Title
HLA variation and antigen presentation in COVID-19 and SARS-CoV-2 infection

Authors
Augusto, Danillo G
Hollenbach, Jill A

Publication Date
2022-06-01

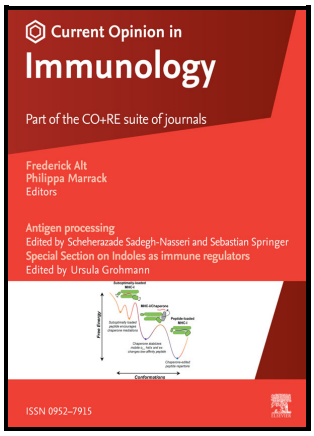
DOI
10.1016/j.coi.2022.102178

Peer reviewed

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

HLA variation and antigen presentation in COVID-19 and SARS-CoV-2 infection

Danillo G. Augusto, Jill A. Hollenbach



PII: S0952-7915(22)00025-5

DOI: https://doi.org/10.1016/j.coi.2022.102178

Reference: COIMMU102178

To appear in: *Current Opinion in Immunology*

Received date: 6 January 2022
Revised date: 26 February 2022
Accepted date: 16 March 2022

Please cite this article as: Danillo G. Augusto and Jill A. Hollenbach, HLA variation and antigen presentation in COVID-19 and SARS-CoV-2 infection, *Current Opinion in Immunology,* (2022) doi:https://doi.org/10.1016/j.coi.2022.102178

HLA variation and antigen presentation in COVID-19 and SARS-CoV-2 infection

Danillo G. Augusto[1,2] and Jill A. Hollenbach[1,3]

1 Department of Neurology, University of California, San Francisco, CA, USA

2 Programa de Pós-Graduação em Genética, Universidade Federal do Paraná, Curitiba, Brazil

3 Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, US

Corresponding Author:

Jill A. Hollenbach, PhD, MPH

University of California, San Francisco

Department of Neurology

675 Nelson Rising Lane, NS-221A

San Francisco, CA 94158

jill.hollenbach@ucsf.edu

Highlights

- Variation in both HLA and SARS-CoV-2 impact antigen presentation to T cells
- HLA variation influences COVID-19 outcomes and T-cell responses to SARS-CoV-2
- Mutations in viral variants alter predicted peptide binding affinities to HLA

Abstract

The extraordinary variation of the human leukocyte antigen (HLA) molecules is critical for diversifying antigen presentation to T cells. Coupled with the rise of novel strains and rapidly evolving immune evasion SARS-CoV-2 proteins, HLA-mediated immunity in COVID-19 is critically important but far from being fully understood. A growing number of studies have found association of HLA variants with different COVID-19 outcomes and that HLA genotypes associate with differential immune responses against SARS-CoV-2. Prediction studies have shown that mutations in multiple viral strains, most concentrated in the Spike protein, affect the affinity between these mutant peptides and HLA molecules. Understanding the impact of this variation on T-cell responses is critical for comprehending immunogenic mechanisms in both natural immunity and vaccine development.

Introduction

Since nearly the start of the global COVID-19 pandemic, there has been an intensive effort to deduce the genetic and immune features that may underlie variation in susceptibility to infection with SARS-CoV-2, as well as disease outcomes in COVID-19. Most COVID-19 patients are either asymptomatic or experience relatively mild symptoms, including fever and cough, but some individuals develop severe pneumonia; a subset of those individuals will progress to develop acute respiratory distress syndrome (ARDS). Acutely ill patients may further develop shock and multiple organ failure[1]. Many specific demographic, medical, and behavioral risk factors have been identified as contributory to more severe disease and poor outcomes. Advanced age, comorbidities such as diabetes and hypertension, smoking history, and African American ancestry have all been associated with increased morbidity and mortality in COVID-19, for example[2]

Infection with SARS-CoV-2 activates both innate and adaptive immune responses. The adaptive immune response is a critical component of protection from viral pathogens. CD4 T cells promote virus-specific antibodies through their action on T-dependent B cells, while CD8 T cells' cytotoxic capacity may kill virally infected cells. T cell responses to SARS-CoV-2 are crucial factors for recognizing and killing infected cells[3]. At the same time, T cells from patients with severe disease have been shown to have phenotypic characteristics associated with differential cytokine secretion, and that these differences are antigen dependent[4]. Because of their pivotal

role in antigen presentation to T cells, the genes encoding the human leukocyte antigen (HLA) molecules have been a primary focus in genetic association studies across a multitude of infectious and immune-mediated disease. We and others have been actively engaged in studies to pinpoint specific *HLA* alleles associated with risk for SARS-CoV-2 infection and/or specific COVID-19 disease outcomes, and more recently, vaccine response. Proceeding in tandem, numerous efforts have sought to identify viral antigens presented either by specific HLA or broadly across allotypes, and in particular, those acting as T cell epitopes. By synthesizing information pertinent to antigen presentation and T cell reactivity with *HLA* genetic associations, we are beginning to develop a more comprehensive picture of the immune mechanisms underpinning differential host response to SARS-CoV-2 exposure and infection.

The extraordinary variation of HLA molecules is determinative in antigen presentation

The *HLA* genes are located within the human major histocompatibility complex (MHC), located on the short arm of chromosome 6 (p21.3). HLA molecules are critical components of the adaptive immune system, which mediates the specific destruction of infected cells and the production of antibodies. Classical HLA class I molecules (HLA-A, HLA-B, and HLA-C) are expressed on all nucleated cells and contain two non-covalently bound polypeptide chains. The polymorphic alpha chain is encoded by the *HLA* gene within the MHC region, while the gene for the non-polymorphic beta-2 microglobulin chain is located on chromosome 15. HLA class I molecules present endogenous peptides, including those derived from intracellular pathogens such as viruses. Foreign peptides presented on class I antigens are recognized by cytotoxic CD8 T lymphocytes[5]. Classical HLA class II molecules (HLA-DR, HLA-DQ, and HLA-DP) are heterodimers composed of an alpha and beta chain encoded by genes within the MHC and present peptides generated in endosomes from protein sources both inside and outside the presenting cell to CD4 T lymphocytes[6]. In contrast to the constitutive expression of class I molecules, the expression of class II molecules is limited to cells of specialized function in immunity, collectively known as professional antigen-presenting cells, including primarily dendritic cells, macrophages, and mature B lymphocytes. In addition, HLA class II are highly expressed on the surface of epithelial cells in both the lung and intestine[7], which is potentially relevant in the context of SARS-CoV-2 infection.

