The HLA-II immunopeptidome of SARS-CoV-2

Shira Weingarten-Gabbay[1,2,3,*,#], Da-Yuan Chen[4,5,*], Siranush Sarkizova[1,*], Hannah B. Taylor[1], Matteo Gentili[1], Leah R. Pearlman[1], Matthew R. Bauer[6], Charles M. Rice[3], Karl R. Clauser[1], Nir Hacohen[1,7,8], Steven A. Carr[1], Jennifer G. Abelin[1,†], Mohsan Saeed[4,5,†], Pardis C. Sabeti[1,2,9,10,11,†]

[1]Broad Institute of MIT and Harvard, Cambridge, MA, USA
[2]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA
[3]Laboratory of Virology and Infectious Disease, The Rockefeller University, New York, NY, USA
[4]Department of Biochemistry & Cell Biology, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, USA
[5]National Emerging Infectious Diseases Laboratories, Boston University, Boston, MA, USA
[6]Harvard Program in Biological and Biomedical Sciences, Division of Medical Sciences, Harvard Medical School, Boston, Massachusetts, USA
[7]Department of Medicine, Massachusetts General Hospital, Boston, MA, USA
[8]Center for Cancer Research, Massachusetts General Hospital, Boston, MA, USA
[9]Massachusetts Consortium on Pathogen Readiness, Boston, MA, USA
[10]Department of Immunology and Infectious Disease, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[11]Howard Hughes Medical Institute, Chevy Chase, MD, USA

* These authors contributed equally

† Senior author

Corresponding author: shirawg@broadinstitute.org (S.W-G)

ABSTRACT

Targeted synthetic vaccines have the potential to transform our response to viral outbreaks; yet the design of these vaccines requires a comprehensive knowledge of viral immunogens, including T-cell epitopes. Having previously mapped the SARS-CoV-2 HLA-I landscape, here we report viral peptides that are naturally processed and loaded onto HLA-II complexes in infected cells. We identified over 500 unique viral peptides from canonical proteins, as well as from overlapping internal open reading frames (ORFs), revealing, for the first time, the contribution of internal ORFs to the HLA-II peptide repertoire. Most HLA-II peptides co-localized with the known CD4+ T cell epitopes in COVID-19 patients. We also observed that two reported immunodominant regions in the SARS-CoV-2 membrane protein are formed at the level of HLA-II presentation. Overall, our analyses show that HLA-I and HLA-II pathways target distinct viral proteins, with the structural proteins accounting for most of the HLA-II peptidome and non-structural and non-canonical proteins accounting for the majority of the HLA-I peptidome. These findings highlight the need for a vaccine design that incorporates multiple viral elements harboring CD4+ and CD8+ T cell epitopes to maximize the vaccine effectiveness.

1

INTRODUCTION

Recent success with synthetic vaccines against viral diseases has demonstrated their promise for future use, but also highlighted the need for a deeper understanding of how the adaptive immune system recognizes and responds to viral infections(Sette & Crotty, 2021; Wherry & Barouch, 2022). In contrast to traditional attenuated vaccines that mimic the natural interaction between viruses and the immune system, rationally designed synthetic vaccines deliver only a small, selected portion of the viral genome, often a single protein, to stimulate an immune response. This allows rationally designed vaccines to be developed quickly, produced in large quantities with low-cost manufacturing, and administered safely (Pardi et al., 2018). However, the development of rationally designed vaccines necessitates a comprehensive knowledge of viral immunology and the parts of the viral genome that are recognized by the different arms of the immune system.

Limitations of the COVID-19 vaccines developed during the course of the pandemic highlighted the importance of eliciting a more comprehensive immune response through T cells(Neale et al., 2023). Early vaccines mostly focused on eliciting humoral immune response through utilizing the viral spike protein (S) (Baden et al., 2021; Martínez-Flores et al., 2021; Sahin et al., 2021). However, challenges with waning immunity led to a greater focus on providing durable immunity by eliciting CD8+ cytotoxic and CD4+ helper T cell responses. Together with the high frequency of mutations in spike and the associated immune evasion of emerging variants, these findings motivated researchers to incorporate T cell targets, such as the nucleocapsid (N), in the next generation of COVID-19 vaccines (Arieta et al., 2023; Dutta Noton K. et al., 2020; Hajnik et al., 2022; Matchett et al., 2021; Oronsky et al., 2022). CD8+ and CD4+ T cells recognize viral peptides that are endogenously processed and presented on the surface via HLA-I and HLA-II complexes, respectively. These T cell epitopes originate from all viral proteins, in contrast to neutralizing antibodies that are mostly confined to external structural viral proteins. Although the presence of a large number of potential T cell epitopes in the viral genome offer a wide range of candidates, it can also present a challenge in identifying the most effective targets for T cell-based vaccines.

To understand the full range of T cell epitopes, it is important to implement non-targeted, comprehensive approaches in addition to conventional T cell assays. Targeted T cell assays require researchers to decide a priori which peptides to screen in the experiment, based on assumptions about which viral proteins are expressed and processed for HLA presentation. While T cell responses to SARS-CoV-2 have been extensively studied, many studies have focused on only a subset of viral proteins, particularly the structural proteins S, N, membrane (M), and envelope (E), rather than considering the full range of proteins in the SARS-CoV-2 proteome (Chen et al., 2021; Joag et al., 2021; Keller et al., 2020; Le Bert et al., 2021; Lee et al., 2021; Mahajan et al., 2021; Nielsen et al., 2021; Rha et al., 2021; Shomuradova et al., 2020). A smaller number of studies have looked at T cell responses to peptides from all annotated viral proteins and found significant responses to both structural and non-structural proteins (Ferretti et al., 2020; Gangaev et al., 2021; Kared et al., 2021; Mateus et al., 2020; Nelde et al., 2021; Prakash et al., 2021; Tarke et al., 2021). Although these studies provide a broader view, they have still generally only included peptides from annotated canonical proteins and have not considered peptides from

2

non-canonical ORFs that were identified through experimental translation measurements (Finkel et al., 2020).

Taking an untargeted approach of HLA immunopeptidome profiling, we recently revealed surprising insights about HLA-I presentation in SARS-CoV-2 infected cells. Immunopeptidome profiling utilizes mass spectrometry to detect peptides that are endogenously presented on the HLA complex in different disease contexts (Abelin et al., 2017; Bassani-Sternberg & Gfeller, 2016; Chong et al., 2018; Croft et al., 2013; McMurtrey et al., 2008; Rucevic et al., 2016; Sarkizova et al., 2020; Schellens et al., 2015; Ternette et al., 2016). By applying it to SARS-CoV-2 infected cells we uncover HLA-I peptides from canonical proteins as well as overlapping ORFs in the coding region of N and S that were overlooked by previous T cell studies (Weingarten-Gabbay et al., 2021). Strikingly, some of the non-canonical peptides were more immunogenic in COVID-19 patients and humanized mice when compared to the strongest canonical protein derived antigens reported to date. Two of the non-canonical peptides were further supported by an independent HLA-I peptidome study from another group (Nagler et al., 2021). We also uncovered that, in contrast to earlier studies focusing mostly on structural proteins, non-structural proteins constitute a significant portion of the HLA-I peptidome, and that the timing of viral protein expression is a key determinant of HLA-I presentation and immunogenicity. The list of SARS-CoV-2 HLA-I peptides resulting from our study was recently used to characterize a T cell-directed mRNA vaccine (BNT162b4) (Arieta et al., 2023). Of the 19 HLA-I peptides observed in cells expressing a multi epitope vaccine, 3 were identical to peptides that we detected in infected cells. Together, HLA immunopeptidome profiling can identify new, highly potent T cell epitopes, inform vaccine design, and deepen our understanding of the determinants of viral antigen presentation.

Immunopeptidome profiling can be utilized to directly characterize another critical process in the antiviral immune response: HLA-II presentation to CD4+ helper T cells. In contrast to HLA-I presentation that samples endogenous cytosolic proteins, HLA-II presents peptides from proteins that have been taken up from outside the cell via endocytosis. Hence, different viral proteins may differentially access the HLA-I and HLA-II pathways. Understanding these differences can enable researchers to design vaccines that elicit both CD8+ and CD4+ T cells responses. Previous HLA-II peptidome studies of SARS-CoV-2 researched HLA-II peptides derived only from the spike protein, using a purified recombinant protein (Knierman et al., 2020; Parker et al., 2021), or from four viral proteins (N, M, E and nsp6) using plasmid overexpression (Nagler et al., 2021). Thus, we have not yet achieved a systematic view of HLA-II peptides from the full SARS-CoV-2 genome in the context of authentic virus infection.

In this study, we set out to achieve a comprehensive map of SARS-CoV-2 peptides that are processed and presented by HLA-II complexes. Using these data, we dissected the source viral proteins and the processed regions within each viral protein that are presented by HLA-II. We contextualize these findings by using thousands of reported CD4+ T cell epitopes, inferring the contribution of antigen processing and presentation steps to T cell responses observed in COVID-19 patients. We then compared the immunopeptidome of HLA-I and HLA-II complexes, revealing important differences between SARS-CoV-2 presentation to CD4+ and CD8+ T cells. The newly identified CD4+ T cell targets and the insights our study rendered about viral HLA-II presentation

will enable a more precise selection of peptides for the next generation of COVID-19 vaccines that aim to target multiple viral proteins. The concepts learned from this study can also be applied to the design of synthetic vaccines against other viral pathogens.

RESULTS

Immunopeptidome profiling of HLA-II peptides in SARS-CoV-2 infected cells

To interrogate the HLA-II immunopeptidome of SARS-CoV-2, we infected A549 and HEK293T cells with the virus after inducing the HLA-II presentation pathway, immunoprecipitated (IP) HLA-II-peptide complexes, and identified bound peptides by liquid chromatography-tandem mass spectrometry (LC-MS/MS)(**Figure 1A**). Both cell lines stably expressed ACE2 and TMPRSS2, two important SARS-CoV-2 entry factors. To induce the HLA-II pathway, we transduced the cells with the MHC class II transactivator (CIITA) using a lentiviral vector. Overexpression of CIITA, a master transcriptional regulator, facilitates cell-surface expression of HLA-II complexes and the peptide-loading machinery and has been previously used to interrogate the HLA-II immunopeptidome of virus-infected cells (Becerra-Artiles et al., 2019, 2022) and tumors (Forlani et al., 2021; Hos et al., 2022). In addition, some lung epithelial cells express HLA-II (Neuwelt et al., 2020; Wosen et al., 2018) and thus, studying the HLA-II immunopeptidome of infected cells mimics HLA-II presentation that can occur *in-vivo* during the course of infection with a respiratory virus.

We characterized the CIITA-transduced cells to ensure proper induction of proteins in the HLA-II pathway. First, we compared cell surface levels of HLA-II in A549 cells with those in a positive control human melanoma A375 cell line that endogenously expresses HLA-II (Deffrennes et al., 2001). The cell surface flow cytometry revealed strong induction (~80-fold) of HLA-II expression in A549/ACE2/TMPRSS2 (A549/AT) cells upon CIITA transduction, with a similar fluorescence intensity as in A375 cells (**Figure 1B**). To monitor the expression of additional proteins in the HLA-II pathway, we examined the whole proteome of A549/AT and HEK293T/AT cells by analyzing lysates after the HLA-II IP using LC-MS/MS. We observed the expected increase in CIITA-induced proteins that are localized in the MHC-II region of the MHC locus including HLA-DM, HLA-DO, and TAP1 (**Figure S1A**).

We also ensured that CIITA overexpression did not affect the cell susceptibility to SARS-CoV-2 infection, since a recent study reported that CIITA can restrict SARS-like coronaviruses in U2OS osteosarcoma cells (Bruchez et al. 2020). We quantified SARS-CoV-2 infection by immunofluorescence (IF) staining of cells with an anti-N antibody at 24 hours post infection (hpi) (**Figure 1C**, **Figure S1B**). When infected at different multiplicities of infection (MOIs), and A549/AT cells exhibited similar infection levels regardless of CIITA expression. In contrast, HEK293T/AT cells showed reduced infection upon CIITA expression, although a substantial number of cells were still positive, with ~50% infected cells at MOI of 3 (the MOI we used for our HLA-II IP experiment).

4

To gauge the technical performance of our experimental system, we examined if the peptides detected by LC-MS/MS match known characteristics of HLA-II peptides. We performed HLA-II IP of non-infected and infected cells in two biological replicates at 24 hpi using a mixture of antibodies targeting the three HLA-II loci (HLA-DR, HLA-DP, and HLA-DQ). We recovered 21,541 and 29,600 unique HLA-II peptides from A549/ACE2/TMPRSS2/CIITA (A549/ATC) and HEK293T/ACE2/TMPRSS2/CIITA (HEK293T/ATC) cells, respectively (**Table S1B**). Of those, 0.5% (n=119) of A549 and 1.3% (n=389) of HEK293T peptides were derived from SARS-CoV-2 proteins (**Table S1C,D**). This result aligns with the overall low percentage of viral peptides reported in previous studies that investigated the HLA-II immunopeptidome of EBV-B cells infected with measles virus (Ovsyannikova et al., 2003) and Monocyte-derived DCs (moDC) pulsed with recombinant H1-HA protein of influenza virus (Cassotta et al., 2020). Furthermore, the recovered HLA-II peptides exhibited the expected 12-25 amino acid length distribution (**Figure 1D**). We examined the amino acid sequences of detected peptides and checked if these sequences conformed to the binding preferences of the HLA-II alleles expressed in A549 and HEK293T cells. The binding motifs, deconvoluted by Gibbs Cluster (Andreatta et al., 2017), agreed with known preferences of HLA-DR heterodimers (Abelin et al., 2019) expressed in these cell lines (**Figure 1E, Table S1A**). As expected, the deconvoluted motifs matched more to HLA-DR heterodimers, and less so to HLA-DP and HLA-DQ heterodimers, in both cell lines, as HLA-DR is often expressed at higher levels when compared to HLA-DP and HLA-DQ (Taylor et al., 2021). Notably, some of the HLA-II alleles expressed by the two cell lines, and confirmed by motif deconvolution are highly prevalent in the European (EUR) and United States (USA) populations including DRB1*07:01 (~14% EUR, 12% ~USA) expressed by A549 and DRB1*15:01 (~14% EUR, ~11% USA) and DRB5*01:01 (~16% EUR) expressed by HEK293T (**Table S1A**).

