COVID-19: Attacks Immune Cells and Interferences with Antigen Presentation through MHC-Like Decoy System

Wenzhong Liu [1,2,*], Hualan Li[2]

[1] School of Computer Science and Engineering, Sichuan University of Science & Engineering, Zigong, 643002, China;

[2] School of Life Science and Food Engineering, Yibin University, Yibin, 644000, China;

[*] Correspondence.Wenzhong Liu, liuwz@suse.edu.cn.

Abstract

The high mortality of COVID-19 is related to poor antigen presentation and lymphopenia. In this present study, domain search results showed that many proteins of the SARS-COV-2 virus had MHC-like domains, which were similar to decoys for the human immune system. MHC-like structures could bind to MHC receptors of immune cells, interfering with antigen presentation. Then the oxygen-free radicals generated by E protein destroyed immune cells after MHC-like of S protein could bind to them. Mutations in the MHC-like region of the viral proteins such as S promoted weaker immune resistance and more robust transmission. S 127-194 were the primary reason for the robust transmission of delta variants. The S 144-162 regulated the formation of S trimer. The mutations of RdRP: G671S and N: D63G of delta variant caused high viral load. S 62-80 of alpha, beta, lambda variants were the important factor for fast-spreading. S 616-676 and 1014-1114 were causes of high mortality for gamma variants infections. These sites were in the MHC-like structure regions.

Keywords: $CD4^+$ T cell; $CD8^+$ T cell; NK cell; Lymphopenia; Delta variant; Neutralizing antibody; N-terminal supersite.

1. Background

The lymphopenia of COVID-19 patients includes $CD4^+$ T cells, $CD8^+$ T cells, B cells, and natural killer (NK) cells, with the damage of $CD8^+$ T cells being more significant(1, 2). No obvious virus infection is detected in lymphocytes and mesenchymal cell(3). Lymphopenia at the initial appearance of COVID-19 is associated with poor prognosis(4). Lymphopenia and its severity are reliable predictors of the clinical outcome of COVID-19, including mortality, intensive care needs, and oxygen requirements(4). Besides, the high fatality rate of COVID-19 is related to the poor performance of MHC II and the low coverage of MHC II(5). The quality of MHCII presented by T cells is an essential prerequisite for T cell-dependent antibody production. The binding capacity of MHC-I epitope load and SARS-COV-2 peptide affects the immunity of T cells to infection(6). Therefore, MHC presentation is closely related to lymphopenia of COVID-19.

Lymphopenia, cell degeneration, necrosis, and atrophy are found in SARS and COVID-19 patients(7). The sort of lymphopenia and apoptosis between SARs and COVID-19 patients seems different. Namely, Lymphopenia in SARs patients precedes apoptosis, while apoptosis in COVID-19 patients precedes lymphopenia(8). The frequency and activation of SARS-COV-2 specific $CD8^+$ T cells increase during severe illness, highlighting differences in T cell responses

associated with disease progression(9). The number of regulatory T cells (Treg) has nothing to do with the severity of the disease, suggesting that T cell exhaustion occurs in a process independent of Treg(10). SARS-COV-2 generates ROS through the combination of E protein and heme(11), in which hydroxyl free radicals can directly destroy the cell membrane and cause damage to immune cells. Immune cells would be directly attacked along the route of antigen recognition and antigen presentation. So, $CD4^+$ T cells, $CD8^+$ T cells, and NK cells are most likely to be destroy and apoptosis in this link of binding to MHC molecules.

Mononuclear macrophages ingest antigens and process them into antigen peptides. Then antigen peptides are combined with surface MHC molecules and are expressed on the cell surface, effectively presenting antigens to helper T lymphocytes. B lymphocytes also have a similar antigen presentation effect. T cells combine with MHC II/antigen to activate B cells. While the BCR of the memory B cell binds to a specific antigen, the antigen is endocytosed by the B cell. After these antigens are cut into fragments, they return to the cell membrane in a state combined with MHC molecules(12). T cells express CD4 or CD8 co-receptors. They recognize non-polymorphic regions of MHC protein on target cells and can bind to partial MHC protein regions(13). Helper T cells express CD4 and recognize MHC class II proteins, while cytotoxic T cells express CD8 and recognize MHC class I proteins(13). NK cells express inhibitory receptors (KIR) of MHC class I molecules. These inhibitory receptors include the human KIR (killer cell Ig-like receptor)(14). Another function of NK cells is recognizing and eliminating cells that cannot express their major histocompatibility complex (MHC) class I molecules(15). Interestingly, individuals with specific MHC alleles are less susceptible to severe forms of malaria(13). It means that the combination of immune cells and MHC is closely related to the susceptibility of certain diseases.

The MHC class II transactivator CIITA induces cell resistance to the Ebola virus and SARS-like coronavirus(16). However, the apparent CD4 conserved residues at the RBD-S1 site of SARS-COV-2 interrupt the CD4-MHC-II interaction for adaptive immune activation(17). The immunity of CD8+ T cells to SARS-COV-2 is related to the severity of COVID-19 and virus control. SARS-COV-2 evades CD8+ T cell surveillance by mutation of the MHC-I restricted epitope of $CD8^+$ T cells(18). Mutant peptides exhibit reduced or abolished MHC-I binding, which is related to the loss of recognition and functional response of $CD8^+$ T cells isolated from HLA-matched COVID-19 patients. However, the proportion of IFN-γ-producing cells in SARS-COV-2 specific $CD8^+$ T cells expressing PD-1 is higher than that of PD-1 cells in multimer+ cells. The SARS-COV-2 specific $CD8^+$ T cells expressing PD-1 are not exhausted and function normally(19). It meant SARS-COV-2 had evolved an MHC-like structure that could bind to the MHC receptor of immune cells. Immune cells that could not attach to the MHC-like form of the virus had survived.

Some viruses have acquired inhibitors that target the MHC class I antigen presentation pathway(20). The cytomegalovirus (CMV) and herpes family encode a series of key molecules required for a targeted immune response(21). All aspects of acute and chronic CMV disease may be controlled by antibodies, NK, and other cells of the innate immune system, as well as $CD8^+$ T and $CD4^+$ T cell(22). About half of the identified genes in cytomegalovirus(23) and beta herpes virus(24) have HCMV homologs(22). The m144(25) and m145 gene families of cytomegalovirus (m17, m145 to m158)(22), m157(26), UL37(27) are all MHC-I-like molecules. The Ly49H NK cell activation receptor recognizes m157(28). Ly49 receptor binds m157 glycoprotein encoded

by mouse CMV (MCMV)(21). Human CMV(HCMV) UL18 binds inhibitory leukocyte immunoglobulin -like receptor R-1(29). Human cytomegalovirus express and distribute a complete library of immune evasion factors for a single MHC class I target(30). Human cytomegalovirus encodes glycoproteins homologous to MHC class I(31). The MHC class I homologs encoded by human cytomegalovirus binds to endogenous peptides(32).

Many viruses have evolved surprising strategies to interfere with the MHC class I antigen presentation pathway(33). After the initial NK response(34), the host will produce adaptive $CD8^+$ T(35) and $CD4^+$ T(36) cellular responses. Viral MHC class I molecules allow evasion of NK cell effector responses in the body(26) and contribute to immune evasion(37). Many studies have shown that MHC class I virus proteins interfere with infected cells recognizing, antigen processing, and presentation(22). The specific recognition of MHC by inhibitory KIR provides excellent protection against a decoy molecule of virus evolution(38). The diversity of the receptor system may be the result of this specific interaction between MHC and KIR molecules. However, NK cells in severely ill patients with COVID-19 are severely depleted. The protective function of inhibitory KIR shows signs of failure. It shows that some regions of human MHC have an irreplaceable role. The MHC-like structures of the virus were precisely in these areas. If a mutation site was in the MHC-like domain, the mutation enhanced the MHC decoy function. In other words, the human immune system hard to neutralize these MHC-like sites by producing antibodies. Otherwise, the antibodies could also bind to MHC proteins. Then the antibodies would affect normal MHC antigen presentation function, causing autoimmune diseases.

The N-terminal domain (NTD) of the S protein and the S2 membrane fusion region may be MHC-like structural sites for the challenging battle between the immune system and the virus. Most antibodies that recognize the SARS-CoV-2 S protein are directed against the receptor-binding domain (RBD)(39). Analysis of the human monoclonal antibody library in the sera of convalescent patients showed that most anti-S antibodies recognize RBD, and a small portion of antibodies recognizes NTD(40). Some NTD-targeted mAbs can effectively inhibit SARS-CoV-2 infection in vitro; in vivo, the immune system uses neutralization and Fc-mediated effector function activities(41). Fc receptor cells generally include B cells, killer cells, and macrophages. Compared with neutralizing RBD targeting antibodies that recognize multiple non-overlapping epitopes, effective NTD targeting neutralizing antibodies appear to target a single supersite(42): N17, N74, N122, and N149. However, popular variants will partially or completely escape the neutralization mediated by human monoclonal antibodies (mAb) targeting the antigen supersite (site i)(43). The variants include B.1.1.7, B. .1.35, and P.1 pedigree. It is difficult for immune system antibodies to neutralize part of the mutation sites in the N-terminal domain and the fusion region of the S2 membrane.