The *HLA* genes are the most polymorphic region of the human genome[8]. More than 30,000 *HLA* alleles have been identified to date, which encode more than 18,000 unique proteins (allotypes) [9]. The majority of nucleotide substitutions in *HLA* alleles are concentrated in the

exons encoding the peptide-binding groove and the region of interaction with T-cell receptors, and, importantly, the most polymorphic positions are those that affect peptide binding[10]. This remarkable variation in the peptide binding groove affects its geometry, charge distribution, and hydrophobicity, determining interaction with individual peptides. Diverse HLA molecules may exhibit distinctive peptide-binding repertoires, while individuals with different *HLA* genotypes may exhibit a differential ability to present specific peptides and elicit immune responses. These variable elements in antigen presentation underlie the many known HLA associations with human disease[11]

Association of HLA polymorphism with COVID-19 disease course

HLA class I and class II alleles have been previously associated with the severe acute respiratory syndrome caused by SARS-CoV[12]. The most robust genetic association studies to-date for SARS-CoV-2 infection have primarily focused on disease outcomes, given the inherent difficulty in assessing infection risk and controlling for exposure. In many cases, these studies have specifically examined severe outcomes in disease (e.g., need for mechanical ventilation or death). Thus far, numerous *HLA* class I and II alleles have been associated with disease outcomes, but without clear consensus. Indeed, some large studies, either in the context of genome wide-association studies (GWAS)[13] or large HLA databases[14], have failed to show a significant influence of *HLA* genotype on disease.

Nevertheless, some interesting results have emerged and are summarized in **Table 1**. For example, *HLA-C*04:01* was found to be associated with severe clinical course of COVID-19 in European patients, with carriers of this allele having twice the risk of requiring mechanical ventilation[15]. In contrast, a different class I allele, *HLA-A*11:01,* was associated with severe disease in one Japanese cohort[16], while a class II allele, *HLA-DRB1*09:01,* was identified in another[17]. Of note, *HLA-A*11:01* was also associated with severe outcome in a Chinese cohort[18], further supporting that observation. In less severe disease, asymptomatic infection is particularly interesting, as it suggests the capacity for early viral clearance. Langton et al.[19] found a significant association of *HLA-DRB1*04:01* in asymptomatic patients with European ancestry relative to those with severe disease. Interestingly, this allele was recently associated with milder disease in an Iranian patient population[20]. Likewise, in our own work, we identified *HLA-B*15:01*, a class I allele in strong linkage disequilibrium with *HLA-DRB1*04:01*, as strongly associated with asymptomatic infection in patients with European ancestry[21]. Discrepancies across studies may be attributed to differences in the definition of disease phenotypes, study

population, and often limited sample sizes. Thus, while the results are mixed, some consistent patterns are beginning to emerge with respect to HLA associations in SARS-CoV-2 infection, and may serve as a basis for interpreting studies related to antigen presentation.

Immunoinformatics prediction of SARS-CoV-2 peptide binding affinity to HLA

A series of early in silico analyses pointed to HLA as relevant molecules for SARS-CoV-2 risk and important targets for vaccine development[25–28]. Interestingly, it was shown that *HLA-B*46:01* has a low predicted binding of peptides for SARS-CoV-2. This observation suggests that individuals expressing this molecule may be more vulnerable to COVID-19[27], which corroborated previous results showing *HLA-B*46:01* association with SARS risk[29]. In contrast, *HLA-15:03* was predicted to protect against COVID-19 by having the greatest ability to present highly conserved SARS-CoV-2 peptides to T cells[27].

Most protein-level mutations in all recently discovered SARS-CoV-2 strains concentrate in the Spike protein (**Figure 1**), the main target for COVID-19 vaccines[30,31] due to its high antigenicity and capacity to induce robust immune responses[32,33]. Not surprisingly, variation in SARS-CoV-2 strains can also affect HLA binding and antigen presentation, whereby a given HLA allotype that may efficiently present a wild-type peptide has differential capacity to present a mutant strain (**Figure 2**). NetMHCpan-4.1 and NetMHCIIpan-4.0 are tools that use tailored machine learning strategies to integrate predictors trained on binding affinity data and mass spectrometry experiments[34]. Leveraging these tools, Nersisyan et al. have created a tool that comprehensively tracks how SARS-CoV-2 mutations are predicted to affect HLA binding[35]. By curating their data, we can observe distinct patterns of HLA affinity to different strains (**Figure 3A**). Delta, highly contagious[36], and Omicron, heavily mutated and associated with increased risk of reinfection[37,38], are the two most prevalent SARS-CoV-2 variant strains to date. Interestingly, Omicron is the variant that encodes the largest number of epitopes predicted to strongly bind both HLA class I and class II (**Figure 3B**), with even more pronounced differences for the Spike protein (**Figure 3C**). Although predictions do not necessarily correlate with T-cell responses, these data allow us to speculate that Omicron mutations, particularly in the spike protein, may not have a detrimental overall effect on HLA mediated T-cell immunity. HLA class II accounts for more than 90% of the stronger binding predictions for all SARS-CoV-2 variants. Among HLA class I, HLA-A accounts for most of the stronger binding predictions, while HLA-B has the most significant number of weak binding predictions (**Figure 3D**). In HLA class II, HLA-DQ is predicted to bind strongly to a small number of peptides and accounts for most of the