Finally, we confirmed that CIITA-transduced cells presented peptides derived from extracellular proteins, as should be expected for HLA-II presentation. To this end, we quantified the fraction of HLA-II peptides derived from bovine serum albumin (BSA), a non-human protein present in the cell growth medium, as done previously (Forlani et al., 2021). As a negative control, we examined the HLA-I immunopeptidome of the same cells (Weingarten-Gabbay et al., 2021). Since HLA-I peptides are mostly processed from endogenous proteins, we expected low representation of the exogenous BSA protein in these data. Indeed, we observed 5.2- and 4.6-fold more BSA-derived peptides in the HLA-II peptidome than the HLA-I peptidome in A549/ATC and HEK293T/ATC cells, respectively (**Figure 1G**). Moreover, BSA-derived HLA-II peptides across both A549/ATC and HEK293T/ATC experiments had the expected lengths of 12-25 amino acids (~87% of peptides), matching the distribution of human-derived HLA-II peptides in the same samples (**Figure 1F**). In contrast, the BSA-derived peptides in HLA-I samples had longer lengths than canonical HLA-I peptides, suggesting that these peptides might have arisen from exogenous peptidase trimming and binding to empty surface HLA-I. Altogether, these data indicate that CIITA-transduced A549/AT and HEK293T/AT cells can present HLA-II peptides generated from endocytosed proteins.
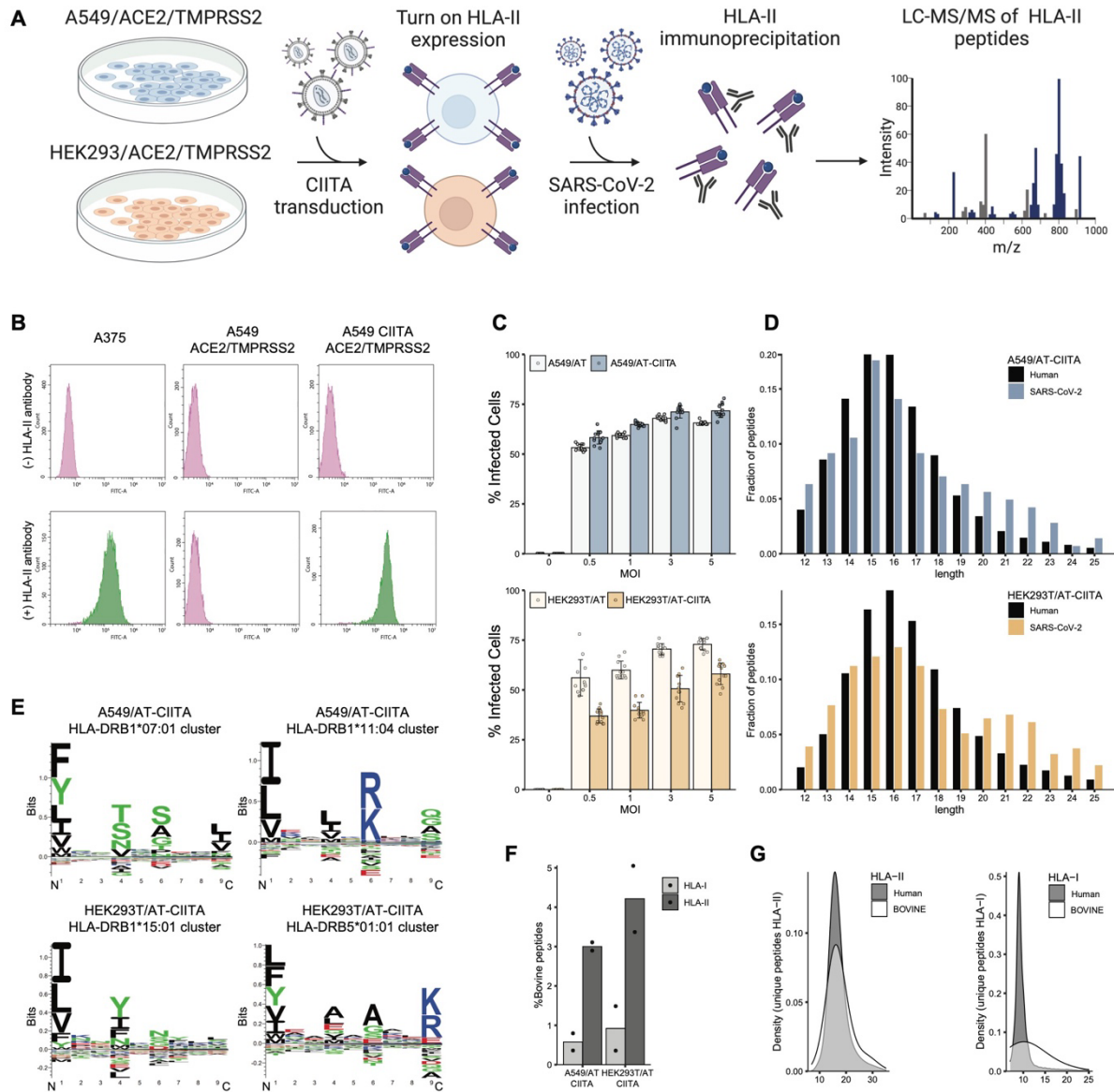
Figure 1. HLA-II immunopeptidome profiling of SARS-CoV-2 infected cells. (A) Schematic of the experimental workflow. **(B)** HLA-II expression measured by flow cytometry using a FITC-conjugated anti HLA-DR/DP/DQ antibody. Gating was based on unstained cells. **(C)** SARS-CoV-2 infection levels in A549/AT and HEK293T/AT cells with and without CIITA transduction. The infection was quantified using immunofluorescence staining for the nucleocapsid protein. MOI = Multiplicity Of Infection. **(D)** Length distribution of the eluted peptides in SARS-CoV-2 infected cells. **(E)** Gibbs Cluster deconvoluted motifs of the eluted peptides. **(F)** The percentage of peptides originating from Bovine proteins detected in the HLA-I and HLA-II immunopeptidomes. HLA-I data are from Weingarten-Gabbay et al. (Weingarten-Gabbay et al., 2021) **(G)** Length distribution of human and bovine peptides in the A549 and HEK293T HLA-I and the HLA-II immunopeptidomes.

6

SARS-CoV-2 HLA-II peptides from canonical and out-of-frame overlapping ORFs

We next analyzed the HLA-II presented peptides derived from SARS-CoV-2 proteins, detecting 469 unique peptides from canonical viral proteins: N, S, M, ORF3a, ORF6, non-structural protein 3 (nsp3) and nsp4 (**Table S1C,D, Figure 2A**). Examining the distribution of HLA-II peptides across SARS-CoV-2 proteins, we observed the expected clustering of peptides into nested sets, with shared core sequences but different N- or C- terminal ends (**Figures 2B and S2**). This pattern is a hallmark of the HLA class II pathway and differentiates it from the class I presentation. While MHC class I molecules structurally constrain the length of loaded peptides, the open structure of the binding groove of MHC-II molecules allows interaction with peptides of variable length, with parts of the peptides protruding out of the binding groove (Lippolis et al., 2002). Interestingly, although A549/ATC and HEK293T/ATC cells express different HLA-II alleles (**Supp Table 1A**) with distinct binding preferences, some clusters contained peptides from both cell lines. This observation suggests that viral antigen processing steps upstream of HLA-II peptide loading play a key role in shaping the HLA-II immunopeptidome.

The untargeted nature of our analysis allowed us to search for peptides originating from non-canonical ORFs (Finkel et al., 2020) in addition to the canonical SARS-CoV-2 proteins. Overall, we detected 11 peptides from two internal overlapping ORFs: ORF9b (overlapping with N, also called N.iORF1) and ORF3c (overlapping with ORF3a, also called 3a.iORF1). ORF9b gave rise to 10 peptides, all of which clustered into two nested sets in the first half of the protein (**Table S1C,D, Figure 2C**). From each nested set, at least one peptide was predicted to bind one of the HLA-II alleles expressed in HEK293T cells: DRB5*01:01 and DQA1*01:02/DQB1*06:02 in the first and second sets, respectively. Interestingly, HLA-II peptides arose from a different region of the ORF9b protein compared to HLA-I peptides, which originated from the C-terminal region (Weingarten-Gabbay et al., 2021). We identified one HLA-II peptide (ALHFLLFFRALPKS) from ORF3c. Although it was only one peptide, it was observed in the same fraction (fxn 3) in two biological replicates and was a high scoring peptide, which supports its authenticity. To the best of our knowledge, this is the first experimental evidence for ORF3c expression at the protein level since it was originally detected using ribosome profiling (Finkel et al., 2020) and computational predictions (Cagliani et al., 2020; Firth, 2020; Jungreis et al., 2021).
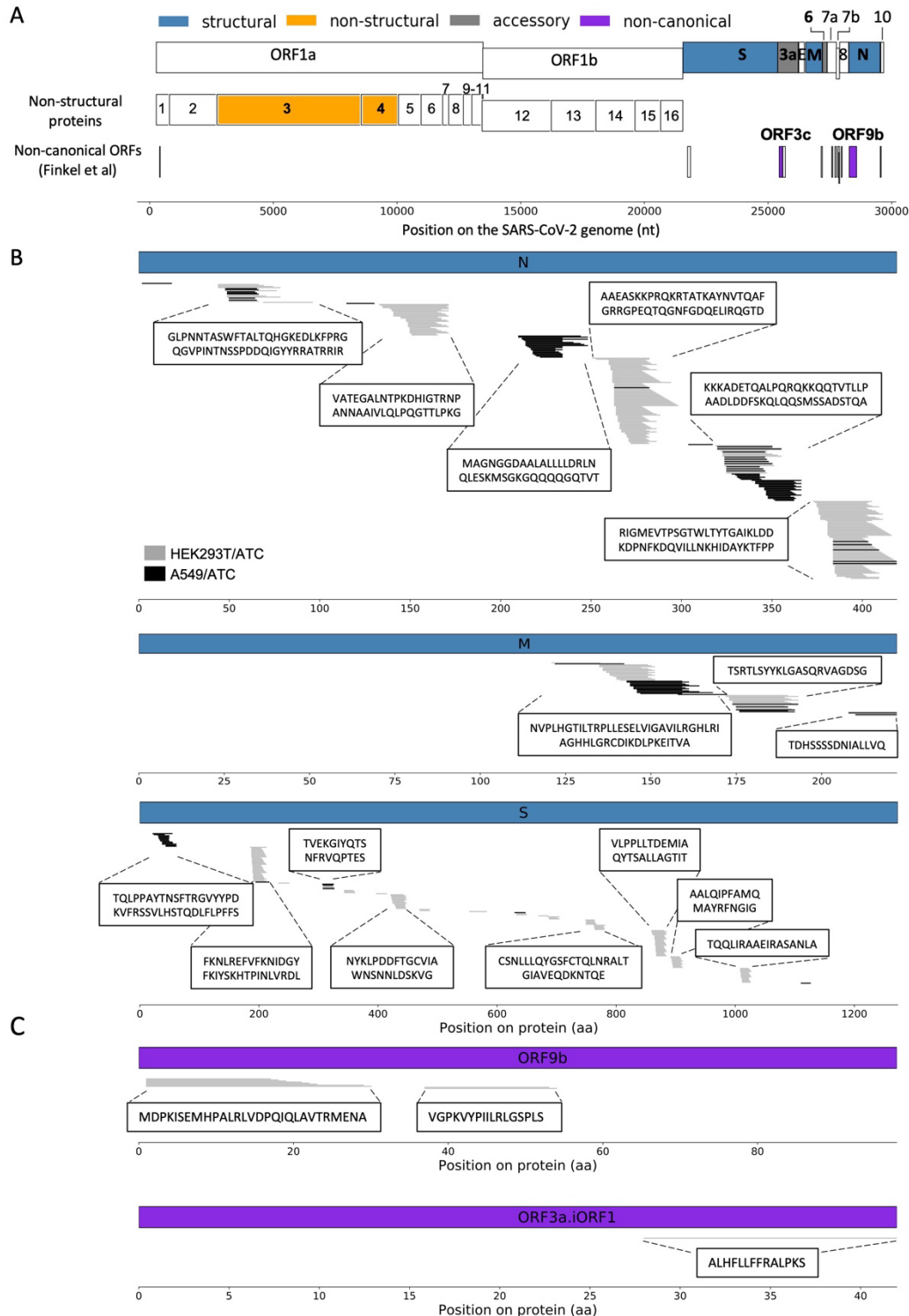
Figure 2. SARS-CoV-2 HLA-II immunopeptidome. (A) Summary of viral proteins presented on HLA-II complexes in infected 549/ATC and HEK293T/ATC cells. **(B)** The location of HLA-II peptides in canonical structural proteins M, N and S. Peptides detected in A549/ATC and HEK293T/AT/C cells are depicted in black and gray, respectively. **(C)** The location of HLA-II peptides in two non-canonical internal ORFs: ORF9b and ORF3c.