In this present study, we used the domain search method to find that many proteins of the SARS-COV-2 virus have MHC-like structures. It indicates that SARS-COV-2 interferes with antigen presentation and attacks immune cells through the MHC-like systems. The SARS-COV-2 virus protein with MHC-like forms could interfere with the antigen presentation response by binding to the MHC receptor of immune cells. The SARS-COV-2 virus employees the MHC-like structures of the S protein as bait. After the SARS-COV-2 S protein binds to $CD4^+$ T, $CD8^+$ T, and NK cells, the oxygen-free radicals(ROS) generated by the E protein destroys these immune cells, resulting in a decrease in the number of lymphocytes. Through the analysis of the MHC-like enhanced regions of existing popular variants, we found that: 127-194 and 144-162 areas of S

MHC-like of delta variants were in the N-terminal domain (NTD); the 62-80 regions of S MHC-like of alpha, beta, lambda variants were also in the N-terminal domain (NTD); the 616-676 and 1014-1114 regions of S MHC-like of gamma variants were in the S2 membrane fusion region.

2. Method

2.1 Data set

1. The sequences of SARS-COV-2 proteins. The SARS-COV-2 protein sequences came from the NCBI database. Including: S, E, N, M, ORF3a, ORF8, ORF7a, ORF7b, ORF6, ORF10, ORF1ab.

2. MHC-related sequence. We downloaded 18,112 protein sequences of MHC-related from the UniProt data set and searched keyword was "MHC". The MHC-related sequences were compared with the viral proteins to search for the conserved domains.

2.2 Localized MEME tool to scan for conserved domains.

The analysis steps are listed as follows:

1. Download MEME from the official website and subsequently install in the virtual machine ubuntu operating system. The virtual machine was VM 15.2.

2. Download the SARS-COV-2 protein sequence from NCBI official website.

3. Download the fasta format sequence of MHC-related from Uniprot official website, respectively. The search keyword was “MHC”.

4. For each sequence in all MHC-related protein, paired with each SARS-COV-2 protein sequence to generate fasta format files for MEME analysis.

5. For the files generated in Step 4, a batch of 50000 was used to create several batches. It was considered as the limited space of the virtual ubuntu system.

6. In ubuntu, searched the conserved domains (E-value<=0.05) of SARS-COV-2 protein and MHC-related with MEME tools in batches.

7. Collected the result files of conserved domains. Find the domain name corresponding to the motif from the uniprot database.

8. We analyzed the domains’ activity of the each SARS-COV-2 protein according to the characteristics of the MHC-related protein domains.

3. Results

We downloaded MHC-related sequences from the UniProt database. Then compared these sequences with the SARS-COV-2 protein sequences to find the domains related to MHC function. We merged the motif sequences according to the domains of the search results. Both MHC-1 and MHC-2 structures include Ig-like and MHC domains. If a viral protein could bind to the antigen peptide like the MHC protein, the viral protein would have both domains.

3.1 SARS-COV-2 virus proteins had Ig-like domains

Ig-like domains are involved in multiple functions, including cell-to-cell recognition, cell surface receptors, muscle structure, and the immune system. We first listed Ig-like domains of

viral proteins in Table 1. Table 1 shows the structural proteins (S, E, N, M) and non-structural proteins (ORF3a, ORF6, ORF7a, ORF7b, ORF8, ORF10, ORF1ab) of SARS-COV-2 all have Ig-like domains. Ig-like (IPR032165) is a domain composed of approximately 100 residues. Smaller domains (74-90 residues) are observed in several Ig-related molecules (CD2, CD4). The Ig-like motifs of ORF10, E , some subprotein of ORF1ab are the short. ORF7a Ig-like A, ORF8 Ig-like A, ORF3a Ig-like C, N Ig-like B and C, M Ig-like A and C, S Ig-like B and H, 3'-to-5' exonuclease C, 3'-to-5' exonuclease C, helicase B motifs are longer. The Ig-like structures may help the receptor of $CD4^+$ T, $CD8^+$ T, and NK cell recognize the MHC-like area of the viral proteins.

Table 1. Motifs of Ig-like domains of SARS-COV-2 virus proteins

Protein	Alias	Motif	Start	End
S	A	WFHAIH	64	69
	B	KVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEY	129	170
	C	DCTMYIC	737	743
	D	MQMAYR	900	905
	E	YHLMSFPQSAPHG	1047	1059
	F	HVTYVPAQEKNFTTAPAICHDGKAHFPRE	1064	1092
	G	THWFVTQRNFYEPQI	1100	1114
	H	DLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCC SCLKGCCSCGSCCKFDEDDSEPVLKGVKLHY	1199	1272
E	A	AILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRVPD	32	72
M	A	WICLLQFAYANRNRFLYIIKLIFLWLLWPVTLACFVLAAVYRI NWITGGIAIAMACLV	31	88
	B	MWSFNPE	109	115
	C	HHLGRCDIKDLPKEITVATSRTLSYYKLGASQRVAGDSGFAAY SRYRIGNYKLNTDHSSSSDNIA	154	218
N	A	QGLPNNTASWFTALTQHGKED	43	63
	B	DQIGYYRRATRRIRGGDGKMKDLSPRWYFYYLGTGPEAGLPY GANKDGIIWVATEGALNTPKDHIG	82	147
	C	LIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTW	291	330
	D	FKDQVILLNKHIDAYKTFPPTE	346	367
ORF3a	A	MDLFMR	1	6
	B	ASKIITLKKRWQ	59	70
	C	YLYALVYFLQSINFVRIIMRLWLCWKCRSKNPLLYDANYFLC WHTNCYDYCIPYNS	107	162
	D	EHDYQIGGYTEKWESGVKDCVVLHSYFTSDYYQ	181	213
	E	HVTFFIYNKIVDEPEEHVQIHTIDGSSGVVNPVMEPIYD	227	265
ORF6	A	MFHLVDFQVTIAEILLIIMRTFKVSIWNLDYIINLIIKNLSKSLTE NKYSQLDEEQPMEID	1	61
ORF7a	A	MKIILFLALITLATCELYHYQECVRGTTVLLKEPCSSGTYEGNS PFHPLADNKFALTCFSTQFAFACPDGVKHVYQLRARSVSPKLF IRQEEVQELYSPIFLIVAAIVFITLCFTL	1	116
ORF7b	A	MIELSLIDFYLCFLAFLLFLVLIMLIIFWFSLELQDHNETCHA	1	43

ORF8	A	MKFLVFLGIITTVAAFHQECSLQSCTQHQPYVVDDPCPIHFYSK WYIRVGARKSAPLIELCVDEAGSKSPIQYIDIGNYTVSCLPFTIN CQEPKLGSLVVRCSFYEDFLEYHDVRV	1	116
ORF10	A	MGYINVFAFPFTIYSLLLCRMNSRNYIAQVDVVNFNLT	1	38
nsp2	A	IDTKRGVYCCREHEHEIAWYTERSEKSYELQTPF	42	75
	B	CDHCGETSWQTGDFVKATCE	143	162
nsp3	A	SHMYCSFY	100	107
	B	EDDYQGKPLEFGATSAALQPEEEQEEDW	134	161
	C	SEYTGNYQCGHYKHITSKE	1007	1025
	D	HKPIVWH	1169	1175
	E	HFISNSWLMWLIINLVQM	1539	1556
	F	YYVWKSYVHVVDGCNSSTCMMCYKRNRATRVE	1573	1604
	G	SHNIALIWNVKDFMSLSEQLRKQIRSAAKKNNLPF	1888	1922
nsp4	A	MRFRRAFGEYSH	302	313
	B	FLAHIQWMVMFTPLVPFWITIAYIICISTKHFYWFFSNYLKRRV	359	402
nsp6	A	YFNMVYMPASWVMRIMTWLDM	80	100
nsp10	A	SCCLYCRCHIDHPNPKGFCDLKGKYVQIPTTC	72	103
	B	CTVCGMWKGYGCSCDQ	117	132
RNA-depe	A	RYFKYWDQTYHPNCVNCLDDRCI	285	307
ndent RNA	B	FYGGWHNMLKTVYSDVENPHLMGWDYPKCDRAMPNMLRI	594	633
polymerase		M		
	C	SRYWEPEFYEAMYTPH	913	928
2'-O-ribose	A	EHSWNADLYKLMGHFAWWT	173	191
methyltrans				
ferase				
3C-like	A	YDCVSFCYMHHMELP	154	168
proteinase				
3'-to-5'	A	DMTYRRLISMMGFKMNYQVNGYPNMFITREEAIRHVRAWIG	48	88
exonucleas	B	PPPGDQFKHLIP	140	151
e	C	TYACWHHSIGFDYVYNPFMIDVQQWGFTGNLQSNHDLYCQV HGNAHVASCDAIMTRCLAVHECFVKRVDWTIEYPIIG	223	300
	D	RHHANEYRLYLDAYNM	485	500
helicase	A	MPLSAPTLVPQEHYVRITG	233	251
	B	SAQCFKMFYKGVITHDVSSAINRPQIGVVREFLTRNPAWRKA VFISPYN	468	516

3.2 SARS-COV-2 virus proteins had MHC domains

We listed MHC-like domains of viral proteins in Table 2. Table 2 shows that the structural proteins (S, E, N, M) and non-structural proteins (ORF3a, ORF6, ORF7a, ORF7b, ORF8, ORF10, ORF1ab) of SARS-COV-2 have MHC_I-like_Ag-recog domains. Many proteins have the MHC_II_alpha and MHC_II_beta domains. N, ORF10, ORF3a, 2'-O-ribose methyltransferase, nsp10, nsp6, S and have MHC2-interact domains. N, ORF10, ORF3a, ORF8, ORF7b, 2'-O-ribose methyltransferase, nsp6, and S have MHCassoc_trimer domain. ORF10, ORF3a, 3'-to-5' exonuclease, nsp4 has MHC_I_2 domain. S has the MHC_I_C domain.