weak binding predictions. Omicron's 29 protein-level mutations in the Spike protein are collectively predicted to change the affinity of 143 peptide-HLA class I and 85 peptide-HLA class II pairs **(Supplementary Table 1).** On the other hand, the mutations observed in Delta have overall less impact in affecting allele-specific HLA peptide binding affinity than those observed in Omicron. **Supplementary Table 2** gives the prediction of the most significant interactions between specific HLA molecules and Spike proteins lost due to the mutations observed in Omicron and Delta. Numerous bioinformatic predictions for SARS-CoV-2 T cell epitopes have also been undertaken to narrow the search space for relevant epitopes [40–42]. Many of these efforts were designed and validated to predict dominant, promiscuous epitopes, independent of ancestry and HLA polymorphism[43], identifying the most parsimonious set of 25-mers to cover the largest population[40]. These immunoinformatic approaches generally compare multiple SARS-CoV-2 sequences and use different algorithms to predict HLA affinity for epitopes from the 10 unique SARS-CoV-2 proteins.

HLA variation in antigen presentation and SARS-CoV-2 T-cell immunity

Although in silico prediction is extremely useful for identifying antigenic peptides, several steps in the antigen presentation pathway may represent a limitation for this strategy. Examples are the protein degradation in the endosomal pathway, degradation of peptides by aminopeptidases in the cytosol, and translocation into the endoplasmic reticulum [44], which are not always integrated into the predictive algorithms. In addition, there are other infection-related variables, such as the changes that the virus causes in the expression of host proteins relevant for antigen presentation machinery, including HLA expression itself.

However, despite these limitations, studies have demonstrated that multiple SARS-CoV-2 peptides predicted to bind HLA can elicit T cell responses. Early investigations demonstrated that while there is some overlap, many SARS-CoV-2 epitopes for CD8 T cells are *HLA* specific[45]. Saini et al. performed a genome-wide T cell epitope mapping and identified 122 immunogenic and a subset of immunodominant SARS-CoV-2 T cell epitopes[46]. Another study applied peptide-loaded HLA tetramers to perform an ex-vivo analysis of pre-existing induced SARS-CoV-2-specific CD8+ T cells and identified a set of immunodominant peptides[47]. Analyzing 31 patients with COVID-19, Gangaev et al.[48] analyzed peptide-HLA class I complexes restricted to 10 common HLA molecules and identified 18 recognized by CD8+ T cells. They further analyzed CD8 T responses and observed gene expression patterns of constrained T cell re-activation, in addition to high expression of the gene *NKG2A* and lack of

cytokine production. More recently, mass spectrometry-based HLA-I immunopeptidomics revealed that SARS-CoV-2 peptides presented by HLA class I also derive from internal out-of-frame open reading frames in spike and nucleocapsid not captured by current vaccines. In addition, this study has shown that early expressed SARS-CoV-2 proteins have a larger contribution to HLA-mediated immunogenicity[49]. Finally, it has been shown that immune response to SARS-CoV-2 is distinguished by *HLA* genotypes, in which the dominant immune response in HLA-B*07 was associated with a more diverse TCR repertoire compared to the response in HLA-A*02, HLA-A*24, and HLA-A*01[50].

