁹⁷ VYIGDPAQL ⁴⁰⁵	0.91	100	100	77
NSP14	A*02:01	³²¹ LLADKFPVL ³²⁹	0.94	11	11	33
NSP14	A*02:01	¹⁷⁶ NLSDRVVVFV ¹⁸⁴	0.93	66	66	77
NSP14	A*02:01	¹⁸⁴ VLWAHGFEL ¹⁹²	0.92	66	44	66
NSP14	A*02:01	⁴⁹⁴ YLDAYNMMI ⁵⁰²	0.91	44	44	44
NSP15	A*02:06	²⁴³ SQLGGLHLL ²⁵¹	0.96	66	66	66
NSP15	A*02:01	²⁹⁷ LLDDDFVEI ³⁰⁵	0.95	66	88	88
NSP15	A*02:06	³¹² SVVSKVVKV ³²⁰	0.94	66	66	55
NSP15	A*02:06	¹⁸¹ KVDGVVQQL ¹⁸⁹	0.92	22	22	11

Table 1. Cont.

10mer Peptides with a Peptide Biding Score of ≥ 0.90						
Protein	HLA-A Allele	Sequence	Sore	OC43 % Identity with SARS-CoV2	HKU1 % Identity with SARS-CoV2	NL63 % Identity with SARS-CoV2
Spike	A*24:02	¹⁰⁶⁶ TYVPAQEKNF ¹⁰⁷⁵	0.94	30	30	10
Spike	A*24:02	¹⁵⁹ VYSSANNCTF ¹⁶⁸	0.90	10	30	20
NSP3	A*24:02	⁷¹⁷ VYYTSNPTTF ⁷²⁶	0.98	30	40	20
NSP6	A*24:02	²⁴² YDYLVSTQEF ²⁵¹	0.91	50	50	50

3.2. Identification of Possible CD8⁺ Epitopes of the SARS-CoV2 Virus Restricted through HLA-B Alleles of the SARS-CoV2

As for HLA-A alleles, we considered a predicted score ≥ 0.90 for a peptide as indicative of a stronger binder. None of the 8mer peptides were found to give a score of ≥ 0.90 and therefore, were not predicted to be restricted through HLA-B alleles. However, 38 9mer peptides were identified, which had high binding scores and were predicted to be restricted through HLA-B alleles (Table 2). The highest number of epitopes were predicted from spike protein (5/38) and NSP13 (4/38). Three epitopes were predicted from each of the following proteins: namely the nucleocapsid, NSP2, NSP3, NSP4, and NSP12. No epitopes were identified from envelope, membrane, and NSP11 proteins. Nine epitopes gave a score of 0.99 and were ⁸⁹⁵IPFAMQMAY⁹⁰³ from spike, ³²⁵TPSGTWLTY³³³ from nucleocapsid, ¹⁹⁵SEVGPEHSL²⁰³ and ⁵⁶²GETLPTEVL⁵⁷⁰ both from NSP2, ¹²⁰EEFEPSTQY¹²⁸ and ⁵⁴⁶QEILGTVSW⁵⁵⁴ both from NSP3, ⁷²LPSLATVAY⁸⁰ from NSP6, ⁴SEFSSLPSY¹² from NSP8, and ⁶⁰⁸VENPHLMGW⁶¹⁶ from NSP12.

Table 2. Predicted CD8⁺ epitopes of the SARS-CoV2 virus restricted through HLA-B alleles.

9mer Peptides with a Peptide Binding Score of ≥ 0.90						
Protein	HLA-B Allele	Sequence	Score	OC43 % Identity with SARS-CoV2	HKU1 % Identity with SARS-CoV2	NL63 % Identity with SARS-CoV2
Spike	B*35:01	⁸⁹⁵ IPFAMQMAY ⁹⁰³	0.99	33	44	33
Spike	B*35:01	⁸³ LPFNDGVYF ⁹¹	0.98	22	33	11
Spike	B*44:03	¹²⁰⁰ QELGKYEYQ ¹²⁰⁸	0.98	44	44	33
Spike	B*35:01	⁶⁸⁶ VASQSIIAY ⁶⁹⁴	0.98	0	0	0
Spike	B*40:01	¹⁰¹⁵ AEIRASANL ¹⁰²³	0.98	22	22	44
Nucleocapsid	B*35:01	³²⁵ TPSGTWLTY ³³³	0.99	22	22	22
Nucleocapsid	B*44:03	³²² MEVTPSGTW ³³⁰	0.96	11	11	0
Nucleocapsid	B*35:01	⁷⁹ SPDDQIGYY ⁸⁷	0.92	22	22	55
NSP1	B*40:02	⁵⁶ VEKGVLPQL ⁶⁴	0.98	22	22	11
NSP1	B*35:01	¹¹⁰ HVGEIPVAY ¹¹⁸	0.95	22	11	11
NSP2	B*40:01	¹⁹⁵ SEVGPEHSL ²⁰³	0.99	11	11	0
NSP2	B*40:01	⁵⁶² GETLPTEVL ⁵⁷⁰	0.99	0	0	0
NSP2	B*44:02	⁵² REHEHEIAW ⁶⁰	0.98	22	22	22
NSP3	B*44:03	¹²⁰ EEFEPSTQY ¹²⁸	0.99	11	11	0
NSP3	B*44:03	⁵⁴⁶ QEILGTVSW ⁵⁵⁴	0.99	22	22	0
NSP3	B*40:01	¹⁷⁹⁹ AELAKNVSL ¹⁸⁰⁷	0.98	0	0	0
NSP4	B*40:02	³⁰⁹ GEYSHVVAF ³¹⁷	0.98	44	44	33
NSP4	B*35:01	³⁷³ VPFWITIAIY ³⁸¹	0.98	33	44	00
NSP4	B*35:01	¹⁷⁴ NVLEGSVAY ¹⁸²	0.97	11	11	44
NSP5	B*35:01	⁹³ TANPKTPKY ¹⁰¹	0.95	66	66	44
NSP6	B*35:01	⁷² LPSLATVAY ⁸⁰	0.99	44	44	33

Table 2. Cont.

9mer Peptides with a Peptide Binding Score of ≥ 0.90						
Protein	HLA-B Allele	Sequence	Score	OC43 % Identity with SARS-CoV2	HKU1 % Identity with SARS-CoV2	NL63 % Identity with SARS-CoV2
NSP7	B*35:01	⁴¹ LAKDTTEAF ⁴⁹	0.91	33	33	22
NSP8	B*44:03	⁴ SEFSSLPSY ¹²	0.99	44	44	66
NSP8	B*40:01	⁴⁷ SEFDRDAAM ⁵⁵	0.93	44	44	44
NSP9	B*44:03	⁶⁷ TELEPPCRF ⁷⁵	0.95	66	66	66
NSP10	B*35:01	¹⁹ FAVDAAKAY ²⁷	0.98	55	55	55
NSP12	B*44:02	⁶⁰⁸ VENPHLMGW ⁶¹⁶	0.99	77	77	88
NSP12	B*40:01	⁸⁷⁴ QEYADVFLH ⁸⁸²	0.98	44	44	44
NSP12	B*35:01	³³⁷ VPFVVSTGY ³⁴⁵	0.97	88	88	55
NSP13	B*35:01	²⁹¹ FAIGLALYY ²⁹⁹	0.96	66	77	66
NSP13	B*44:03	¹⁴¹ TEETFKLSY ¹⁴⁹	0.96	77	77	55
NSP13	B*40:01	¹⁵⁵ REVLSDREL ¹⁶³	0.95	55	55	44
NSP13	B*40:02	¹⁶¹ RELHLSWEV ¹⁶⁹	0.91	77	77	66
NSP14	B*35:01	⁴²⁸ TPAFDKSAF ⁴³⁶	0.96	44	44	77
NSP14	B*35:03	¹⁹ HPTQAPTHL ²⁷	0.94	55	55	55
NSP15	B*35:03	⁴⁹ LPVNVAFEL ⁵⁷	0.96	66	66	88
NSP15	B*44:03	²⁰⁰ QEFKPRSQM ²⁰⁸	0.92	33	33	55
NSP16	B*44:02	¹⁴¹ KENDSKEGF ¹⁴⁹	0.94	55	55	66
10mer Peptides with a Peptide Binding Score of ≥ 0.90						
Protein	HLA-B Allele	Sequence	Score	OC43 % Identity with SARS-CoV2	HKU1 % Identity with SARS-CoV2	NL63 % Identity with SARS-CoV2
Spike	B*44:02	⁹⁵ TEKSNIIRGW ¹⁰⁴	0.95	10	0	10
Membrane	B*44:03	¹¹ EELKKLLEQW ²⁰	0.94	40	20	30
NSP2	B*44:03	⁴⁸⁹ KEIKESVQTF ⁴⁹⁸	0.95	0	0	10
NSP3	B*44:03	¹²⁰ EEEFEPSTQY ¹²⁹	0.98	20	10	0
NSP3	B*44:03	¹⁰⁷² TEIDPKLDNY ¹⁰⁸¹	0.96	40	30	10
NSP3	B*35:01	⁵⁰² VPTDNYITTY ⁵¹¹	0.94	0	0	0
NSP3	B*44:03	⁹⁴ GEFKLASHMY ¹⁰³	0.93	50	50	0
NSP7	B*40:01	⁴⁶ TEAFEKMSVL ⁵⁵	0.91	50	40	20
NSP8	B*44:03	³ ASEFSSLPSY ¹²	0.94	40	40	60
NSP10	B*40:01	⁵ TEVPANSTVL ¹⁴	0.92	50	60	50
NSP12	B*44:03	⁸⁷⁵ QEYADVFLH ⁸⁸⁴	0.99	50	50	40
NSP12	B*44:03	¹⁶⁶ VENPDILRVY ¹⁷⁵	0.95	80	70	80
NSP12	B*44:03	⁶⁰⁸ DVENPHLMGW ⁶¹⁷	0.90	80	70	80
NSP13	B*40:01	⁴⁴⁶ AEIVDTVSAL ⁴⁵⁵	0.96	90	80	80
NSP14	B*44:02	⁷⁷ EEAIRHVRAW ⁸⁶	0.96	70	60	60
NSP14	B*35:01	⁴² IPGIPKDMTY ⁵¹	0.92	20	20	30
NSP15	B*35:01	²⁶⁹ IPMDSTVKNY ²⁷⁸	0.96	30	30	10
NSP15	B*40:01	⁴⁰ VELFENKTTL ⁴⁹	0.95	50	30	60

17/38 of these epitopes were predicted to be restricted through HLA-B*35 11/38 through HLA-B*40 and 10/38 through HLA-B*44. While most of these peptides showed <45% homology with OC43, NL63 and HKU-1, ¹⁴¹TEETFKLSY¹⁴⁹ and ⁶⁰⁸VENPHLMGW⁶¹⁶ restricted through HLA-B*44, ³³⁷VPFVVSTGY³⁴⁵ restricted through HLA-B*35 and ¹⁶¹RELHLSWEV¹⁶⁹

restricted through HLA-B*40 showed a homology of >75% with OC43 and HKU-1, which are two other beta coronaviruses.

Only 18 10mers peptides were predicted to have a score of ≥ 0.90 that were restricted through HLA-B (Table 2), 4/18 were identified NSP3. 10mer peptides with high binding scores were not predicted from envelope, nucleocapsid, NSP1, NSP4, NSP5, NSP6, NSP9, NSP11, and NSP16 9 proteins. 3/18 of these predicted 10mers were shown to be restricted through HLA-B*35, 4/18 through HLA-B*40, and 11/18 through HLA-B*44. Therefore, there may be a higher probability for 10mers to bind to HLA-B*44 alleles.

3.3. Identification of Possible CD8⁺ Epitopes of the SARS-CoV2 Virus Restricted through HLA-C Alleles of the SARS-CoV2

None of the 8mer and 10mer peptides were found to give a score of ≥ 0.90 and therefore, were not predicted to be restricted through HLA-C alleles. However, 21 9mer peptides that were identified had high binding scores and were predicted to be restricted through HLA-C alleles (Table 3). The highest number of epitopes were predicted from NSP3 protein (6/21) and spike (4/21). Epitopes were not predicted from each of the following proteins: namely the nucleocapsid, envelop, NSP2, NSP4, NSP6, NSP9, NSP11, NSP12, NSP15 and NSP16.

Table 3. Predicted CD8⁺ epitopes of the SARS-CoV2 virus restricted through HLA-C alleles.

9mer Peptides with a Peptide Binding Score of ≥ 0.90						
Protein	HLA-B Allele	Sequence	Score	OC43 % Identity with SARS-CoV2	HKU1 % Identity with SARS-CoV2	NL63 % Identity with SARS-CoV2
Spike	HLA-C*04:01	¹¹³⁷ VYDPLQPEL ¹¹⁴⁵	0.99	22	22	11
Spike	HLA-C*06:02	³²⁷ VRFPNITNL ³³⁵	0.98	11	11	0
Spike	HLA-C*03:02	⁶⁸⁷ VASQSIAY ⁶⁹⁵	0.97	11	11	0
Spike	HLA-C*03:02	¹⁰⁵⁴ QSAPHGVPF ¹⁰⁶²	0.93	55	55	55
Membrane	HLA-C*03:02	³⁷ FAYANRRNF ⁴⁵	0.96	44	44	22
Membrane	HLA-C*03:02	¹⁷⁰ VATSRTLSY ¹⁷⁸	0.96	33	33	44
NSP1	HLA-C*03:02	⁷⁷ RTAPHGVM ⁸⁵	0.90	11	11	11
NSP3	HLA-C*03:02	¹⁷⁷⁶ YVNTFSSTF ¹⁷⁸⁴	0.96	66	66	33
NSP3	HLA-C*04:01	⁷³⁴ TFDNLKTL ⁷⁴²	0.96	11	11	11
NSP3	HLA-C*03:02	¹⁶⁵¹ VARDLSLQF ¹⁶⁵⁹	0.95	44	33	33
NSP3	HLA-C*04:01	¹⁷⁷² MFDAYVNTF ¹⁷⁸⁰	0.95	77	77	33
NSP3	HLA-C*04:01	³⁶⁴ LYDKLVSSF ³⁷²	0.94	33	33	11
NSP3	HLA-C*03:02	¹⁷³⁵ SAKSASVYY ¹⁷⁴³	0.90	33	33	22
NSP5	HLA-C*03:02	⁹³ TANPKTPKY ¹⁰¹	0.96	66	66	44
NSP7	HLA-C*03:02	⁴¹ LAKDTTEAF ⁴⁹	0.92	33	33	22
NSP8	HLA-C*03:02	¹³⁰ VVIPDYNTY ¹³⁸	0.92	44	55	44
NSP10	HLA-C*03:02	¹⁹ FAVDAAKAY ²⁷	0.99	55	55	55
NSP13	HLA-C*03:02	²⁰⁹ VVYRGTTY ²¹⁷	0.95	77	77	44
NSP13	HLA-C*03:02	²⁹¹ FAIGLALY ²⁹⁹	0.94	77	77	55
NSP13	HLA-C*03:02	²²⁵ FVLTSHVTM ²³³	0.90	77	77	77
NSP14	HLA-C*06:02	¹⁶² VRIKIVQML ¹⁷⁰	0.93	77	77	77

3.4. Identification of CD8⁺ T Cell Epitopes of SARS-CoV2, Which Show $\geq 75\%$ Homology with OC43, HKU1, and NL63

After identification of peptides that had high predicted values to be restricted through the common HLA-A, B, and C alleles present in the Sri Lankan population, we proceeded to identify the regions of the SARS-CoV2 virus, which had a >75% homology with the

HCoVs. We then proceeded to identify CD8⁺ T cell epitopes within these regions, which were candidates to be restricted through these HLA alleles. This was to determine if we could identify CD8⁺ T cell epitopes of the SARS-CoV2, which were likely to cross-react with the other HCoVs. None of the predicted CD8⁺ 8mer epitopes identified within the SARS-CoV2 virus gave a high binding score and therefore, only predicted 9mer and 10mer CD8⁺ T cell epitopes of the SARS-CoV2 virus were analyzed for the degree of homology with OC43, HKU1, and NL63.

Epitopes that were identified to have a $\geq 75\%$ homology with more than two HCoV viruses are shown in Table 4. Thirty-four 9mer epitopes and 18 10mer peptides identified within the SARS-CoV2 virus had $\geq 75\%$ homology with ≥ 2 HCoV viruses. Of the 9mer peptides, 22/34 epitopes gave a peptide binding score of ≥ 0.90 . 11/34 of these CD8⁺ T cell epitopes within these cross-reactive regions were predicted to be restricted through HLA-A, 6/34 were predicted to be restricted through HLA-B alleles and 5/34 were predicted to be restricted through HLA-C alleles. Six highly cross-reactive CD8⁺ T cell epitopes (9mers) with high HLA-A (A*201 and A*206) binding scores were identified from the NSP12 and NSP13 (³³⁴FVDGVPFVV³⁴²). Two of these peptides showed 100% homology with OC43 and HKU1. The Envelope, nucleocapsid proteins and the other non-structural proteins (apart from NSP12 and NSP13) did not have regions with $>75\%$ homology with the other HCoVs. The alignment of SARS-CoV2 NSP12 and NSP13, in which most of the cross-reactive epitopes were identified from and their position are shown in Figures S1 and S2.

Table 4. Predicted CD8⁺ epitopes of the SARS-CoV2 virus, which show $\geq 75\%$ homology with OC43, HKU1, and NL63.

9mer Peptides with $\geq 75\%$ Homology with OC43, HKU1, and NL63						
Protein	HLA Allele	Sequence	Score	OC43 % Identity with SARS-CoV2	HKU1 % Identity with SARS-CoV2	NL63 % Identity with SARS-CoV2
Spike	A*24:02	¹²⁰⁸ QYIKWPWYI ¹²¹⁶	0.95	77	77	77
Membrane	A*24:02	⁹⁵ YFIASFRLF ¹⁰³	0.91	77	77	77
NSP3	C*04:01	¹⁷⁷² MFDAYVNTF ¹⁷⁸⁰	0.95	77	77	33
NSP5	A*02:06	¹⁵⁹ FVYMHQLEL ¹⁶⁷	0.76	100	88	66
NSP5	B*35:01	⁹⁵ NPKTPKYKF ¹⁰³	0.47	77	77	55
NSP10	B*35:03	³⁶ QPITNCVKM ⁴⁴	0.74	88	88	66
NSP12	A*02:01	¹²³ TMADLVYAL ¹³¹	0.93	77	77	77
NSP12	A*02:06	³³⁴ FVDGVPFVV ³⁴²	0.93	100	100	88
NSP12	A*02:01	⁸⁵⁴ LMIERFVSL ⁸⁶²	0.91	88	88	55
NSP12	A*02:01	³³⁴ FVDGVPFVV ³⁴²	0.90	100	100	77
NSP12	B*44:03	⁶⁰⁸ VENPHLMGW ⁶¹⁶	0.99	77	77	88
NSP12	B*35:01	³³⁷ VPFVSTGY ³⁴⁵	0.97	88	88	55
NSP12	C*03:02	⁵³⁴ NVIPTITQM ⁵⁴²	0.82	77	77	66
NSP12	C*03:02	³⁴⁰ FVVSTGYHF ³⁴⁸	0.80	77	77	55
NSP13	A*02:01	²³⁹ TLVPQEHYV ²⁴⁷	0.95	77	77	44
NSP13	A*24:02	³⁹⁷ VYIGDPAQL ⁴⁰⁵	0.91	100	100	77
NSP13	A*02:06	²³⁹ TLVPQEHYV ²⁴⁷	0.91	77	77	44
NSP13	B*35:01	²⁹¹ FAIGLALY ²⁹⁹	0.96	77	77	66
NSP13	B*44:03	¹⁴¹ TEETFKLSY ¹⁴⁹	0.96	77	77	55
NSP13	B*40:02	¹⁶¹ RELHLSWEV ¹⁶⁹	0.91	77	77	66
NSP13	C*03:02	²⁰⁹ VVYRGTTY ²¹⁷	0.95	77	77	44
NSP13	C*03:02	²⁹¹ FAIGLALY ²⁹⁹	0.94	77	77	55
NSP13	C*03:02	²²⁵ FVLTSHTVM ²³³	0.90	77	77	77
NSP13	C*06:02	²¹¹ YRGTTYKL ²¹⁹	0.78	88	88	55

Table 4. Cont.

9mer Peptides with $\geq 75\%$ Homology with OC43, HKU1, and NL63						
Protein	HLA Allele	Sequence	Score	OC43 % Identity with SARS-CoV2	HKU1 % Identity with SARS-CoV2	NL63 % Identity with SARS-CoV2
NSP14	A*02:01	¹⁷⁶ NLSDRVVFV ¹⁸⁴	0.93	77	66	77
NSP14	B*35:01	⁵⁰⁹ WVYKQFDTY ⁵¹⁷	0.65	77	77	55
NSP14	C*06:02	¹⁶² VRIKIVQML ¹⁷⁰	0.93	77	77	77
NSP14	C*03:02	⁴⁸⁷ HANEYRLYL ⁴⁹⁵	0.83	77	77	55
NSP15	A*02:01	²⁹⁷ LLDDDFVEI ³⁰⁵	0.95	66	88	88
NSP15	B*35:03	⁴⁹ LPVNVAFEL ⁵⁷	0.96	77	66	88
NSP16	A*02:01	⁵³ YLNTLTLAV ⁶¹	0.89	88	88	55
NSP16	A*24:02	⁴⁶ KYTQLCQYL ⁵⁴	0.82	100	100	100
NSP16	A*33:03	²⁴⁷ MSKFPLKLR ²⁵⁵	0.78	77	77	44
NSP16	C*04:01	¹³¹ MYDPKTKNV ¹³⁹	0.83	77	77	44
10mer Peptides with $\geq 75\%$ Homology with OC43, HKU1, and NL63						
Protein	HLA Allele	Sequence	Score	OC43 % Identity with SARS-CoV2	HKU1 % Identity with SARS-CoV2	NL63 % Identity with SARS-CoV2
NSP4	C*03:02	⁴⁷⁸ FSNSGSDVLY ⁴⁸⁷	0.41	80	70	20
NSP12	A*02:06	³³² KIFVDGVPFV ³⁴¹	0.88	90	90	70
NSP12	A*02:01	³³² KIFVDGVPFV ³⁴¹	0.88	90	90	70
NSP13	A*24:02	²¹⁶ TYKLVNGDYF ²²⁵	0.77	80	70	50
NSP13	A*02:01	⁴⁰ KLVLSVNPYV ⁴⁹	0.58	80	80	50
NSP13	A*33:03	³⁸¹ NYDLSVNNAR ³⁹⁰	0.54	80	80	80
NSP13	A*33:03	⁵⁵¹ ETAHSCNVNR ⁵⁶⁰	0.52	90	90	80
NSP13	B*40:01	⁴⁴⁶ AEIVDTVSAL ⁴⁵⁵	0.96	90	80	80
NSP13	B*40:02	⁴⁴⁶ AEIVDTVSAL ⁴⁵⁵	0.87	90	80	80
NSP14	A*33:03	⁵¹⁶ TYNLWNTFTR ⁵²⁵	0.58	80	80	50
NSP14	A*24:02	⁵¹⁰ VYKQFDTYNL ⁵¹⁹	0.55	80	80	60
NSP15	A*33:03	⁵² NVAFELWAKR ⁶¹	0.66	80	60	90
NSP15	A*02:06	²⁴³ SQLGGLHLI ²⁵²	0.52	70	70	80
NSP16	A*24:02	²⁴¹ SYSLFDMSKF ²⁵⁰	0.75	80	80	50
NSP16	A*33:03	²⁴⁶ DMSKFPLKLR ²⁵⁵	0.59	80	70	50
NSP16	A*24:02	²²¹ GYVMHANYIF ²³⁰	0.55	80	80	70
NSP16	A*02:01	²⁴³ SLFDMSKFPL ²⁵²	0.47	90	80	50