SARS-CoV-2 HLA-II peptides co-localize with epitopes that elicit CD4+ T cell responses in COVID-19 patients

To evaluate if the HLA-II peptides that we detected by mass spectrometry contribute to T cell responses in COVID-19 patients, we compared our HLA-II viral peptides to reported CD4+ epitopes derived from SARS-CoV-2 proteins. We used a curated dataset of T cell epitopes reported by Grifoni et al. (Grifoni et al., 2021). This dataset combines 9 studies that tested CD4+ T cell responses using various assays including ELISpot, Intracellular Cytokine Staining (ICS) and Activation Induced Markers (AIM) (**Figure 3A**) (Keller et al., 2020; Le Bert et al., 2020, 2021; Mahajan et al., 2021; Mateus et al., 2020; Nelde et al., 2021; Peng et al., 2020; Prakash et al., n.d.; Tarke et al., 2021).

First, we checked the overlap between viral proteins that generated HLA-II peptides and those known to contain CD4+ T cell epitopes. We computed the fraction of total CD4+ epitopes that were derived from each of the SARS-CoV-2 proteins. To avoid biases stemming from over-representation of highly characterized viral proteins, such as S, N and M, we limited our analysis to four studies that surveyed the entire canonical SARS-CoV-2 proteome (highlighted in asterisk in **Figure 3A**) (Mateus et al., 2020; Nelde et al., 2021; Prakash et al., n.d.; Tarke et al., 2021). As expected, viral proteins for which we observed HLA-II peptides elicited T cell responses significantly more frequently than those with no detectable HLA-II peptides (Wilcoxon rank-sum p $< 10^{-3}$, **Figure 3B**).

We then considered individual viral proteins and tested if the HLA-II peptides co-localize with regions that elicited CD4+ T cell responses in COVID-19 patients. We focused on the three structural proteins that gave rise to the majority of the detected HLA-II peptides: N (n=281), S (n=143), and M (n=56). These three proteins were also the most abundant source of epitopes in the compiled T cell data, accounting for 54% of the total CD4+ epitopes detected in the canonical SARS-CoV-2 proteome. To determine if HLA-II peptides were derived from regions that were more immunogenic in patients, we counted the number of amino acids that were covered by HLA-II peptides, CD4+ epitopes and both, and computed a hypergeometric p-value that estimates the overlap between these two groups. Since we compared peptide localization within individual proteins, we accounted for epitopes reported in all 9 CD4+ T cell studies listed in **Figure 3A**, including studies that examined only a few SARS-CoV-2 proteins. We found a statistically significant enrichment of HLA-II peptides in regions from which CD4+ epitopes were derived with $p<10^{-29}$, $p<10^{-5}$, and $p<10^{-5}$ for M, N and S, respectively (**Figure 3C-E**).

Furthermore, our analysis showed that HLA-II peptides greatly overlap the two immunodominant regions reported in the M protein. These regions, M:144-163 and M:173-192, were recently identified as hotspots of CD4-restricted epitopes that elicited T cell responses in eight and six COVID-19 convalescent samples, respectively (Keller et al., 2020). The HLA-II peptides that we detected by mass spectrometry were also confined to the same region of M with highest density in the two reported hot spots: M:121-171 (n=34 peptides) and M:172-192 (n=16 peptides), both detected in HEK293T/ATC and A549/ATC cells (**Figure 2B and 3C**, **Table S1C,D,F**). Moreover, the predicted HLA restriction of the two T cell epitopes from M:144-163 described by Keller et al.

(Keller et al., 2020) matches with two of the HLA-II alleles expressed in A549: DRB1*11:04 and DRB4*01:01. Altogether, these results indicate that the HLA-II immunopeptidome of SARS-CoV-2 co-localizes with reported CD4+ T cell epitopes and defines the immunogenic regions in the M protein.
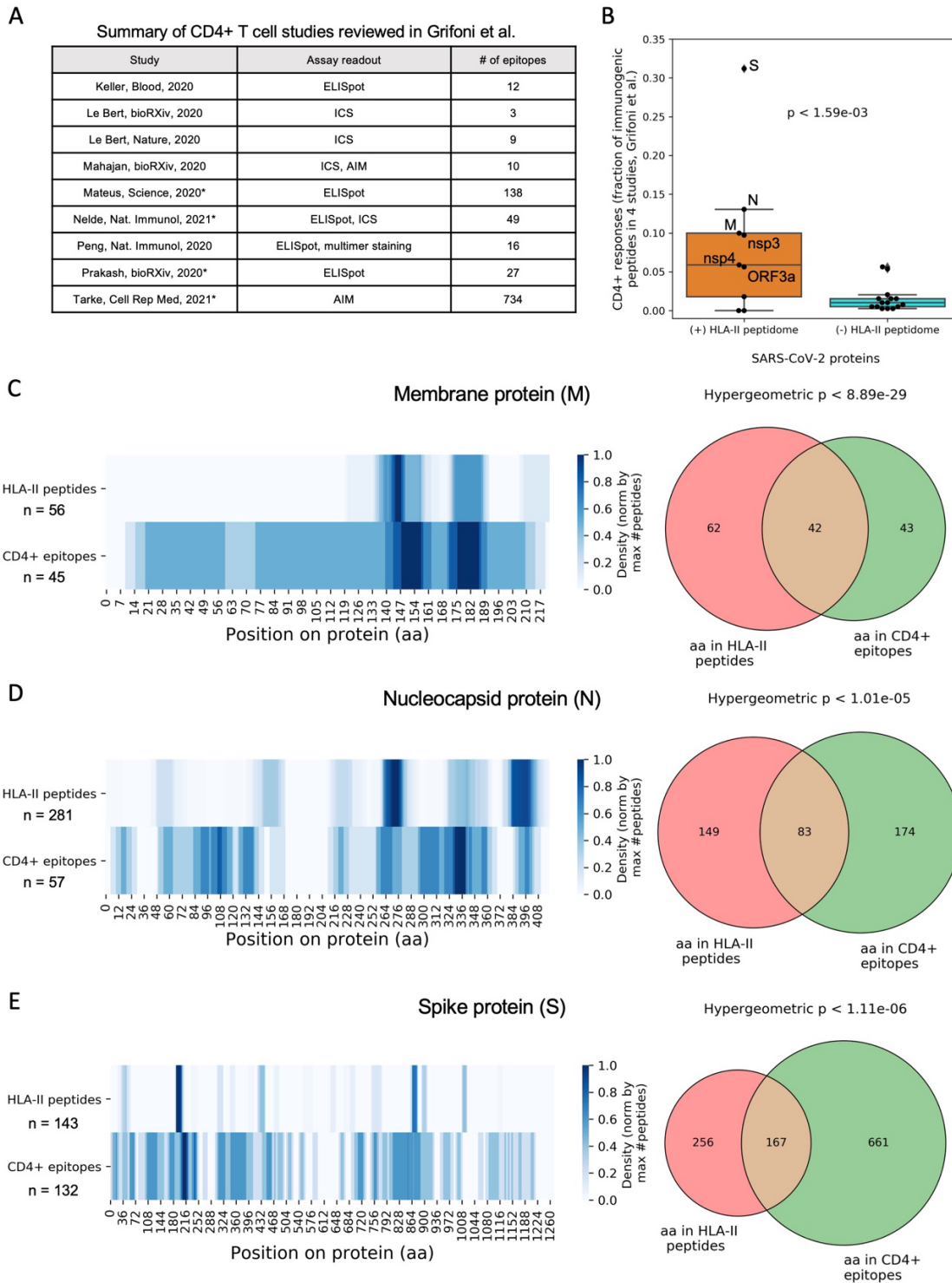

A Summary of CD4+ T cell studies reviewed in Grifoni et al.

Study	Assay readout	# of epitopes
Keller, Blood, 2020	ELISpot	12
Le Bert, bioRXiv, 2020	ICS	3
Le Bert, Nature, 2020	ICS	9
Mahajan, bioRXiv, 2020	ICS, AIM	10
Mateus, Science, 2020*	ELISpot	138
Nelde, Nat. Immunol, 2021*	ELISpot, ICS	49
Peng, Nat. Immunol, 2020	ELISpot, multimer staining	16
Prakash, bioRXiv, 2020*	ELISpot	27
Tarke, Cell Rep Med, 2021*	AIM	734

Figure 3. Systematic comparison between the SARS-CoV-2 HLA-II immunopeptidome and known CD4+ T cell epitopes. (A) A list of nine studies that identified CD4+ T cells epitopes in COVID-19 patients as described in Table 1 of Grifoni et al (Grifoni et al., 2021)**.** Studies that included peptides from all canonical SARS-CoV-2 proteins are highlighted with an asterisk. **(B)** Comparing the immunogenicity of SARS-CoV-2 proteins that were detected on the HLA-II complex to proteins that were not. For each protein, we plotted the fraction of CD4+ T cell epitopes from all the detected epitopes in four genome-wide studies (denoted with asterisk in panel A). Wilcoxon-rank sum p-value is shown. **(C-E)** Comparing the location of HLA-II peptides that were detected by mass spectrometry in infected A459/AT/CIITA and HEK293T/AT/CIITA cells to previously reported CD4+ T cell epitopes. (left) heatmap showing the density of HLA-II peptides and the reported CD4+ T cell epitopes across individual viral proteins. Rows were normalized according to the maximal density in each row. (right) venn diagram showing the number of amino acids that are covered by HLA-II peptides, CD4+ T cell epitopes and both. To reduce background levels, we counted amino acids with a greater coverage than the mean of each row. Hypergeometric p-value is shown for each viral protein.

Distinct subsets of viral proteins are presented on the HLA-I and HLA-II complexes

In the context of vaccine design, it is important to decipher both CD4+ and CD8+ T cell epitopes, since effective vaccines require potent induction of both arms of the adaptive cellular response. Many of the current vaccines strategies rely on the delivery of a single viral protein to invoke both CD8+ and CD4+ T cell responses and mainly focus on the structural proteins S or N (Barouch, 2022; Creech et al., 2021; Dai & Gao, 2021; Krammer, 2020; Kyriakidis et al., 2021; Oronsky et al., 2022; Silva et al., 2022). However, HLA-I and HLA-II presentation, which facilitate CD4+ and CD8+ responses, respectively, have distinct antigen processing steps, raising the likelihood that each of these pathways samples a different subset of viral proteins.