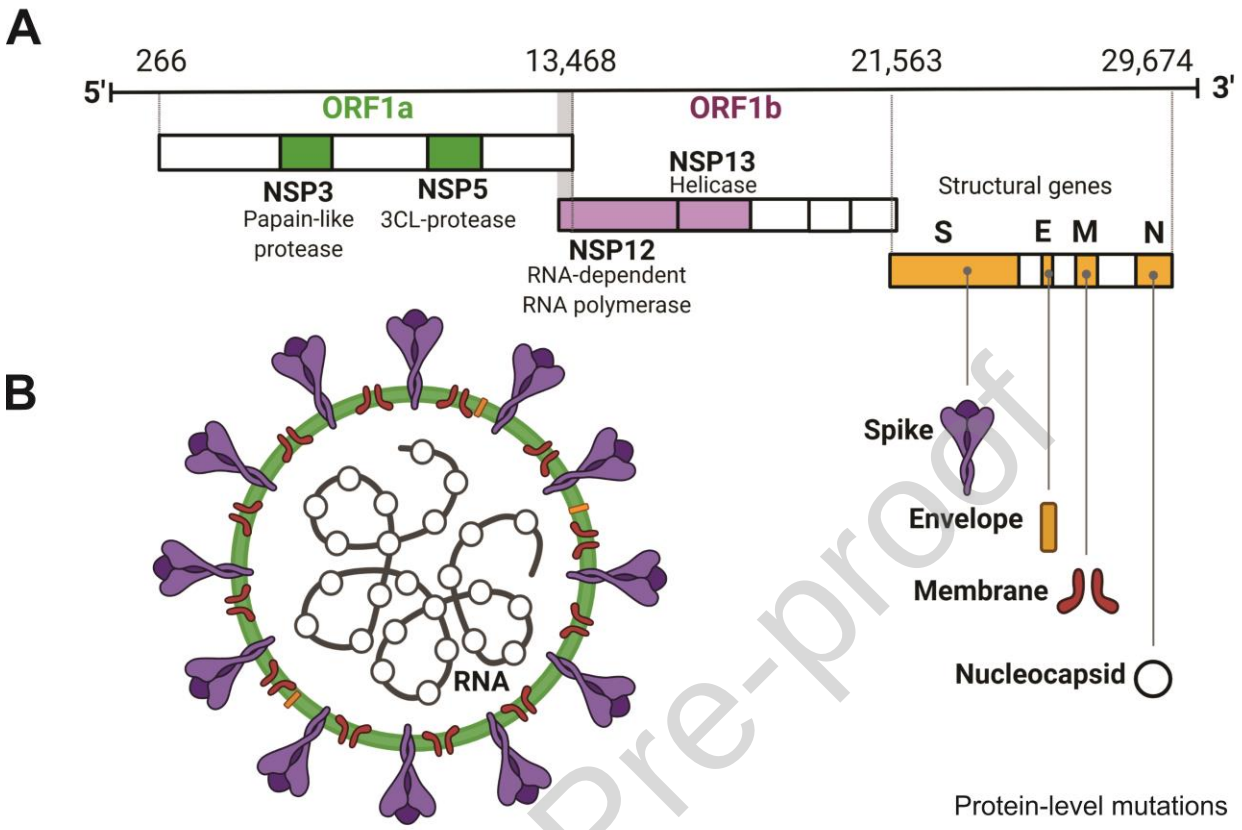
Conclusion

Two years into the global pandemic, much has been learned about the relationship between HLA polymorphism, variation in SARS-CoV-2, and the impact that this variability has on antigen presentation and immunity. However, the rapid appearance of immune evasive strains has rendered a more complete understanding of these relationships something of a moving target. Likewise, examination of more and diverse populations is needed to fully comprehend the role of HLA in disease outcomes and vaccine response and efficacy. In the future, additional large-scale studies will be needed to more fully detail the finely tuned relationship of the HLA system and continually novel SARS-CoV-2 strains to improve our understanding of the impact of antigen presentation in disease.

Acknowledgements

This work was supported by the National Institutes of Health, USA (NIH grant R01AI159260).

Figure legends



C

Protein-level mutations

	Spike	E	M	N	NS3	NS6	NS7a	NS8	NS9b	NS9c	NSP2	NSP3	NSP4	NSP5	NSP6	NSP7	NSP9	NSP12	NSP13	NSP14	NSP15	NSP16	Total
Omicron GR/484A	29	1	3	3	-	-	-	2	1	-	4	1	1	2	-	-	1	-	-	-	-	**48**
Delta G/478K.V1	5	-	1	3	1	-	2	-	1	1	-	2	-	-	-	-	-	1	-	1	-	-	**18**
Alpha 202012/01 GRY	9	-	-	3	-	-	-	1	-	1	-	3	-	-	1	-	1	1	-	-	-	-	**20**
Beta GH/501Y.V2	9	1	-	1	2	-	-	-	-	1	1	1	-	-	1	-	-	1	-	-	-	-	**18**
Gamma GR/501Y.V3	12	-	-	2	1	-	-	1	1	1	-	2	-	-	1	-	-	1	1	-	-	-	**23**
Lambda GR/452Q.V1	2	-	-	3	-	-	-	-	-	2	-	4	1	-	1	-	-	1	-	-	-	1	**15**
Mu GH	9	-	-	2	1	-	-	3	1	1	-	3	1	-	1	-	-	1	2	1	-	-	**26**
Epsilon GH/452R.V1	4	-	-	1	2	-	-	-	-	1	1	-	-	-	-	-	-	1	1	-	-	-	**11**
Zeta GR/484K.V2	3	-	-	4	-	-	-	-	-	1	-	-	-	1	-	1	-	1	-	-	-	-	**11**
Eta G/484K.V3	8	1	1	4	-	1	-	-	1	1	-	1	-	-	1	-	-	1	-	-	-	-	**20**
Theta GR/1092K.V1	10	-	-	1	-	-	-	-	-	1	-	2	2	-	-	1	-	1	2	-	-	-	**20**
Iota GH/253G.V1	6	-	-	2	2	-	-	2	-	1	1	-	3	-	2	-	-	1	-	-	-	-	**20**
Kappa G/452R.V3	8	-	1	2	1	-	1	-	-	2	-	1	-	-	1	-	-	1	2	-	3	-	**23**

Figure 1. Genome organization, virion structure, and overview of protein-level mutations in SARS-CoV-2 strains.

A) Genomic organization of SARS-CoV-2. ORF1a and ORF1b encode 16 non-structural proteins (NSP1– NSP16). NSP3 and NSP5 (ORF1a) encode papain-like protease and 3CL-

protease, respectively. NSP12 encodes RNA-dependent RNA polymerase (RdRp) and NSP13 encodes RNA helicase (ORF1b). The structural genes encode the structural proteins: (S) Spike, (E) Envelope, (M) Membrane, and (N) Nucleocapsid. **B)** SARS-CoV-2 virion structure. **C)** Overview of protein-level mutations of SARS-CoV-2 strains[35]. Figure created with Biorender.com.