We downloaded the functional descriptions of the relevant domains from the interpro

database.

Table 2. MHC domains' motifs of SARS-COV-2 virus proteins

Protein	Domain	Alias	Motif	Start	End
S	MHC_I_2	A	CEFQFCNDPFLGVYYHKNNKSWMESE	131	156
		B	WPWYIW	1212	1217
	MHC_I_C	A	KWPWYIWLGFIAGLIAIVMVTIMLCCM	1211	1237
	MHC_II_alpha	A	CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSS	131	162
		B	QIPFAMQMAYR	895	905
		C	LGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSC	1203	1243
	MHC_II_beta	A	WFHAIHVSGTNGTKRFD	64	80
		B	VIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCT FEYVSQPFLMD	127	178
		C	FAMQMAYRFN	898	907
		D	KMSECV	1028	1033
		E	YVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQR	1067	1107
		F	DLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTS CCSCLKGCCSCGSCCKFDEDDSEPV	1199	1264
	MHC_I-like_Ag-recog	A	VTWFHAIH	62	69
		B	VIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCT FEYVSQPFLMDLEGKQGNFKNLREFVF	127	194
		C	RFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGV	346	382
		D	SNKKFLPFQQFGRDIADTTDAVRDPQTLE	555	583
		E	NCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHV NNSYECDIPIGAGICASYQT	616	676
		F	IPFAMQMAYR	896	905
		G	RAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQ SAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVF VSNGTHWFVTQRNFYEPQI	1014	1114
		H	QPELDSFKEELDKY	1142	1155
		I	ESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCC MTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT	1195	1273
	MHC2-interact	A	APAICHDGKAHFPRE	1078	1092
		B	WPWYIW	1212	1217
		C	IVMVTIMLCCMTSCCSCLKGCC	1227	1248
	MHCassoc_trimer	A	YYHKNNKSWMESEFRVYSS	144	162
		B	AHFPREGVFVSNGTHW	1087	1102
		C	ELGKYEQYIKWPWYIW	1202	1217
E	MHC_II_alpha	A	TLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLN	30	66
	MHC_II_beta	A	FVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKN LNSSRVPD	23	72
	MHC_I-like_Ag-recog	A	ALRLCAYCCNI	36	46

	MHCassoc_trimer	A	CAYCCNI	40	46
M	MHC_II_alpha	A	RCDIKDLPKE	158	167
	MHC_II_beta	A	EELKKLLEQWN	11	21
		B	SMWSFNPETN	108	117
		C	HHLGRCDIKDLPKEITVATSRTLSYYKLGASQRVAGDSGFAAYSRYRIGNYKLNTDHSSSSDN	154	216
	MHC_I-like_Ag-recog	A	GHHLGRCDIKD	153	163
N	MHC_II_alpha	A	DQIGYYRRATRRIRGGDGKMKDLSPRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHI	82	146
		B	TDYKHWPQIAQFAPSASAFFGMSRIGMEVT	296	325
	MHC_II_beta	A	SWFTALTQHGKEDLKFPRGQGVPIN	51	75
		B	QIGYYRRATRRIRGGDGKMKDLSPRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIV	83	158
		C	EQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDDKDPNFKDQVILLNKHIDAYKTFPPTEPK	280	369
	MHC_I-like_Ag-recog	A	RRPQGLPNNTASWFTALTQHGKEDL	40	64
		B	DDQIGYYRRATRRIRGGDGKMKDLSPRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQGTTLPKGFY	81	172
		C	LIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDDKDPNFKDQVILLNKHIDAYKTFPPTEPKKD	291	371
	MHC2-interact	A	RWYFYYL	107	113
	MHCassoc_trimer	A	WYFYYL	108	113
ORF3a	MHC_I_2	A	NFVRIIMRLWLCW	119	131
	MHC_II_alpha	A	RIIMRLWLCWKCRSKNPLLYDANYFLCWHTNCYDYCIPYN	122	161
	MHC_II_beta	A	MDLFMR	1	6
		B	NFVRIIMRLWLCWKCRSKNPLLYDANYFLCWHTNCYDYCIP	119	159
	MHC_I-like_Ag-recog	A	INFVRIIMRLWLCWKCRSKNPLLYDANYFLCWHTNCYDYCIPY	118	160
		B	YNKIVDEPEEHVQIH	233	247
		C	NPVMEP	257	262
	MHC2-interact	A	YFLCWHTNC	145	153
	MHCassoc_trimer	A	IMRLWLCWKCRSKNP	124	138
		B	DANYFLCWHTNCYDYCIPYN	142	161
ORF6	MHC_II_alpha	A	DFQVTIAEILLIIMRTFKVSIWNLDYIINLIIKNLSKSLTENKYSQ	6	51
	MHC_II_beta	A	MFHLVDFQVTIAEILLIIMRTFKVSIWNLDYIINLIIKNLSKSLTENKY	1	49

	MHC_I-like_Ag-recog	A	MFHLVD	1	6
		B	TFKVSIWNLDYIINLIIKNLSKSLTENKYSQ	21	51
ORF7a	MHC_II_alpha	A	YEGNSPFH	40	47
	MHC_II_beta	A	GTTVLLKEPCSSGTYEGNSPFHPLADNKFALTCFSTQFAFAC PDGVKHVYQLRARSVSPKLFIRQEEVQELYSPIFLIVAAIVFI	26	110
	MHC_I-like_Ag-recog	A	CELYHYQECVRG	15	26
		B	YEGNSPFHPLADNK	40	53
		C	CPDGVKHVY	67	75
ORF7b	MHC_II_alpha	A	DHNETCHA	36	43
	MHC_II_beta	A	CFLAFLLFLVLIMLIIFWFSLELQDHNETCH	12	42
	MHC_I-like_Ag-recog	A	FYLCFLAFLLFLVLIMLIIFWFSLELQDHNETCHA	9	43
	MHCassoc_trimer	A	MLIIFWFSLELQDHNETCH	24	42
ORF8	MHC_II_alpha	A	TTVAAFHQECSLQSCTQHQPYVVDDPCPIHFYSKWYIRVGA RKSAPLI	11	58
	MHC_II_beta	A	TVAAFHQECSLQSCTQHQPYVVDDPCPIHFYSKWYIRVGAR KSAPLIELCVDEAGSKSPIQYIDIGNYTVSCLPFTINC	12	90
	MHC_I-like_Ag-recog	A	ITTVAAFHQECSLQSCTQHQPYVVDDPCPIHFYSKWYIRVGA RKS	10	54
		B	DEAGSKSPIQYIDI	63	76
		C	NYTVSCLPFTINCQEPK	78	94
	MHCassoc_trimer	A	DDPCPIHFYSKW	34	45
ORF10	MHC_I_2	A	AFPFTIYSLLLCRMNSRNYIAQVDVVN	8	34
	MHC_II_beta	A	MGYINVFAFPFTIYSLLLCRMNSRNYIAQVDVVNFNLT	1	38
	MHC_I-like_Ag-recog	A	MGYINVFAFPFTIYSLLLCRMNSRNYIAQVDVVNFN	1	36
	MHC2-interact	A	MGYINVFAFPFTIYSLLLC	1	19
	MHCassoc_trimer	A	MGYINVFAFPFTIYSLLLCRMNSRNYIA	1	28
nsp2	MHC_II_alpha	A	RGVYCCREHEHEIAW	46	60
	MHC_II_beta	A	DTKRGVYCCREHEHEIAWYTERSEKSYELQTPF	43	75
	MHC_II_beta	B	DGFMGRIRSVYPVASPNECNQMCLSTLMKCDHCGETSWQT	114	153
	MHC_I-like_Ag-recog	A	FIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEI	41	77
	MHC_I-like_Ag-recog	B	KLDGFMGRIRSVYPVASPNECNQMCLSTLMKCDHCGETSW QTGDFVKATCEFCGTENLTKEGATTCGYLPQNA	112	184
	MHC_I-like_Ag-recog	C	CPACHNSEVGPEHSLAEYHN	190	209
nsp3	MHC_II_beta	A	DYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNTWCI RCLWS	1153	1197