To compare viral proteins presented on HLA-I versus HLA-II complexes, we examined the HLA-I(Weingarten-Gabbay et al., 2021) and HLA-II immunopeptidomes of SARS-CoV-2 infected HEK293T/ATC and A549/ATC cells. We computed the number of peptides observed from each source protein and assessed the representation of proteins in four groups: Non-structural proteins, structural proteins, accessory proteins, and non-canonical ORFs. HLA-II presentation was dominated by the structural proteins N, S, and M, accounting for 95.2% of the detected peptides in HEK293T/ATC and A549/ATC cells, with negligible contribution of the non-structural (1.2%), accessory (1.4%) and non-canonical proteins (2.2%) (**Figure 4A**). In contrast, only 27% of the detected HLA-I peptides were derived from structural proteins, with a large fraction of peptides from non-structural proteins (45.9%) and non-canonical ORFs (24.3%) (**Figure 4B**). Together, these results point to a different presentation profile of SARS-CoV-2 on the HLA-I and HLA-II complexes and suggest that CD4+ and CD8+ T cells engage with different parts of the virus.
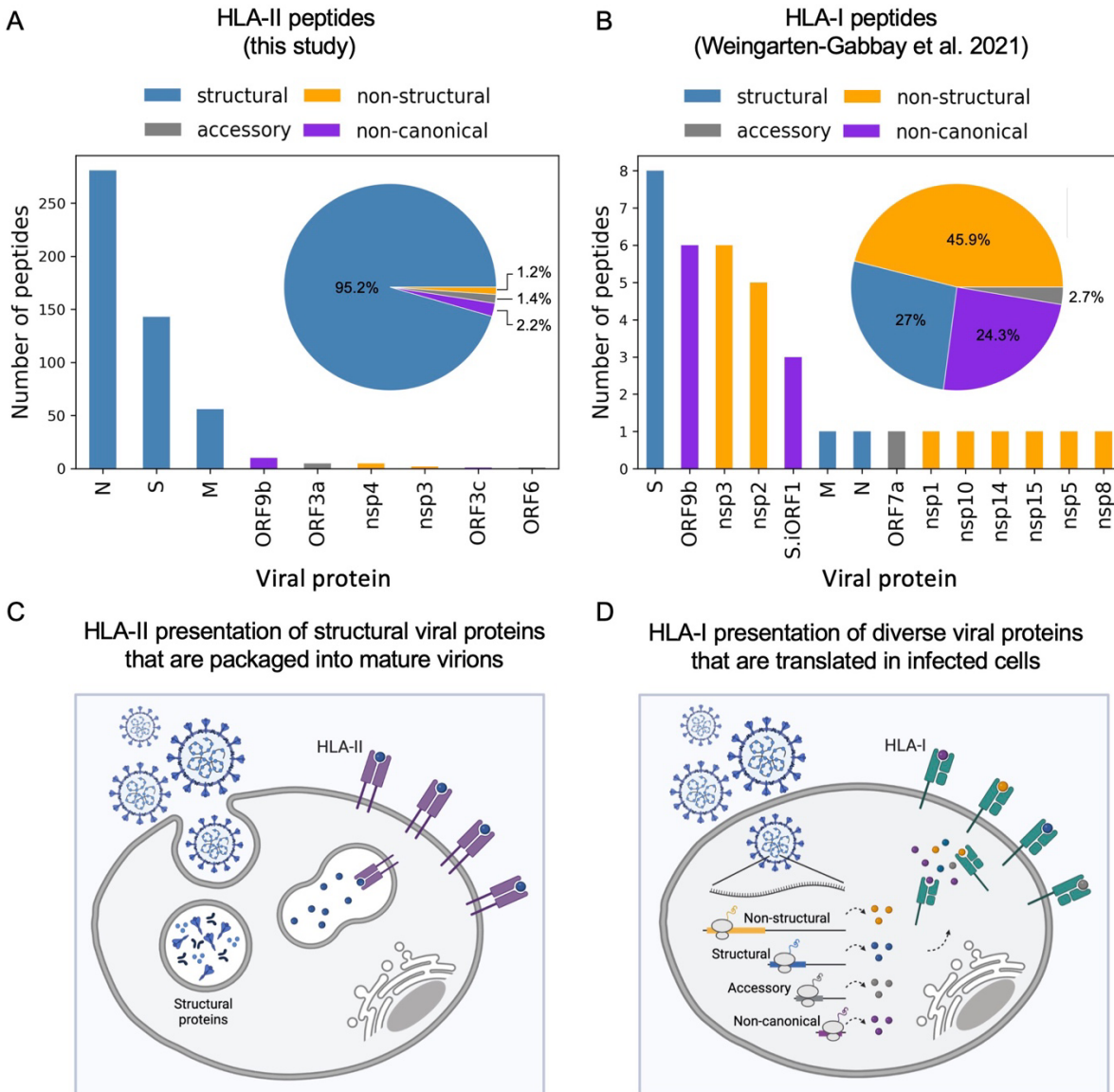
Figure 4. SARS-CoV-2 protein representation on the HLA-I and HLA-II complexes. (A) A bar chart showing the number of HLA-II peptides detected in SARS-CoV-2 infected A549/AT/CIITA and HEK293T/AT/CIITA cells for each viral protein. Inside the frame is a pie chart showing the relative abundance of peptides derived from structural proteins, non-structural proteins, accessory proteins and non-canonical ORFs. **(B)** Similar to (A) for HLA-I peptides reported in Weingarten-Gabbay et al. (Weingarten-Gabbay et al., 2021). **(C)** A cartoon illustrating the source proteins for the HLA-II processing and presentation pathway. Viruses are endocytosed by an antigen presenting cell; The viral structural proteins are cleaved within the endosomal–lysosomal compartment and loaded onto HLA-II complexes. **(D)** A cartoon illustrating the source proteins for the HLA-I processing and presentation pathway. Viral proteins are produced from the translation of genomic and subgenomic viral RNAs in infected cells. These proteins are cleaved and loaded onto HLA-I complexes.

DISCUSSION

We provide the first genomic landscape of SARS-CoV-2 peptides that are naturally processed and presented on the HLA-II complex. This genome-wide view allowed us to systematically compare the HLA-II immunopeptidome to the T cell epitopes in COVID-19 patients and to the HLA-I immunopeptidome of SARS-CoV-2 and to uncover new insights into antigen presentation.

Our work adds to a growing list of studies that employ overexpression of the CIITA master regulator to infer the HLA-II immunopeptidome of cancer cells and viruses (Becerra-Artiles et al., 2019, 2022; Forlani et al., 2021; Hos et al., 2022). Although the cells we profiled were not professional antigen-presenting cells (APCs), we believe that our measurements uncovered genuine HLA-II peptides: The peptides we detected were in the expected length range and contained sequence motifs that matched with the HLA-II alleles of respective cell lines. In addition, we detect presentation of exogenous proteins derived from extracellular bovine serum. Most importantly, our immunopeptidome data successfully captured the SARS-CoV-2 CD4+ T cell epitopes that were detected in a wide range of independent studies using conventional targeted T cell assays. The source proteins of HLA-II peptides in our study were also found to be the most immunogenic proteins in convalescent COVID-19 patients. Moreover, the HLA-II peptides co-localized with regions in the SARS-CoV-2 proteins that are known to elicit strong CD4+ T cell responses.

In the context of highly pathogenic viruses, CIITA induction in HLA-II null cell lines provides an efficient platform to probe the entire viral genome within the restrictions of high-containment facilities. To date, HLA-II immunopeptidome studies of SARS-CoV-2 were limited to a single protein, using pulse experiments with a recombinant protein (Knierman et al., 2020; Parker et al., 2021), or four exogenously expressed proteins using plasmid overexpression (Nagler et al., 2021). By inducing the HLA-II pathway in SARS-CoV-2 infected cells, we could detect HLA-II peptides from the entire viral genome. Although HLA-II peptides are predominantly presented by APCs, such as dendritic cells and macrophages, they can also be presented by non-immune cells upon induction of HLA-II expression, such as lung epithelial cells exposed to IFN-gamma (Neuwelt et al., 2020; Wosen et al., 2018). Thus, the CIITA induction system can also recapitulate naturally occurring HLA-II presentation events of infected cells *in-vivo*. Overexpressing CIITA in the same cell lines in which we profiled the HLA-I immunopeptidome allowed us to use the same reagents and inactivation protocol that we optimized for Biosafety Level 3 (BSL3) laboratory (Weingarten-Gabbay et al., 2022). Thus, we believe that this approach can readily be extended to additional high-containment viruses.

Our analysis uncovers striking differences in the subset of viral proteins that are presented on HLA-I versus HLA-II complexes. We hypothesize that the observed differences stem from the different stages of the viral life cycle at which viral proteins are processed and loaded onto HLA molecules (**Figure 4C,D**). The HLA class II pathway can present peptides from exogenous sources, e.g. mature virions and infected cell debris, that are cleaved within endosomal–lysosomal compartments and loaded onto HLA-II complexes (**Figure 4C**). Thus, the repertoire of HLA-II peptides reflects the viral proteins that are part of mature virus particles. In contrast, the

13

HLA-I pathway samples viral proteins that are actively translated in the cytoplasm of infected cells. Thus, the HLA-I immunopeptidome mirrors the translatome of the virus, including non-structural proteins, accessory proteins, and non-canonical proteins, which are not necessarily incorporated into mature virions (**Figure 4D**). These differences may extend to other viruses as well, many of which encode proteins that are required for viral replication in infected cells but are not packaged into mature virions. The distinct repertoire of viral peptides that are presented on HLA-I versus HLA-II complexes suggests that CD4+ and CD8+ T cells recognize different parts of the viral genome. This observation warrants revision of the current "one-protein-fits-all" approach that forms the basis of most synthetic vaccines, including the broadly administered COVID-19 mRNA vaccines. Rather, incorporating targets from different classes of viral proteins, including structural, non-structural and non-canonical proteins, has the potential to invoke a more holistic response of both CD4+ and CD8+ T cells.

An intriguing question in T cell immunology is what determines the immunodominance of a specific protein or a region within a protein. Observing T cell reactivity in convalescent patients to immunodominant epitopes represents the final outcome in a chain of events that determine which peptides will elicit a T cell response. Some factors defining the immunodominance of a given epitope include protein expression levels, accessibility to proteolytic cleavage and antigen processing, loading onto the HLA complex, the presence of a matching T cell receptor (TCR) in the repertoire of naive T cells, and the binding affinity of an HLA-peptide complex with its matching TCR. Since the immunopeptidome represents antigen processing and presentation steps occurring prior to interaction with TCR, it distinguishes between epitope selectivity at the level of presentation versus T cell recognition. A recent study identified that the HLA-II peptidome of influenza virus defines H1-HA immunodominant regions targeted by memory CD4+ T cells (Cassotta et al., 2020). Similarly to this observation, we found that the two immunodominant hotspots in the M protein (Keller et al., 2020) greatly overlap the HLA-II immunopeptidome of infected cells, suggesting that antigen processing and presentation steps define the immunodominant regions of M. Together, these observations demonstrate the robustness of immunopeptidome analysis for identifying the most relevant protein regions for designing T cell assays and selecting candidates for vaccines.

Our work uncovers HLA-II peptides from two non-canonical overlapping ORFs: ORF9b and ORF3c, and provides the first experimental evidence for ORF3c expression at the peptide level. In contrast to HLA-I presentation, in which non-canonical peptides were enriched on the HLA-I complex, non-canonical peptides represented only a small fraction of the HLA-II immunopeptidome. However, the detection of peptides from non-canonical ORFs in both the HLA-I and the HLA-II immunopeptidomes emphasizes the importance of incorporating non-canonical ORFs into T cell studies to achieve a complete picture of the antiviral immune profile. The discovery of a peptide from ORF3c highlights another important aspect of immunopeptidome studies and its potential impact on our understanding of non-canonical ORFs. Although non-canonical ORFs have numerous roles in the viral life cycle, including regulating viral gene expression and modulating virus infection, their detection in tryptic proteome experiments is often challenging due to their small size, shorter half-life, and in some cases, the lack of observable tryptic peptides. Of the 23 non-canonical ORFs that were identified in the SARS-CoV-2 genome

14

(Finkel et al., 2020), only one (ORF9b) has so far been detected in global tryptic proteomic experiments (Weingarten-Gabbay et al., 2021). The longer half-life of the HLA peptide complex compared to the non-canonical ORF translation product in the cell may increase the probability of detection by mass spectrometry (Ruiz Cuevas et al., 2021). As an example, although we and others identified three peptides from S.iORF1 (an internal overlapping ORF in spike) on the HLA-I complex (Nagler et al., 2021; Weingarten-Gabbay et al., 2021), this protein was not detected in whole tryptic proteome analysis of the same infected cell lysates. ORF3c was recently shown to inhibit innate immunity by restricting IFN-β production, exposing an important mechanism of SARS-CoV-2 immune evasion (Stewart et al., 2022). Our study provides evidence that this important non-canonical protein is expressed in cells that are infected with SARS-CoV-2. Thus, in addition to enhancing our understanding of viral antigen presentation, immunopeptidome studies contribute to our basic understanding of viruses by illuminating the complete set of canonical and non-canonical viral proteins.

There are limitations to this study. Although CIITA overexpression activated the HLA-II pathway in the cell lines used in our study, allowing the investigation of SARS-CoV-2 HLA-II immunopeptidome, these cells may not capture the unique biology of how APCs subtypes uptake and process mature virions and virus infected cells *in vivo*. In contrast to the cells profiled in this study, APCs are not naturally infected by SARS-CoV-2. It is possible that productive infection impacts the antigen processing and presentation steps and provides an internal source of viral proteins, in addition to the endocytosed particles, for production of HLA-II peptides. For instance, Ghosh and colleagues (Ghosh et al. 2020) showed that β-Coronaviruses can traffic to lysosomes and egress by Arl8b-dependent lysosomal exocytosis. In this context, lysosomes are deacidified, which can inactivate proteolytic enzymes in infected cells and impair antigen presentation. Further, our study only uncovered the peptides that are presented by the HLA-II alleles expressed in A549 and HEK293T cells. As such, we might have missed SARS-CoV-2 HLA-II peptides that were incompatible with A549 and HEK293T cells. Nonetheless, our study identifies HLA-II peptides that can be presented on virally infected cells and provides valuable insights into which viral proteins are more likely to be presented by SARS-CoV-2 infected cells.

METHODS

Plasmid
Lentiviral vectors pLOC_hACE2_PuroR and pLOC_hTMPRSS2_BlastR, harboring human ACE2 and TMPRSS2, respectively, have been described (Chen et al., JVI). To generate a lentiviral vector containing human CIITA, we amplified the CIITA cDNA from pcDNA3 myc CIITA (Addgene #14650) and cloned it into pTRIP-SFFV-Hygro-2A (previously described(Gentili et al., 2023)) via Gibson assembly. The resultant plasmid was named pTRIP-SFFV-Hygro-2A-myc-CIITA.

Cell culture

Human embryonic kidney HEK293T cells (female), human lung A549 cells (male), and African green monkey kidney epithelial Vero E6 cells (female) were maintained at 37ºC and 5% CO_2 in DMEM containing 10% FBS. To generate HEK293T and A549 cells overexpressing human ACE2,

TMPRSS2, and CIITA, we transduced these cells with lentiviral vectors pLOC_hACE2_PuroR, pLOC_hTMPRSS2_BlastR, and pTRIP-SFFV-Hygro-2A-myc-CIITA, and selected for the triple-transduced cells in culture medium supplemented with 1 µg/ml each of puromycin and blasticidin and 320 µg/ml of hygromycin. A375 cells were obtained from ATCC (ATCC ® CRL-1619). A375 cells were grown in ATCC-formulated Dulbecco's Modified Eagle's Medium (Catalog No. 30-2002) with fetal bovine serum to a final concentration of 10% using ATCC guidelines. A375 cells were harvested by trypsinization (Trypsin-EDTA 0.25%, Gibco™ 25200056), pelleted and rinsed in PBS twice. Pellets were snap frozen and stored at −80 °C.

Flow cytometry

1.5*10^6 A549 or A375 cells were incubated with 2ul FITC anti HLA-DR, DP, DQ antibody (BD Pharmingen #562008) in 100ul PBS at 4ºC for 45 minutes. Cells were washed three times with PBS, resuspended in 400ul PBS and analyzed using a CytoFLEX flow cytometer.