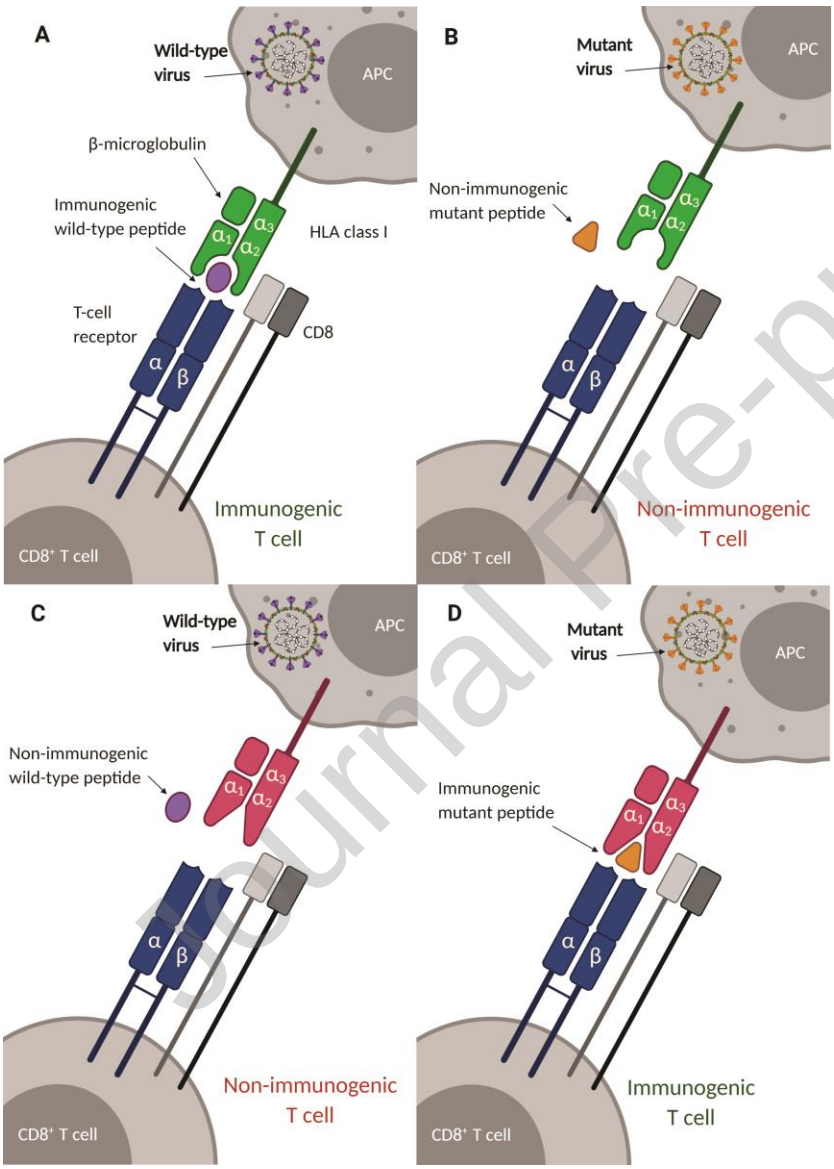


Figure 2. Variation in both HLA and viral peptides determine antigen presentation and immunogenicity in SARS-CoV-2 infection. Distinct viral strains may result in peptides with differential affinity to specific HLA molecules. Variation in the HLA peptide groove is equally important for determining binding affinity. Non-immunogenic peptides do not elicit immunogenic T-cell responses. In this example, we represent different scenarios suggesting how variation in

both the HLA class I molecule and viral peptide may affect binding. **A)** Wild-type peptide binds to HLA and is presented to T-cell receptor. **B)** Mutant peptide does not bind to HLA and is not presented to T-cell receptor. **C)** Wild-type peptide does not bind to variant HLA and is not presented to T-cell receptor. **D)** Mutant variant binds to variant HLA and is presented to T-cell receptor. Figure created with Biorender.com.