	MHC_II_beta	B	IMQLFFSYFAVHFISNSWLMWLIINLVQMAPISAMVRMYIFF ASFYYVWKSYVHVVDGCNSSTCMMCYKRNRATRVECT	1528	1606
	MHC_II_beta	C	CSARHIN	1876	1882
	MHC_I-like_Ag-recog	A	ASHMYCSFYPPDEDEEEGDCEEEEF	99	123
	MHC_I-like_Ag-recog	B	QPEEEQEEDW	152	161
	MHC_I-like_Ag-recog	C	NEKQEILGTVSWNLREMLAHAEETR	544	568
	MHC_I-like_Ag-recog	D	WCIRCLW	1190	1196
	MHC_I-like_Ag-recog	E	SWLMWLIINLVQMAPISAMVRMYIFFASFYYVW	1544	1576
	MHC_I-like_Ag-recog	F	RRSFYVYANGGKGFCKLHNWNCVNCDT	1613	1639
nsp4	MHC_I_2	A	QWMVMFTPLVPFWITIAYIICISTKHFYWFFSNYLKRR	364	401
	MHC_II_alpha	A	QWMVMFTPLVPFWI	364	377
	MHC_II_beta	A	EYCRHGTCER	219	228
	MHC_II_beta	B	HIQWMVMFTPLVPFWITIAYIICISTKHFYWFFSNYLKRR	362	401
	MHC_I-like_Ag-recog	A	PVHVMSKHTDFSSEIIGYKAIDGGVTRDIASTDTCFANKHAD FDTWFSQR	29	78
	MHC_I-like_Ag-recog	B	FYLTNDVSFLAHIQWMVMFTPLVPFWITIAYIICISTKHFYWF	351	393
nsp6	MHC_II_beta	A	QSTQWSLFFFLYENAFLPFAMGIIAMSAFAMMFVKH	27	62
	MHC_II_beta	B	WVMRIMTWLDM	90	100
	MHC_I-like_Ag-recog	A	SWVMRIMTWLDM	89	100
	MHC2-interact	A	MVYMPASWVMRIMTWLDM	83	100
	MHCassoc_trimer	A	GTHHWL	9	14
nsp7	MHC_I-like_Ag-recog	A	WAQCVQLHND	29	38
nsp8	MHC_I-like_Ag-recog	B	KSEFDRDAAMQRKLEKMADQAMTQMYKQARSEDKRAKVT SAMQTM	46	90
nsp10	MHC_II_beta	A	NMDQESFGGASCCLYCRCHIDHPNP	62	86
	MHC_II_beta	B	WKGYGCSCDQLREPMLQ	123	139
	MHC_I-like_Ag-recog	A	TPEANMDQESFGGASCCLYCRCHIDHPN	58	85
	MHC2-interact	A	YCRCHIDHPNPKGFCD	76	91
RNA-dependent RNA polymerase	MHC_II_alpha	A	RKHTTCCSLSHRFYR	640	654
	MHC_II_alpha	B	YWEPEF	915	920
	MHC_II_beta	A	ERLKLFDRYFKYWDQTYHPNCVNCLDDRCILH	278	309
	MHC_II_beta	B	YSDVENPHLMGWDYPKCDRAMPNMLRIMA	606	634
	MHC_II_beta	C	HPNQEYADVFHLYLQYIRKLHDELTGHMLDMYSVM	872	906

	MHC_II_beta	D	SRYWEPEFYEAMYT	913	926
	MHC_I-like_Ag-recog	A	SNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRL	78	119
	MHC_I-like_Ag-recog	B	FDRYFKYWDQTYHPNCVNCLDDRCILH	283	309
	MHC_I-like_Ag-recog	C	KFYGGWHNMLKTVYSDVENPHLMGWDYPKCDRAMPNMLRIMASLVLARKHTTCCSLSHRFYRLANECAQVLSEMVMCGGS	593	672
	MHC_I-like_Ag-recog	D	KCWTETDLTKGPHEFCSQHTMLVKQGDDY	798	826
	MHC_I-like_Ag-recog	E	LMIERFVSLAIDAYPLTKHPNQEYADVFHLYLQYIRKLHDELTGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHT	854	929
2'-O-ribose methyltransferase	MHC_II_alpha	A	TEHSWNADLYKLMGHFAWW	172	190
	MHC_II_beta	A	PREQIDGYVMHANYIFWRNT	215	234
	MHC_I-like_Ag-recog	A	HSWNADLYKLMGHFAWWT	174	191
	MHC_I-like_Ag-recog	B	PREQIDGYVMHANYIFWR	215	232
	MHC2-interact	A	MGHFAWWTAF	184	193
	MHCassoc_trimer	A	MGHFAWW	184	190
3C-like proteinase	MHC_II_beta	A	YMHHMEL	161	167
	MHC_I-like_Ag-recog	A	YDCVSFCYMHHME	154	166
3'-to-5' exonuclease	MHC_I_2	A	CWHHSIGFDYVYNPFMIDVQQW	226	247
	MHC_II_beta	A	EGLCVDIPGIPKDMTYRRLISMMGFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEGCHATRE	36	99
	MHC_II_beta	B	CWHHSIGFDYVYNPFMIDVQQW	226	247
	MHC_II_beta	C	AVCRHHANEYRLYLDAYNMMISAGFSLWVYKQ	482	513
	MHC_I-like_Ag-recog	A	IPGIPKDMTYRRLISMMGFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEGCHATR	42	98
	MHC_I-like_Ag-recog	B	DTYACWHHSIGFDYVYNPFMIDVQQWGFTGNLQSNHDLYCQVHGNAHVASCDAIMTRCLAVHECFVKRVDWTIEYPIIGDELKINAACRKVQHM	222	315
	MHC_I-like_Ag-recog	C	CRHHANEYRLYLDAYNMMISAGFSLWVYKQFDTYNLWNTF	484	523
endoRNAse	MHC_I-like_Ag-recog	A	RNLQEFKPRSQMEIDFLELAMDEFIERYKLEGYAFEHI	198	235
helicase	MHC_I-like_Ag-recog	A	RPFLCCKCCYDHVISTSH	22	39
	MHC_I-like_Ag-recog	B	EPEYFNSVCRLMKTIGPDMFLGTCRR	418	443
	MHC_I-like_Ag-recog	C	REFLTRNPAWRKAVFISPYNSQNA	497	520

The members of MHC_I_2 (PF14586) are called retinoic acid-inducible proteins. They are ligands that activate the immune receptor NKG2D. NKG2D is widely expressed on natural killer cells, T cells, and macrophages. MHC_I_C (PF06623) represents the C-terminal region of MHC class I antigen. MHC_I-like_Ag-recog (IPR011161) is an MHC class I antigen recognition sample. Class I MHC glycoproteins are expressed on the surface of all somatic nucleated cells, except neurons. MHC class I receptors present peptide antigens synthesized in the cytoplasm, including self-peptides (offered for self-tolerance) and foreign peptides (such as viral proteins). These antigens are produced by degraded protein fragments transported by the TAP protein (antigenic peptide transporter) to the endoplasmic reticulum, where they can bind to MHC I molecules and then transport them to the cell surface through the Golgi apparatus. MHC class I Receptors display antigens recognized by cytotoxic T cells that can destroy virus-infected or malignant (self-peptide excess) cells. $CD8^+$ T toxic cells and NK cells can recognize class I MHC proteins.

MHC_II_alpha (SM00920) is the alpha domain of class II histocompatibility antigen. MHC_II_beta (SM00921) is the beta domain of class II histocompatibility antigen. Class II MHC glycoproteins are expressed on the surface of antigen-presenting cells (APC), including macrophages, dendritic cells, and B cells. MHC II protein presents extracellular peptide antigens derived from foreign substances such as bacteria. Proteins from pathogens are degraded into peptide fragments within the APC. These fragments are sequestered into endosomes to bind to MHC class II proteins before being transported to the cell surface. MHC class II receptors display antigens for recognition by helper T cells and Inflammatory T cells. CD4+T helper cells recognize MHC class II proteins.

MHC2-interact (PF09307) is the interaction domain of CLIP and MHC2. Members of this family are found in Class II Invariant Chain Related Peptides (CLIP). They are required for binding to the Class II Major Histocompatibility Complex (MHC) in the MHC Class II processing pathway. MHCassoc_trimer (PF08831) is an invariant chain trimerization domain related to class II MHC. The folding and positioning of MHC class II heterodimers require class II-related consistent chain peptides. This domain participates in the trimerization of the ectoderm and interferes with DM/Class II binding. The trimeric protein forms a cylindrical shape, which is considered necessary for the interaction between the invariant and class II molecules.

We noticed that S protein could form a trimer structure and had three MHCassoc_trimer domains: MHCassoc_trimer A, B, and C. MHCassoc_trimer A is in S1 protein, but MHCassoc_trimer B and C in S2 protein. It represents that MHCassoc_trimer plays an important role in the formation of S protein trimer.

3.3 MHC-like structures had a decoy function against the immune system

Above analysis shows that structural proteins and non-structural proteins can bind to T ($CD4^+$ T and $CD8^+$ T) and NK immune cells through MHC-like structures. The binding prevented the MHC receptors of immune cells from securing to MHC, causing interference in antigen presentation. In addition, E protein generates oxygen free radicals (ROS) after attaching to heme, and the hydroxyl free radicals directly damaged cell membranes(11). After CD4+T, CD8+ T and NK cells were bound to the MHC-like structure of S protein, the hydroxyl free radicals generated by E protein destroyed these immune cell membranes. It caused immune cells to die due to oxidative stress. For these two reasons, the MHC-like structure of the SARS-COV-2 virus protein had a decoy function against immune cells.

We noticed that N protein could form a multimer, S protein could form a trimer structure, and E could form a pentameric channel structure. The ORF3a protein and ORF8 protein can form a dimer structure, respectively. The dimer of ORF3a protein (or ORF8 protein) has a groove structure. The Ig-like sites of the ORF3a protein do not fully overlap with the MHC-like areas (Table 1 and Table 2). However, Table 2 shows that the MHC II and MHC I of ORF3a are located at the "CWKCR" heme-binding area(44) and upstream and downstream. But the sites of MHC structures are not near the groove structure on the ORF3a crystal structure view (PDBID: 6xdc). Therefore, ORF3a is unlikely to have the ability to bind antigen peptides.