SARS-CoV-2 virus stock preparation and titration

The SARS-CoV-2 USA-WA1/2020 isolate (NCBI accession number: MT246667) was deposited by the Centers for Disease Control and Prevention and obtained through BEI Resources, NIAID, NIH (NR-52281). We then passaged the virus twice onto Vero E6 cells to obtain the P2 stock, as previously described (Chen et al., JVI). The virus titration was performed on Vero E6 cells. All experiments in this study utilized the P2 stock.

Quantification of virus infectivity using immunofluorescence

A549 and 293T cells stably overexpressing ACE2, TMPRSS2, and CIITA were infected with SARS-CoV-2 at an MOI of 0.5, 1, or 3 for 12, 18, 24, 36, or 48 hours. At indicated times, the culture medium was removed, and the cells were fixed with 4% paraformaldehyde for 60 minutes at room temperature. The cells were then permeabilized with 0.1% of Triton X-100 in PBS for 10 minutes and hybridized with anti-SARS-CoV nucleocapsid (rabbit polyclonal) antibody (1:2000, Rockland, #200-401-A50) at 4ºC overnight. AlexaFluor 568 goat anti-rabbit antibody (Invitrogen, #A11011) was used as the secondary antibody. Finally, DAPI was used to stain cell nuclei. Images were captured with an EVOS microscope using a 10x lens, and the percentage of infected cells was calculated with ImageJ.

Immunoprecipitation of HLA-II complexes from cells

Cells from 3x15cm dishes were scraped into 2.5ml/dish of cold lysis buffer (20mM Tris, pH 8.0, 100mM NaCl, 6mM MgCl2, 1mM EDTA, 60mM Octyl β-d-glucopyranoside, 0.2mM Iodoacetamide, 1.5% Triton X-100, 50xC0mplete Protease Inhibitor Tablet-EDTA free and PMSF) obtaining a total of ~9 mL lysate. This lysate was split into 6 eppendorf tubes, with each tube receiving 1.5 mL volume, and incubated on ice for 15 min with 1ul of Benzonase (Thomas Scientific, E1014-25KU) to degrade nucleic acid. The lysates were then centrifuged at 4,000 rpm for 22 min at 4°C and the supernatants were transferred to another set of 6 eppendorf tubes

16

containing a mixture of pre-washed beads (Millipore Sigma, GE17-0886-01) and 12.5 uL (12.5 ug) of MHC class II antibodies in a 3:1:1 mixture of TAL-1B5 (Abcam, ab20181), EPR11226 (Abcam, ab157210) and B-K27 (Abcam, ab47342). The immune complexes were captured on the beads by incubating on a rotor at 4°C for 3hr in the BSL3 lab. Virus inactivation was confirmed before subsequent samples processing outside the BSL3 using plaque assay (Weingarten-Gabbay et al., 2021, 2022). In total, nine washing steps were performed; one wash with 1mL of cold lysis wash buffer (20mM Tris, pH 8.0, 100mM NaCl, 6mM MgCl2, 1mM EDTA, 60mM Octyl β-d-glucopyranoside, 0.2mM Iodoacetamide, 1.5% Triton X-100), four washes with 1mL of cold complete wash buffer (20mM Tris, pH 8.0, 100mM NaCl, 1mM EDTA, 60mM Octyl β-d-glucopyranoside, 0.2mM Iodoacetamide), and four washes with 20mM Tris pH 8.0 buffer. Dry beads were stored at −80°C until mass-spectrometry analysis was performed.

HLA-II peptidome desalting and LC-MS/MS data generation

HLA peptides were eluted and desalted from beads as follows: wells of the tC18 40mg Sep-Pak desalting plate (Waters, Milford, MA) were activated with 2x 1 mL of methanol (MeOH) and 500 µL of 99.9% acetonitrile (ACN)/0.1% formic acid (FA), then washed with 4x 1 mL of 1% FA. A 10µmPE fritted filter plate containing the HLA-IP beads was placed on top of the Sep-Pak plate. To dissociate peptides from HLA molecules and facilitate peptides binding to the tC18 solid phase, 200 µL of 3% ACN/5% FA was added to the beads in the filter plate. 100 fmol internal retention time (iRT) standards (Biognosys SKU: Ki-3002-2) was spiked into each sample as a loading control and pushed through both the filter plate and 40 mg Sep-Pak plate. Following sample loading there was one wash with 400 µL of 1% FA. Beads were then incubated with 500 µL of 10% acetic acid (AcOH) three times for 5 min to further dissociate bound peptides from the HLA molecules. The beads were rinsed once with 1 mL 1% FA and the filter plate was removed. The Sep-Pak desalt plate was rinsed with 1 mL 1% FA an additional three times. The peptides were eluted from the Sep-Pak desalt plate using 250 µL of 15% ACN/1% FA and 2x 250 µL of 50% ACN/1% FA. HLA peptides were eluted into 1.5 mL micro tubes (Sarstedt, Nümbrecht, Germany), frozen, and dried down via vacuum centrifugation. Dried peptides were stored at -80ºC until microscaled basic reverse phase separation.

Briefly, peptides were loaded on Stage-tips with 2 punches of SDB-XC material (Empore 3M). HLA-II peptides were eluted in three fractions with increasing concentrations of ACN (5%, 15%, and 40% in 0.1% NH_4OH, pH 10). Peptides were reconstituted in 3% ACN/5% FA prior to loading onto an analytical column (35 cm, 1.9 µm C18 (Dr. Maisch HPLC GmbH), packed in-house PicoFrit 75 µm inner diameter, 10 µm emitter (New Objective)). Peptides were eluted with a linear gradient (EasyNanoLC 1200, Thermo Fisher Scientific) ranging from 6–30% Solvent B (0.1% FA in 90% ACN) over 84 min, 30–90% B over 9 min and held at 90% B for 5 min at 200 nl/min. MS/MS data were acquired on a Thermo Orbitrap Exploris 480 equipped with (HLA-I) and without (HLA-II) FAIMS (Thermo Fisher Scientific) in data-dependent acquisition. FAIMS compensation voltages (CVs) were set to -50 and -70 with a cycle time of 1.5 s per FAIMS experiment. MS2 fill time was set to 100 ms; collision energy was 30, 34 or 36 CE see **Table S2** for file mappings and LC-MS/MS acquisition variations.

Proteome analysis from HLA enrichment flow-through

Flow-throughs of the HLA-II IP that were stored as flash-frozen native protein lysates were briefly thawed on ice for ~15 min. Once thawed, 10% SDS was added for a final concentration of 5% SDS to denature the lysate, resulting in a final volume of ~1.5 mL lysate which was prepared for S-Trap digestion (Abelin et al., 2023).

Protein concentration was estimated using a BCA assay for scaling of digestion enzymes. Disulfide bonds were reduced in 5 mM DTT for 30 min at 25°C and 1000 rpm shaking and cysteine residues were alkylated in 10 mM IAA in the dark for 45 min at 25°C and 1000 rpm shaking. Lysates were then transferred to a 15 mL conical tube to prepare for protein precipitation. 27% phosphoric acid was added at a 1:10 ratio of lysate volume to acidify and proteins were precipitated with 6x sample volume of ice cold S-Trap buffer (90% methanol, 100 mM TEAB). The precipitate was transferred in successive loads of 3 mL to a S-Trap Midi (Protifi) and loaded with 1 min centrifugation at 4000 x g, mixing the remaining precipitate thoroughly between transfers. The precipitated proteins were washed 4x with 3 mL S-Trap buffer at 4000 x g for 1 min. To digest the deposited protein material, 350 µL digestion buffer (50 mM TEAB) containing both trypsin and endopeptidase C (LysC), each at 1:50 enzyme:substrate, was passed through each S-Trap column with 1 min centrifugation at 4000 x g. The digestion buffer was then added back atop the S-Trap and the cartridges were left capped overnight at 25°C.

Peptide digests were eluted from the S-Trap, first with 500 µL 50 mM TEAB and next with 500 µL 0.1% FA, each for 30 sec at 1000 x g. The final elution of 500 µL 50% ACN/0.1% FA was centrifuged for 1 min at 4000 x g to clear the cartridge. Peptide concentration of the pooled elutions was estimated with a BCA assay, and 10 µg peptide was used for stagetip fractionation. Each 25ug proteome sample was reconstituted in 4.5 mM ammonium formate (pH 10) in 2% (vol/vol) acetonitrile and separated into four fractions using basic reversed phase fractionation on a C-18 Stage-tip. Fractions were eluted at 5%, 12.5%, 15%, and 50% ACN/4.5 mM ammonium formate buffer (pH 10) and dried. Fractions were reconstituted in 3%ACN/5%FA, and 1 ug was used for LC-MS/MS analysis.

Data-dependent acquisition was performed using Orbitrap Exploris 480 V2.0 software in positive ion mode at a spray voltage of 1.8 kV. MS1 spectra were measured with a resolution of 60,000, a normalized AGC target of 300% for, a maximum injection time of 10 ms, and a mass range from 350 to 1800 m/z. The data-dependent mode cycle was set to trigger MS/MS on up to the top 20 most abundant precursors per cycle at an MS2 resolution of 45,000, an AGC target of 30%, an isolation window of 0.7 m/z, a maximum injection time of 105 ms for proteome, and an HCD collision energy of 34%. Peptides that triggered MS/MS scans were dynamically excluded from further MS/MS scans for 20 s in proteome/phosphoproteome/ubiquitylome and for 30 s in acetylome, with a ±10 ppm mass tolerance. Theoretical precursor envelope fit filter was enabled with a fit threshold of 50% and window of 1.2 m/z. Monoisotopic peak determination was set to peptide and charge state screening was enabled to only include precursor charge states 2–6 with an intensity threshold of 5.0e3. Advanced peak determination (APD) was enabled. "Perform dependent scan on single charge state per precursor only" was disabled.

LC-MS/MS data interpretation

Peptide sequences were interpreted from MS/MS spectra using Spectrum Mill (SM) v 7.08 (proteomics.broadinstitute.org) against a RefSeq-based sequence database containing 41,457 proteins mapped to the human reference genome (hg38) obtained via the UCSC Table Browser (https://genome.ucsc.edu/cgi-bin/hgTables) on June 29, 2018, with the addition of 13 proteins encoded in the human mitochondrial genome, 264 common laboratory contaminant proteins, 553 human non-canonical small open reading frames, 28 SARS-CoV2 proteins obtained from RefSeq derived from the original Wuhan-Hu-1 China isolate NC_045512.2 (https://www.ncbi.nlm.nih.gov/nuccore/1798174254;(Wu et al., 2020)), and 23 novel unannotated virus ORFs whose translation is supported by Ribo-seq(Finkel et al., 2020) for a total of 42,337 proteins. Among the 28 annotated SARS-CoV2 proteins we opted to omit the full-length polyproteins ORF1a and ORF1ab, to simplify peptide-to-protein assignment, and instead represented ORF1ab as the mature 16 individual non-structural proteins that result from proteolytic processing of the 1a and 1ab polyproteins. We added the D614G variant of the SARS-Cov2 Spike protein that is commonly observed in European and American virus isolates, and also added 2036 entries from 6-frame translation of the SARS-Cov2 genome for all possible ORFs longer than 6 amino acids.

Parameters for the SM MS/MS search module for HLA-II immunopeptidomes included: no enzyme specificity; precursor and product mass tolerance of ±10 ppm; minimum matched peak intensity of 30%; ESI-QEXACTIVE-HCD-HLA-v3 scoring; fixed modification: carbamidomethylation of cysteine; variable modifications: cysteinylation of cysteine, oxidation of methionine, deamidation of asparagine, acetylation of protein N-termini, and pyroglutamic acid at peptide N-terminal glutamine; and precursor mass shift range of -18 to 81 Da. For tryptic proteomes, parameters included: ''trypsin allow P'' enzyme specificity with up to 4 missed cleavages, precursor and product mass tolerance of ± 20 ppm, and 30% minimum matched peak intensity (40% for acetylome). Scoring parameters were ESI-QEXACTIVE- HCD-v2. Allowed fixed modifications included carbamidomethylation of cysteine and selenocysteine.Allowed variable modifications for whole proteome datasets were acetylation of protein N-termini, oxidized methionine, deamidation of asparagine, hydroxylation of proline in PG motifs, pyro-glutamic acid at peptide N-terminal glutamine, and pyro-carbamidomethylation at peptide N-terminal cysteine with a precursor MH+ shift range of -18 to 97 Da.

ACKNOWLEDGEMENT

AUTHORS CONTRIBUTION

S.W.-G. and J.G.A. conceptualized the study. S.W.-G., J.G.A., and M.S. designed the experiments. D.-Y.C., H.B.T., M.G., L.R.P., M.R.B. performed experiments. S.W.-G., S.S., H.B.T., K.R.C., and J.G.A. performed data analysis. N.H., S.A.C., M.S., and P.C.S. supervised the work. S.W.-G., J.G.A., M.S., and P.C.S. wrote the manuscript with contributions from all the authors.