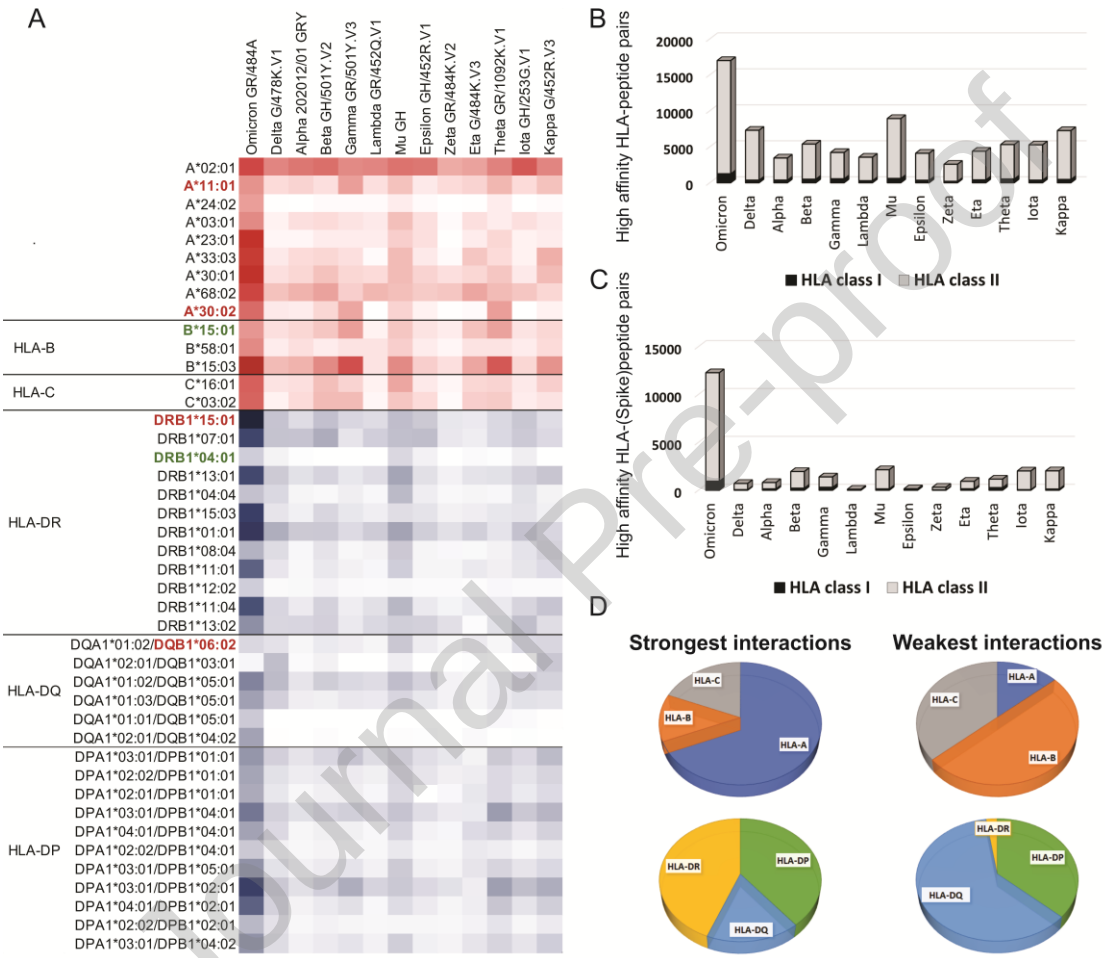


Figure 3. Peptides from different SARS-CoV-2 variants have distinctive affinities for HLA molecules. Data extracted from T-cell COVID-19 Atlas [35] determined using netMHCpan-4.1 and netMHCIIpan-4.0 [34]. The affinity scores were not directly compared across genes; all plots show absolute numbers of strong or weak interactions. **A)** Most relevant HLA-peptide interactions across all SARS-CoV-2 variants. The color intensity in each box represents the absolute number of strong interactions (IC_{50} affinity ≤ 50 nM) predicted between specific HLA allotypes and peptides from each SARS-CoV-2 variant, varying from white (zero strong interaction) to dark red (67 strong interactions; HLA class I) and dark blue (656 strong interactions; HLA class II). We included only the allotypes with the strongest interactions based

on the affinity scores and excluded those HLA variants observed in low frequencies ($f < 0.05$) in three reference populations from 1000 Genomes Dataset (CEU, YRI, and CHB)[39]. Allotypes from each locus are ordered from the most frequent to the least frequent, according to the maximum frequency observed in these three reference populations. HLA variants that have been previously associated with COVID-19 are shown in bold, with those associated with risk or severe disease shown in red and those associated with asymptomatic or mild infection, in green. Omicron is the variant predicted to exhibit the highest number of peptides strongly interacting with HLA class I and class II molecules, considering the mutations in **(B)** all viral proteins and also **(C)** only the Spike protein. **D)** Distribution of allotypes predicted to have strong and weak interaction for SARS-CoV-2 stratified by locus. On the left, the plot represents the stratification of top 30% of the strongest interactions with SARS-CoV-2 peptides; on the right, the distribution of allotypes in the bottom 30%, representing the HLA molecules with weak or no interaction with SARS-CoV-2 peptides.

References

1. Goyal P, Choi JJ, Pinheiro LC, Schenck EJ, Chen R, Jabri A, Satlin MJ, Campion TR, Nahid M, Ringel JB, et al.: Clinical Characteristics of Covid-19 in New York City. *New Engl J Medicine* 2020, 382:NEJMc2010419.

2. Zheng Z, Peng F, Xu B, Zhao J, Liu H, Peng J, Li Q, Jiang C, Zhou Y, Liu S, et al.: Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis. *J Infect* 2020, 81:e16–e25.

3. Weiskopf D, Schmitz KS, Raadsen MP, Grifoni A, Okba NMA, Endeman H, Akker JPC van den, Molenkamp R, Koopmans MPG, Gorp ECM van, 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:eabd2071.

4. Altmann DM, Boyton RJ: SARS-CoV-2 T cell immunity: Specificity, function, durability, and role in protection. *Sci Immunol* 2020, 5:eabd6160.

5. Salter RD, Benjamin RJ, Wesley PK, Buxton SE, Garrett TPJ, Clayberger C, Krensky AM, Norment AM, Littman DR, Parham P: A binding site for the T-cell co-receptor CD8 on the α3 domain of HLA-A2. *Nature* 1990, 345:41–46.

6. Brodsky FM, Guagliardi LE: The Cell Biology of Antigen Processing and Presentation. *Annu Rev Immunol* 1991, 9:707–744.

7. Wosen JE, Mukhopadhyay D, Macaubas C, Mellins ED: Epithelial MHC Class II Expression and Its Role in Antigen Presentation in the Gastrointestinal and Respiratory Tracts. *Front Immunol* 2018, 9:2144.

8. Shiina T, Hosomichi K, Inoko H, Kulski JK: The HLA genomic loci map: expression, interaction, diversity and disease. *J Hum Genet* 2009, 54:15–39.

9. Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SGE: The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids Res* 2015, 43:D423–D431.

10. Deutekom HWM van, Keşmir C: Zooming into the binding groove of HLA molecules: which positions and which substitutions change peptide binding most? *Immunogenetics* 2015, 67:425–436.

11. Lenz TL, Spirin V, Jordan DM, Sunyaev SR: Excess of Deleterious Mutations around HLA Genes Reveals Evolutionary Cost of Balancing Selection. *Mol Biol Evol* 2016, 33:2555–2564.

12. Aguiar VRC, Augusto DG, Castelli EC, Hollenbach JA, Meyer D, Nunes K, Petzl-Erler ML: An immunogenetic view of COVID-19. *Genet Mol Biol* 2021, 44:e20210036.

13. Initiative C-19 HG, InitiativeLeadership C-19 HG, Niemi MEK, Karjalainen J, Liao RG, Neale BM, Daly M, Ganna A, group W, leaders W group, et al.: Mapping the human genetic architecture of COVID-19. *Nature* 2021, 600:472–477.

14. Shachar SB, Barda N, Manor S, Israeli S, Dagan N, Carmi S, Balicer R, Zisser B, Louzoun Y: MHC Haplotyping of SARS-CoV-2 Patients: HLA Subtypes Are Not Associated with the Presence and Severity of COVID-19 in the Israeli Population. *J Clin Immunol* 2021, 41:1–8.

15. Weiner J, Suwalski P, Holtgrewe M, Rakitko A, Thibeault C, Müller M, Patriki D, Quedenau C, Krüger U, Ilinsky V, et al.: Increased risk of severe clinical course of COVID-19 in carriers of HLA-C*04:01. *Eclinicalmedicine* 2021, 40:101099–101099.

•This study identified a strong association of HLA-C*04:01 with severe outcome of COVID-19, showing that individuals carrying this allele had twice the risk of intubation after infection with SARS-CoV-2.