The Ig-like structure of the transmembrane protein ORF8 overlaps with the MHC-like system. The dimer of ORF8 has no rod-like structure. Table 2 and the crystal structure view of ORF8 (PDBID: 7jtl) show that the MHC II and MHC I structures of ORF8 include sites near the groove structure. Therefore, ORF8 may trapp antigen peptides through the MHC structure and interfering with antigen presentation. Besides, ORF8 captures MHC-1 and reroutes to autophagosomes for degradation(45). Table 2 indicates that the MHC I-like domain of ORF8 is MHC_I-like_Ag-recog. ORF8 also has the MHCassoc_trimer domain. The MHC II-like domain of ORF8 overlaps with the MHC_I-like_Ag-recog and MHCassoc_trimer structures. So ORF8 may trap MHC I by MHC_I-like_Ag-recog and MHCassoc_trimer domains. Therefore, the MHC-like system of ORF8 has a decoy function for MHC I or antigen peptides.

3.4 MHC-like enhanced regions of S protein mutation

If the S mutation site was in the MHC-like domain, the mutation enhanced the MHC decoy function. Then the human immune system hard to neutralize these MHC-like sites by producing antibodies. Otherwise, it would affect the normal MHC antigen presentation function by combing MHC and the antibodies. Based on this principle, we analyzed several significant variants that were now popular to determine the MHC-like enhanced region of the S protein.

SARS-COV-2 Delta variant. The B.1.617.2/Delta variant is highly confluent, especially in infected hamsters more pathogenic than the prototype SARS-COV-2(46). The virus is more infectious and directly reduces the efficacy of antibodies produced by infection and vaccines. It is the most prevalent and difficult mutant virus strain in the world.

B.1.617.2/Delta variant mutation sites include(47): T19R, G142D, E156G, F157Δ, R158Δ, L452R, T478K, D614G, P681R, D950N. Table 3 shows that G142D, E156G, F157Δ, and R158Δ are all in the MHC_II_alpha A, MHC_II_beta B, MHC_I-like_Ag-recog B domains. Both G142D and E156G are in the MHC_I_2 A domain. E156G, F157Δ, and R158Δ are all located in the MHCassoc_trimer A domain. Other mutation sites are not in the MHC-like domain. These four MHC-like domains are highly overlapping. Combining these four MHC-like domain sites, the MHC-like enhanced distribution area of B.1.617.2/Delta variant S protein is 127-194. It is in the N-terminal domain (14-305 residues, the S1 protein region) (48). Among them, MHCassoc_trimer (144-162) plays an essential role in forming S trimer. The MHC-like enhanced distribution area of the S protein has MHC-I_like and MHC-II_like functions, so it can also bind to CD4+ T, CD8+ T, NK cells.

Table 3 The S1 mutation sites of the B.1.617.2/Delta variant are in the MHC-like domains

Domain	Alias	Start	End	T19R	G142D	E156G	F157Δ	R158Δ	L452R	T478K	D614G	P681R	D950N
MHC_I_2	A	131	156		V	V							
	B	1212	1217										
MHC_I_C	A	1211	1237										
MHC_II_alpha	A	131	162		V	V	V	V					
	B	895	905										
	C	1203	1243										
MHC_II_beta	A	64	80										
	B	127	178		V	V	V	V					
	C	898	907										
	D	1028	1033										
	E	1067	1107										
	F	1199	1264										
MHC_I-like_Ag-recog	A	62	69										
	B	127	194		V	V	V	V					
	C	346	382										
	D	555	583										
	E	616	676										
	F	896	905										
	G	1014	1114										
	H	1142	1155										
	I	1195	1273										
MHC2-interact	A	1078	1092										
	B	1212	1217										
	C	1227	1248										
MHCassoc_trimer	A	144	162			V	V	V					
	B	1087	1102										
	C	1202	1217										

SARS-COV-2 Gamma variant. The seropositivity rate of SARS-COV-2 antibody is very high. There is a greater chance of infectivity and death. The S mutation sites of Gamma variants are(49, 50): L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I, V1176F. Table 4 shows that D138Y is in the MHC_I_2 A, MHC_II_alpha A, MHC_II_beta B, and MHC_I-like_Ag-recog B domains. R190S, H655Y and T1027I are in MHC_I-like_Ag-recog B, E, G domains, respectively. Compared with the B.1.617.2/Delta variant, the MHC-like domain of the S protein of the SARS-COV-2 Gamma variant has two mutation points, H655Y, and T1027I. It shows that Gamma variant S participates in receptor binding H655Y and participates in membrane fusion T1027I in the MHC-LIKE region. Therefore, the infection rate and mortality of Gamma variants are high. However, the Gamma variant does not have a mutation site located in the MHCassoc_trimer domain. It may not enhance the immune escape of the regulatory region of the S trimer. Therefore, the MHCassoc_trimer domain (144-162) is an important reason why the Delta variant spreads infection faster than the Gamma variant.

Table 4 P.1 /Gamma Variant S1 and S2 mutation sites are in the MHC-like domains

Domain	Alias	Start	End	L18F	T20N	P26S	D138Y	R190S	K417T	E484K	N501Y	D614G	H655Y	T1027I	V1176F
MHC_I_2	A	131	156				V								
	B	1212	1217												
MHC_I_C	A	1211	1237												
MHC_II_alpha	A	131	162				V								
	B	895	905												
	C	1203	1243												
MHC_II_beta	A	64	80												
	B	127	178				V								
	C	898	907												
	D	1028	1033												
	E	1067	1107												
	F	1199	1264												
MHC_I-like_Ag-recog	A	62	69												
	B	127	194				V	V							
	C	346	382												
	D	555	583												
	E	616	676										V		
	F	896	905												
	G	1014	1114											V	
	H	1142	1155												
	I	1195	1273												
MHC2-interact	A	1078	1092												
	B	1212	1217												
	C	1227	1248												
MHCassoc_trimer	A	144	162												
	B	1087	1102												
	C	1202	1217												

SARS-COV-2 Alpha variant. B.1.1.7 /Alpha variant has a more tremendous increase in the transmission rate than the earlier SARS-COV-2 virus. However, there is no significant difference in overall mortality. The mutation site of S in the B.1.1.7/Alpha variant is(51): Δ69- 70, Δ144, Δ145, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H. Table 5 shows that Δ69-70 is at MHC_II_beta A and MHC_I-like_Ag-recog A domains. Δ144 and Δ145 is in MHC_I_2 A, MHC_II_alpha A, MHC_II_beta B, MHC_I-like_Ag-recog B, MHCassoc_trimer A domains. A570D is in MHC_I-like_Ag-recog D domain. Δ144 and A570D are all in the S1 protein. There are no MHC-like domain mutations in the S2 protein. The mutation at position 144-145 is in the Alpha variant S. The mutation at position 156 is in the delta variant S. They are in the MHCassoc_trimer A domain. The mutations at position 144-162 may enhance the immune escape of the MHC-like region involved in receptor binding and regulate trimer's formation.

Table 5. B.1.1.7/Alpha variant S1 mutation sites are in the MHC-like domains

Domain	Alias	Start	End	Δ69-70	Δ144-Δ145	N501Y	A570D	D614G	P681H	T716I	S982A	D1118H
MHC_I_2	A	131	156		V							
	B	1212	1217									
MHC_I_C	A	1211	1237									
MHC_II_alpha	A	131	162		V							
	B	895	905									
	C	1203	1243									
MHC_II_beta	A	64	80	V								
	B	127	178		V							
	C	898	907									
	D	1028	1033									
	E	1067	1107									
	F	1199	1264									
MHC_I-like_Ag-recog	A	62	69	V								
	B	127	194		V							
	C	346	382									
	D	555	583				V					
	E	616	676									
	F	896	905									
	G	1014	1114									
	H	1142	1155									
	I	1195	1273									
MHC2-interact	A	1078	1092									
	B	1212	1217									
	C	1227	1248									
MHCassoc_trimer	A	144	162		V							
	B	1087	1102									
	C	1202	1217									

SARS-COV-2 Beta variant. The vaccine is effective against the B.1.351/Beta variant. The mutation sites of B.1.351/Beta variant S are(52): L18F, D80A, D215G, LAL241-243Δ, K417N, E484K, N501Y, D614G, A701V. Table 6 shows that D80A is in the MHC_II_beta A domains. Most of the other mutation sites are not in the MHC-like domains. It shows that most of the mutation sites do not affect the MHC-like domain.

Mutation sites G75V, T76I of the SARS-COV-2 C.37/lambda(53) variant are both in the MHC_II_beta A domain. So S 62-80 of SARS-COV-2 alpha, beta, lambda variants were the first MHC-like enhanced distribution area. The second MHC-like enhanced distribution area of S protein is 127-194, located in the N-terminal domain (14-305 residues) of S1 protein. MHCassoc_trimer (144-162) is a trimer of S Formation, which plays an important role. The third and fourth MHC-like enhanced distribution areas of S protein are MHC_I-like_Ag-recog E (616-676), MHC_I-like_Ag-recog G (1014-1114).

Table 6. The S1 mutation site of the B.1.351/Beta variant is located in the MHC-like domain

Domain	Alias	Start	End	L18F	D80A	D215G	LAL241-243Δ	K417N	E484K	N501Y	D614G	A701V
MHC_I_2	A	131	156									
	B	1212	1217									
MHC_I_C	A	1211	1237									
MHC_II_alpha	A	131	162									
	B	895	905									
	C	1203	1243									
MHC_II_beta	A	64	80		V							
	B	127	178									
	C	898	907									
	D	1028	1033									
	E	1067	1107									
	F	1199	1264									
MHC_I-like_Ag-recog	A	62	69									
	B	127	194									
	C	346	382									
	D	555	583									
	E	616	676									
	F	896	905									
	G	1014	1114									
	H	1142	1155									
	I	1195	1273									
MHC2-interact	A	1078	1092									
	B	1212	1217									
	C	1227	1248									
MHCassoc_trimer	A	144	162									
	B	1087	1102									
	C	1202	1217									

3.5 Mutations in MHC-like regions of RNA-dependent RNA polymerase and N protein causing high viral load of Delta variant

The viral load of Delta variant patients is very high. It shows that the replication activity of the Delta variant of the SARS-COV-2 virus is very active. The viral proteins directly related to viral replication activities are orf1ab and N proteins. We searched the orf1ab and N protein mutation sites of five variants of Alpha, Beta, Gama, Delta, and Lambda from “CORONAVIRUS CORONAVIRUS ANTIVIRAL & RESISTANCE DATABASE” (https: // covdb. stanford. edu/ page/ mutation-viewer). Then compared with Table 2 to find the MHC-like enhanced region sites of orf1ab and N proteins (Table 7). Table 7 shows that the orf1ab and N protein mutation sites of Alpha, Beta, and Gama variants are not in the MHC-like region. The N protein mutation site of the Lambda variant is also not in the MHC-like area. However, the mutation sites of Delta (orf1ab and N protein) and Lambda (orf1ab protein) are in the MHC-like region. The RdRP: G671S mutation site of Delta variant orf1ab is at MHC_I-like_Ag-recog C. The nsp3:F1569V of Delta variant orf1ab is at MHC_II_beta B and MHC_I-like_Ag-recog E. The D63G of Dalta variant N is in

MHC_II_beta A, MHC_I-like_Ag-recog A. Both N and RNA-dependent RNA polymerase are directly related to virus replication. In addition, the interaction between N and Nsp3 is essential for connecting the viral genome for processing. So, table 7 indicates that the mutations of RdRP: G671S and N: D63G enhanced the immune escape ability of the Delta variant virus during the replication process.Therefore, the replication activity for this variant is very active

Table 7. MHC-like enhancement sites of orf1ab and N proteins

Protein	Variant	Code	Mutation site	MHC-like enhancement site
orf1ab	Alpha	B.1.1.7	nsp3:T183I, nsp3:A890D, nsp3:I1412T, nsp6:SGF106-108, RdRp:P323L	-
	Beta	B.1.351	nsp2:T85I, nsp3:K837N, 3CL:K90R, nsp6:SGF106-108, RdRP:P323L	-
	Gama	P.1	nsp3:S370L, nsp3:K977Q, nsp6:SGF106-108, RdRP:P323L, nsp13: E341D	-
	Delta	B.1.617.2	nsp3:A488S, nsp3:P1228L, nsp3:P1469S, nsp4:V167L, nsp4:T492I, nsp6:T77A, RdRP:P323L, RdRP: G671S, nsp13:P77L, nsp14:A394V	RdRP: G671S (MHC_I-like_Ag-recog C)
	Lambda	C.37	nsp3:T428I, nsp3:P1469S, nsp3:F1569V, nsp4:L438P, nsp4:T492I, 3CL:G15S, nsp6:SGF106-108, RdRP:P323L	nsp3:F1569V(MHC_II_beta B,MHC_I-like_Ag-recog E)
N	Alpha	B.1.1.7	D3L, R203K, G204R, S235F	-
	Beta	B.1.351	T205I	-
	Gama	P.1	P80R, R203K, G204R	-
	Delta	B.1.617.2	D63G, R203M, G215G,D377Y	D63G (MHC_II_beta A, MHC_I-like_Ag-recog A)
	Lambda	C.37	P13L, R203K, G204R, G214C	-

4. Discussion

4.1 Genetic variation of MHC protected immune cells

The major histocompatibility complex (MHC) molecule is a cell surface protein complex encoded in the human leukocyte antigen (HLA) locus(54). The genetic variation of the three major histocompatibility complex (MHC) class I genes (human leukocyte antigen A [HLA-A], -B, and -C genes) affect the susceptibility and severity of COVID-19 disease(55). The HLA gene complex is closely linked to genes, and there is little exchange between homologous chromosomes. HLA loci located on the same chromosome constitute a closely linked gene group (including HLA-I and II genes), called haplotype or haplotype. A haplotype is inherited as a unit. To the offspring, it is called haplotype genetics. However, HLA haplotypes are not the main risk/protective factor for SARS-COV-2 infection or severity in the Israeli population(56). It indicates that the genetic variation of MHC structure may protect immune cells that can bind to MHC to a certain extent.

In this present study, we found that many proteins of the SARS-COV-2 virus have MHC-like structures recognized by MHC receptors. CD4 or CD8 co-receptors expressed by T cells can bind

to part of MHC proteins (13). The inhibitory receptors of NK cells can also bind to the MHC-1 receptor recognition structure. Therefore, the S protein of the SARS-COV-2 virus could bind to $CD4^+$T, $CD8^+$T, and NK cells through MHC-like structures. Then the ROS generated by the E protein destroyed these immune cells (11), resulting in a decrease in lymphocytes. The genetic variation of HLA may produce MHC molecules that could not bind to the viral MHC-like structure. It was helpful for immune cells to evade the attachment and positioning of SARS-COV-2 MHC-like proteins. In this situation, the antigen presentation response would not be disturbed, and immune cells (such as $CD4^+$T, $CD8^+$T, NK cells) would be protected from the virus's ROS damage.

4.2 S mutations in the MHC-like regions promoted weaker immune resistance and more robust transmission.

If a mutation site was in the MHC-like domain, the mutation enhanced the MHC decoy function. It challenged the production of antibodies to neutralize these MHC decoy sites. If antibodies could attach to the MHC-like proteins, the antibodies could also bind to MHC proteins. Then the normal MHC antigen presentation function would be affected, and the body would appear autoimmune diseases. It is not a piece of good news for CD4+ T, CD8+ T, NK cells, and other immune cells that can bind to MHC. These immune cells could indiscriminately bind to MHC-like structures of the S protein, and were attacked by ROS from the E protein. Then the immune system could not effectively perform the antigen presentation for the SARS-COV-2 virus protein. It also could not produce the neutralizing antibody effectively. Moreover, the probability of infected cells was killed by immune cells would be significantly reduced.

This present study found that neutralizing antibodies were challenging to generated for mutations in S MHC-like regions 127-194 and 144-162. It occurred with delta variant infections. The delta variant was the SARS-COV-2 virus with a robust transmission. The S 62-80 mutations of SARS-COV-2 alpha, beta, lambda variants had a similar situation.McCallum, M. et al. found (43): R246A substitution reduces the binding of S2L28, S2M28, and S2X333. This substitution significantly affected the binding of S2X28 and mAb 4A8. The L18F, D80A, D253G/Y, or S255F variants only abolish the combination of S2L28 and NTD. The L18F substitution exists in B.1.351 and P. 1 pedigree. The Y144 deletion abolished the binding to S2M28, S2X28, S2X333, and 4A8 instead of S2L28. It explains that these mAbs have lost the ability to neutralize the B.1.1.7 S pseudovirus, which contains this deletion. The H146Y mutant reduces S2M28 , S2X28, especially the combination of 4A8. The binding of all site i-specific NTD mAbs to B.1.351 NTD is abolished, and 4A8 does not recognize this NTD variant. The evidence indicates that the NTD variants located in MHC-like regions 127-194 and 144-162 enhance the immune escape of the virus and increase the efficiency of virus transmission.

We also found that the immune system was challenging to generate neutralizing antibodies against mutations in S MHC-like 616-676 and 1014-1114 regions. It happened to gamma variant infections. The gamma variant was the SARS-COV-2 virus that caused high mortality. Rita E. Chen et al. finnd that specific monoclonal antibodies have reduced or weakened neutralizing activity against B.1.351, B.1.1.28, B.1.617.1, and B.1.526 viruses in cell culture(57). And the neutralizing effect of antibodies against H655Y and T1027I mutation sites is not apparent(57). It shows that the variants in MHC-like regions 616-676 and 1014-1114 also strengthen the immune escape of the virus, and enhance the virus's receptor engagement and membrane fusion ability.

5. Conclusion

The high mortality rate of COVID-19 is related to poor antigen presentation and lymphopenia. MHC genetic variations may protect immune cells. Cytomegalovirus (CMV) and the herpes family encode a series of MHC-like molecules required for targeted immune responses to achieve immune escape. This present study used bioinformatics methods to study whether the SARS-COV-2 virus proteins also had MHC-like structures. The domain search results indicate that MHC receptors could recognize many proteins of the SARS-COV-2 virus because of their MHC-like domains. The MHC-like structures were equivalent to bait against the human immune system. We believed that the SARS-COV-2 virus proteins with MHC-like structures could bind to the MHC receptor of immune cells to interfere with the antigen presentation response. After the S protein was bound to CD4+T, CD8+T, and NK cells through MHC-like structures, ROS generated by the E protein destroyed these immune cells, decreasing the number of lymphocytes. Mutations in the MHC-like region of the proteins such as S protein promoted weaker immune resistance and more robust transmission. The mutations in the S MHC-like 127-194 and 144-162 regions were the reason for the entire transmission of delta variant. It is worth noting that the 144-162 region regulates the formation of S trimer. Mutations in S MHC-like 62-80 of SARS-COV-2 alpha, beta, lambda variants were one important factor for fast-spreading. The mutations in the S MHC-like 616-676 and 1014-1114 regions were causes of high mortality for gamma variants infections. The mutations of RdRP: G671S and N: D63G of delta variant caused high viral load.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

The datasets and results supporting the conclusions of this article are available at : https://pan.baidu.com/s/1A0DlmP0po3QK_guBOzoWpw, code：x7c9
or: https://mega.nz/folder/M3p0nASC#CO8HgROV9Ydzr9d7RVEMEQ

Competing interests

The authors declare that they have no competing interests.

Funding

This work was funded by a grant from the Talent Introduction Project of Sichuan University of Science and Engineering (award number: 2018RCL20, grant recipient: WZL).

Author’s contribution

Funding was obtained by WZL. Besides, design, analysis and writing are finished by WZL, while data curation and manuscript check are undertaken by HLL. Both authors have read and agreed to the published version of the manuscript.

Acknowledgements

Not applicable.

Author details

[1] School of Computer Science and Engineering, Sichuan University of Science & Engineering, Zigong, 643002, China.
[2] School of Life Science and Food Engineering, Yibin University, Yibin, 644000, China.

References

1. Huang, C., Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, and X. Gu. 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The lancet* 395: 497-506.
2. Xu, Z., L. Shi, Y. Wang, J. Zhang, L. Huang, C. Zhang, S. Liu, P. Zhao, H. Liu, and L. Zhu. 2020. Pathological findings of COVID-19 associated with acute respiratory distress syndrome. *The Lancet respiratory medicine* 8: 420-422.
3. Sami, R., F. Fathi, N. Eskandari, M. Ahmadi, R. ArefNezhad, and H. Motedayyen. 2021. Characterizing the immune responses of those who survived or succumbed to COVID-19: Can immunological signatures predict outcome? *Cytokine* 140: 155439.
4. Lee, J., S.-S. Park, T. Y. Kim, D.-G. Lee, and D.-W. Kim. 2021. Lymphopenia as a Biological Predictor of Outcomes in COVID-19 Patients: A Nationwide Cohort Study. *Cancers (Basel)* 13: 471.
5. Liang, C., E. Bencurova, E. Psota, P. Neurgaonkar, M. Prelog, C. Scheller, and T. Dandekar. 2021. Population-Predicted MHC Class II Epitope Presentation of SARS-CoV-2 Structural Proteins Correlates to the Case Fatality Rates of COVID-19 in Different Countries. *Int J Mol Sci* 22: 2630.
6. Wilson, E. A., G. Hirneise, A. Singharoy, and K. S. Anderson. 2021. Total predicted MHC-I epitope load is inversely associated with population mortality from SARS-CoV-2. *Cell Reports Medicine* 2: 100221.
7. Jafarzadeh, A., S. Jafarzadeh, P. Nozari, P. Mokhtari, and M. Nemati. 2021. Lymphopenia an important immunological abnormality in patients with COVID-19: Possible mechanisms. *Scandinavian Journal of Immunology* 93: e12967.
8. Yan, S., and G. Wu. 2021. Is lymphopenia different between SARS and COVID-19 patients? *The FASEB Journal* 35: e21245.
9. Saini, S. K., D. S. Hersby, T. Tamhane, H. R. Povlsen, S. P. A. Hernandez, M. Nielsen, A. O. Gang, and S. R. Hadrup. 2021. SARS-CoV-2 genome-wide T cell epitope mapping reveals immunodominance and substantial CD8+ T cell activation in COVID-19 patients. *Science Immunology* 6.
10. Tavakolpour, S., T. Rakhshandehroo, E. X. Wei, and M. Rashidian. 2020. Lymphopenia during the COVID-19 infection: What it shows and what can be learned. *Immunology letters* 225: 31-32.
11. Wenzhong, L., and L. Hualan. 2021. COVID-19: captures iron and generates reactive oxygen species to damage the human immune system. *Autoimmunity* 54: 213-224.
12. McHeyzer-Williams, M., S. Okitsu, N. Wang, and L. McHeyzer-Williams. 2012. Molecular programming of B cell memory. *Nature Reviews Immunology* 12: 24-34.
13. Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. 2002. T cells and MHC proteins. In *Molecular Biology of the Cell. 4th edition*. Garland Science.
14. Anfossi, N., P. André, S. Guia, C. S. Falk, S. Roetynck, C. A. Stewart, V. Breso, C. Frassati, D. Reviron, D. Middleton, F. Romagné, S. Ugolini, and E. Vivier. 2006. Human NK Cell Education by Inhibitory Receptors for MHC Class I. *Immunity* 25: 331-342.

15. Ljunggren, H.-G., and K. Kärre. 1990. In search of the 'missing self': MHC molecules and NK cell recognition. *Immunology today* 11: 237-244.
16. Bruchez, A., K. Sha, J. Johnson, L. Chen, C. Stefani, H. McConnell, L. Gaucherand, R. Prins, K. A. Matreyek, and A. J. Hume. 2020. MHC class II transactivator CIITA induces cell resistance to Ebola virus and SARS-like coronaviruses. *Science* 370: 241-247.
17. Kumar, S. A., D. Dasgupta, and H. Wei. 2020. An in-silico based clinical insight on the effect of noticeable CD4 conserved residues of SARS-CoV-2 on the CD4-MHC-II interactions. *bioRxiv*.
18. Agerer, B., M. Koblischke, V. Gudipati, M. Smyth, A. Popa, J.-W. Genger, L. Endler, D. M. Florian, V. Muehlgrabner, and A. Lercher. 2020. SARS-CoV-2 escapes CD8 T cell surveillance via mutations in MHC-I restricted epitopes. *BioRxiv*.
19. Rha, M.-S., H. W. Jeong, J.-H. Ko, S. J. Choi, I.-H. Seo, J. S. Lee, M. Sa, A. R. Kim, E.-J. Joo, J. Y. Ahn, J. H. Kim, K.-H. Song, E. S. Kim, D. H. Oh, M. Y. Ahn, H. K. Choi, J. H. Jeon, J.-P. Choi, H. B. Kim, Y. K. Kim, S.-H. Park, W. S. Choi, J. Y. Choi, K. R. Peck, and E.-C. Shin. 2021. PD-1-Expressing SARS-CoV-2-Specific CD8+ T Cells Are Not Exhausted, but Functional in Patients with COVID-19. *Immunity* 54: 44-52.e43.
20. van de Weijer, M. L., R. D. Luteijn, and E. J. H. J. Wiertz. 2015. Viral immune evasion: Lessons in MHC class I antigen presentation. *Seminars in Immunology* 27: 125-137.
21. Arase, H., E. S. Mocarski, A. E. Campbell, A. B. Hill, and L. L. Lanier. 2002. Direct recognition of cytomegalovirus by activating and inhibitory NK cell receptors. *Science* 296: 1323-1326.
22. Revilleza, M. J., R. Wang, J. Mans, M. Hong, K. Natarajan, and D. H. Margulies. 2011. How the virus outsmarts the host: function and structure of cytomegalovirus MHC-I-like molecules in the evasion of natural killer cell surveillance. *Journal of Biomedicine and Biotechnology* 2011.
23. Rawlinson, W. D., H. E. Farrell, and B. G. Barrell. 1996. Analysis of the complete DNA sequence of murine cytomegalovirus. *Journal of virology* 70: 8833-8849.
24. Brocchieri, L., T. N. Kledal, S. Karlin, and E. S. Mocarski. 2005. Predicting coding potential from genome sequence: application to betaherpesviruses infecting rats and mice. *Journal of virology* 79: 7570-7596.
25. Natarajan, K., A. Hicks, J. Mans, H. Robinson, R. Guan, R. A. Mariuzza, and D. H. Margulies. 2006. Crystal Structure of the Murine Cytomegalovirus MHC-I Homolog m144. *Journal of Molecular Biology* 358: 157-171.
26. Pyzik, M., A. A. Dumaine, B. Charbonneau, N. Fodil-Cornu, S. Jonjic, and S. M. Vidal. 2014. Viral MHC Class I–like Molecule Allows Evasion of NK Cell Effector Responses In Vivo. *The Journal of Immunology* 193: 6061-6069.
27. Wyrwicz, L. S., and L. Rychlewski. 2008. Cytomegalovirus immediate early gene UL37 encodes a novel MHC-like protein. *Acta Biochimica Polonica* 55: 67-74.
28. Tripathy, S. K., H. R. C. Smith, E. A. Holroyd, J. T. Pingel, and W. M. Yokoyama. 2006. Expression of m157, a Murine Cytomegalovirus-Encoded Putative Major Histocompatibility Class I (MHC-I)-Like Protein, Is Independent of Viral Regulation of Host MHC-I. *Journal of Virology* 80: 545-550.
29. Wilkinson, G. W., P. Tomasec, R. J. Stanton, M. Armstrong, V. Prod'homme, R. Aicheler, B. P. McSharry, C. R. Rickards, D. Cochrane, and S. Llewellyn-Lacey. 2008. Modulation of natural killer cells by human cytomegalovirus. *Journal of clinical virology* 41: 206-212.
30. Halenius, A., C. Gerke, and H. Hengel. 2015. Classical and non-classical MHC I molecule manipulation by human cytomegalovirus: so many targets—but how many arrows in the quiver? *Cellular & Molecular Immunology* 12: 139-153.

31. Beck, S., and B. G. Barrell. 1988. Human cytomegalovirus encodes a glycoprotein homologous to MHC class-I antigens. *Nature* 331: 269-272.
32. Fahnestock, M. L., J. L. Johnson, R. M. Renny Feldman, J. M. Neveu, W. S. Lane, and P. J. Bjorkman. 1995. The MHC class I homolog encoded by human cytomegalovirus binds endogenous peptides. *Immunity* 3: 583-590.
33. Ambagala, A. P. N., J. C. Solheim, and S. Srikumaran. 2005. Viral interference with MHC class I antigen presentation pathway: The battle continues. *Veterinary Immunology and Immunopathology* 107: 1-15.
34. Wiertz, E., A. Hill, D. Tortorella, and H. Ploegh. 1997. Cytomegaloviruses use multiple mechanisms to elude the host immune response. *Immunology letters* 57: 213-216.
35. Reddehase, M. J., W. Mutter, K. Münch, H.-J. Bühring, and U. H. Koszinowski. 1987. CD8-positive T lymphocytes specific for murine cytomegalovirus immediate-early antigens mediate protective immunity. *Journal of virology* 61: 3102-3108.
36. He, X., H. Yoshida, and Y. Minamishima. 1995. Analysis of the role of CD4+ T-cells during murine cytomegalovirus infection in different strains of mice. *Virus research* 36: 233-245.
37. Revilleza, M. J., R. Wang, J. Mans, M. Hong, K. Natarajan, and D. H. Margulies. 2011. How the Virus Outsmarts the Host: Function and Structure of Cytomegalovirus MHC-I-Like Molecules in the Evasion of Natural Killer Cell Surveillance. *Journal of Biomedicine and Biotechnology* 2011: 724607.
38. Carrillo-Bustamante, P., C. Keşmir, and R. J. De Boer. 2013. Virus encoded MHC-like decoys diversify the inhibitory KIR repertoire. *PLoS computational biology* 9: e1003264.
39. Lok, S.-M. 2021. An NTD supersite of attack. *Cell Host & Microbe* 29: 744-746.
40. McCallum, M., A. De Marco, F. A. Lempp, M. A. Tortorici, D. Pinto, A. C. Walls, M. Beltramello, A. Chen, Z. Liu, and F. Zatta. 2021. N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell* 184: 2332-2347. e2316.
41. Suryadevara, N., S. Shrihari, P. Gilchuk, L. A. VanBlargan, E. Binshtein, S. J. Zost, R. S. Nargi, R. E. Sutton, E. S. Winkler, E. C. Chen, M. E. Fouch, E. Davidson, B. J. Doranz, R. E. Chen, P.-Y. Shi, R. H. Carnahan, L. B. Thackray, M. S. Diamond, and J. E. Crowe. 2021. Neutralizing and protective human monoclonal antibodies recognizing the N-terminal domain of the SARS-CoV-2 spike protein. *Cell* 184: 2316-2331.e2315.
42. Cerutti, G., Y. Guo, T. Zhou, J. Gorman, M. Lee, M. Rapp, E. R. Reddem, J. Yu, F. Bahna, J. Bimela, Y. Huang, P. S. Katsamba, L. Liu, M. S. Nair, R. Rawi, A. S. Olia, P. Wang, B. Zhang, G.-Y. Chuang, D. D. Ho, Z. Sheng, P. D. Kwong, and L. Shapiro. 2021. Potent SARS-CoV-2 neutralizing antibodies directed against spike N-terminal domain target a single supersite. *Cell Host & Microbe* 29: 819-833.e817.
43. McCallum, M., A. De Marco, F. A. Lempp, M. A. Tortorici, D. Pinto, A. C. Walls, M. Beltramello, A. Chen, Z. Liu, F. Zatta, S. Zepeda, J. di Iulio, J. E. Bowen, M. Montiel-Ruiz, J. Zhou, L. E. Rosen, S. Bianchi, B. Guarino, C. S. Fregni, R. Abdelnabi, S.-Y. C. Foo, P. W. Rothlauf, L.-M. Bloyet, F. Benigni, E. Cameroni, J. Neyts, A. Riva, G. Snell, A. Telenti, S. P. J. Whelan, H. W. Virgin, D. Corti, M. S. Pizzuto, and D. Veesler. 2021. N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell* 184: 2332-2347.e2316.
44. Liu, W., and H. Li. 2020. COVID-19: attacks the 1-beta chain of hemoglobin and captures the porphyrin to inhibit human heme metabolism. *ChemRxiv. Preprint.* chemrxiv.11938173.v9.
45. Zhang, Y., J. Zhang, Y. Chen, B. Luo, Y. Yuan, F. Huang, T. Yang, F. Yu, J. Liu, and B. Liu. 2020. The ORF8 protein of SARS-CoV-2 mediates immune evasion through potently downregulating MHC-I. *biorxiv*.
46. Saito, A., T. Irie, R. Suzuki, T. Maemura, H. Nasser, K. Uriu, Y. Kosugi, K. Shirakawa, K. Sadamasu, I. Kimura, J. Ito, J. Wu, K. Iwatsuki-Horimoto, M. Ito, S. Yamayoshi, S. Ozono, E. P. Butlertanaka, Y. L.

Tanaka, R. Shimizu, K. Shimizu, K. Yoshimatsu, R. Kawabata, T. Sakaguchi, K. Tokunaga, I. Yoshida, H. Asakura, M. Nagashima, Y. Kazuma, R. Nomura, Y. Horisawa, K. Yoshimura, A. Takaori-Kondo, M. Imai, T. G. t. P. J. Consortium, S. Nakagawa, T. Ikeda, T. Fukuhara, Y. Kawaoka, and K. Sato. 2021. SARS-CoV-2 spike P681R mutation, a hallmark of the Delta variant, enhances viral fusogenicity and pathogenicity. *bioRxiv*: 2021.2006.2017.448820.

47. Arora, P., A. Kempf, I. Nehlmeier, A. Sidarovich, N. Krüger, L. Graichen, A.-S. Moldenhauer, M. S. Winkler, S. Schulz, H.-M. Jäck, M. V. Stankov, G. M. N. Behrens, S. Pöhlmann, and M. Hoffmann. 2021. Increased lung cell entry of B.1.617.2 and evasion of antibodies induced by infection and BNT162b2 vaccination. *bioRxiv*: 2021.2006.2023.449568.
48. Xia, S., Y. Zhu, M. Liu, Q. Lan, W. Xu, Y. Wu, T. Ying, S. Liu, Z. Shi, and S. Jiang. 2020. Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein. *Cellular & molecular immunology*: 1-3.
49. Naveca, F., V. Nascimento, V. Souza, A. Corado, F. Nascimento, G. Silva, A. Costa, D. Duarte, K. Pessoa, and L. Gonçalves. 2021. Phylogenetic relationship of SARS-CoV-2 sequences from Amazonas with emerging Brazilian variants harboring mutations E484K and N501Y in the Spike protein. *Virological. org*.
50. Faria, N. R., T. A. Mellan, C. Whittaker, I. M. Claro, D. d. S. Candido, S. Mishra, M. A. Crispim, F. C. Sales, I. Hawryluk, and J. T. McCrone. 2021. Genomics and epidemiology of the P. 1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* 372: 815-821.
51. Rambaut, A., N. Loman, O. Pybus, W. Barclay, J. Barrett, A. Carabelli, T. Connor, T. Peacock, D. L. Robertson, and E. Volz. 2020. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. *Genom. Epidemiol*: 1-5.
52. Tegally, H., E. Wilkinson, and M. Giovanetti. 2021. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* 592: 438-443.
53. Romero, P. E., A. Dávila-Barclay, G. Salvatierra, L. González, D. Cuicapuza, L. Solis, P. Marcos-Carbajal, J. Huancachoque, L. Maturrano, and P. Tsukayama. 2021. The Emergence of SARS-CoV-2 Variant Lambda (C.37) in South America. *medRxiv*: 2021.2006.2026.21259487.
54. Sanchez-Mazas, A. 2020. A review of HLA allele and SNP associations with highly prevalent infectious diseases in human populations. *Swiss medical weekly* 150.
55. Nguyen, A., J. K. David, S. K. Maden, M. A. Wood, B. R. Weeder, A. Nellore, and R. F. Thompson. 2020. Human leukocyte antigen susceptibility map for severe acute respiratory syndrome coronavirus 2. *Journal of virology* 94: e00510-00520.
56. Ben Shachar, S., N. Barda, S. Manor, S. Israeli, N. Dagan, S. Carmi, R. Balicer, B. Zisser, and Y. Louzoun. 2021. MHC Haplotyping of SARS-CoV-2 Patients: HLA Subtypes Are Not Associated with the Presence and Severity of COVID-19 in the Israeli Population. *Journal of Clinical Immunology*.
57. Chen, R. E., E. S. Winkler, J. B. Case, I. D. Aziati, T. L. Bricker, A. Joshi, T. L. Darling, B. Ying, J. M. Errico, S. Shrihari, L. A. VanBlargan, X. Xie, P. Gilchuk, S. J. Zost, L. Droit, Z. Liu, S. Stumpf, D. Wang, S. A. Handley, W. B. Stine, P.-Y. Shi, M. E. Davis-Gardner, M. S. Suthar, M. G. Knight, R. Andino, C. Y. Chiu, A. H. Ellebedy, D. H. Fremont, S. P. J. Whelan, J. E. Crowe, L. Purcell, D. Corti, A. C. M. Boon, and M. S. Diamond. 2021. In vivo monoclonal antibody efficacy against SARS-CoV-2 variant strains. *Nature*.