DECLARATION OF INTEREST

S.W.-G., D-Y.C, S.S., K.R.C, N.H., S.A.C., J.G.A., M.S., and P.C.S. are named co-inventors on a patent application related to this work filed by The Broad Institute that is being made available in accordance with the COVID-19 technology licensing framework to maximize access to university innovations. N.H. is a founder of Neon Therapeutics, Inc. (now BioNTech US), was a member of its scientific advisory board, and holds shares. N.H. is also an advisor for IFM Therapeutics. S.A.C is a member of the scientific advisory boards of Kymera, PTM BioLabs, Seer and PrognomIQ. J.G.A. is a past employee of Neon Therapeutics, Inc. (now BioNTech US). P.C.S. is a co-founder of and consultant to Sherlock Biosciences and Delve Biosciences and a board member of Danaher Corporation and holds equity in the companies. The remaining authors declare no competing interests.
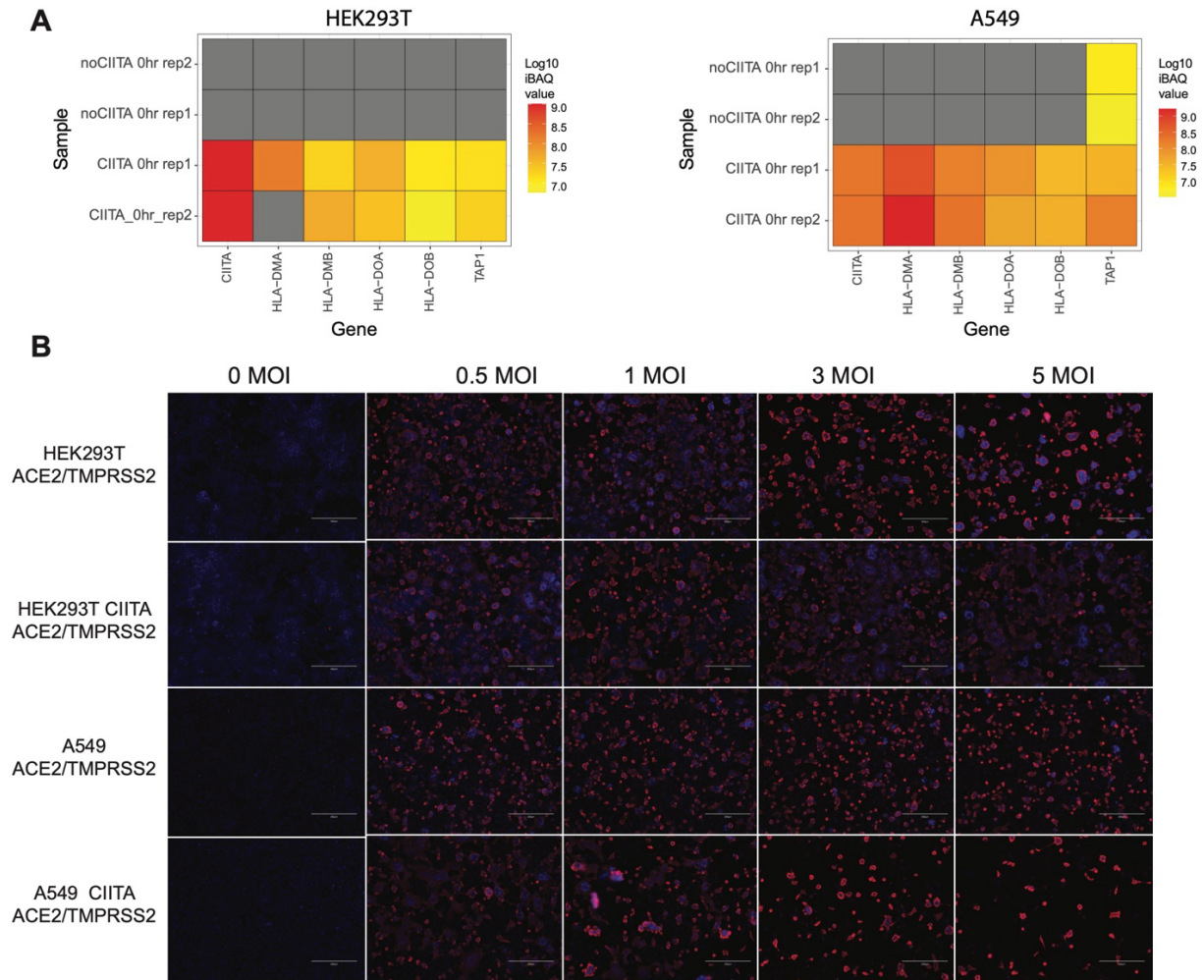
SUPPPLEMENTAL FIGURES

Figure S1. Proper induction of the MHC-II locus and SARS-CoV-2 infectivity in CIITA overexpressing cells. (A) Heatmap of log10 iBAQ values for key proteins in the MHC-II locus, observed in whole proteome datasets of A549/AT and HEK293T/AT cells with and without CIITA transduction. Shown are two biological replicates for each condition. **(B)** SARS-CoV-2 infectivity assay comparing A549/AT and HEK293T/AT with and without CIITA transduction. Representative immunofluorescence images at 24 hpi. Red, nucleocapsid; Blue, DAPI. Images were captured with an EVOS microscope using a 10x objective lens.
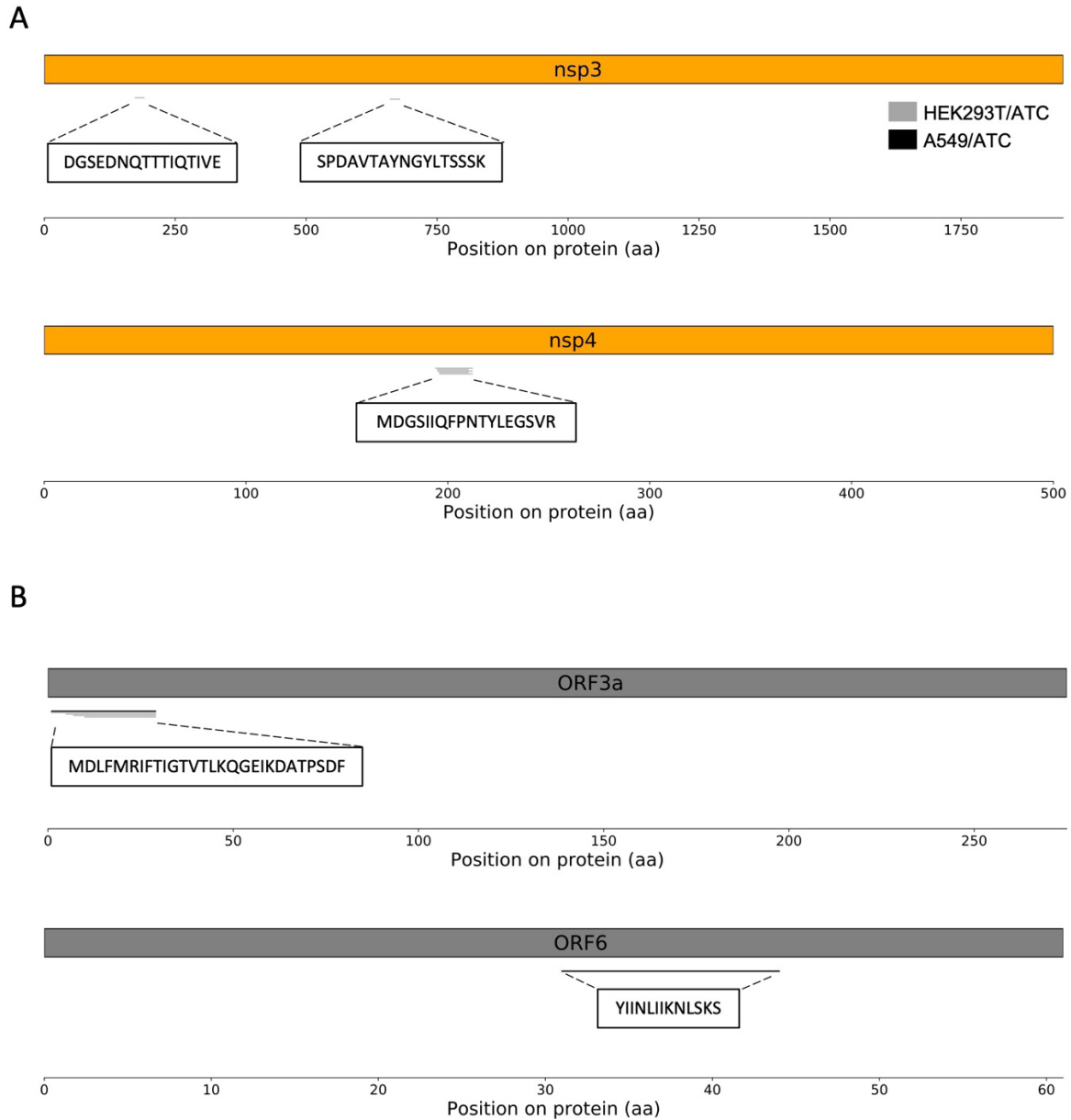
Figure S2. HLA-II peptides in SARS-CoV-2 non-structural and accessory proteins. (A) The location of HLA-II peptides across the non-structural proteins nsp3 and nsp4. **(B)** The location of HLA-II peptides across the accessory proteins ORF3a and ORF6. Peptides detected in A549/ATC and HEK293T/ATC cells are depicted in black and gray, respectively.

REFERENCES

Abelin, J. G., Bergstrom, E. J., Rivera, K. D., Taylor, H. B., Klaeger, S., Xu, C., Verzani, E. K., Jackson White, C., Woldemichael, H. B., Virshup, M., Olive, M. E., Maynard, M., Vartany, S. A., Allen, J. D., Phulphagar, K., Harry Kane, M., Rachimi, S., Mani, D. R., Gillette, M. A., … Carr, S. A. (2023). Workflow enabling deepscale immunopeptidome, proteome, ubiquitylome, phosphoproteome, and acetylome analyses of sample-limited tissues. *Nature Communications*, *14*(1), 1851.

Abelin, J. G., Harjanto, D., Malloy, M., Suri, P., Colson, T., Goulding, S. P., Creech, A. L., Serrano, L. R., Nasir, G., Nasrullah, Y., McGann, C. D., Velez, D., Ting, Y. S., Poran, A., Rothenberg, D. A., Chhangawala, S., Rubinsteyn, A., Hammerbacher, J., Gaynor, R. B., … Rooney, M. S. (2019). Defining HLA-II Ligand Processing and Binding Rules with Mass Spectrometry Enhances Cancer Epitope Prediction. *Immunity*, *51*(4), 766–779.e17.

Abelin, J. G., Keskin, D. B., Sarkizova, S., Hartigan, C. R., Zhang, W., Sidney, J., Stevens, J., Lane, W., Zhang, G. L., Eisenhaure, T. M., Clauser, K. R., Hacohen, N., Rooney, M. S., Carr, S. A., & Wu, C. J. (2017). Mass Spectrometry Profiling of HLA-Associated Peptidomes in Mono-allelic Cells Enables More Accurate Epitope Prediction. *Immunity*, *46*(2), 315–326.

Andreatta, M., Alvarez, B., & Nielsen, M. (2017). GibbsCluster: unsupervised clustering and alignment of peptide sequences. *Nucleic Acids Research*, *45*(W1), W458–W463.

Arieta, C. M., Xie, Y. J., Rothenberg, D. A., Diao, H., Harjanto, D., Meda, S., Marquart, K., Koenitzer, B., Sciuto, T. E., Lobo, A., & Others. (2023). The T-cell-directed vaccine BNT162b4 encoding conserved non-spike antigens protects animals from severe SARS-CoV-2 infection. *Cell*. https://www.cell.com/cell/pdf/S0092-8674(23)00403-8.pdf

Baden, L. R., El Sahly, H. M., Essink, B., Kotloff, K., Frey, S., Novak, R., Diemert, D., Spector, S. A., Rouphael, N., Creech, C. B., McGettigan, J., Khetan, S., Segall, N., Solis, J., Brosz, A., Fierro, C., Schwartz, H., Neuzil, K., Corey, L., … COVE Study Group. (2021). Efficacy

and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *The New England Journal of Medicine*, *384*(5), 403–416.

Barouch, D. H. (2022). Covid-19 Vaccines — Immunity, Variants, Boosters. In *New England Journal of Medicine* (Vol. 387, Issue 11, pp. 1011–1020). https://doi.org/10.1056/nejmra2206573

Bassani-Sternberg, M., & Gfeller, D. (2016). Unsupervised HLA Peptidome Deconvolution Improves Ligand Prediction Accuracy and Predicts Cooperative Effects in Peptide-HLA Interactions. *Journal of Immunology* , *197*(6), 2492–2499.

Becerra-Artiles, A., Cruz, J., Leszyk, J. D., Sidney, J., Sette, A., Shaffer, S. A., & Stern, L. J. (2019). Naturally processed HLA-DR3-restricted HHV-6B peptides are recognized broadly with polyfunctional and cytotoxic CD4 T-cell responses. *European Journal of Immunology*, *49*(8), 1167–1185.

Becerra-Artiles, A., Nanaware, P. P., Muneeruddin, K., Weaver, G. C., Shaffer, S. A., Calvo-Calle, J. M., & Stern, L. J. (2022). Immunopeptidome profiling of human coronavirus OC43-infected cells identifies CD4 T cell epitopes specific to seasonal coronaviruses or cross-reactive with SARS-CoV-2. *bioRxiv : The Preprint Server for Biology*. https://doi.org/10.1101/2022.12.01.518643

Cagliani, R., Forni, D., Clerici, M., & Sironi, M. (2020). Coding potential and sequence conservation of SARS-CoV-2 and related animal viruses. *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, *83*, 104353.