16. Khor S-S, Omae Y, Nishida N, Sugiyama M, Kinoshita N, Suzuki T, Suzuki M, Suzuki S, Izumi S, Hojo M, et al.: HLA-A*11:01:01:01, HLA-C*12:02:02:01-HLA-B*52:01:02:02, Age and Sex Are Associated With Severity of Japanese COVID-19 With Respiratory Failure. *Front Immunol* 2021, 12:658570.

17. Anzurez A, Naka I, Miki S, Nakayama-Hosoya K, Isshiki M, Watanabe Y, Nakamura-Hoshi M, Seki S, Matsumura T, Takano T, et al.: Association of HLA-DRB1*09:01 with severe COVID-19. *Hla* 2021, 98:37–42.

18. Wang F, Huang S, Gao R, Zhou Y, Lai C, Li Z, Xian W, Qian X, Li Z, Huang Y, et al.: Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility. *Cell Discov* 2020, 6:83.

19. Langton DJ, Bourke SC, Lie BA, Reiff G, Natu S, Darlay R, Burn J, Echevarria C: The influence of HLA genotype on the severity of COVID-19 infection. *Hla* 2021, 98:14–22.

20. Ebrahimi S, Ghasemi-Basir HR, Majzoobi MM, Rasouli-Saravani A, Hajilooi M, Solgi G: HLA-DRB1*04 may predict the severity of disease in a group of Iranian COVID-19 patients. *Hum Immunol* 2021, 82:719–725.

21. Augusto DG, Yusufali T, Peyser ND, Butcher X, Marcus GM, Olgin JE, Pletcher MJ, Maiers M, Hollenbach JA: HLA-B*15:01 is associated with asymptomatic SARS-CoV-2 infection. *Medrxiv* 2021, doi:10.1101/2021.05.13.21257065.

22. Schindler E, Dribus M, Duffy BF, Hock K, Farnsworth CW, Gragert L, Liu C: HLA genetic polymorphism in patients with Coronavirus Disease 2019 in Midwestern United States. *Hla* 2021, 98:370–379.

23. Abdelhafiz AS, Ali A, Fouda MA, Sayed DM, Kamel MM, Kamal LM, Khalil MA, Bakry RM: HLA-B*15 predicts survival in Egyptian patients with COVID-19. *Hum Immunol* 2021, doi:10.1016/j.humimm.2021.09.007.

24. Novelli A, Andreani M, Biancolella M, Liberatoscioli L, Passarelli C, Colona VL, Rogliani P, Leonardis F, Campana A, Carsetti R, et al.: HLA alleles frequencies and susceptibility to COVID-19 in a group of 99 Italian patients. *Hla* 2020, 96:10.1111/tan.14047.

25. Ahmed SF, Quadeer AA, McKay MR: Preliminary Identification of Potential Vaccine Targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies. *Viruses* 2020, 12:254.

26. Lee CH, Koohy H: In silico identification of vaccine targets for 2019-nCoV. *F1000research* 2020, 9:145.

27. Nguyen A, David JK, Maden SK, Wood MA, Weeder BR, Nellore A, Thompson RF: Human Leukocyte Antigen Susceptibility Map for Severe Acute Respiratory Syndrome Coronavirus 2. *J Virol* 2020, 94:e00510-20.

28. Enayatkhani M, Hasaniazad M, Faezi S, Guklani H, Davoodian P, Ahmadi N, Einakian MA, Karmostaji A, Ahmadi K: Reverse vaccinology approach to design a novel multi-epitope vaccine candidate against COVID-19: an in silico study. *J Biomol Struct Dyn* 2020, 39:1–19.

29. Lin M, Tseng H-K, Trejaut JA, Lee H-L, Loo J-H, Chu C-C, Chen P-J, Su Y-W, Lim KH, Tsai Z-U, et al.: Association of HLA class I with severe acute respiratory syndrome coronavirus infection. *Bmc Med Genet* 2003, 4:9–7.

30. Heinz FX, Stiasny K: Distinguishing features of current COVID-19 vaccines: knowns and unknowns of antigen presentation and modes of action. *Npj Vaccines* 2021, 6:104.

31. Ndwandwe D, Wiysonge CS: COVID-19 vaccines. *Curr Opin Immunol* 2021, 71:111–116.

32. Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire AT, Veesler D: Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 2020, 181:281-292.e6.

33. Sternberg A, Naujokat C: Structural features of coronavirus SARS-CoV-2 spike protein: Targets for vaccination. *Life Sci* 2020, 257:118056.

34. 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:gkaa379-.

35. Nersisyan S, Zhiyanov A, Shkurnikov M, Tonevitsky A: T-CoV: a comprehensive portal of HLA-peptide interactions affected by SARS-CoV-2 mutations. *Nucleic Acids Res* 2021, doi:10.1093/nar/gkab701.

••This paper presents the T-cell COVID-19 Atlas (https://t-cov.hse.ru), an online portal that allows users to comprehensively analyze how mutations in multiple SARS-CoV-2 strains are predicted to affect HLA peptide binding affinities

36. Yang W, Shaman J: COVID-19 pandemic dynamics in India, the SARS-CoV-2 Delta variant, and implications for vaccination. *Medrxiv* 2021, doi:10.1101/2021.06.21.21259268.

37. Pulliam JRC, Schalkwyk C van, Govender N, Gottberg A von, Cohen C, Groome MJ, Dushoff J, Mlisana K, Moultrie H: Increased risk of SARS-CoV-2 reinfection associated with emergence of the Omicron variant in South Africa. *Medrxiv* 2021, doi:10.1101/2021.11.11.21266068.

38. Wang L, Cheng G: Sequence analysis of the emerging SARS-CoV-2 variant Omicron in South Africa. *J Med Virol* 2021, doi:10.1002/jmv.27516.

39. Gourraud P-A, Khankhanian P, Cereb N, Yang SY, Feolo M, Maiers M, Rioux JD, Hauser S, Oksenberg J: HLA Diversity in the 1000 Genomes Dataset. *Plos One* 2014, 9:e97282.