Of the 10mer peptides analyzed, 18 were identified within SARS-CoV2 to have $\geq 75\%$ homology with ≥ 2 HCoVs (Table 4). Only one 10mer peptide identified within NSP-13 (⁴⁴⁶AEIVDTVSAL⁴⁵⁵) gave a score of ≥ 0.90 score. 14/18 of these 10mer CD8⁺ T cell epitopes were predicted to be restricted through HLA-A and 3/18 were predicted to be restricted through HLA-B alleles. Only one epitope out of 18 was predicted to be restricted through HLA-C. As with the 9mer peptides, 6 of the 10mer peptides, which were highly homologous with OC43, HKU1 and NL63 were found within the NSP12 and NSP13 region.

3.5. Identification of CD8⁺ T cell Epitopes of SARS-CoV2, Which Show $\leq 25\%$ Homology with OC43, HKU1, and NL63

After identification of highly cross reactive CD8⁺ T cell epitopes within the SARS-CoV2, we proceeded to identify regions, which were specific to the virus and did not cross react with other HCoV2, and therefore, are likely to be SARS-CoV2 specific CD8⁺ T cell epitopes. 9mer peptides of the representing different regions of the SARS-CoV2 virus,

which have $\leq 25\%$ homology with >2 HCoV viruses were analyzed and 60 such potential CD8⁺ T cell epitopes were identified. (Table 5). 19/60 CD8⁺ T cell epitopes were predicted to be restricted through HLA-A alleles, 20/60 epitopes were predicted to be restricted through HLA-B alleles and 21/62 epitopes were predicted to be restricted through HLA-C alleles. 31/60 9mer peptides gave a binding score of ≥ 0.90 . 12/31 of these CD8⁺ T cell epitopes were predicted to be restricted through HLA-A alleles, 14/31 predicted to be restricted through HLA- B alleles and 5/31 predicted to be restricted through HLA- C alleles. A region within the spike protein (⁶⁸⁶VASQSIAY⁶⁹⁴) had no homology with the other HCoVs but had a high binding score of >0.95 to HLA-B*3501 and two other 9mer peptides within the nucleocapsid (³²⁵TPSGTWLTY³³³ and ³²²MEVTPSGTW³³⁰) had $<22\%$ homology and were predicted to be restricted through HLA-B*3501 and HLA-B*4403. Three other CD8⁺ T cell epitopes within NSP2, NSP3 and NSP6, which had high binding scores but had 0% homology were also identified.

Table 5. Predicted CD8⁺ epitopes of the SARS-CoV2 virus, which show $\leq 25\%$ homology with OC43, HKU1, and NL63.

9mer Peptides with $\leq 25\%$ Homology with OC43, HKU1, and NL63						
Protein	HLA Allele	Sequence	Score	OC43 % Identity with SARS-CoV2	HKU1 % Identity with SARS-CoV2	NL63 % Identity with SARS-CoV2
Spike	A*24:02	⁶³⁵ VYSTGSNVF ⁶⁴³	0.93	22	22	22
Spike	A*02:01	¹⁰⁹ TLDSKTQSL ¹¹⁷	0.91	22	22	33
Spike	B*35:01	⁸³ LPFNDGVYF ⁹¹	0.98	22	33	11
Spike	B*35:01	⁶⁸⁶ VASQSIAY ⁶⁹⁴	0.98	0	0	0
Spike	B*40:01	¹⁰¹⁵ AEIRASANL ¹⁰²³	0.98	22	22	44
Spike	C*04:01	¹¹³⁷ VYDPLQPEL ¹¹⁴⁵	0.99	22	22	11
Spike	C*06:02	³²⁷ VRFPNITNL ³³⁵	0.98	11	11	0
Spike	C*03:02	⁶⁸⁷ VASQSIAY ⁶⁹⁵	0.97	11	11	0
Spike	C*07:01	³²⁷ VRFPNITNL ³³⁵	0.87	11	11	0
Spike	C*07:02	³²⁷ VRFPNITNL ³³⁵	0.84	11	11	0
Spike	C*06:02	⁴⁰² IRGDEVRQI ⁴¹⁰	0.83	11	11	0
Spike	C*04:01	⁷⁸ RFDNPVLPF ⁸⁶	0.83	11	11	11
Membrane	A*33:03	¹³⁸ LVIGAVILR ¹⁴⁶	0.72	22	22	22
Membrane	B*40:01	¹³⁶ SELVIGAVI ¹⁴⁴	0.73	11	11	11
Envelope	A*33:03	⁶¹ RVKNLNSSR ⁶⁹	0.61	11	11	0
Envelope	A*33:03	³⁰ TLAILTALR ³⁸	0.60	22	22	11
Envelope	B*40:01	⁶ SEETGTLIV ¹⁴	0.55	11	22	11
Envelope	B*35:03	⁴ FVSEETGTL ¹²	0.43	11	11	11
Nucleocapsid	B*35:01	³²⁵ TPSGTWLTY ³³³	0.99	22	22	22
Nucleocapsid	B*44:03	³²² MEVTPSGTW ³³⁰	0.96	11	11	0
Nucleocapsid	C*03:03	⁴⁰³ FSKQLQQSM ⁴¹¹	0.73	11	11	11
NSP1	A*24:02	¹³⁵ SYGADLKSF ¹⁴³	0.89	0	0	0
NSP1	A*02:01	⁸⁴ VMVELVAEL ⁹²	0.85	22	22	0
NSP1	B*40:02	⁵⁶ VEKGVLPQL ⁶⁴	0.98	22	22	11
NSP1	B*35:01	¹¹⁰ HVGEIPVAY ¹¹⁸	0.95	22	11	11
NSP1	B*35:01	⁸⁹ VAELEGIQY ⁹⁷	0.63	11	0	11
NSP1	C*03:02	⁷⁷ RTAPHGHEVM ⁸⁵	0.90	11	11	11
NSP1	C*03:02	¹¹⁰ HVGEIPVAY ¹¹⁸	0.81	22	11	11
NSP1	C*03:02	¹⁶⁵ HSSGVTRREL ¹⁷³	0.71	11	11	0
NSP2	A*24:02	⁴⁹⁷ TFFKLVNKF ⁵⁰⁵	0.90	33	22	22

Table 5. Cont.

9mer Peptides with ≤25% Homology with OC43, HKU1, and NL63						
Protein	HLA Allele	Sequence	Score	OC43 % Identity with SARS-CoV2	HKU1 % Identity with SARS-CoV2	NL63 % Identity with SARS-CoV2
NSP2	A*02:06	⁴²⁰ YITGGVVQL ⁴²⁸	0.86	22	22	22
NSP2	A*02:06	⁴³⁹ TVYEKLKPV ⁴⁴⁷	0.83	11	11	11
NSP2	B*40:01	¹⁹⁵ SEVGPEHSL ²⁰³	0.99	11	11	0
NSP2	B*40:01	⁵⁶² GETLPTEVL ⁵⁷⁰	0.99	0	0	0
NSP2	B*44:03	⁵² REHEHEIAW ⁶⁰	0.98	22	22	22
NSP2	C*06:02	³⁶³ VRSIFSRTL ³⁷¹	0.88	11	11	11
NSP2	C*03:02	³⁸⁷ TILDGISQY ³⁹⁵	0.73	22	22	22
NSP3	A*24:02	⁷²⁶ YYTSNPTTF ⁷³⁴	0.99	22	22	22
NSP3	A*24:02	¹³⁴⁹ NYMPYFFTL ¹³⁵⁷	0.98	33	22	11
NSP3	A*24:02	⁸¹⁶ YYHTTDPSP ⁸²⁴	0.96	11	11	0
NSP3	B*44:03	¹²⁰ EEFEPSTQY ¹²⁸	0.99	11	11	0
NSP3	B*44:03	⁵⁴⁶ QEILGTVSW ⁵⁵⁴	0.99	22	22	0
NSP3	B*40:01	¹⁷⁹⁹ AELAKNVSL ¹⁸⁰⁷	0.98	0	0	0
NSP3	C*04:01	⁷³⁴ TFDNLKTL ⁷⁴²	0.96	11	11	11
NSP3	C*07:02	⁷¹⁸ YYTSNPTTF ⁷²⁶	0.89	22	22	22
NSP3	C*03:02	⁷⁶⁸ MSMTYGQQF ⁷⁷⁶	0.87	22	22	33
NSP3	C*03:02	³³⁶ FGADPIHSL ³⁴⁴	0.87	22	22	22
NSP3	C*03:02	¹⁴³⁶ MSNLGMPSY ¹⁴⁴⁴	0.80	11	11	11
NSP4	A*24:02	³⁵¹ FYLTNDVSE ³⁵⁹	0.92	22	22	22
NSP4	B*35:01	¹⁷⁴ NVLEGSVAY ¹⁸²	0.97	11	11	44
NSP4	C*03:02	²⁵ YLITPVHVM ³³	0.87	11	11	11
NSP4	C*03:02	¹⁷⁴ NVLEGSVAY ¹⁸²	0.73	22	22	22
NSP6	A*02:01	¹⁴¹ TLMNVLTLV ¹⁴⁹	0.92	0	0	0
NSP6	A*24:02	⁸⁴ VYMPASWVM ⁹²	0.91	0	0	11
NSP6	A*24:02	¹¹⁵ MYASAVVLL ¹²³	0.90	22	22	22
NSP6	B*35:01	¹⁵⁶ NALDQAISM ¹⁶⁴	0.88	22	22	11
NSP6	B*35:01	¹⁶⁷ LIISVTSNY ¹⁷⁵	0.56	11	11	22
NSP6	C*04:01	¹³¹ VYDDGARRY ¹³⁹	0.89	11	11	11
NSP14	A*02:01	³²¹ LLADKFPVL ³²⁹	0.94	11	11	33
NSP15	A*02:06	¹⁸¹ KVDGVVQQL ¹⁸⁹	0.92	22	22	11
10mer Peptides with ≤25% Homology with OC43, HKU1, and NL63						
Protein	HLA Allele	Sequence	Score	OC43 % Identity with SARS-CoV2	HKU1 % Identity with SARS-CoV2	NL63 % Identity with SARS-CoV2
Spike	A*24:02	³⁶⁸ LYNSASFSTF ³⁷⁷	0.89	20	20	30
Spike	A*24:02	⁷⁸⁸ IYKTPPIKDF ⁷⁹⁷	0.88	10	20	10
Spike	B*44:02	⁹⁵ TEKSNIIRGW ¹⁰⁴	0.95	10	0	10
Spike	B*35:01	²²⁹ LPIGINITRF ²³⁸	0.82	10	20	20
Spike	C*04:01	⁹³² TVYDPLQPEL ⁹⁴¹	0.59	20	20	10
Spike	C*07:01	⁷⁷ KRFDNPVLPF ⁸⁶	0.47	10	10	10
Membrane	A*33:03	¹³⁷ ELVIGAVILR ¹⁴⁶	0.65	20	20	20
Membrane	A*33:03	¹⁷⁷ SYKLGASQR ¹⁸⁶	0.55	40	20	20
Membrane	B*40:01	¹³⁶ SELVIGAVIL ¹⁴⁵	0.76	10	10	20
Envelope	A*33:03	⁶⁰ SRVKNLNSRR ⁶⁹	0.48	10	30	0

Table 5. Cont.

10mer Peptides with $\leq 25\%$ Homology with OC43, HKU1, and NL63						
Protein	HLA Allele	Sequence	Score	OC43 % Identity with SARS-CoV2	HKU1 % Identity with SARS-CoV2	NL63 % Identity with SARS-CoV2
Envelope	A*33:03	²⁹ VTLAILTALR ³⁸	0.41	30	20	0
Nucleocapsid	A*33:03	¹⁴⁰ NTPKDHIGTR ¹⁴⁹	0.60	20	10	20
Nucleocapsid	A*02:01	³⁹⁸ ADLDDFSKQL ⁴⁰⁷	0.44	10	30	0
Nucleocapsid	B*44:03	³²¹ GMEVTPSGTW ³³⁰	0.87	0	0	0
Nucleocapsid	B*35:01	³²⁴ VTPSGTWLTY ³³³	0.68	20	20	20
Nucleocapsid	B*40:01	³²² MEVTPSGTWL ³³¹	0.67	10	10	0
NSP1	A*33:03	¹⁶² NTKHSSGVTR ¹⁷¹	0.74	0	0	0
NSP1	A*33:03	⁶⁸ YVFIKRSAR ⁷⁷	0.49	30	20	10
NSP1	A*02:06	¹⁴ VQLSLPVLQV ²³	0.45	20	10	10
NSP1	A*33:03	¹⁵ QLSLPVLQVR ²⁴	0.40	10	10	10
NSP1	B*35:01	⁶¹ LPQLEQPYVF ⁷⁰	0.81	20	20	30
NSP1	B*40:02	¹¹² GEIPVAYRKV ¹²¹	0.54	30	20	20
NSP1	B*40:02	⁹ NEKTHVQLSL ¹⁸	0.49	20	20	20
NSP1	B*35:03	⁶¹ LPQLEQPYVF ⁷⁰	0.43	20	20	30
NSP1	B*35:03	¹⁸ LPVLQVRDVL ²⁷	0.41	20	20	10
NSP1	C*15:02	⁷ RTAPHGHVMV ¹⁶	0.52	30	10	10
NSP2	A*02:01	³⁸⁹ ILDGISQYSL ³⁹⁸	0.65	0	0	10
NSP2	A*33:03	⁵²⁹ FVTHSKGLYR ⁵³⁸	0.61	0	0	30
NSP2	A*02:01	²⁸⁸ KLNEEIAIIL ²⁹⁷	0.59	10	10	30
NSP2	A*24:02	¹ AYTRYVDNNF ¹⁰	0.59	30	20	20
NSP2	B*44:03	⁴⁸⁹ KEIKESVQTF ⁴⁹⁸	0.95	0	0	10
NSP2	B*44:03	⁴⁵² EKFKEGVEF ⁴⁶¹	0.89	10	10	30
NSP2	B*40:01	³⁴⁴ GEQKSILSPL ³⁵³	0.85	10	10	20
NSP2	C*03:02	²²⁴ IAFGGCVFSY ²³³	0.59	10	10	20
NSP3	A*24:02	¹⁶ QGYKSVNITF ²⁵	0.79	10	20	20
NSP3	A*24:02	¹⁰⁴⁰ EYKGPITDVF ¹⁰⁴⁹	0.75	20	20	0
NSP3	B*44:03	¹²⁰ EEEFEPSTQY ¹²⁹	0.98	20	10	0
NSP3	B*35:01	⁵⁰² VPTDNYITY ⁵¹¹	0.94	0	0	0
NSP4	A*02:06	¹⁰¹ FVVPGLPGTI ¹¹⁰	0.65	20	10	50
NSP4	B*40:01	⁹⁷ REVGFFVPGI ¹⁰⁶	0.73	20	10	50
NSP5	A*24:02	¹²⁵ VYQCAMRPNF ¹³⁴	0.50	20	20	40
NSP6	A*33:03	⁸⁴ VYMPASWVMR ⁹³	0.53	0	0	10
NSP6	B*35:01	⁴³ LPFAMGIAM ⁵²	0.71	0	0	10
NSP7	B*44:03	⁷³ EEMLDNRATL ⁸²	0.58	10	10	40
NSP8	A*02:01	¹⁵¹ SALWEIQQVV ¹⁶⁰	0.68	20	30	20
NSP8	C*03:02	¹³ AAFATAQEAY ²²	0.45	10	10	40
NSP14	B*35:01	⁴² IPGIPKDMTY ⁵¹	0.92	20	20	30
NSP15	A*33:03	²¹⁵ ELAMDEFIER ²²⁴	0.57	40	20	20
NSP15	B*44:03	¹⁶⁹ GEAVKTQFNY ¹⁷⁸	0.74	10	10	10

We identified 49 10mers as CD8⁺ T cell epitopes, which had $\leq 25\%$ homology with two HCoV viruses (Table 5). 5/44 epitopes gave a score of ≥ 0.90 . 22/49 of these CD8⁺ T cell epitopes were predicted to be restricted through HLA-A, 22/49 were predicted to be

restricted through HLA-B alleles and 5/49 were predicted to be restricted through HLA-C alleles. Again, the peptides that had the highest binding scores and least percentage identified were predicted to be restricted through HLA-B*3501 and HLA-B*4403. The highest binders, which were specific to SARS-CoV2, were identified within the spike protein (⁹⁵TEKSNIIRGW¹⁰⁴), NSP2 (⁴⁸⁹KEIKESVQTF⁴⁹⁸), and NSP3 (¹²⁰EEEEFEPSTQY¹²⁹ and ⁵⁰²VPTDNYITTY⁵¹¹).

3.6. Conservation Analysis of the Candidate CD8⁺ T Cell Epitopes with Binding Scores of ≥ 0.90

We proceeded to investigate if the 18 candidate CD8⁺ T cell epitopes that had a percentage identity of >75% with other HCoV2s and the 31 SARS-CoV2 specific (<25% percentage identity) were conserved within the SARS-CoV2. We found that these candidate epitopes were highly conserved (Tables S2 and S3) and these regions were highly conserved within the new SARS-CoV2 variants as well (Figures S3 and S4).

3.7. Similarity of Candidate Peptides with Published CD8⁺ T Cell SARS-CoV2 Epitopes

Several CD8⁺ T cell epitopes that are restricted through different HLA-A and B alleles have been published [11,18–21]. We proceeded to find out if any of the candidate CD8⁺ T cell epitopes were already identified in patients who were naturally infected with the SARS-CoV2 virus. We found that 20/31 candidate highly conserved T cell epitopes which were found to be specific to the SARS-CoV2 (<25% homology with other HCoV2s) had been identified in infected individuals (Table S4). In our HLA allele prediction analysis using the dominant HLA alleles in Sri Lanka, although some of the epitopes were predicted to be restricted through HLA-B*3501 and HLA-B*4403, some of these epitopes were found to be restricted through HLA-A*0201, A*1101 and HLA-A*0301.

7/18 of the candidate T cell epitopes, which were found to be cross reactive (>75% homology with the other HCoV2s) were also identified from those who were naturally infected. For the candidate CD8⁺ T cell epitopes that were found to be cross reactive with other HCoV2, the predicted HLA allele by us and the HLA allele restriction identified following natural infection were similar in 4/7 epitopes (Table S5).

4. Discussion

In this study we have identified candidate CD8⁺ T cell epitopes, which were highly conserved within SARS-CoV2, and some which show >75% percentage homology with the HCoV2s OC43, HKU1 and NL63, and therefore, are candidates to give rise to cross-reactive T cell responses. The majority of the predicted CD8⁺ T cell epitopes (binding score of ≥ 0.90), which had a high degree of homology with the other three HCoV2s were within the NSP12 and NSP13 proteins. They were predicted to be restricted through HLA-A*2402, HLA-A*0201, HLA-A*0206 and HLA-B alleles B*3501. Therefore, the presence of SARS-CoV2 cross reactive CD8⁺ T cell responses could depend on the frequency of the above HLA-A alleles in a population, as the most cross-reactive candidate CD8⁺ T cell epitopes are restricted through these alleles.

The frequency of HLA-A*0201 and HLA-A*0206 in the Sri Lankan population are 4.9% to 6.6% and 2.1% to 2.4% respectively, while the frequency of HLA-A*2402 is 20.8% to 30.3% [17,22]. HLA-B*3501 frequency in the Sri Lankan population is 21% to 23% [17,22]. In contrast, the most frequent HLA-A alleles in the European, US and Brazilian populations are HLA-A*0201 and A*0206 (24.5% to 27.5%), which are several fold higher than in the Sri Lankan population, while HLA-A*2402 and B*3501 are lower (7.9% to 9.5%) [23–25]. In silico analysis recently showed that HLA-A*0201 was associated with a higher risk of COVID-19, while HLA-A*2402 was shown to associate with higher capacity to present SARS-CoV2 antigens [26]. SARS-CoV2 specific HLA-A*0201 CD8⁺ T cell epitopes were shown to have suboptimal antiviral response and of a reduced frequency when compared to other viral infections such as influenza and Epstein-Barr viral infection [27]. The in-silico analysis showed that countries (Italy, France, Germany, Brazil) in which the most frequent HLA-A allele was A*0201 had the highest COVID-19 case fatality rates (CFRs), whereas

those where HLA-A*2402 allele was the most frequent (India, Iran) had lower CFRs [26]. Therefore, it would be important to investigate if certain HLA alleles, presented CD8⁺ T cell epitopes that associate with protection, whereas if certain other alleles present epitopes that are associated with immunopathology and poor antiviral capacity.

In contrast, the candidate CD8⁺ T cell epitopes, which were highly conserved identified within SARS-CoV2 virus that are likely to be specific, were predominantly identified within the structural proteins (spike, envelope, membrane, and nucleocapsid) and the NSP1, NSP2, and NSP3. 6/31 of these candidate CD8⁺ T cell epitopes (binding score of ≥ 0.90), that were specific to SARS-CoV2 (<25% homology with the HCoVs) were predicted to be restricted through HLA-B*3501, 6/31 through HLA-B*4001, 7/31 through HLA-B*4403 and 8/31 through HLA-A*2402. Only 3/31 of the SARS-CoV2 specific candidate T cell epitopes were predicted to be restricted through HLA-A*0201 or A*0206 alleles, common in Europe, USA, and Brazil. Therefore, HLA-A*2402 and HLA-B*3501, HLA-B*4001, and HLA-B*4403, which are predominant HLA alleles in Sri Lanka and India, may present both highly cross-reactive and SARS-CoV2 specific CD8⁺ T cell epitopes. Indeed, our analysis showed that 20/31 highly conserved, SARS-CoV2 specific candidate CD8⁺ T cell epitopes were already identified in those with natural infection. Although our HLA allele prediction using the dominant HLA alleles in Sri Lanka predicted these epitopes to be restricted through HLA-B*3501 and HLA-B*4403, some were found to be presented through different HLA alleles in those who were naturally infected in Europe and USA [11,18,21]. However, given that these epitopes are highly conserved within the virus, it is possible that they could be presented by multiple HLA alleles, which should be further investigated.

T cell responses of higher magnitude and breadth have been observed in patients who had more severe COVID-19 [12,28]. However, it is not yet known if a higher magnitude and breadth of T cell responses in COVID-19 is associated with protection or immunopathology. There is a debate if cross-reactive T cells cause immunopathology in certain viral infections such as in dengue [29,30] and if such cross-reactive T cells in SARS-CoV2 are protective should be investigated. It is hoped that the candidate epitopes presented here, will be of help in subsequent functional T cell analyses, particularly as viral variants emerge as they were found to be highly conserved within the new UK SARS-CoV2 variant, B.1.1.7 and the new South African variant B.1.351. By focusing on highly conserved regions within and between each coronavirus, the candidate epitopes may be of value in understanding immune responses across populations and for future vaccine design. In addition, since many vaccines for COVID-19 are currently being rolled out, it would be crucial to understand T cell responses that associate with protection and the differences in the functionality of epitope specific T cell responses, presented through different HLA alleles.

5. Conclusions

In summary, we have identified candidate SARS-CoV2 CD8⁺ T cell responses that are highly cross reactive with other HCoVs and that also are specific to SARS-CoV2. Cross-reactive epitopes were predominantly identified from NSP12 and NSP13, while specific epitopes were identified within the structural proteins and NSP1–3. It would be crucial to understand the CD8⁺ T cell epitopes presented through the most frequent HLA alleles, their phenotype and functionality to better understand the immune responses to SARS-CoV2 and possible implications for vaccines.

Supplementary Materials: <https://www.mdpi.com/article/10.3390/v13060972/s1>. Figure S1: Multiple alignment of 9mer and 10mer peptide sequences identified from SARS-CoV2 NSP12 with OC43, HKU1, and NL63. All peptides shown have a predicted binding score of ≥ 0.90 , Figure S2: Multiple alignment of 9mer and 10mer peptide sequences identified from SARS-CoV2 NSP13 with OC43, HKU1 and NL63. All peptides shown have a predicted binding score of ≥ 0.90 , Figure S3: Conservation analysis of six candidate CD8⁺ T cell epitopes that had a high degree of homology (>75%) with HCoVs and a binding score of ≥ 0.90 with the SARS-CoV2 variants in different countries, and with the new UK variant (B.1.1.7) and the new South African variant B.1.351, Figure S4: Conservation analysis of six candidate CD8⁺ T cell epitopes that were specific to SARS-CoV2 (<25% homology

with HCoV) and a binding score of ≥ 0.90 with the SARS-CoV2 variants in different countries, and with the new UK variant (B.1.1.7) and the new South African variant B.1.351. Table S1: Homology of different proteins of the SARS-CoV2 with OC43, HKU1, and NL63 corona viruses, Table S2: Highly conserved candidate CD8+ T cell epitopes that had a high degree of homology ($>75\%$) with HCoV) and a binding score of ≥ 0.90 , Table S3: Highly conserved candidate CD8+ T cell epitopes that were specific to SARS-CoV2 ($<25\%$ homology with HCoV) and a binding score of ≥ 0.90 with the SARS-CoV2 variants in different countries, Table S4: The candidate CD8+ T cell epitopes ($<25\%$ homology with other HCoV) and their predicted HLA allele, the published CD8+ T cell epitopes (marked in red) and their HLA restriction, Table S5: The candidate CD8+ T cell epitopes ($>75\%$ homology with other HCoV) and their predicted HLA allele, the published CD8+ T cell epitopes (marked in red) and their HLA restriction.

Author Contributions: Conceptualization, P.D.P., G.N.M. and G.S.O.; methodology, P.D.P., G.N.M. and G.S.O.; formal analysis, P.D.P. and G.N.M.; data analysis, D.M. and S.D.; funding, C.J., G.N.M. and G.S.O.; writing, P.D.P. and G.N.M.; review and editing, G.N.M., C.J. and G.S.O. All authors have read and agreed to the published version of the manuscript.

Funding: We are grateful to the Centre for Dengue Research, UK Medical Research Council and the Foreign and Commonwealth Office for support.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data is available in the manuscript and its supporting files.

Conflicts of Interest: The authors declare that they have no competing interest.

References

1. Worldometer COVID-19 Coronavirus Outbreak. Available online: <https://www.worldometers.info/coronavirus/> (accessed on 15 January 2021).
2. Wang, G.; Wu, C.; Zhang, Q.; Yu, B.; Lu, J.; Zhang, S.; Wu, G.; Wu, Y.; Zhong, Y. Clinical characteristics and the risk factors for severe events of elderly coronavirus disease 2019 patients. *Zhong Nan Da Xue Xue Bao Yi Xue Ban* **2020**, *45*, 542–548. [PubMed]
3. Albitar, O.; Ballouze, R.; Ooi, J.P.; Sheikh Ghadzi, S.M. Risk factors for mortality among COVID-19 patients. *Diabetes Res. Clin. Pract.* **2020**, *166*, 108293. [CrossRef]
4. Tian, W.; Jiang, W.; Yao, J.; Nicholson, C.J.; Li, R.H.; Sigurslid, H.H.; Wooster, L.; Rotter, J.I.; Guo, X.; Malhotra, R. Predictors of mortality in hospitalized COVID-19 patients: A systematic review and meta-analysis. *J. Med. Virol.* **2020**, *92*, 1875–1883. [CrossRef]
5. Centre, C.R. Mortality Analysis. Available online: <https://coronavirus.jhu.edu/data/mortality> (accessed on 23 December 2020).
6. WHO. COVID-19 Weekly Situation Report; SEARO WHO: WHO SEARO Situation Reports, 14th May 2021; WHO South-East Asia Regional Office: New Delhi, India, 2021; p. 10.
7. Sorci, G.; Faivre, B.; Morand, S. Explaining among-country variation in COVID-19 case fatality rate. *Sci. Rep.* **2020**, *10*, 18909. [CrossRef] [PubMed]
8. Escobar, L.E.; Molina-Cruz, A.; Barillas-Mury, C. BCG vaccine protection from severe coronavirus disease 2019 (COVID-19). *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 17720–17726. [CrossRef]
9. Ng, K.W.; Faulkner, N.; Cornish, G.H.; Rosa, A.; Harvey, R.; Hussain, S.; Ulferts, R.; Earl, C.; Wrobel, A.G.; Benton, D.J.; et al. Preexisting and de novo humoral immunity to SARS-CoV-2 in humans. *Science* **2020**, *370*, 1339–1343. [CrossRef]
10. Lin, A.; He, Z.B.; Zhang, S.; Zhang, J.G.; Zhang, X.; Yan, W.H. Early Risk Factors for the Duration of Severe Acute Respiratory Syndrome Coronavirus 2 Viral Positivity in Patients with Coronavirus Disease 2019. *Clin. Infect. Dis.* **2020**, *71*, 2061–2065. [CrossRef]
11. Peng, Y.; Mentzer, A.J.; Liu, G.; Yao, X.; Yin, Z.; Dong, D.; Dejnirattisai, W.; Rostron, T.; Supasa, P.; Liu, C.; et al. Broad and strong memory CD4⁺ and CD8⁺ T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. *Nat. Immunol.* **2020**, *21*, 1336–1345. [CrossRef] [PubMed]
12. Grifoni, A.; Weiskopf, D.; Ramirez, S.I.; Mateus, J.; Dan, J.M.; Moderbacher, C.R.; Rawlings, S.A.; Sutherland, A.; Premkumar, L.; Jardi, R.S.; et al. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* **2020**, *181*, 1489–1501.e15. [CrossRef] [PubMed]
13. Le Bert, N.; Tan, A.T.; Kunasegaran, K.; Tham, C.Y.L.; Hafezi, M.; Chia, A.; Chng, M.H.Y.; Lin, M.; Tan, N.; Linster, M.; et al. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* **2020**, *584*, 457–462. [CrossRef] [PubMed]

14. Weiskopf, D.; Schmitz, K.S.; Raadsen, M.P.; Grifoni, A.; Okba, N.M.A.; Endeman, H.; van den Akker, J.P.C.; Molenkamp, R.; Koopmans, M.P.G.; van Gorp, E.C.M.; et al. Phenotype and kinetics of SARS-CoV-2-specific T cells in COVID-19 patients with acute respiratory distress syndrome. *Sci. Immunol.* **2020**, *5*, 48. [[CrossRef](#)] [[PubMed](#)]
15. Kim, Y.; Ponomarenko, J.; Zhu, Z.; Tamang, D.; Wang, P.; Greenbaum, J.; Lundegaard, C.; Sette, A.; Lund, O.; Bourne, P.E.; et al. Immune epitope database analysis resource. *Nucleic Acids Res.* **2012**, *40*, W525–W530. [[CrossRef](#)]
16. 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](#)] [[PubMed](#)]
17. Malavige, G.N.; Rostron, T.; Seneviratne, S.L.; Fernando, S.; Sivayogan, S.; Wijewickrama, A.; Ogg, G.S. HLA analysis of Sri Lankan Sinhalese predicts North Indian origin. *Int. J. Immunogenet.* **2007**, *34*, 313–315. [[CrossRef](#)]
18. Shomuradova, A.S.; Vagida, M.S.; Sheetikov, S.A.; Zornikova, K.V.; Kiryukhin, D.; Titov, A.; Peshkova, I.O.; Khmelevskaya, A.; Dianov, D.V.; Malasheva, M.; et al. SARS-CoV-2 Epitopes Are Recognized by a Public and Diverse Repertoire of Human T Cell Receptors. *Immunity* **2020**, *53*, 1245–1257.e5. [[CrossRef](#)] [[PubMed](#)]
19. Nelde, A.; Bilich, T.; Heitmann, J.S.; Maringer, Y.; Salih, H.R.; Roerden, M.; Lubke, M.; Bauer, J.; Rieth, J.; Wacker, M.; et al. SARS-CoV-2-derived peptides define heterologous and COVID-19-induced T cell recognition. *Nat. Immunol.* **2021**, *22*, 74–85. [[CrossRef](#)]
20. Ferretti, A.P.; Kula, T.; Wang, Y.; Nguyen, D.M.V.; Weinheimer, A.; Dunlap, G.S.; Xu, Q.; Nabilsi, N.; Perullo, C.R.; Cristofaro, A.W.; et al. Unbiased Screens Show CD8⁺ T Cells of COVID-19 Patients Recognize Shared Epitopes in SARS-CoV-2 that Largely Reside outside the Spike Protein. *Immunity* **2020**, *53*, 1095–1107.e3. [[CrossRef](#)]
21. Schulien, I.; Kemming, J.; Oberhardt, V.; Wild, K.; Seidel, L.M.; Killmer, S.; Sagar, D.; Daul, F.; Salvat Lago, M.; Decker, A.; et al. Characterization of pre-existing and induced SARS-CoV-2-specific CD8⁺ T cells. *Nat. Med.* **2020**. [[CrossRef](#)]
22. Grifoni, A.; Weiskopf, D.; Lindestam Arlehamn, C.S.; Angelo, M.; Leary, S.; Sidney, J.; Frazier, A.; Phillips, E.; Mallal, S.; Mack, S.J.; et al. Sequence-based HLA-A, B, C, DP, DQ, and DR typing of 714 adults from Colombo, Sri Lanka. *Hum. Immunol.* **2018**, *79*, 87–88. [[CrossRef](#)]
23. Ayo, C.M.; da Silveira Camargo, A.V.; Xavier, D.H.; Batista, M.F.; Carneiro, O.A.; Brandao de Mattos, C.C.; Ricci, O., Jr.; de Mattos, L.C. Frequencies of allele groups HLA-A, HLA-B and HLA-DRB1 in a population from the northwestern region of Sao Paulo State, Brazil. *Int. J. Immunogenet.* **2015**, *42*, 19–25. [[CrossRef](#)]
24. Neville, M.J.; Lee, W.; Humburg, P.; Wong, D.; Barnardo, M.; Karpe, F.; Knight, J.C. High resolution HLA haplotyping by imputation for a British population bioresource. *Hum. Immunol.* **2017**, *78*, 242–251. [[CrossRef](#)] [[PubMed](#)]
25. Requena, D.; Medico, A.; Chacon, R.D.; Ramirez, M.; Marin-Sanchez, O. Identification of Novel Candidate Epitopes on SARS-CoV-2 Proteins for South America: A Review of HLA Frequencies by Country. *Front. Immunol.* **2020**, *11*, 2008. [[CrossRef](#)] [[PubMed](#)]
26. Tomita, Y.; Ikeda, T.; Sato, R.; Sakagami, T. Association between HLA gene polymorphisms and mortality of COVID-19: An in silico analysis. *Immun. Inflamm. Dis.* **2020**, *8*, 684–694. [[CrossRef](#)] [[PubMed](#)]
27. Habel, J.R.; Nguyen, T.H.O.; van de Sandt, C.E.; Juno, J.A.; Chaurasia, P.; Wragg, K.; Koutsakos, M.; Hensen, L.; Jia, X.; Chua, B.; et al. Suboptimal SARS-CoV-2-specific CD8⁺ T cell response associated with the prominent HLA-A*02:01 phenotype. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 24384–24391. [[CrossRef](#)] [[PubMed](#)]
28. Peng, Y.; Mentzer, A.J.; Liu, G.; Yao, X.; Yin, Z.; Dong, D.; Dejnirattisai, W.; Rostron, T.; Supasa, P.; Liu, C.; et al. Broad and strong memory CD4⁺ and CD8⁺ T cells induced by SARS-CoV-2 in UK convalescent COVID-19 patients. *bioRxiv* **2020**. [[CrossRef](#)]
29. Mongkolsapaya, J.; Dejnirattisai, W.; Xu, X.N.; Vasanawathana, S.; Tangthawornchaikul, N.; Chairunsri, A.; Sawasdivorn, S.; Duangchinda, T.; Dong, T.; Rowland-Jones, S.; et al. Original antigenic sin and apoptosis in the pathogenesis of dengue hemorrhagic fever. *Nat. Med.* **2003**, *9*, 921–927. [[CrossRef](#)]
30. Mongkolsapaya, J.; Duangchinda, T.; Dejnirattisai, W.; Vasanawathana, S.; Avirutnan, P.; Jairungsri, A.; Khemnu, N.; Tangthawornchaikul, N.; Chotiyarnwong, P.; Sae-Jang, K.; et al. T cell responses in dengue hemorrhagic fever: Are cross-reactive T cells suboptimal? *J. Immunol.* **2006**, *176*, 3821–3829. [[CrossRef](#)]