Cassotta, A., Paparoditis, P., Geiger, R., Mettu, R. R., Landry, S. J., Donati, A., Benevento, M., Foglierini, M., Lewis, D. J. M., Lanzavecchia, A., & Sallusto, F. (2020). Deciphering and predicting CD4+ T cell immunodominance of influenza virus hemagglutinin. *The Journal of Experimental Medicine*, *217*(10). https://doi.org/10.1084/jem.20200206

Chen, Z., Ruan, P., Wang, L., Nie, X., Ma, X., & Tan, Y. (2021). T and B cell Epitope analysis of

SARS-CoV-2 S protein based on immunoinformatics and experimental research. *Journal of Cellular and Molecular Medicine*, *25*(2), 1274–1289.

Chong, C., Marino, F., Pak, H., Racle, J., Daniel, R. T., Müller, M., Gfeller, D., Coukos, G., & Bassani-Sternberg, M. (2018). High-throughput and Sensitive Immunopeptidomics Platform Reveals Profound Interferonγ-Mediated Remodeling of the Human Leukocyte Antigen (HLA) Ligandome. *Molecular & Cellular Proteomics: MCP*, *17*(3), 533–548.

Creech, C. B., Walker, S. C., & Samuels, R. J. (2021). SARS-CoV-2 Vaccines. *JAMA: The Journal of the American Medical Association*, *325*(13), 1318–1320.

Croft, N. P., Smith, S. A., Wong, Y. C., Tan, C. T., Dudek, N. L., Flesch, I. E. A., Lin, L. C. W., Tscharke, D. C., & Purcell, A. W. (2013). Kinetics of antigen expression and epitope presentation during virus infection. *PLoS Pathogens*, *9*(1), e1003129.

Dai, L., & Gao, G. F. (2021). Viral targets for vaccines against COVID-19. *Nature Reviews. Immunology*, *21*(2), 73–82.

Deffrennes, V., Vedrenne, J., Stolzenberg, M. C., Piskurich, J., Barbieri, G., Ting, J. P., Charron, D., & Alcaïde-Loridan, C. (2001). Constitutive expression of MHC class II genes in melanoma cell lines results from the transcription of class II transactivator abnormally initiated from its B cell-specific promoter. *Journal of Immunology* , *167*(1), 98–106.

Dutta Noton K., Mazumdar Kaushiki, & Gordy James T. (2020). The Nucleocapsid Protein of SARS–CoV-2: a Target for Vaccine Development. *Journal of Virology*, *94*(13), e00647–20.

Ferretti, A. P., Kula, T., Wang, Y., Nguyen, D. M. V., Weinheimer, A., Dunlap, G. S., Xu, Q., Nabilsi, N., Perullo, C. R., Cristofaro, A. W., Whitton, H. J., Virbasius, A., Olivier, K. J., Jr, Buckner, L. R., Alistar, A. T., Whitman, E. D., Bertino, S. A., Chattopadhyay, S., & MacBeath, G. (2020). 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*, *53*(5), 1095–1107.e3.

Finkel, Y., Mizrahi, O., & Nachshon, A. (2020). The coding capacity of SARS-CoV-2. *bioRxiv*.

https://www.biorxiv.org/content/10.1101/2020.05.07.082909v1.abstract

Firth, A. E. (2020). A putative new SARS-CoV protein, 3c, encoded in an ORF overlapping ORF3a. *The Journal of General Virology*, *101*(10), 1085–1089.

Forlani, G., Michaux, J., Pak, H., Huber, F., Marie Joseph, E. L., Ramia, E., Stevenson, B. J., Linnebacher, M., Accolla, R. S., & Bassani-Sternberg, M. (2021). CIITA-Transduced Glioblastoma Cells Uncover a Rich Repertoire of Clinically Relevant Tumor-Associated HLA-II Antigens. *Molecular & Cellular Proteomics: MCP*, *20*, 100032.

Gangaev, A., Ketelaars, S. L. C., Isaeva, O. I., Patiwael, S., Dopler, A., Hoefakker, K., De Biasi, S., Gibellini, L., Mussini, C., Guaraldi, G., Girardis, M., Ormeno, C. M. P. T., Hekking, P. J. M., Lardy, N. M., Toebes, M., Balderas, R., Schumacher, T. N., Ovaa, H., Cossarizza, A., & Kvistborg, P. (2021). Identification and characterization of a SARS-CoV-2 specific CD8+ T cell response with immunodominant features. *Nature Communications*, *12*(1), 2593.

Gentili, M., Liu, B., Papanastasiou, M., Dele-Oni, D., Schwartz, M. A., Carlson, R. J., Al'Khafaji, A. M., Krug, K., Brown, A., Doench, J. G., Carr, S. A., & Hacohen, N. (2023). ESCRT-dependent STING degradation inhibits steady-state and cGAMP-induced signalling. *Nature Communications*, *14*(1), 611.

Grifoni, A., Sidney, J., Vita, R., Peters, B., Crotty, S., Weiskopf, D., & Sette, A. (2021). SARS-CoV-2 human T cell epitopes: Adaptive immune response against COVID-19. *Cell Host & Microbe*, *29*(7), 1076–1092.

Hajnik, R. L., Plante, J. A., Liang, Y., Alameh, M.-G., Tang, J., Bonam, S. R., Zhong, C., Adam, A., Scharton, D., Rafael, G. H., Liu, Y., Hazell, N. C., Sun, J., Soong, L., Shi, P.-Y., Wang, T., Walker, D. H., Sun, J., Weissman, D., … Hu, H. (2022). Dual spike and nucleocapsid mRNA vaccination confer protection against SARS-CoV-2 Omicron and Delta variants in preclinical models. *Science Translational Medicine*, *14*(662), eabq1945.

Hos, B. J., Tondini, E., Camps, M. G. M., Rademaker, W., van den Bulk, J., Ruano, D., Janssen, G. M. C., de Ru, A. H., van den Elsen, P. J., de Miranda, N. F. C. C., van Veelen,

P. A., & Ossendorp, F. (2022). Cancer-specific T helper shared and neo-epitopes uncovered by expression of the MHC class II master regulator CIITA. *Cell Reports*, *41*(2), 111485.

Joag, V., Wijeyesinghe, S., Stolley, J. M., Quarnstrom, C. F., Dileepan, T., Soerens, A. G., Sangala, J. A., O'Flanagan, S. D., Gavil, N. V., Hong, S.-W., Bhela, S., Gangadhara, S., Weyu, E., Matchett, W. E., Thiede, J., Krishna, V., Cheeran, M. C.-J., Bold, T. D., Amara, R., … Masopust, D. (2021). Cutting Edge: Mouse SARS-CoV-2 Epitope Reveals Infection and Vaccine-Elicited CD8 T Cell Responses. *Journal of Immunology* , *206*(5), 931–935.

Jungreis, I., Sealfon, R., & Kellis, M. (2021). SARS-CoV-2 gene content and COVID-19 mutation impact by comparing 44 Sarbecovirus genomes. *Nature Communications*, *12*(1), 2642.

Kared, H., Redd, A. D., Bloch, E. M., Bonny, T. S., Sumatoh, H. R., Kairi, F., Carbajo, D., Abel, B., Newell, E. W., Bettinotti, M., Benner, S. E., Patel, E. U., Littlefield, K., Laeyendecker, O., Shoham, S., Sullivan, D., Casadevall, A., Pekosz, A., Nardin, A., … Quinn, T. C. (2021). SARS-CoV-2-specific CD8+ T cell responses in convalescent COVID-19 individuals. *The Journal of Clinical Investigation*. https://doi.org/10.1172/JCI145476

Keller, M. D., Harris, K. M., Jensen-Wachspress, M. A., Kankate, V. V., Lang, H., Lazarski, C. A., Durkee-Shock, J., Lee, P.-H., Chaudhry, K., Webber, K., Datar, A., Terpilowski, M., Reynolds, E. K., Stevenson, E. M., Val, S., Shancer, Z., Zhang, N., Ulrey, R., Ekanem, U., … Bollard, C. M. (2020). SARS-CoV-2–specific T cells are rapidly expanded for therapeutic use and target conserved regions of the membrane protein. *Blood*, *136*(25), 2905–2917.

Knierman, M. D., Lannan, M. B., Spindler, L. J., McMillian, C. L., Konrad, R. J., & Siegel, R. W. (2020). The Human Leukocyte Antigen Class II Immunopeptidome of the SARS-CoV-2 Spike Glycoprotein. *Cell Reports*, *33*(9), 108454.

Krammer, F. (2020). SARS-CoV-2 vaccines in development. *Nature*, *586*(7830), 516–527.

Kyriakidis, N. C., López-Cortés, A., González, E. V., Grimaldos, A. B., & Prado, E. O. (2021).

SARS-CoV-2 vaccines strategies: a comprehensive review of phase 3 candidates. *NPJ Vaccines*, *6*(1), 28.

Le Bert, N., Clapham, H. E., Tan, A. T., Chia, W. N., Tham, C. Y. L., Lim, J. M., Kunasegaran, K., Tan, L. W. L., Dutertre, C.-A., Shankar, N., Lim, J. M. E., Sun, L. J., Zahari, M., Tun, Z. M., Kumar, V., Lim, B. L., Lim, S. H., Chia, A., Tan, Y.-J., … Tam, C. C. (2021). Highly functional virus-specific cellular immune response in asymptomatic SARS-CoV-2 infection. *The Journal of Experimental Medicine*, *218*(5). https://doi.org/10.1084/jem.20202617

Le Bert, N., Tan, A. T., Kunasegaran, K., Tham, C. Y. L., Hafezi, M., Chia, A., Chng, M. H. Y., Lin, M., Tan, N., Linster, M., Chia, W. N., Chen, M. I.-C., Wang, L.-F., Ooi, E. E., Kalimuddin, S., Tambyah, P. A., Low, J. G.-H., Tan, Y.-J., & Bertoletti, A. (2020). SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature*. https://doi.org/10.1038/s41586-020-2550-z

Lee, E., Sandgren, K., Duette, G., Stylianou, V. V., Khanna, R., Eden, J.-S., Blyth, E., Gottlieb, D., Cunningham, A. L., & Palmer, S. (2021). Identification of SARS-CoV-2 Nucleocapsid and Spike T-Cell Epitopes for Assessing T-Cell Immunity. *Journal of Virology*, *95*(6). https://doi.org/10.1128/JVI.02002-20

Lippolis, J. D., White, F. M., Marto, J. A., Luckey, C. J., Bullock, T. N. J., Shabanowitz, J., Hunt, D. F., & Engelhard, V. H. (2002). Analysis of MHC class II antigen processing by quantitation of peptides that constitute nested sets. *Journal of Immunology* , *169*(9), 5089–5097.

Mahajan, S., Kode, V., Bhojak, K., Karunakaran, C., Lee, K., Manoharan, M., Ramesh, A., Hv, S., Srivastava, A., Sathian, R., Khan, T., Kumar, P., Gupta, R., Chakraborty, P., & Chaudhuri, A. (2021). Immunodominant T-cell epitopes from the SARS-CoV-2 spike antigen reveal robust pre-existing T-cell immunity in unexposed individuals. *Scientific Reports*, *11*(1), 13164.

Martínez-Flores, D., Zepeda-Cervantes, J., Cruz-Reséndiz, A., Aguirre-Sampieri, S., Sampieri,

A., & Vaca, L. (2021). SARS-CoV-2 Vaccines Based on the Spike Glycoprotein and Implications of New Viral Variants. *Frontiers in Immunology*, *12*. https://doi.org/10.3389/fimmu.2021.701501

Matchett, W. E., Joag, V., Stolley, J. M., Shepherd, F. K., Quarnstrom, C. F., Mickelson, C. K., Wijeyesinghe, S., Soerens, A. G., Becker, S., Thiede, J. M., Weyu, E., O'Flanagan, S. D., Walter, J. A., Vu, M. N., Menachery, V. D., Bold, T. D., Vezys, V., Jenkins, M. K., Langlois, R. A., & Masopust, D. (2021). Cutting Edge: Nucleocapsid Vaccine Elicits Spike-Independent SARS-CoV-2 Protective Immunity. *Journal of Immunology* , *207*(2), 376–379.

Mateus, J., Grifoni, A., Tarke, A., Sidney, J., Ramirez, S. I., Dan, J. M., Burger, Z. C., Rawlings, S. A., Smith, D. M., Phillips, E., Mallal, S., Lammers, M., Rubiro, P., Quiambao, L., Sutherland, A., Yu, E. D., da Silva Antunes, R., Greenbaum, J., Frazier, A., … Weiskopf, D. (2020). Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science*, *370*(6512), 89–94.

McMurtrey, C. P., Lelic, A., Piazza, P., Chakrabarti, A. K., Yablonsky, E. J., Wahl, A., Bardet, W., Eckerd, A., Cook, R. L., Hess, R., Buchli, R., Loeb, M., Rinaldo, C. R., Bramson, J., & Hildebrand, W. H. (2008). Epitope discovery in West Nile virus infection: Identification and immune recognition of viral epitopes. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(8), 2981–2986.

Nagler, A., Kalaora, S., Barbolin, C., Gangaev, A., Ketelaars, S. L. C., Alon, M., Pai, J., Benedek, G., Yahalom-Ronen, Y., Erez, N., Greenberg, P., Yagel, G., Peri, A., Levin, Y., Satpathy, A. T., Bar-Haim, E., Paran, N., Kvistborg, P., & Samuels, Y. (2021). Identification of presented SARS-CoV-2 HLA class I and HLA class II peptides using HLA peptidomics. *Cell Reports*, *35*(13), 109305.

Neale, I., Ali, M., Kronsteiner, B., Longet, S., Abraham, P., Deeks, A. S., Brown, A., Moore, S. C., Stafford, L., Dobson, S. L., Plowright, M., Newman, T. A. H., Wu, M. Y., Carr, E. J., Beale, R., Otter, A. D., Hopkins, S., Hall, V., Tomic, A., … Crick COVID Immunity Pipeline.

(2023). CD4+ and CD8+ T cell and antibody correlates of protection against Delta vaccine breakthrough infection: A nested case-control study within the PITCH study. In *medRxiv*. https://doi.org/10.1101/2023.02.16.23285748

Nelde, A., Bilich, T., Heitmann, J. S., Maringer, Y., Salih, H. R., Roerden, M., Lübke, M., Bauer, J., Rieth, J., Wacker, M., Peter, A., Hörber, S., Traenkle, B., Kaiser, P. D., Rothbauer, U., Becker, M., Junker, D., Krause, G., Strengert, M., … Walz, J. S. (2021). SARS-CoV-2-derived peptides define heterologous and COVID-19-induced T cell recognition. *Nature Immunology*, *22*(1), 74–85.

Neuwelt, A. J., Kimball, A. K., Johnson, A. M., Arnold, B. W., Bullock, B. L., Kaspar, R. E., Kleczko, E. K., Kwak, J. W., Wu, M.-H., Heasley, L. E., Doebele, R. C., Li, H. Y., Nemenoff, R. A., & Clambey, E. T. (2020). Cancer cell-intrinsic expression of MHC II in lung cancer cell lines is actively restricted by MEK/ERK signaling and epigenetic mechanisms. *Journal for Immunotherapy of Cancer*, *8*(1). https://doi.org/10.1136/jitc-2019-000441

Nielsen, S. S., Vibholm, L. K., Monrad, I., Olesen, R., Frattari, G. S., Pahus, M. H., Højen, J. F., Gunst, J. D., Erikstrup, C., Holleufer, A., Hartmann, R., Østergaard, L., Søgaard, O. S., Schleimann, M. H., & Tolstrup, M. (2021). SARS-CoV-2 elicits robust adaptive immune responses regardless of disease severity. *EBioMedicine*, *68*, 103410.

Oronsky, B., Larson, C., Caroen, S., Hedjran, F., Sanchez, A., Prokopenko, E., & Reid, T. (2022). Nucleocapsid as a next-generation COVID-19 vaccine candidate. *International Journal of Infectious Diseases: IJID: Official Publication of the International Society for Infectious Diseases*, *122*, 529–530.

Ovsyannikova, I. G., Johnson, K. L., Naylor, S., Muddiman, D. C., & Poland, G. A. (2003). Naturally processed measles virus peptide eluted from class II HLA-DRB1*03 recognized by T lymphocytes from human blood. *Virology*, *312*(2), 495–506.

Pardi, N., Hogan, M. J., Porter, F. W., & Weissman, D. (2018). mRNA vaccines — a new era in vaccinology. In *Nature Reviews Drug Discovery* (Vol. 17, Issue 4, pp. 261–279).

https://doi.org/10.1038/nrd.2017.243

Parker, R., Partridge, T., Wormald, C., Kawahara, R., Stalls, V., Aggelakopoulou, M., Parker, J., Powell Doherty, R., Ariosa Morejon, Y., Lee, E., Saunders, K., Haynes, B. F., Acharya, P., Thaysen-Andersen, M., Borrow, P., & Ternette, N. (2021). Mapping the SARS-CoV-2 spike glycoprotein-derived peptidome presented by HLA class II on dendritic cells. *Cell Reports*, *35*(8), 109179.

Peng, Y., Mentzer, A. J., Liu, G., Yao, X., Yin, Z., Dong, D., Dejnirattisai, W., Rostron, T., Supasa, P., Liu, C., López-Camacho, C., Slon-Campos, J., Zhao, Y., Stuart, D. I., Paesen, G. C., Grimes, J. M., Antson, A. A., Bayfield, O. W., Hawkins, D. E. D. P., … Dong, T. (2020). Broad and strong memory CD4+ and CD8+ T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. *Nature Immunology*, *21*(11), 1336–1345.

Prakash, S., Srivastava, R., Coulon, P.-G., Dhanushkodi, N. R., Chentoufi, A. A., Tifrea, D. F., Edwards, R. A., Figueroa, C. J., Schubl, S. D., Hsieh, L., Buchmeier, M. J., Bouziane, M., Nesburn, A. B., Kuppermann, B. D., & BenMohamed, L. (2021). Genome-Wide B Cell, CD4+, and CD8+ T Cell Epitopes That Are Highly Conserved between Human and Animal Coronaviruses, Identified from SARS-CoV-2 as Targets for Preemptive Pan-Coronavirus Vaccines. *Journal of Immunology* , *206*(11), 2566–2582.

Prakash, S., Srivastava, R., Coulon, P.-G., Dhanushkodi, N. R., Chentoufi, A. A., Tifrea, D. F., Edwards, R. A., Figueroa, C., Schubl, S. D., Hsieh, L., Buchmeier, M. J., Bouziane, M., Nesburn, A. B., Kuppermann, B. D., & Benmohamed, L. (n.d.). Genome-Wide Asymptomatic B-Cell, CD4 and CD8 T-Cell Epitopes, that are Highly Conserved between Human and Animal Coronaviruses, Identified from SARS-CoV-2 as Immune Targets for Pre-Emptive Pan-Coronavirus Vaccines. In *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3712675

Rha, M.-S., Jeong, H. W., Ko, J.-H., Choi, S. J., Seo, I.-H., Lee, J. S., Sa, M., Kim, A. R., Joo, E.-J., Ahn, J. Y., Kim, J. H., Song, K.-H., Kim, E. S., Oh, D. H., Ahn, M. Y., Choi, H. K.,

Jeon, J. H., Choi, J.-P., Kim, H. B., … Shin, E.-C. (2021). PD-1-Expressing SARS-CoV-2-Specific CD8+ T Cells Are Not Exhausted, but Functional in Patients with COVID-19. *Immunity*, *54*(1), 44–52.e3.

Rucevic, M., Kourjian, G., Boucau, J., Blatnik, R., Garcia Bertran, W., Berberich, M. J., Walker, B. D., Riemer, A. B., & Le Gall, S. (2016). Analysis of Major Histocompatibility Complex-Bound HIV Peptides Identified from Various Cell Types Reveals Common Nested Peptides and Novel T Cell Responses. *Journal of Virology*, *90*(19), 8605–8620.

Ruiz Cuevas, M. V., Hardy, M.-P., Hollý, J., Bonneil, É., Durette, C., Courcelles, M., Lanoix, J., Côté, C., Staudt, L. M., Lemieux, S., Thibault, P., Perreault, C., & Yewdell, J. W. (2021). Most non-canonical proteins uniquely populate the proteome or immunopeptidome. *Cell Reports*, *34*(10), 108815.

Sahin, U., Muik, A., Vogler, I., Derhovanessian, E., Kranz, L. M., Vormehr, M., Quandt, J., Bidmon, N., Ulges, A., Baum, A., Pascal, K. E., Maurus, D., Brachtendorf, S., Lörks, V., Sikorski, J., Koch, P., Hilker, R., Becker, D., Eller, A.-K., … Türeci, Ö. (2021). BNT162b2 vaccine induces neutralizing antibodies and poly-specific T cells in humans. *Nature*, *595*(7868), 572–577.

Sarkizova, S., Klaeger, S., Le, P. M., Li, L. W., Oliveira, G., Keshishian, H., Hartigan, C. R., Zhang, W., Braun, D. A., Ligon, K. L., Bachireddy, P., Zervantonakis, I. K., Rosenbluth, J. M., Ouspenskaia, T., Law, T., Justesen, S., Stevens, J., Lane, W. J., Eisenhaure, T., … Keskin, D. B. (2020). A large peptidome dataset improves HLA class I epitope prediction across most of the human population. *Nature Biotechnology*, *38*(2), 199–209.

Schellens, I. M., Meiring, H. D., Hoof, I., Spijkers, S. N., Poelen, M. C. M., van Gaans-van den Brink, J. A. M., Costa, A. I., Vennema, H., Keşmir, C., van Baarle, D., & van Els, C. A. C. M. (2015). Measles Virus Epitope Presentation by HLA: Novel Insights into Epitope Selection, Dominance, and Microvariation. *Frontiers in Immunology*, *6*, 546.

Sette, A., & Crotty, S. (2021). Adaptive immunity to SARS-CoV-2 and COVID-19. *Cell*, *184*(4),

861–880.

Shomuradova, A. S., Vagida, M. S., Sheetikov, S. A., Zornikova, K. V., Kiryukhin, D., Titov, A., Peshkova, I. O., Khmelevskaya, A., Dianov, D. V., Malasheva, M., Shmelev, A., Serdyuk, Y., Bagaev, D. V., Pivnyuk, A., Shcherbinin, D. S., Maleeva, A. V., Shakirova, N. T., Pilunov, A., Malko, D. B., … Efimov, G. A. (2020). SARS-CoV-2 Epitopes Are Recognized by a Public and Diverse Repertoire of Human T Cell Receptors. *Immunity*, *53*(6), 1245–1257.e5.

Silva, E. K. V. B., Bomfim, C. G., Barbosa, A. P., Noda, P., Noronha, I. L., Fernandes, B. H. V., Machado, R. R. G., Durigon, E. L., Catanozi, S., Rodrigues, L. G., Pieroni, F., Lima, S. G., Teodoro, W. R., Queiroz, Z. A. J., Silveira, L. K. R., Charlie-Silva, I., Capelozzi, V. L., Guzzo, C. R., & Fanelli, C. (2022). Immunization with SARS-CoV-2 Nucleocapsid protein triggers a pulmonary immune response in rats. *PloS One*, *17*(5), e0268434.

Stewart, H., Lu, Y., O'Keefe, S., Valpadashi, A., Cruz-Zaragoza, L. D., Michel, H. A., Nguyen, S. K., Carnell, G. W., Lukhovitskaya, N., Milligan, R., Jungreis, I., Lulla, V., Davidson, A. D., Matthews, D. A., High, S., Rehling, P., Emmott, E., Heeney, J. L., Edgar, J. R., … Firth, A. E. (2022). The SARS-CoV-2 protein ORF3c is a mitochondrial modulator of innate immunity. In *bioRxiv* (p. 2022.11.15.516323). https://doi.org/10.1101/2022.11.15.516323

Tarke, A., Sidney, J., Kidd, C. K., Dan, J. M., Ramirez, S. I., Yu, E. D., Mateus, J., da Silva Antunes, R., Moore, E., Rubiro, P., Methot, N., Phillips, E., Mallal, S., Frazier, A., Rawlings, S. A., Greenbaum, J. A., Peters, B., Smith, D. M., Crotty, S., … Sette, A. (2021). Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases. *Cell Reports. Medicine*, *2*(2), 100204.

Taylor, H. B., Klaeger, S., Clauser, K. R., Sarkizova, S., Weingarten-Gabbay, S., Graham, D. B., Carr, S. A., & Abelin, J. G. (2021). MS-Based HLA-II Peptidomics Combined With Multiomics Will Aid the Development of Future Immunotherapies. *Molecular & Cellular Proteomics: MCP*, *20*, 100116.

Ternette, N., Yang, H., Partridge, T., Llano, A., Cedeño, S., Fischer, R., Charles, P. D., Dudek, N. L., Mothe, B., Crespo, M., Fischer, W. M., Korber, B. T. M., Nielsen, M., Borrow, P., Purcell, A. W., Brander, C., Dorrell, L., Kessler, B. M., & Hanke, T. (2016). Defining the HLA class I-associated viral antigen repertoire from HIV-1-infected human cells. *European Journal of Immunology*, *46*(1), 60–69.

Weingarten-Gabbay, S., Klaeger, S., Sarkizova, S., Pearlman, L. R., Chen, D.-Y., Gallagher, K. M. E., Bauer, M. R., Taylor, H. B., Dunn, W. A., Tarr, C., Sidney, J., Rachimi, S., Conway, H. L., Katsis, K., Wang, Y., Leistritz-Edwards, D., Durkin, M. R., Tomkins-Tinch, C. H., Finkel, Y., … Sabeti, P. C. (2021). Profiling SARS-CoV-2 HLA-I peptidome reveals T cell epitopes from out-of-frame ORFs. *Cell*, *184*(15), 3962–3980.e17.

Weingarten-Gabbay, S., Pearlman, L. R., Chen, D.-Y., Klaeger, S., Taylor, H. B., Welch, N. L., Keskin, D. B., Carr, S. A., Abelin, J. G., Saeed, M., & Sabeti, P. C. (2022). HLA-I immunopeptidome profiling of human cells infected with high-containment enveloped viruses. *STAR Protocols*, *3*(4), 101910.

Wherry, E. J., & Barouch, D. H. (2022). T cell immunity to COVID-19 vaccines. *Science*, *377*(6608), 821–822.

Wosen, J. E., Mukhopadhyay, D., Macaubas, C., & Mellins, E. D. (2018). Epithelial MHC Class II Expression and Its Role in Antigen Presentation in the Gastrointestinal and Respiratory Tracts. *Frontiers in Immunology*, *9*, 2144.

Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., Hu, Y., Tao, Z.-W., Tian, J.-H., Pei, Y.-Y., Yuan, M.-L., Zhang, Y.-L., Dai, F.-H., Liu, Y., Wang, Q.-M., Zheng, J.-J., Xu, L., Holmes, E. C., & Zhang, Y.-Z. (2020). A new coronavirus associated with human respiratory disease in China. *Nature*, *579*(7798), 265–269.