40. Poran A, Harjanto D, Malloy M, Arieta CM, Rothenberg DA, Lenkala D, Buuren MM van, Addona TA, Rooney MS, Srinivasan L, et al.: Sequence-based prediction of SARS-CoV-2 vaccine targets using a mass spectrometry-based bioinformatics predictor identifies immunogenic T cell epitopes. *Genome Med* 2020, 12:70.

41. Kiyotani K, Toyoshima Y, Nemoto K, Nakamura Y: Bioinformatic prediction of potential T cell epitopes for SARS-Cov-2. *J Hum Genet* 2020, 65:569–575.

42. Grifoni A, Sidney J, Zhang Y, Scheuermann RH, Peters B, Sette A: A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host Microbe* 2020, 27:671-680.e2.

••This paper analyzed the homology of multiple coronaviruses to identify potential T and B cell epitopes for SARS-CoV-2, aiming to better understand antigen presentation facilitate vaccine design.

43. Paul S, Arlehamn CSL, Scriba TJ, Dillon MBC, Oseroff C, Hinz D, McKinney DM, Pro SC, Sidney J, Peters B, et al.: Development and validation of a broad scheme for prediction of HLA class II restricted T cell epitopes. *J Immunol Methods* 2015, 422:28–34.

44. Neefjes J, Jongsma MLM, Paul P, Bakke O: Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol* 2011, 11:823–836.

45. Ferretti AP, Kula T, Wang Y, Nguyen DMV, Weinheimer A, Dunlap GS, Xu Q, Nabilsi N, Perullo CR, Cristofaro AW, 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.

46. Saini SK, Hersby DS, Tamhane T, Povlsen HR, Hernandez SPA, Nielsen M, Gang AO, Hadrup SR: SARS-CoV-2 genome-wide T cell epitope mapping reveals immunodominance and substantial CD8+ T cell activation in COVID-19 patients. *Sci Immunol* 2021, 6:eabf7550.

47. Schulien I, Kemming J, Oberhardt V, Wild K, Seidel LM, Killmer S, Sagar, Daul F, Lago MS, Decker A, et al.: Characterization of pre-existing and induced SARS-CoV-2-specific CD8+ T cells. *Nat Med* 2021, 27:78–85.

••This study analyzed epitopes that were predicted in silico to bind to HLA allotypes that are common in worldwide populations and confirmed T-cell responses to 62% of the predicted peptides.

48. Gangaev A, Ketelaars SLC, Isaeva OI, Patiwael S, Dopler A, Hoefakker K, Biasi SD, Gibellini L, Mussini C, Guaraldi G, et al.: Identification and characterization of a SARS-CoV-2 specific CD8+ T cell response with immunodominant features. *Nat Commun* 2021, 12:2593.

••This study identified 18 SARS-CoV-2 epitopes recognized by CD8+ T cells and performed in-depth characterization of SARS-CoV-2-specific CD8+ T cell responses of patients with acute critical and severe disease.

49. Weingarten-Gabbay S, Klaeger S, Sarkizova S, Pearlman LR, Chen D-Y, Gallagher KME, Bauer MR, Taylor HB, Dunn WA, Tarr C, et al.: Profiling SARS-CoV-2 HLA-I peptidome reveals T cell epitopes from out-of-frame ORFs. *Cell* 2021, 184:3962-3980.e17.

••This study describes the first HLA-class I immunopeptidome of SARS-CoV-2 in two cell lines at different times post infection using mass spectrometry

50. Francis JM, Leistritz-Edwards D, Dunn A, Tarr C, Lehman J, Dempsey C, Hamel A, Rayon V, Liu G, Wang Y, et al.: Allelic variation in class I HLA determines CD8+ T cell repertoire shape and cross-reactive memory responses to SARS-CoV-2. *Sci Immunol* 2021.

••Applying single-cell, multi-omic technology, this study characterized T-cell to SARS-CoV-2 across different *HLA* class I alleles, showing that *HLA* genotype influences CD8+ T cell repertoire shape and utilization of immune recall.

Table 1. Summary of *HLA* associations with COVID-19 outcomes. OR = odds ratio; CI = confidence interval. *African American** Hispanic

	Allele	OR (CI 95%)	*p*-value	corrected *p*-value	Country	Association with	Reference
Class I	*HLA-A*11:01*	3.41 (1.50–7.73)		3.34E-03	Japan	Severe disease	[16]
	*HLA-A*11:01*	2.33	8.51E-03		Japan	Severe disease	[18]
	*HLA-A*30:02*	2.2 (1.4–3.6)	1.70E-03	1.00E-02	USA*	Infection	[22]
	*HLA-B*15*		1.00E-02		Egypt	Survival	[23]
	*HLA-B*15:01*	2.4 (1.54-3.64)	5.67E-05	1.70E-03	USA	Asymptomatic infection	[21]
	*HLA-B*51:01*	3.38	0.007017		Japan	Severe disease	[18]
	*HLA-C*04:01*	5.4 (1.9–15.1)	1.10E-04	7.40E-03	Germany	Severe disease	[15]
	*HLA-C*14:02*	4.75	3.03E-03		Japan	Severe disease	[18]
	*HLA-DQB1*06:02*		1.00E-04	1.60E-03	Italy	Infection	[24]
Class II	*HLA-DRB1*04*	0.289	5.00E-03		Iran	Mild disease	[20]
	*HLA-DRB1*04:01*		3.00E-03		UK	Asymptomatic infection	[19]

*HLA-DRB1*08:02*	9.0 (2.2–37.9)	1.00E-02	3.00E-02	USA**	Infection	[22]
*HLA-DRB1*09:01*	3.62(1.57–8.35)	2.51E-03		Japan	Severe disease	[17]
*HLA-DRB1*15:01*		1.50E-03	4.80E-02	Italy	Infection	[24]

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: