bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# Negligible impact of SARS-CoV-2 variants on CD4$^+$ and CD8$^+$ T cell reactivity in COVID-19 exposed donors and vaccinees.

## AUTHORS

Alison Tarke[1,2], John Sidney[1], Nils Methot[1], Yun Zhang[4], Jennifer M. Dan[1,3], Benjamin Goodwin[1], Paul Rubiro[1], Aaron Sutherland[1], Ricardo da Silva Antunes[1], April Frazier[1], Stephen A. Rawlings[3], Davey M. Smith[3], Bjoern Peters[,1,3], Richard H. Scheuermann[1,4,5], Daniela Weiskopf[1], Shane Crotty[1,3], Alba Grifoni[1*], Alessandro Sette[1,3*]

## AFFILIATIONS

[1]Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology (LJI), La Jolla, CA 92037, USA
[2]Department of Internal Medicine and Center of Excellence for Biomedical Research (CEBR), University of Genoa, Genoa, 16132, Italy.
[3] Department of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego (UCSD), La Jolla, CA 92037, USA
[4]J. Craig Venter Institute, La Jolla, CA 92037, USA.
[5]Department of Pathology, University of California, San Diego, CA 92093,

* indicates equal contributions

Correspondence: alex@lji.org (A.S.); agrifoni@lji.org (A.G.)
Lead Contact: alex@lji.org (A.S.)

## SUMMARY

The emergence of SARS-CoV-2 variants highlighted the need to better understand adaptive immune responses to this virus. It is important to address whether also CD4+ and CD8+ T cell responses are affected, because of the role they play in disease resolution and modulation of COVID-19 disease severity. Here we performed a comprehensive analysis of SARS-CoV-2-specific CD4+ and CD8+ T cell responses from COVID-19 convalescent subjects recognizing the ancestral strain, compared to variant lineages B.1.1.7, B.1.351, P.1, and CAL.20C as well as recipients of the Moderna (mRNA-1273) or Pfizer/BioNTech (BNT162b2) COVID-19 vaccines. Similarly, we demonstrate that the sequences of the vast majority of SARS-CoV-2 T cell epitopes are not affected by the mutations found in the variants analyzed. Overall, the results demonstrate that CD4+ and CD8+ T cell responses in convalescent COVID-19 subjects or COVID-19 mRNA vaccinees are not substantially affected by mutations found in the SARS-CoV-2 variants.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## INTRODUCTION

The emergence of several SARS-CoV-2 variants of concern (VOC) with multiple amino acid replacements has implications for the future control of the COVID-19 pandemic (Davies et al., 2020; Kirby, 2021; Tegally et al., 2020; Volz et al., 2021). Variants of concern include the UK (United Kingdom) variant 501Y.V1 lineage B.1.1.7 (Davies et al., 2020), the SA (South Africa) variant 501Y.V2 lineage B.1.351 (Tegally et al., 2020), the BR (Brazilian) variant 501Y.V3 lineage P.1 (Voloch et al., 2020) and the CA (California) variant CAL.20C lineage B.1.427 (Zhang et al., 2021). The B.1.1.7 variant is associated with increased transmissibility (Rambaut et al., 2020; Washington et al., 2021), and similar epidemiological observations have been reported for the SA and BR variants (Tegally et al., 2020; Voloch et al., 2020).

Mutations of greatest concern are present in the viral Spike (S) protein, and include notable mutations in the receptor binding domain (RBD), N-terminal domain (NTD), and furin cleavage site region. Several of these mutations directly affect ACE2 receptor binding affinity, which may impact infectivity, viral load, or transmissibility (Greaney et al., 2021; Starr et al., 2021; Wang et al., 2021a; Zahradník et al., 2021). Several of the mutations were also noted to be in regions bound by neutralizing antibodies, so it is crucial to address to what extent the mutations associated with the variants impact immunity induced by either SARS-CoV-2 infection or vaccination.

Several reports address the effect of these mutations on antibody binding and function, by either monoclonal or polyclonal antibody responses, and considering both natural infection or vaccination (Edara et al., 2021; Greaney et al., 2021; Muik et al., 2021; Shen et al., 2021; Skelly et al., 2020; Stamatatos et al., 2021; Supasa et al., 2021; Wang et al., 2021a; Wang et al., 2021b; Wibmer et al., 2021; Wu et al., 2021). In general, the impact of the B.1.1.7 variant mutations on neutralizing antibody titers is moderate (Emary et al., 2021; Muik et al., 2021; Shen et al., 2021; Skelly et al., 2020; Supasa et al., 2021; Wu et al., 2021). In contrast, the mutations associated with the B.1.351 and P.1. variants are associated with more pronounced loss of neutralizing capacity (Cele et al., 2021; Skelly et al., 2020; Wang et al., 2021a; Wibmer et al., 2021; Wu et al., 2021). Concerning vaccination responses, the AstraZeneca ChAdOx1 vaccine has been associated with a partial loss of neutralizing antibody activity against B.1.1.7 (Skelly et al., 2020), and a large loss of neutralizing activity against B.1.351 (Voysey et al., 2021). Consistent with these reports, ChAdOx1 maintains efficacy against B.1.1.7 (Emary et al., 2021; Hall et al., 2021), but has a major loss in efficacy against mild COVID-19 with the B.1.351 variant (Voysey et al., 2021). Current epidemiological evidence is that the BNT162b2 Pfizer/BioNTech COVID-19 vaccine retains its efficacy against B.1.1.7 in the UK and in reports from Israel (Amit et al., 2021). Novavax (NVX-CoV2373) has reported differential protective immunity against the parental strain, B.1.1.7, and B.1.351 in vaccine clinical trials (96%, 86%, and 60%) (Novavax Inc., 2021), whereas the Janssen Ad26.COV2.S 1-dose COVID-19 vaccine, which elicits lower neutralizing antibody titers (Sadoff et al., 2021), has relatively similar protection for moderate COVID-19 against both the ancestral strain and B.1.351 (72% and 64%)(FDA, 2021a, b).

Several lines of evidence suggest that $CD4^+$ and $CD8^+$ T cell responses play important roles in resolution of SARS-CoV-2 infection and COVID-19 (Sette and Crotty, 2021), including modulating disease severity in humans (Rydyznski Moderbacher et al., 2020; Tan et al., 2021) and reducing viral loads in non-human primates (Munoz-Fontela et al., 2020). Further, persons with agammaglobulinemia or pharmaceutical depletion of B cells generally experience an uncomplicated COVID-19 disease course (Sette and Crotty, 2021; Soresina et al., 2020). Robust $CD4^+$ and $CD8^+$ T cell memory is induced after COVID-19 (Breton et al., 2021; Dan et al., 2021; Peng et al., 2020; Wang et al., 2021b), and multiple COVID-19 vaccines elicit $CD4^+$ and $CD8^+$ T cell responses (Baden et al., 2021; Dowd et al., 2020; Keech et al., 2020; Sadoff et al., 2021; Voysey et al., 2021). It is therefore key to address the potential impact of SARS-CoV-2 variants mutations on T cell reactivity; however, little data is currently available on this topic (Skelly et al., 2020).

Here, we take a combined experimental and bioinformatics approach to address how SARS-CoV-2 variants of concern impact T cell reactivity. We directly assess T cell responses from persons recovered from COVID-19 obtained before the emergence of the variants, and from recent Moderna mRNA-1273 or Pfizer/BioNTech BNT162b2 vaccinees, for their capacity to recognize peptides derived from the ancestral reference sequence and the B.1.1.7, B1.351, P.1 and the CAL.20C variants. Bioinformatic analyses were used to predicted the impact of mutations in the various variants with sets of previously

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

reported $CD4^+$ and $CD8^+$ T cell epitopes derived from the ancestral reference sequence (Tarke et al., 2021).

## RESULTS

### Sequence analysis, peptide pool generation and selection of cohorts of COVID-19 convalescent and recent vaccinees

As a first step, we mapped the specific mutations (amino acid replacements and deletions) associated with several of the current variants of concern, including the SARS-CoV-2 B.1.1.7, B1.351, P.1 and the CAL.20C variants, as compared to the SARS-CoV-2 Wuhan ancestral sequence (NCBI acc. no. NC_045512.2). Briefly, the genomic sequences were downloaded from GISAID, translated using the VIGOR4 tool available on the Virus Pathogen Resource (ViPR) (Pickett et al., 2012), and then compared with the protein sequence of the Wuhan ancestral strain to identify all the possible amino acid changes, as listed in **Table S1**.

Next, we synthesized the corresponding peptides associated with the different variants and generated new peptide pools spanning the full genome sequences of the ancestral Wuhan strain and the respective B.1.1.7, B1.351, P.1 and the CAL20C variants (**Table S2**). As described below, the resulting peptide pools were assessed for their capacity to be recognized by memory T cells responses derived from natural infection in convalescents and vaccinees, and responses to the variant and ancestral genome antigen-specific pools were compared.

Our convalescent donors were adults with ages ranging from 21 to 57 years of age (median 39); 27% were male and 73% female (**Table 1**). SARS-CoV-2 infection in these donors was determined by PCR-based testing during the acute phase of their infection, if available (55% of the cases), and/or seropositivity determined by plasma SARS-CoV-2 S protein RBD IgG ELISA (Stadlbauer et al., 2020)(**Fig. S1**). From these donors, PBMC samples were collected between July to October 2020 period, when the dominant local strain was the ancestral reference virus.

From vaccinated donors, we collected PBMC after recent vaccination with the Moderna mRNA-1273 or the Pfizer/BioNTech BNT162b2 vaccines, approximately 14 days following the second dose administration (**Table 1**). These donors ranged in age from 22 to 67 years (median 43) and 26% were male and 74% were female. All vaccinees had significant RBD IgG titers in the 1843 to 16365 range, consistent with recent vaccination **(Fig. S1**).

### $CD4^+$ and $CD8^+$ T cell antigenicity against Spike variant sequences in convalescent samples

We previously described the use of Activation Induced Marker (AIM) assays (Dan et al., 2021; Grifoni et al., 2020b; Mateus et al., 2020; Rydyznski Moderbacher et al., 2020; Tarke et al., 2021) to measure $CD4^+$ and $CD8^+$ T cell responses to pools of overlapping peptides spanning the entire sequence of the SARS-CoV-2 antigens. Here, we utilized the same AIM assays using $OX40^+CD137^+$ and $CD69^+CD137^+$ markers for $CD4^+$ and $CD8^+$ T cells reactivity, respectively (Grifoni et al., 2020b; Mateus et al., 2020). As shown in **Fig. 1A-B**, good $CD4^+$ and $CD8^+$ T cell reactivity was observed in convalescent donors with pools of overlapping peptides spanning the S protein of the ancestral Wuhan sequence, but also for each of the corresponding variant S pools. Geomean reactivity ranged from 0.09 to the 0.10 for $CD4^+$ T cells, and 0.08 to the 0.11 for $CD8^+$ T cells; No significant difference was observed between the pool of S peptides corresponding to the ancestral sequence and those corresponding to the different variants (CD4: UK p=0.90; SA p=0.50; BR p=0.49; CA p=0.85 and CD8: UK p=0.16; SA p=0.07; BR p=0.18; CA p=0.20 by the Wilcoxon test). These values (here and in subsequent graphs) are not corrected for multiple comparisons, as the correction would only decrease the statistical power for detecting a significant difference; therefore, not performing multiple comparison corrections is the more conservative and stringent test.

These T cell analyses were extended using a FluoroSPOT assay system, to measure the capacity of the various pools to elicit functional responses in terms of secretion of IFNγ and IL-5 cytokines (**Fig. 1C-D)**. As shown in **Fig. 1C**, reactivity was observed for the S pools in convalescent donors in terms of IFNγ with geomean reactivity ranging from 32 to 45 Spot Forming Cells (SFC) per million PBMCs. Compared to the ancestral strain, mild decreases in the 24–30% range were noted for B.1.1.7, P.1 and CAL.20C variant pools (UK p=0.01; SA p=0.48; BR p=0.05 and CA p=0.01 by the Wilcoxon test), while

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

no difference was observed for B.1.351. As expected, no IL-5 reactivity was observed for any of the pools (**Fig. 1D).**

To further expand these findings, we considered the dose response of the various S pools in terms of stimulation of $CD4^+$ and $CD8^+$ T cell specific responses. As shown in **Fig. 1E**, $CD4^+$ T cell dose dependent responses for the Wuhan and four variant pools were similar. The same pattern was also observed for $CD8^+$ T cell responses (**Fig. 1F**).

**$CD4^+$ and $CD8^+$ T cell antigenicity of proteome-wide SARS-CoV-2 variant sequences in convalescent samples**

As shown in **Table S1**, mutations found in the variants studied herein were not limited to the Spike protein, but occurred in several additional antigens encoded in the SARS-CoV-2 genome. To address their potential impact on the overall proteome-wide $CD4^+$ and $CD8^+$ T cell reactivity, we tested overlapping peptide pools spanning the entire proteome of the ancestral Wuhan sequence in comparison with corresponding pools representing the different variants.

Overall, reactivity to the peptide pools spanning the variant genomes was found to be similar to that against the ancestral Wuhan strain (**Fig. 2**). When the sum total of reactivity throughout the genome was considered, no differences or decreases in reactivity compared to the ancestral were noted for the variant pools (CD4: UK p=0.58; SA p= 0.46; BR p= 0.27; CA p= 0.08 and CD8: UK p= 0.25; SA p= 0.15; BR p= 0.02; CA p= 0.30 by the Wilcoxon test uncorrected p values) (**Fig. 2A-B**).

We previously showed that in COVID-19 convalescent subjects a set of 10 different antigens (nsp3, nsp4, nsp6, nsp12, nsp13, S, ORF3a, M, ORF8 and N) account for 83 and 81% of the total $CD4^+$ and $CD8^+$ T cell responses, respectively (Tarke et al., 2021). Here a similar overall pattern of dominant antigens was observed. When single proteins are considered, no variant pool showed a decrease in reactivity when a multi-hypothesis testing correction was applied (**Fig. 2C-D).** It is worth noting that this specific comparison is for illustration purposes only, as this study is not fully powered to rule out minor differences that could be observed in the individual antigens.

In conclusion, these experiments suggest that memory $CD4^+$ or $CD8^+$ T cells from individuals that have been infected with the ancestral SARS-CoV-2 strain recognize the ancestral reference strain and the variant genome-wide sequences with similar efficiency.

**$CD4^+$ and $CD8^+$ T cell antigenicity against Spike variant sequences in recent vaccines samples**

We also studied T cell responses by individuals who received authorized mRNA COVID-19 vaccines. We focused our analysis on T cell reactivity to peptide pools spanning the Spike antigen of the ancestral strain, which is the basis of the presently used vaccines. For both $CD4^+$ and $CD8^+$ T cell reactivity, the magnitude of responses to pools encompassing the sequences from the ancestral Wuhan genome and the different variants considered range from a geomean of 0.15 to 0.19 for $CD4^+$ T cells and a geomean of 0.16-0.24 for $CD8^+$ T cells. Comparison of the variant pools to the ancestral sequence showed no significant difference for $CD4^+$ T and $CD8^+$ T cells reactivity, with the exception of the B.1.351 pools, where mild decreases of 29% and 33%, respectively, were observed (CD4: UK p=0.47; SA p=0.01; BR p=0.91; CA p=0.41; CD8: UK p=0.03; SA p=0.001; BR p=0.15 and CA p=0.02 by the Wilcoxon test) (**Fig. 3A-B**).

The results from the FluoroSPOT assay system (**Fig. 3C-E**) showed good reactivity in terms of IFNγ, with geomean reactivity ranging from 54 to the 70 SFC per million PBMCs (**Fig. 3C**). Minimal IL-5 responses were observed, with geomean reactivity ranging from 22 to 25 SFC/$10^6$, which is slightly above the limit of detection (**Fig. 3D**). On a per donor basis, the IFNγ response was found to account for more than 80% of the total response, on average (range 81% to 87%), irrespective of whether the ancestral strain or any of the variants was considered (**Fig. 3E**).

Similar to the experiments in convalescent donors, we also considered the dose response of the various Spike pools in terms of stimulation of $CD4^+$ or $CD8^+$ T cell specific responses for vaccinees. As shown in **Fig. 3F-G**, $CD4^+$ and $CD8^+$ T cell dose responses for the ancestral pools and the four variant pools were similar. Taken together these results indicate that the responses to the ancestral and variant Spike pools are similar for both $CD4^+$ and $CD8^+$ T cells in mRNA vaccinees.

**Conservation analysis of sets of defined $CD4^+$ and $CD8^+$ T cell epitopes**

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We recently reported a comprehensive study of epitopes recognized in convalescent subjects, leading to the identification of 280 different $CD4^+$ T cell epitopes (Tarke et al., 2021). Here, we analyzed how many of those epitopes would be impacted by mutations in the different variants. As shown in **Fig. 4A**, we found that 89.6%, 90%, 94.3% and 97.1% (average 93%) of the $CD4^+$ T cell epitopes identified by Tarke et al. are conserved in the B.1.1.7, B1.351, P.1 and the CAL20C variants. A similar pattern is observed when the magnitude of responses associated with the various epitopes is considered, rather than the simple number of epitopes (**Fig. 4B**). The fully conserved $CD4^+$ T cell epitopes account for 84.4%, 88.1%, 95.7% and 97.8% (average 91.5%) of the total response, when comparing the B.1.1.7, B.1.351, P.1 and the CAL.20C variants, respectively.

The study of Tarke et al. also reported the identification of 523 $CD8^+$ T cell epitope associated with unique HLA restrictions (Tarke et al., 2021). Performing a similar analysis as above, (**Fig. 4C**) we found that 508 (97.1%) of these 523 $CD8^+$ T cell epitopes are totally conserved within the B.1.1.7 variant, 509 (97.3%) are conserved within the B.1.351 variant, 509 (97.3%) are conserved within the P.1 variant and 512 (97.9%) within the CAL.20C variant. Similarly, in terms of magnitude of $CD8^+$ T cell responses associated with the various epitopes, totally conserved $CD8^+$ T cell epitopes account for 98.3%, 98.4%, 97.9% and 97.8% of the total responses (**Fig. 4D**) for the B.1.1.7, B.1.351, P.1 and the CAL.20C variants respectively, with an average of 98.1%.

Finally, we analyzed the degree of $CD4^+$ and $CD8^+$ T cell epitope conservation if we restricted our analysis only to epitopes contained in the Spike antigen. The number of S-derived epitopes conserved at 100% sequence identity was, on average, 84.5% for the $CD4^+$ T cell epitopes (**Fig. 4E),** and 95.3% for the $CD8^+$ T cell epitopes (**Fig. 4G**). Similarly, in terms of magnitude of $CD4^+$ T cell responses associated with the various S epitopes, totally conserved $CD4^+$ T cell epitopes account for 95.5%, 75.3%, 89.8% and 98.3% of the total responses (**Fig. 4F**) for the B.1.1.7, B.1.351, P.1 and the CAL.20C variants respectively, with an average of 89.7%. In terms of the magnitude of $CD8^+$ T cell responses, totally conserved epitopes account for 95.2%, 97.6%, 95.4% and 97.3% of the total responses (**Fig. 4H**) for the B.1.1.7, B.1.351, P.1 and the CAL.20C variants respectively, with an average of 96.4%.

While the restriction of the HLA class II epitopes in the Tarke et al. (Tarke et al., 2021) study was not unequivocally assigned, the restriction of the class I epitopes is implicitly inferred based on HLA allele specific predictions and testing in HLA matched donors. Accordingly, we further analyzed the extent to which the affected epitopes would be impacted by their respective associated mutation by determining, for each epitope/matching epitope variant, their predicted binding affinity for the corresponding putative HLA class I restriction element. Predicted binding capacity was determined using the NetMHCpan BA 4.1 tool provided by the IEDB's analysis resource (Dhanda et al., 2019; Reynisson et al., 2020).

In the case of the B.1.1.7, B.1.351, P.1 and the CAL.20C variants, the % of mutations associated with no decrease in binding capacity, conservatively defined as a 2-fold reduction, was 73.3%, 78.6%, 78.6% and 45.5%, respectively (with the CAL.20.C variant having a smallest number of total mutations, noted above) (**Fig. S2** and **Table S3**). In conclusion, the analyses suggest that the vast majority of $CD8^+$ T cell epitopes are unaffected by mutations found in all the different variants. The corresponding mutations are also predicted to have minor effects on the total T cell response, thus providing a molecular basis for the marginal impact on T cell reactivity by COVID-19 convalescent subjects and recipients of COVID-19 mRNA vaccines.

## DISCUSSION

The present study addresses a key knowledge gap pertaining to the potential of emergent SARS-CoV-2 variants to evade recognition by human immune responses. We focused on T cell responses elicited by either natural infection or vaccination with the Pfizer/BioNTech and Moderna COVID-19 mRNA vaccines. We found negligible effects on both $CD4^+$ or $CD8^+$ T cell responses to all four variants investigated, to include the B.1.1.7, B.1.351, P.1 and CAL.20C variants found in the UK, South Africa, Brazil and California, respectively. To more comprehensively assess T cell functionalities, the comparison between the original Wuhan isolate and the variants was performed utilizing different T cell methodologies, such as the AIM assay (quantifying T cells with a range of functionalities), and the FluoroSPOT assay (quantifying cells with specific cytokine-secreting activity). We also tested whether any of the variant sequences might be associated with an altered cytokine polarization; marginal IL-5

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

production was detected in any of the conditions tested. This is relevant, since it was reported that single amino acid replacements in an epitope sequence can lead to a change in the cytokines produced (Evavold and Allen, 1991; Sloan-Lancaster and Allen, 1996), and a Th2-like response pattern was initially hypothesized to be linked to adverse outcomes in SARS-specific responses (Peeples, 2020).

The data provide some positive news in light of justified concern over the impact of SARS-CoV-2 variants of concern on efforts to control and eliminate the present pandemic. Undoubtedly, several of the variants are associated with increased transmissibility, and also have been associated with decreased susceptibility to neutralizing antibodies from infected or vaccinated individuals. In contrast, the data presented here suggests that T cell responses are largely unaffected by the variants. While it is not anticipated that circulating memory T cells would be effective in preventing SARS-CoV-2 infection, it is plausible that they can reduce COVID-19 severity (Lipsitch et al., 2020; Sette and Crotty, 2021). Several lines of evidence support this notion, such as observations that early SARS-CoV-2 T cell responses are associated with milder COVID-19 (Rydyznski Moderbacher et al., 2020; Tan et al., 2021). Thus, the T cell response may contribute to limiting COVID-19 severity induced by variants that partially or largely escape neutralizing antibodies. This is consistent with T cell mediated immunity observed in humans against a different respiratory pathogen, influenza, for which heterologous immunity against diverse influenza strains is associated with memory T cells to conserved epitopes (Greenbaum et al., 2009; Sridhar et al., 2013; Wilkinson et al., 2012).

Our data also provide a molecular basis for the lack of impact of the mutations associated with the variants analyzed on T cell responses. Prior reports have identified a large number of T cell epitopes recognized throughout the SARS-CoV-2 proteome, including Spike (Ferretti et al., 2020; Keller et al., 2020; Le Bert et al., 2020; Nelde et al., 2020; Peng et al., 2020; Snyder et al., 2020). We furthered this point by an analysis of the Tarke et al. data set, showing that 93% of $CD4^+$ T cell, and 97% of $CD8^+$ T cell, epitopes are completely conserved in the variants. Further, we found that even in the epitopes affected by single mutations, no negative affect HLA binding capacity in the majority of cases is expected. The apparent higher conservation of $CD8^+$ T cell epitopes is to be expected based on the shorter length of HLA class I binding peptides (usually 9-10 amino acids) as compared to their class II counterparts (13-17. This effect is counterbalanced by $CD8^+$ T cells being generally less tolerant of amino substitutions as compared to $CD4^+$ T cells (Grifoni et al., 2020a; Weiskopf et al., 2014). Overall, we observed that the effect of the variant mutations on the global $CD4^+$ and $CD8^+$ T cell responses was negligible.

Mutations associated with the variants could be reflective of adaptation in terms of optimizing replication or binding to ACE2, but also reflective of adaptation to escape immune recognition by antibodies (Andreano et al., 2020; Starr et al., 2020; Wang et al., 2021a; Wang et al., 2021b; Zahradník et al., 2021). Indeed, higher viral binding to a cellular receptor can be a mechanism of compensatory viral evolution in the presence of neutralizing antibodies (Hensley et al., 2009). In that respect, while mutation to escape antibody binding has been well documented for influenza (Andrews et al., 2015; Doud et al., 2018; Krammer et al., 2018) and SARS-CoV-2, immune escape at the level of T cell responses in human populations has not been reported for other acute respiratory infections. Because of HLA polymorphism, the epitope repertoire recognized is likely to be substantially different from one individual to the next, greatly decreasing the likelihood of immune escape by an acute virus. An advantage conferred to the virus by a mutation in a person would not be linked to an immune response escape advantage in a non-HLA matched individual. At the same time, our data does not rule out that each person could be strongly affected by the mutations of specific variants. For SARS-CoV-2, this property of T cell recognition is further enhanced by the fact that the T cell responses against SARS-CoV-2 are highly multi-antigenic and multi-specific, with tens of different epitopes recognized by $CD4^+$ and $CD8^+$ T cells in a given individual (Braun et al., 2020; Ferretti et al., 2020; Nelde et al., 2020; Tarke et al., 2021).

The results here have potential implications for engineering coronavirus vaccines with broader protective immunity against variants of concern. Clearly the most straightforward path is to update the current vaccines to target a variant Spike, given how highly successful several COVID-19 vaccines have proven to be against the parental SARS-CoV-2 strain. Our results suggest that a parallel alternative approach could involve inclusion of additional antigens and epitopes, perhaps selected on the basis of low mutational propensity (Gaiha et al., 2019), to ensure that neutralizing antibodies are complemented with T cell responses to minimize COVID-19 morbidity and mortality.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

### *Limitations and future directions.*

The present study did not assess decreases in antibody reactivity, as several other studies have already investigated this matter (Edara et al., 2021; Greaney et al., 2021; Muik et al., 2021; Shen et al., 2021; Skelly et al., 2020; Stamatatos et al., 2021; Supasa et al., 2021; Wang et al., 2021a; Wang et al., 2021b; Wibmer et al., 2021; Wu et al., 2021). Further, our studies utilized overlapping peptide pools, and as such we could not exclude that some of the mutations might involve alterations in terms of antigen processing for either class I or class II, which would be undetected by using pools of "preprocessed" peptides. The number of donors studied was also limited, although variability in T cell reactivity suggestive of large variant-associated effects were not observed. While we have no reason to suspect that substantial differences might exist between the epitope specificity of responses elicited by different vaccines, our study did not address this point. Our study was designed to test for differences in overall response to the different variants and was not powered or designed to investigate differences between the mRNA vaccines. Furthermore, in our study we tested samples from convalescent donors infected before October 2020; thus, it is unlikely that any of the donors would have been infected by any of the variants, as this date precedes their diffusion to appreciable degree in the US and California. Finally, we have only investigated whether responses induced by the ancestral reference sequence are able to cross-recognize variant sequences, as this is relevant to the current situation. We have not examined whether responses induced by an infection with a variant sequence will be able to cross-recognize the ancestral reference sequence present in the currently approved vaccines.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## ACKNOWLEDGEMENTS

This study has been funded by the NIH NIAID (award AI142742 to S.C., A.S., contract Nr. 75N9301900065 to A.S. and D.W., contract Nr. 75N93019C00001 to A.S. and B.P., NIH grant U01 CA260541-01 to DW, K08 award AI135078 to J.D., and AI036214 to D.S. and HHSN75N93019C00076 to R.H.S). Additional support has been provided by UCSD T32s (AI007036 and AI007384 to S.A.R and S.I.R) and the Jonathan and Mary Tu Foundation (D.S.). A.T. was supported by a PhD student fellowship through the Clinical and Experimental Immunology Course at the University of Genoa, Italy. We thank Gina Levi and the LJI clinical core for assistance in sample coordination and blood processing. We gratefully thank the authors from the originating laboratories responsible for obtaining the specimens, as well as the submitting laboratories where the genome data were generated and shared via GISAID, and on which this research is based. We would like to thank Vamseedhar Rayaprolu and Erica Ollmann Saphire for providing the recombinant SARS-CoV-2 Receptor Binding Domain (RBD) protein used in the ELISA assay.

## AUTHOR CONTRIBUTIONS

Conceptualization: A.T., A.G., S.C. and A.S.; Data curation and bioinformatic analysis, Y.Z. R.H.S. B.P.; Formal analysis: A.T., J.S., A.G.; Funding acquisition: S.C., A.S., D.W., S.I.R., S.A.R., and J.M.D.; Investigation: A.T., N.M., A.Su., B.G. J.S., D.W. A.S and A.G.; Project administration: A.F. Resources: S.I.R., S.A.R., J.D.; Supervision: J.S., S.C., D.W., A.S., R.d.S. and A.G.; Writing: A.T., D.W., S.C., A.S., and A.G.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE LEGENDS

**Figure 1. T cell responses of COVID-19 convalescent individuals against SARS-CoV-2 Spike for the different variants.** PBMCs of COVID-19 convalescent individuals (n=11) were stimulated with the Spike MPs corresponding to the ancestral reference strain (Wu, black) and the B.1.1.7 (UK, grey), B.1.351 (SA, red), P.1 (BR, orange) and CAL.20C (CA, light blue) SARS-CoV-2 variants. **A**) Percentages of AIM$^+$ (OX40$^+$CD137$^+$) CD4$^+$T cells. **B**) Percentages of AIM$^+$ (CD69$^+$CD137$^+$) CD8$^+$ T cells. **C**) IFNγ spot forming cells (SFC) per million PBMCs **D**) IL-5 SFC per million PBMCs. Paired comparisons of Wuhan S MP versus each of the variants were performed by Wilcoxon test and are indicated by the p values in panels **A**-**C**. The data shown in panels **A** and **B** are plotted to show the Spike MPs titration (1 µg/mL, 0.1 µg/mL, 0.01 µg/mL) for CD4$^+$ (**E**) and CD8$^+$ (**F**) T cells in each SARS-CoV-2 variant and the geometric mean of the 0.1ug/mL condition is listed above each titration. In all panels, the bars represent the geometric mean.

**Figure 2. T cell responses of COVID-19 convalescent individuals against SARS-CoV-2 proteome for the different variants.** PBMCs of the COVID-19 convalescent individuals (n=11) were stimulated with the MPs for the entire viral proteome corresponding to the ancestral reference strain (Wu, black) and the B.1.1.7 (UK, grey), B.1.351 (SA, red), P.1. (BR, orange) and CAL.20C (CA, light blue) SARS-CoV-2 variants. **A**) Percentages of AIM$^+$ (OX40$^+$CD137$^+$) CD4$^+$ T cells for the total reactivity. **B**) Percentages of AIM$^+$ (CD69$^+$CD137$^+$) CD8$^+$ T cells for the total reactivity. Bars represent the geometric mean. Paired comparisons of Wuhan versus each of the variants were performed by Wilcoxon tests. **C**) Percentages of AIM$^+$ (OX40$^+$CD137$^+$) CD4$^+$ T cells for each MP **D**) Percentages of AIM$^+$ (OX40$^+$CD137$^+$) CD4+ T cells for each MP.

**Figure 3. T cell responses of COVID-19 vaccinee individuals against SARS-CoV-2 Spike for the different variants.** PBMCs of Pfizer/BioNTech BNT162b2 (n=8, triangles) and Moderna mRNA-1273 COVID-19 vaccines (n=11, circles) were stimulated with the Spike MPs corresponding to the ancestral reference strain (Wu, black) and the B.1.1.7 (UK, grey), B.1.351 (SA, red), P.1. (BR, orange) and CAL.20C (CA, light blue) SARS-CoV-2 variants. **A**) Percentages of AIM$^+$ (OX40$^+$CD137$^+$) CD4$^+$ T cells. **B**) Percentages of AIM$^+$ (CD69$^+$CD137$^+$) CD8$^+$ T cells. **C**) IFNγ spot forming cells (SFC) per million PBMCs **D**) IL-5 Spot forming cells (SFC) per million PBMCs. **E**) Percentages of IFNγ were calculated from the total IFNγ and IL-5 SFC per million PBMCs. Paired comparisons of the ancestral reference strain-based S MP versus each of the variants were performed by Wilcoxon test and are indicated by the p values in panels **A**-**E**. The data shown in panels **A** and **B** are also plotted showing the spike MPs titration (1 µg/mL, 0.1 µg/mL, 0.01 µg/mL) for CD4$^+$ (**F**) and CD8$^+$ (**G**) T cells in each SARS-CoV-2 variant. The geometric mean of the 0.1ug/mL condition is listed above each titration. In all panels, the bars represent the geometric mean.

**Figure 4. SARS-CoV-2 T cell epitope sequences affected by the variants.** CD4$^+$ and CD8$^+$ T cell epitopes of the ancestral strain identified in a previous study (Tarke et al.) are analyzed as a function of the number and percentage of response that are or are not conserved across the B.1.1.7 (UK, grey), B.1.351 (SA, red), P.1. (BR, orange) and CAL.20C (CA, light blue) SARS-CoV-2 variants. The SARS-CoV-2 epitopes for the most immunodominant SARS-CoV-2 proteins in terms of numbers and percentage of response are shown for CD4$^+$ (**A-B**) and CD8$^+$ (**C-D**) T cells. The SARS-CoV-2 epitopes for the Spike protein only in terms of numbers and percentage of response are shown for CD4$^+$ (**E-F**) and CD8$^+$ (**G-H**) T cells.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## TABLES

**Table 1**. Characteristics of donor cohorts.

| | **COVID-19 (n = 11)** | **Vaccinees (n = 19)** |
|---|---|---|
| **Age** (years) | 21-57 [Median = 39, IQR = 33] | 22-67 [Median = 43, IQR = 36] |
| **Gender** | | |
| Male (%) | 27% (3/11) | 26% (5/19) |
| Female (%) | 73% (8/11) | 74% (14/19) |
| **Sample Collection Date** | July–Oct 2020 | Jan–Feb 2021 |
| **SARS-CoV-2 PCR** | Positive = 83% (5/6)<br>Not tested= 45% (5/11) | N/A |
| **S RBD IgG Positive** | 100% (11/11) | 100% (19/19) |
| **Peak disease Severity**[a] | | |
| Mild | 100% (11/11) | N/A |
| Moderate | 0% (0/11) | |
| Severe | 0% (0/11) | |
| Critical | 0% (0/11) | |
| **Race-Ethnicity** | | |
| White- not Hispanic or Latino | 82% (9/11) | 42% (8/19) |
| Hispanic or Latino | 9% (1/11) | 16% (3/19) |
| Asian | 9% (1/11) | 42% (8/19) |
| American Indian/Alaska Native | 0% (0/11) | 0% (0/19) |
| Not reported | 0% (0/11) | 0% (0/19) |
| **Days at Collection** | 38-80 (11/11) [Median =50, IQR = 45][b] | 13-30 (8/19) Pfizer<br>12-15 (11/19) Moderna [Median = 14, IQR = 14][c] |

[a] According to WHO criteria.
[b] Post Symptom Onset
[c] 2nd dose of vaccination

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## SUPPLEMENTAL MATERIALS

## SUPPLEMENTAL FIGURE LEGENDS

**Figure S1. SARS-COV-2 serology of the all the cohorts analyzed in this study.** Related to Figures 1, 2 and 3 and Table 1. Spike RBD serology in COVID-19 convalescents (n=11) and COVID-19 vaccines (n=19).

**Figure S2. Effect of mutations on CD8 epitope predicted HLA class I binding capacity.** Related to Figure 4. For each $CD8^+$ T cell epitope associated with a mutation found in the respective variants, the predicted HLA binding capacity of original sequence and the mutated sequence was calculated. Based on the results, each instance was categorized as a function of whether the binding capability of the mutated peptide is increased (>2-fold), neutral or decreased (<2-fold). Each analysis is done separately for the B.1.1.7 (UK, grey), B.1.351 (SA, red), P.1. (BR, orange) and CAL.20C (CA, light blue) SARS-CoV-2 variants.

**Figure S3. Gating strategy.** Related to Figures 1, 2, 3 and 4. Representative graphs illustrating the gating strategy used in the flow cytometry AIM assays in order to define antigen-specific $CD4^+$ (outlined in blue) and $CD8^+$ (outlined in red) T cells by the expression of $OX40^+CD137^+$ and $CD69^+$ $CD137^+$, respectively. These graphs depict one of the COVID-19 convalescent donors from this study and are representative of the gating strategy utilized with all donors tested.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## SUPPLEMENTAL TABLE LEGENDS

**Table S1**. Related to Figures 1, 2, 3 and 4. List of amino acid positions and relative amino acid changes in the different variants studied with respect to the ancestral Wuhan strain.

**Table S2.** Related to Figures 1, 2, 3 and 4. List of mutated peptides with respect to the ancestral Wuhan strain in the different variants studied.

**Table S3.** Related to Figure 4. Effect of mutations on CD8 epitope HLA class I binding capacity.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# STAR METHODS

## RESOURCE AVAILABILITY

### Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr. Alessandro Sette (alex@lji.org).

### Materials Availability

Aliquots of synthesized sets of peptides utilized in this study will be made available upon request. There are restrictions to the availability of the peptide reagents due to cost and limited quantity.

### Data and Code Availability

The published article includes all data generated or analyzed during this study, and summarized in the accompanying tables, figures and supplemental materials.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Human Subjects

*Convalescent COVID-19 Donors.* Convalescent donors were enrolled at either a UC San Diego Health clinic under the approved IRB protocols of the University of California, San Diego (UCSD; 200236X), or at the La Jolla Institute (LJI; VD-214). All donors were California residents and samples were collected from August to October 2020, before any of the SARS-CoV-2 variants described herein had been detected in California. These donors were referred to the study by a health care provider or were self-referred. The CRO BioIVT provided additional cohorts of COVID-19 convalescent donors who had been confirmed positive for COVID-19 by PCR following the resolution of symptoms. The total cohort of convalescent donors represented both sexes (27% male, 73% female) and ranged from 21 to 57 years of age (median 39 years). All samples were confirmed seropositive against SARS-CoV-2 by ELISA, as described below. Details of this convalescent COVID-19 cohort are listed in Table 1. All convalescent COVID-19 donors provided informed consent to participate in the present and future studies at the time of enrollment.

*COVID-19 vaccinees.* The La Jolla Institute recruited 19 healthy adults who had received the first and second dose of the Pfizer/BioNTech BNT162b2 (n=8) or Moderna mRNA-1273 COVID-19 vaccines (n=11). Blood draws took place under IRB approved protocols two to four weeks after the second dose of the vaccine was administered. All donors had their SARS-CoV-2 antibody titers measured by ELISA, as described below. The cohort of vaccines represented ranged from 22 to 67 years of age (median 43 years) and represented both sexes (26% male, 74% female). At the time of enrollment in the study, all donors gave informed consent.

## METHOD DETAILS

### Isolation of peripheral blood mononuclear cells (PBMCs) and plasma

Collection and processing of blood samples was performed as previously described (Dan et al., 2021; Tarke et al., 2021). Briefly, whole blood was collected in heparin coated blood bags or in ACD tubes and centrifuged for 15 minutes at 1850 rpm to separate the cellular fraction from the plasma. The plasma was then removed and stored at -20°C. The cellular fraction next underwent density-gradient sedimentation using Ficoll-Paque (Lymphoprep, Nycomed Pharma, Oslo, Norway) to separate the PBMCs as previously described (Weiskopf 2013). Isolated PBMCs were cryopreserved in cell recovery media containing 10% DMSO (Gibco), supplemented with 90% heat inactivated fetal bovine serum (FBS; Hyclone Laboratories, Logan UT) and stored in liquid nitrogen until used in the assays.

### SARS-CoV-2 RBD ELISA

Serology to SARS-CoV-2 was determined for all donor cohorts as previously described (Rydyznski Moderbacher et al., 2020). Briefly, 96-well half-area plates (ThermoFisher 3690) were coated with 1 ug/mL SARS-CoV-2 Spike (S) Receptor Binding Domain (RBD) and incubated at 4°C overnight. The next day plates were blocked at room temperature for 2 hours with 3% milk in phosphate buffered saline (PBS)

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

containing 0.05% Tween-20. Heat-inactivated plasma was added to the plates for an additional 90-minute incubation at room temperature followed by incubation with the conjugated secondary antibody, detection, and subsequent data analysis by reading the plates on Spectramax Plate Reader at 450 nm using the SoftMax Pro. The limit of detection (LOD) was defined as 1:3. Limit of sensitivity (LOS) for SARS-CoV-2 infected individuals was established based on uninfected subjects, using plasma from normal healthy donors not exposed to SARS-CoV-2.

**Mutation analysis of SARS-CoV-2 UK, California, South Africa and Brazil variants**

Genome sequences for the variant viruses were downloaded from GISAID. These sequences were screened to select those without ambiguous residues and generated from Illumina sequencing technologies using an in-house sequence QC script. The selected genomic sequences were then translated into protein amino acid sequences using the VIGOR4 tool available on the Virus Pathogen Resource (ViPR)(Pickett et al., 2012). Sequence variations in the variant viruses were derived by comparison with Wuhan-1 (NC_045512.2). One or more representative sequences were considered for the UK (EPI_ISL_601443), Brazilian (EPI_ISL_804823), Californian (EPI_ISL_847619; EPI_ISL_847621; EPI_ISL_847643) and South Africa (EPI_ISL_660629; EPI_ISL_736930; EPI_ISL_736932; EPI_ISL_736944; EPI_ISL_736966; EPI_ISL_736971; EPI_ISL_736973; EPI_ISL_825104; EPI_ISL_825120; EPI_ISL_825131) variants. A summary of all the amino acids mutated in the different variants respect to the Wuhan sequence and considered in this study is available in **Table S1**.

**SARS-CoV-2 Wuhan and variant peptide synthesis and pooling**

Peptides were synthesized that spanned entire SARS-CoV-2 proteins and corresponded to the ancestral Wuhan sequence or the B.1.1.7 (UK), B.1.351 (SA), P.1 (BR) and CAL.20C (CA) SARS-CoV-2 variants. Peptides were 15-mers overlapping by 10 amino acids and were synthesized as crude material (TC Peptide Lab, San Diego, CA). All peptides were individually resuspended in dimethyl sulfoxide (DMSO) at a concentration of 10–20 mg/mL. Megapools (MP) for each antigen were created by pooling aliquots of these individual peptides, undergoing another lyophilization, and resuspending in DMSO at 1 mg/mL.

**Bioinformatic analysis of T cell epitopes**

The binding capacity of SARS-CoV-2 T cell epitopes, and their corresponding variant-derived peptides, for their putative HLA class I restricting allele(s) was determined utilizing the NetMHCpan BA 4.1 algorithm (Reynisson et al., 2020), as implemented by the IEDB's analysis resource (Dhanda et al., 2019; Vita et al., 2019). Predicted binding is expressed in terms of $IC_{50}$ nM. For each epitope-variant pair a ratio of affinities (WT/variant) was determined. Ratios >2, indicating a 2-fold or greater increase in affinity due to the mutation, were categorized as an increase in binding capacity, and <0.5 as a decrease; ratios between 0.5 and 2 were designated as neutral.

**Flow cytometry-based AIM assay**

Activation induced cell marker (AIM) assay has previously been described in detail elsewhere (da Silva Antunes et al., 2021; Dan et al., 2021; Reiss et al., 2017). In summary, PBMCs were cultured for in the presence of SARS-CoV-2 specific (Wuhan or variant) MPs [1 μg/ml] in 96-well U-bottom plates at a concentration of $1x10^6$ PBMC per well. As a negative control, an equimolar amount of DMSO was used to stimulate the cells in triplicate wells and as positive controls phytohemagglutinin (PHA, Roche, 1μg/ml) and a cytomegalovirus MP (CMV, combining CD4 and CD8 MPs, 1μg/ml) were also included. After incubation for 20–24 hours at 37°C, 5% $CO_2$, the cells were stained with CD3 BUV805 or CD3 AF700 (4:100 or 4:100; BD Biosciences Cat# 612895 or Life Technologies Cat# 56-0038-42, respectively), CD4 BV605 (4:100; BD Biosciences Cat# 562658), CD8 BUV496 or BV650 (2:100 or 4:100; BD Biosciences Cat# 612942 or Biolegend Cat# 301042), and Live/Dead eFluor506 (5:1000; eBioscience Cat# 65-0866-14). Cells were also stained to measure activation with the following markers: CD137 APC (4:100; Biolegend Cat# 309810), OX40 PE-Cy7 (2:100; Biolegend Cat#350012), and CD69 PE (10:100; BD Biosciences Cat# 555531). All samples were acquired on a ZE5 5-laser or 4-laser cell analyzer (Bio-rad laboratories) and analyzed with FlowJo software (Tree Star). In the resulting data generated from the AIM assays, the background was removed from the data by subtracting the average of the % of $AIM^+$

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cells plated in triplicate wells stimulated with DMSO. The Stimulation Index (SI) was calculated by dividing the % of $AIM^+$ cells after SARS-CoV-2 stimulation with the average % of $AIM^+$ cells in the negative DMSO control. An SI greater than 2 and a minimum of 0.02 % or 0.03 % $AIM^+$ $CD4^+$ or $CD8^+$ cells, respectively, after background subtraction was considered to be a positive response. The gates for $AIM^+$ cells were drawn relative to the negative and positive controls for each donor. A representative example of the gating strategy is depicted in **Fig. S3.**

### FluoroSPOT assays

96-well FluoroSpot plates were coated with anti-cytokine antibodies for IFNγ and IL-5 (mAbs 1-D1K and TRFK5, respectively; Mabtech, Stockholm, Sweden) at a concentration of 10µg/mL. PBMCs were stimulated in triplicate at a density of $200x10^3$ cells/well with S MPs corresponding to each of the SARS-CoV-2 variants analyzed (1µg/mL), PHA (1µg/mL), and DMSO (0.1%), as positive and negative controls respectively. After 20 hours of incubation at 37°C, 5% $CO_2$, cells were discarded and plates were washed before the addition of cytokine antibodies (mAbs 7-B6-1-BAM and 5A10-WASP; Mabtech, Stockholm, Sweden). After a 2-hour incubation, plates were washed again with PBS/0.05% Tween20 and incubated for 1 hour with fluorophore-conjugated antibodies (Anti-BAM-490 and Anti-WASP-640). An AID iSPOT FluoroSpot reader (AIS-diagnostika, Germany) was used to count the fluorescent spots that resulted from cells secreting IFNγ and IL-5. Each peptide MP was considered positive compared to the DMSO negative control based on the following criteria: 20 or more spot forming cells (SFC) per $10^6$ PBMC after subtraction, a stimulation index (S.I.) greater than 2, and a p value <0.05 by either a Poisson or T test calculated between the triplicates of the MP and the relative negative control.

## QUANTIFICATION AND STATISTICAL ANALYSIS

Data and statistical analyses were performed in FlowJo 10 and GraphPad Prism 8.4, unless otherwise stated. Statistical details of the experiments are provided in the respective figure legends and in each method section pertaining the specific technique applied. Data plotted in logarithmic scales are expressed as geometric mean. Statistical analyses were performed using Wilcoxon matched-pairs signed rank test for paired comparisons. Multi-hypothesis testing corrections (MHTC) have not been applied in the study by design. The study is not designed or powered to address differences across different proteins. The primary hypothesis is that no significant differences are observed across the different variants, and this is more stringently addressed avoiding to correct for MHTC, since a difference that is not significant would remain so even after corrections. Therefore, reporting the data without applying MHTC is a more stringent criterion which is appropriately being applied in this case to avoid false negatives. Details pertaining to significance are also noted in the respective figure legends.

## DECLARATION OF INTEREST

A.S. is a consultant for Gritstone, Flow Pharma, Oxford Immunetech, Caprion, Arcturus and Avalia. S.C. is a consultant for Avalia. All other authors declare no conflict of interest. LJI has filed for patent protection for various aspects of vaccine design and identification of specific epitopes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## REFERENCES

Amit, S., Regev-Yochay, G., Afek, A., Kreiss, Y., and Leshem, E. (2021). Early rate reductions of SARS-CoV-2 infection and COVID-19 in BNT162b2 vaccine recipients. Lancet.

Andreano, E., Piccini, G., Licastro, D., Casalino, L., Johnson, N.V., Paciello, I., Monego, S.D., Pantano, E., Manganaro, N., Manenti, A.*, et al.* (2020). SARS-CoV-2 escape in vitro from a highly neutralizing COVID-19 convalescent plasma. bioRxiv, 2020.2012.2028.424451.

Andrews, S.F., Huang, Y., Kaur, K., Popova, L.I., Ho, I.Y., Pauli, N.T., Henry Dunand, C.J., Taylor, W.M., Lim, S., Huang, M.*, et al.* (2015). Immune history profoundly affects broadly protective B cell responses to influenza. Sci Transl Med *7*, 316ra192.

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.*, et al.* (2021). Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. N Engl J Med *384*, 403-416.

Braun, J., Loyal, L., Frentsch, M., Wendisch, D., Georg, P., Kurth, F., Hippenstiel, S., Dingeldey, M., Kruse, B., Fauchere, F.*, et al.* (2020). SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. Nature.

Breton, G., Mendoza, P., Hagglof, T., Oliveira, T.Y., Schaefer-Babajew, D., Gaebler, C., Turroja, M., Hurley, A., Caskey, M., and Nussenzweig, M.C. (2021). Persistent cellular immunity to SARS-CoV-2 infection. J Exp Med *218*.

Cele, S., Gazy, I., Jackson, L., Hwa, S.-H., Tegally, H., Lustig, G., Giandhari, J., Pillay, S., Wilkinson, E., Naidoo, Y.*, et al.* (2021). Escape of SARS-CoV-2 501Y.V2 variants from neutralization by convalescent plasma. medRxiv, 2021.2001.2026.21250224.

da Silva Antunes, R., Pallikkuth, S., Williams, E., Yu, E.D., Mateus, J., Quiambao, L., Wang, E., Rawlings, S.A., Stadlbauer, D., Jiang, K.*, et al.* (2021). Differential T cell reactivity to seasonal coronaviruses and SARS-CoV-2 in community and health care workers. medRxiv.

Dan, J.M., Mateus, J., Kato, Y., Hastie, K.M., Yu, E.D., Faliti, C.E., Grifoni, A., Ramirez, S.I., Haupt, S., Frazier, A.*, et al.* (2021). Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. Science *371*.

Davies, N.G., Barnard, R.C., Jarvis, C.I., Kucharski, A.J., Munday, J., Pearson, C.A.B., Russell, T.W., Tully, D.C., Abbott, S., Gimma, A.*, et al.* (2020). Estimated transmissibility and severity of novel SARS-CoV-2 Variant of Concern 202012/01 in England. medRxiv, 2020.2012.2024.20248822.

Dhanda, S.K., Mahajan, S., Paul, S., Yan, Z., Kim, H., Jespersen, M.C., Jurtz, V., Andreatta, M., Greenbaum, J.A., Marcatili, P.*, et al.* (2019). IEDB-AR: immune epitope database-analysis resource in 2019. Nucleic Acids Res *47*, W502-W506.

Doud, M.B., Lee, J.M., and Bloom, J.D. (2018). How single mutations affect viral escape from broad and narrow antibodies to H1 influenza hemagglutinin. Nature communications *9*, 1386.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Dowd, S.M., Zalta, A.K., Burgess, H.J., Adkins, E.C., Valdespino-Hayden, Z., and Pollack, M.H. (2020). Double-blind randomized controlled study of the efficacy, safety and tolerability of eszopiclone vs placebo for the treatment of patients with post-traumatic stress disorder and insomnia. World journal of psychiatry *10*, 21-28.

Edara, V.V., Floyd, K., Lai, L., Gardner, M., Hudson, W., Piantadosi, A., Waggoner, J.J., Babiker, A., Ahmed, R., Xie, X.*, et al.* (2021). Infection and mRNA-1273 vaccine antibodies neutralize SARS-CoV-2 UK variant. medRxiv, 2021.2002.2002.21250799.

Emary, K., Golubchik, T., Aley, P., Ariani, C., Angus, B.J., Bibi, S., Blane, B., Bonsall, D., Cicconi, P., Charlton, S.*, et al.* (2021). Efficacy of ChAdOx1 nCoV-19 (AZD1222) Vaccine Against SARS-CoV-2 VOC 202012/01 (B.1.1.7) (SSRN).

Evavold, B.D., and Allen, P.M. (1991). Separation of IL-4 production from Th cell proliferation by an altered T cell receptor ligand. Science *252*, 1308-1310.

FDA (2021a). Vaccines and Related Biological Products Advisory Committee February 26, 2021 Meeting Briefing Document Addendum- Sponsor

FDA (2021b). Vaccines and Related Biological Products Advisory Committee February 26, 2021 Meeting Briefing Document- Sponsor.

Ferretti, A.P., Kula, T., Wang, Y., Nguyen, D.M.V., Weinheimer, A., Dunlap, G.S., Xu, Q., Nabilsi, N., Perullo, C.R., Cristofaro, A.W.*, et al.* (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.

Gaiha, G.D., Rossin, E.J., Urbach, J., Landeros, C., Collins, D.R., Nwonu, C., Muzhingi, I., Anahtar, M.N., Waring, O.M., Piechocka-Trocha, A.*, et al.* (2019). Structural topology defines protective CD8(+) T cell epitopes in the HIV proteome. Science *364*, 480-484.

Greaney, A.J., Loes, A.N., Crawford, K.H.D., Starr, T.N., Malone, K.D., Chu, H.Y., and Bloom, J.D. (2021). Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. Cell host & microbe.

Greenbaum, J.A., Kotturi, M.F., Kim, Y., Oseroff, C., Vaughan, K., Salimi, N., Vita, R., Ponomarenko, J., Scheuermann, R.H., Sette, A.*, et al.* (2009). Pre-existing immunity against swine-origin H1N1 influenza viruses in the general human population. Proc Natl Acad Sci U S A *106*, 20365-20370.

Grifoni, A., Voic, H., Dhanda, S.K., Kidd, C.K., Brien, J.D., Buus, S., Stryhn, A., Durbin, A.P., Whitehead, S., and Diehl, S.A. (2020a). T cell responses induced by attenuated flavivirus vaccination are specific and show limited cross-reactivity with other flavivirus species. Journal of virology *94*.

Grifoni, A., Weiskopf, D., Ramirez, S.I., Mateus, J., Dan, J.M., Moderbacher, C.R., Rawlings, S.A., Sutherland, A., Premkumar, L., Jadi, R.S.*, et al.* (2020b). Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. Cell.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Hall, V.J., Foulkes, S., Saei, A., Andrews, N., ,, Oguti, B., Charlett, A., Wellington, E., Stowe, J., Gillson, N., Atti, A.*, et al.* (2021). Effectiveness of BNT162b2 mRNA Vaccine Against Infection and COVID-19 Vaccine Coverage in Healthcare Workers in England, Multicentre Prospective Cohort Study (the SIREN Study). http://dx.doi.org/10.2139/ssrn.3790399. Ssrn.

Hensley, S.E., Das, S.R., Bailey, A.L., Schmidt, L.M., Hickman, H.D., Jayaraman, A., Viswanathan, K., Raman, R., Sasisekharan, R., Bennink, J.R.*, et al.* (2009). Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. Science *326*, 734-736.

Keech, C., Albert, G., Cho, I., Robertson, A., Reed, P., Neal, S., Plested, J.S., Zhu, M., Cloney-Clark, S., Zhou, H.*, et al.* (2020). Phase 1-2 Trial of a SARS-CoV-2 Recombinant Spike Protein Nanoparticle Vaccine. N Engl J Med *383*, 2320-2332.

Keller, M.D., Harris, K.M., Jensen-Wachspress, M.A., Kankate, V., Lang, H., Lazarski, C.A., Durkee-Shock, J.R., Lee, P.H., Chaudhry, K., Webber, K.*, et al.* (2020). SARS-CoV-2 specific T-cells Are Rapidly Expanded for Therapeutic Use and Target Conserved Regions of Membrane Protein. Blood.

Kirby, T. (2021). New variant of SARS-CoV-2 in UK causes surge of COVID-19. The Lancet Respiratory medicine *9*, e20-e21.

Krammer, F., Garcia-Sastre, A., and Palese, P. (2018). Is It Possible to Develop a "Universal" Influenza Virus Vaccine? Potential Target Antigens and Critical Aspects for a Universal Influenza Vaccine. Cold Spring Harbor perspectives in biology *10*.

Le Bert, N., Tan, A.T., Kunasegaran, K., Tham, C.Y.L., Hafezi, M., Chia, A., Chng, M.H.Y., Lin, M., Tan, N., Linster, M.*, et al.* (2020). SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. Nature.

Lipsitch, M., Grad, Y.H., Sette, A., and Crotty, S. (2020). Cross-reactive memory T cells and herd immunity to SARS-CoV-2. Nat Rev Immunol *20*, 709-713.

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.*, et al.* (2020). Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. Science *370*, 89-94.

Muik, A., Wallisch, A.K., Sanger, B., Swanson, K.A., Muhl, J., Chen, W., Cai, H., Maurus, D., Sarkar, R., Tureci, O.*, et al.* (2021). Neutralization of SARS-CoV-2 lineage B.1.1.7 pseudovirus by BNT162b2 vaccine-elicited human sera. Science.

Munoz-Fontela, C., Dowling, W.E., Funnell, S.G.P., Gsell, P.S., Riveros-Balta, A.X., Albrecht, R.A., Andersen, H., Baric, R.S., Carroll, M.W., Cavaleri, M.*, et al.* (2020). Animal models for COVID-19. Nature *586*, 509-515.

Nelde, A., Bilich, T., Heitmann, J.S., Maringer, Y., Salih, H.R., Roerden, M., Lubke, M., Bauer, J., Rieth, J., Wacker, M.*, et al.* (2020). SARS-CoV-2-derived peptides define heterologous and COVID-19-induced T cell recognition. Nat Immunol.

Novavax Inc. (2021). Novavax COVID-19 Vaccine Demonstrates 89.3% Efficacy in UK Phase 3 Trial.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Peeples, L. (2020). News Feature: Avoiding pitfalls in the pursuit of a COVID-19 vaccine. Proc Natl Acad Sci U S A *117*, 8218-8221.

Peng, Y., Mentzer, A.J., Liu, G., Yao, X., Yin, Z., Dong, D., Dejnirattisai, W., Rostron, T., Supasa, P., Liu, C.*, et al.* (2020). Broad and strong memory CD4(+) and CD8(+) T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. Nat Immunol *21*, 1336-1345.

Pickett, B.E., Sadat, E.L., Zhang, Y., Noronha, J.M., Squires, R.B., Hunt, V., Liu, M., Kumar, S., Zaremba, S., Gu, Z.*, et al.* (2012). ViPR: an open bioinformatics database and analysis resource for virology research. Nucleic Acids Res *40*, D593-598.

Rambaut, A., Loman, N., Pybus, O., Barclay, W., Barrett, J., Carabelli, A., Connor, T., Peacock, T., Robertson, D.L., and E, V. (2020). Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. https://virologicalorg/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563.

Reiss, S., Baxter, A.E., Cirelli, K.M., Dan, J.M., Morou, A., Daigneault, A., Brassard, N., Silvestri, G., Routy, J.P., Havenar-Daughton, C.*, et al.* (2017). Comparative analysis of activation induced marker (AIM) assays for sensitive identification of antigen-specific CD4 T cells. PLoS One *12*, e0186998.

Reynisson, B., Alvarez, B., Paul, S., Peters, B., and Nielsen, M. (2020). 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 *48*, W449-W454.

Rydyznski Moderbacher, C., Ramirez, S.I., Dan, J.M., Grifoni, A., Hastie, K.M., Weiskopf, D., Belanger, S., Abbott, R.K., Kim, C., Choi, J.*, et al.* (2020). Antigen-Specific Adaptive Immunity to SARS-CoV-2 in Acute COVID-19 and Associations with Age and Disease Severity. Cell.

Sadoff, J., Le Gars, M., Shukarev, G., Heerwegh, D., Truyers, C., de Groot, A.M., Stoop, J., Tete, S., Van Damme, W., Leroux-Roels, I.*, et al.* (2021). Interim Results of a Phase 1-2a Trial of Ad26.COV2.S Covid-19 Vaccine. N Engl J Med.

Sette, A., and Crotty, S. (2021). Adaptive immunity to SARS-CoV-2 and COVID-19. Cell *184*, 861-880.

Shen, X., Tang, H., McDanal, C., Wagh, K., Fischer, W., Theiler, J., Yoon, H., Li, D., Haynes, B.F., Sanders, K.O.*, et al.* (2021). SARS-CoV-2 variant B.1.1.7 is susceptible to neutralizing antibodies elicited by ancestral Spike vaccines. bioRxiv.

Skelly, D.T., Harding, A.C., Gilbert-Jaramillo, J., Knight, M.L., Longet, S., Brown, A., Adele, S., Adland, E., Brown, H., Medawar Laboratory Team.*, et al.* (2020). Vaccine-induced immunity provides more robust heterotypic immunity than natural infection to emerging SARS-CoV-2 variants of concern. Research square.

Sloan-Lancaster, J., and Allen, P.M. (1996). Altered peptide ligand-induced partial T cell activation: molecular mechanisms and role in T cell biology. Annu Rev Immunol *14*, 1-27.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Snyder, T.M., Gittelman, R.M., Klinger, M., May, D.H., Osborne, E.J., Taniguchi, R., Zahid, H.J., Kaplan, I.M., Dines, J.N., Noakes, M.N.*, et al.* (2020). Magnitude and Dynamics of the T-Cell Response to SARS-CoV-2 Infection at Both Individual and Population Levels. medRxiv.

Soresina, A., Moratto, D., Chiarini, M., Paolillo, C., Baresi, G., Foca, E., Bezzi, M., Baronio, B., Giacomelli, M., and Badolato, R. (2020). Two X-linked agammaglobulinemia patients develop pneumonia as COVID-19 manifestation but recover. Pediatric allergy and immunology : official publication of the European Society of Pediatric Allergy and Immunology *31*, 565-569.

Sridhar, S., Begom, S., Bermingham, A., Hoschler, K., Adamson, W., Carman, W., Bean, T., Barclay, W., Deeks, J.J., and Lalvani, A. (2013). Cellular immune correlates of protection against symptomatic pandemic influenza. Nat Med *19*, 1305-1312.

Stadlbauer, D., Amanat, F., Chromikova, V., Jiang, K., Strohmeier, S., Arunkumar, G.A., Tan, J., Bhavsar, D., Capuano, C., Kirkpatrick, E.*, et al.* (2020). SARS-CoV-2 Seroconversion in Humans: A Detailed Protocol for a Serological Assay, Antigen Production, and Test Setup. Curr Protoc Microbiol *57*, e100.

Stamatatos, L., Czartoski, J., Wan, Y.-H., Homad, L.J., Rubin, V., Glantz, H., Neradilek, M., Seydoux, E., Jennewein, M.F., MacCamy, A.J.*, et al.* (2021). Antibodies elicited by SARS-CoV-2 infection and boosted by vaccination neutralize an emerging variant and SARS-CoV-1. medRxiv, 2021.2002.2005.21251182.

Starr, T.N., Greaney, A.J., Addetia, A., Hannon, W.W., Choudhary, M.C., Dingens, A.S., Li, J.Z., and Bloom, J.D. (2021). Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. Science.

Starr, T.N., Greaney, A.J., Hilton, S.K., Ellis, D., Crawford, K.H.D., Dingens, A.S., Navarro, M.J., Bowen, J.E., Tortorici, M.A., Walls, A.C.*, et al.* (2020). Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. Cell *182*, 1295-1310 e1220.

Supasa, P., Zhou, D., Dejnirattisai, W., Liu, C., Mentzer, A.J., Ginn, H.M., Zhao, Y., Duyvesteyn, H.M.E., Nutalai, R., Tuekprakhon, A.*, et al.* (2021). Reduced neutralization of SARS-CoV-2 B.1.1.7 variant by convalescent and vaccine sera. Cell.

Tan, A.T., Linster, M., Tan, C.W., Le Bert, N., Chia, W.N., Kunasegaran, K., Zhuang, Y., Tham, C.Y.L., Chia, A., Smith, G.J.D.*, et al.* (2021). Early induction of functional SARS-CoV-2-specific T cells associates with rapid viral clearance and mild disease in COVID-19 patients. Cell reports, 108728.

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.*, et al.* (2021). Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases. Cell reports Medicine *2*, 100204.

Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., Doolabh, D., Pillay, S., San, E.J., Msomi, N.*, et al.* (2020). Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. medRxiv, 2020.2012.2021.20248640.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Vita, R., Mahajan, S., Overton, J.A., Dhanda, S.K., Martini, S., Cantrell, J.R., Wheeler, D.K., Sette, A., and Peters, B. (2019). The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Res *47*, D339-D343.

Voloch, C.M., Silva F, R.d., de Almeida, L.G.P., Cardoso, C.C., Brustolini, O.J., Gerber, A.L., Guimarães, A.P.d.C., Mariani, D., Costa, R.M.d., Ferreira, O.C.*, et al.* (2020). Genomic characterization of a novel SARS-CoV-2 lineage from Rio de Janeiro, Brazil. medRxiv, 2020.2012.2023.20248598.

Volz, E., Mishra, S., Chand, M., Barrett, J.C., Johnson, R., Geidelberg, L., Hinsley, W.R., Laydon, D.J., Dabrera, G., O'Toole, Á.*, et al.* (2021). Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insights from linking epidemiological and genetic data. medRxiv, 2020.2012.2030.20249034.

Voysey, M., Clemens, S.A.C., Madhi, S.A., Weckx, L.Y., Folegatti, P.M., Aley, P.K., Angus, B., Baillie, V.L., Barnabas, S.L., Bhorat, Q.E.*, et al.* (2021). Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. Lancet *397*, 99-111.

Wang, P., Liu, L., Iketani, S., Luo, Y., Guo, Y., Wang, M., Yu, J., Zhang, B., Kwong, P.D., Graham, B.S.*, et al.* (2021a). Increased Resistance of SARS-CoV-2 Variants B.1.351 and B.1.1.7 to Antibody Neutralization. bioRxiv, 2021.2001.2025.428137.

Wang, Z., Schmidt, F., Weisblum, Y., Muecksch, F., Barnes, C.O., Finkin, S., Schaefer-Babajew, D., Cipolla, M., Gaebler, C., Lieberman, J.A.*, et al.* (2021b). mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. bioRxiv, 2021.2001.2015.426911.

Washington, N.L., Gangavarapu, K., Zeller, M., Bolze, A., Cirulli, E.T., Barrett, K.M.S., Larsen, B.B., Anderson, C., White, S., Cassens, T.*, et al.* (2021). Genomic epidemiology identifies emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. medRxiv, 2021.2002.2006.21251159.

Weiskopf, D., Angelo, M.A., Sidney, J., Peters, B., Shresta, S., and Sette, A. (2014). Immunodominance changes as a function of the infecting dengue virus serotype and primary versus secondary infection. J Virol *88*, 11383-11394.

Wibmer, C.K., Ayres, F., Hermanus, T., Madzivhandila, M., Kgagudi, P., Lambson, B.E., Vermeulen, M., van den Berg, K., Rossouw, T., Boswell, M.*, et al.* (2021). SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. bioRxiv, 2021.2001.2018.427166.

Wilkinson, T.M., Li, C.K., Chui, C.S., Huang, A.K., Perkins, M., Liebner, J.C., Lambkin-Williams, R., Gilbert, A., Oxford, J., Nicholas, B.*, et al.* (2012). Preexisting influenza-specific CD4+ T cells correlate with disease protection against influenza challenge in humans. Nat Med *18*, 274-280.

Wu, K., Werner, A.P., Koch, M., Choi, A., Narayanan, E., Stewart-Jones, G.B.E., Colpitts, T., Bennett, H., Boyoglu-Barnum, S., Shi, W.*, et al.* (2021). Serum Neutralizing Activity Elicited by mRNA-1273 Vaccine - Preliminary Report. N Engl J Med.

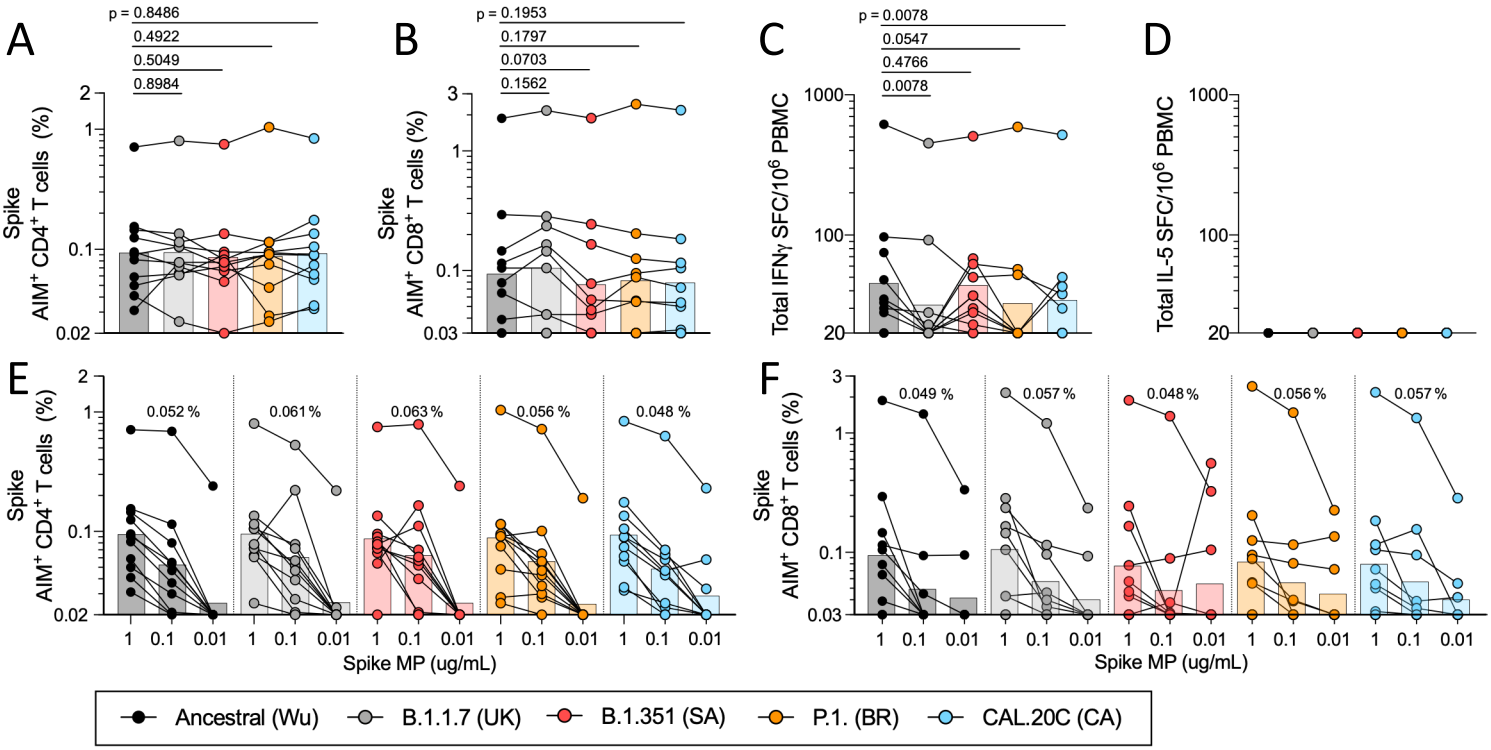
bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Zahradník, J., Marciano, S., Shemesh, M., Zoler, E., Chiaravalli, J., Meyer, B., Dym, O., Elad, N., and Schreiber, G. (2021). SARS-CoV-2 RBD in vitro evolution follows contagious mutation spread, yet generates an able infection inhibitor. bioRxiv, 2021.2001.2006.425392.

Zhang, W., Davis, B.D., Chen, S.S., Martinez, J.M.S., Plummer, J.T., and Vail, E. (2021). Emergence of a novel SARS-CoV-2 strain in Southern California, USA. medRxiv, 2021.2001.2018.21249786.

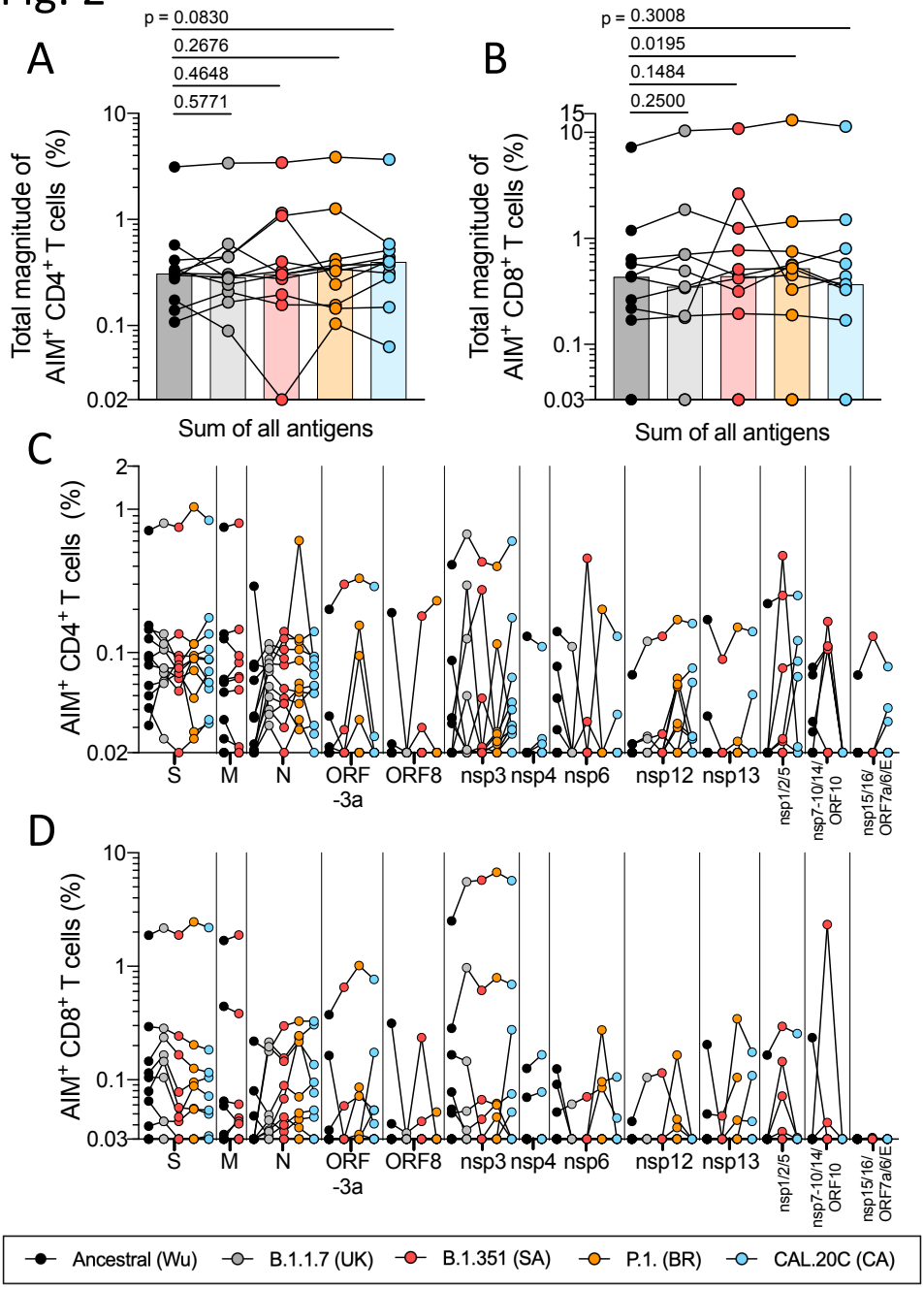
bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# Fig. 1



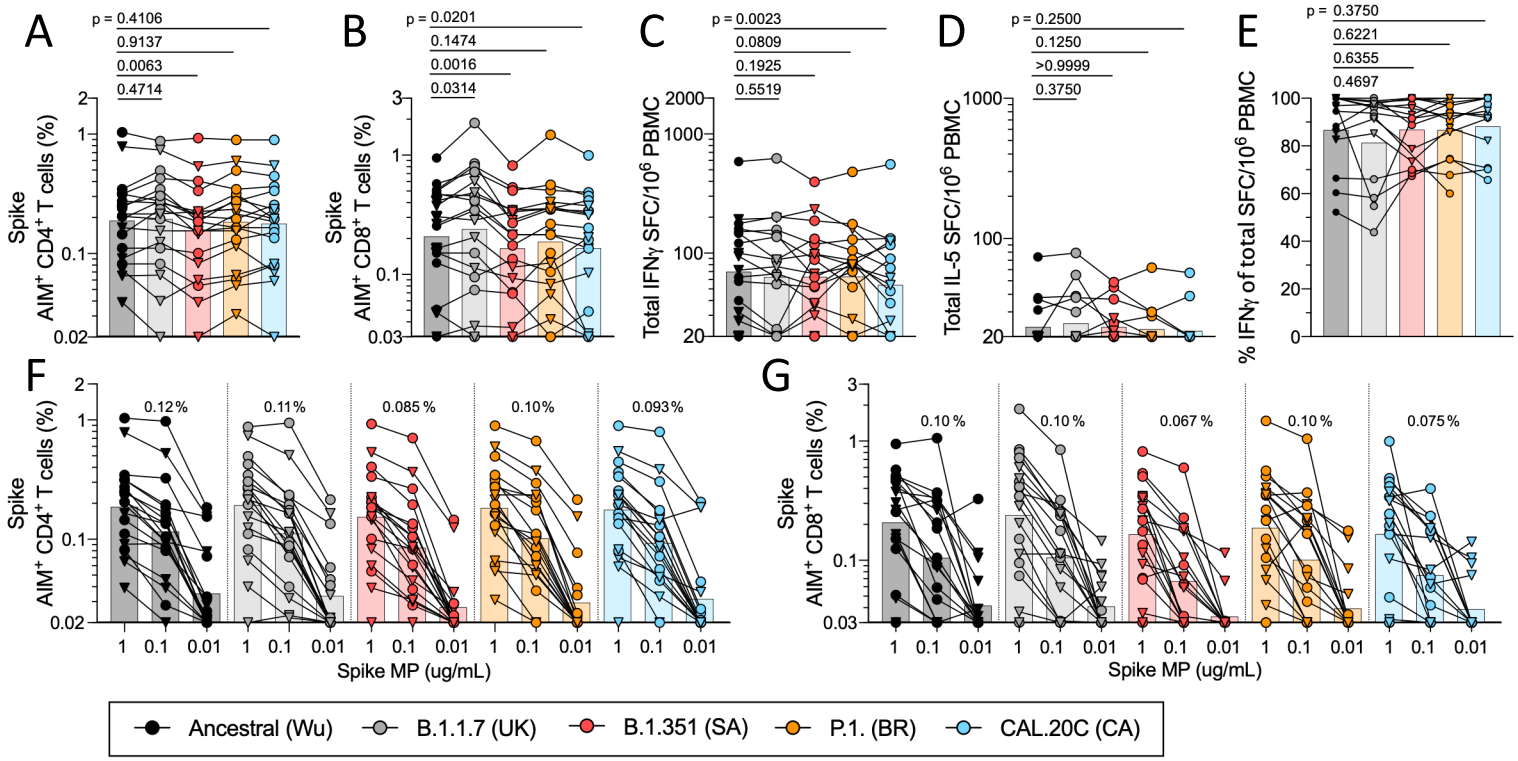
bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. 2



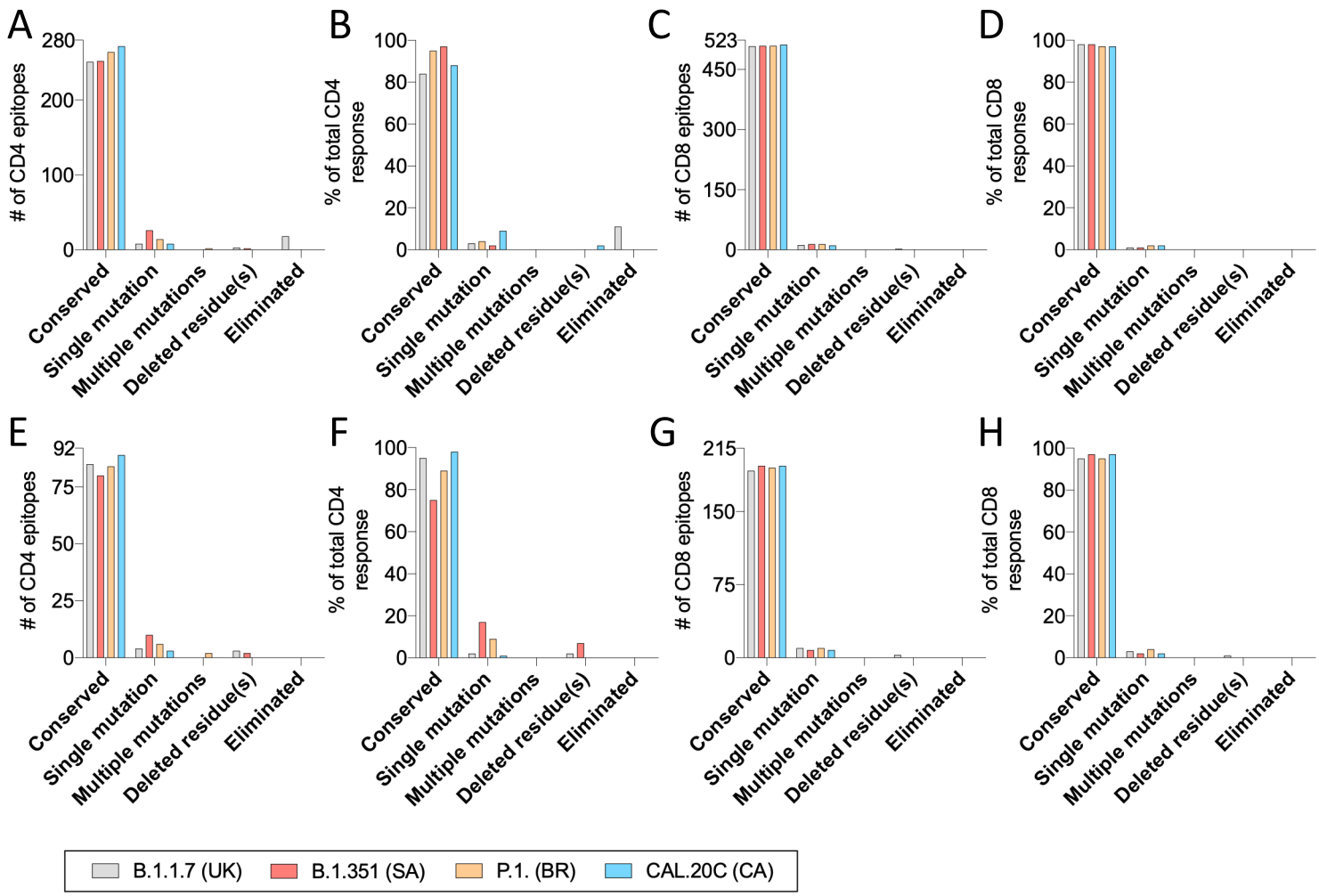
bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. 3



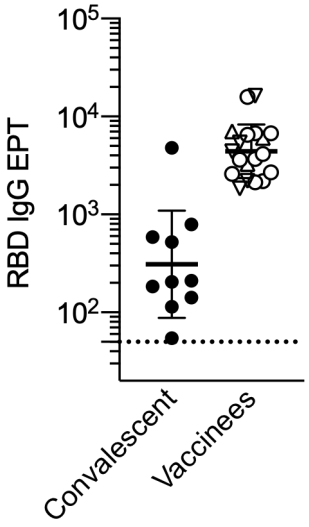
bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. 4



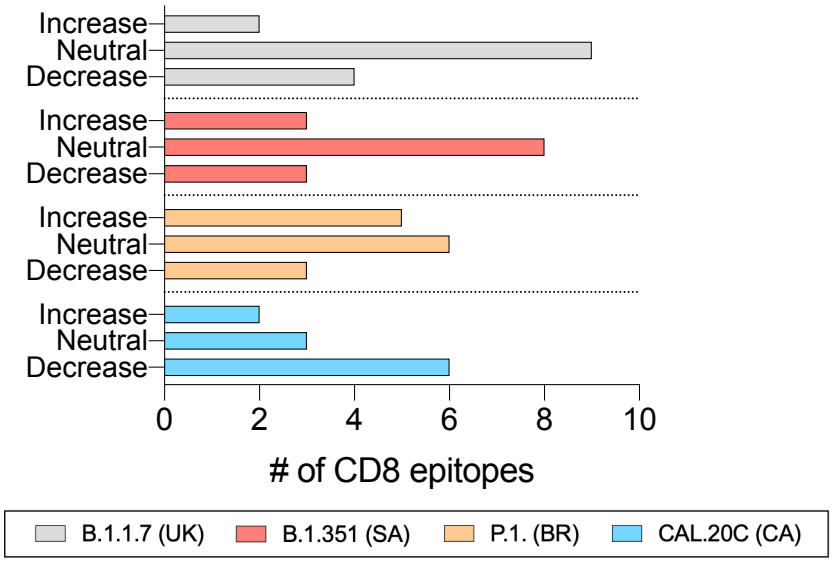
bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. S1



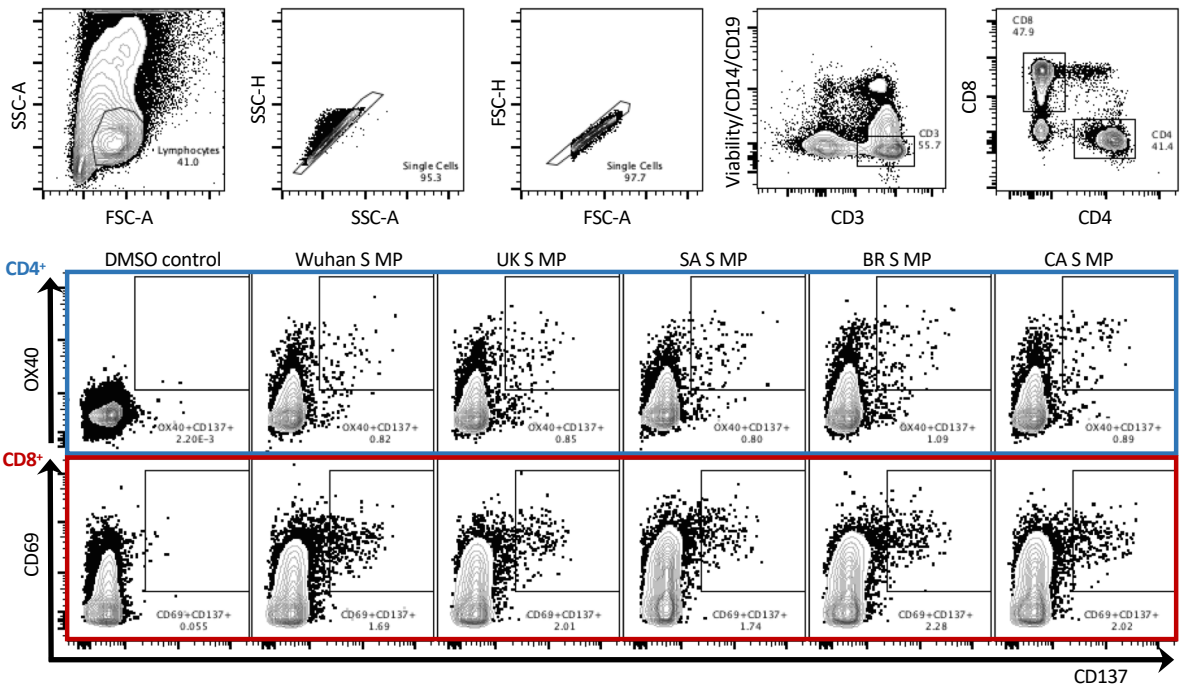
bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. S2



bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# Fig. S3



bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**Table S1**. Related to Figures 1, 2, 3 and 4. List of amino acid positions and relative amino acid changes in the different variants studied with respect to the ancestral Wuhan strain.

| Protein | Amino acid position | Ancestral (Wu) | B.1.1.7 (UK) | B.1.351 (SA) | P.1. (BR) | CAL.20C (CA) |
|---|---|---|---|---|---|---|
| S | 13 | S | | | | I |
| S | 18 | L | | F | F | |
| S | 20 | T | | | N | |
| S | 26 | P | | | S | |
| S | 69 | H | Del | | | |
| S | 70 | V | Del | | | |
| S | 80 | D | | A | | |
| S | 138 | D | | | Y | |
| S | 145 | Y | Del | | | |
| S | 152 | W | | | | C |
| S | 190 | R | | | S | |
| S | 215 | D | | G/H | | |
| S | 241 | L | | Del | | |
| S | 242 | L | | Del | | |
| S | 243 | A | | Del | | |
| S | 417 | K | | N | T | |
| S | 452 | L | | | | R |
| S | 484 | E | | K | K | |
| S | 501 | N | Y | Y | Y | |
| S | 570 | A | D | | | |
| S | 614 | D | G | G | G | G |
| S | 655 | H | | | Y | |
| S | 681 | P | H | | | |
| S | 701 | A | | V | | |
| S | 716 | T | I | | | |
| S | 938 | L | | | | F |
| S | 982 | S | A | | | |
| S | 1027 | T | | | I | |
| S | 1118 | D | H | | | |
| S | 1176 | V | | | F | |
| S | 1191 | K | | | | N |
| M | 162 | K | | N | | |
| N | 3 | D | L | | | |
| N | 13 | P | | S | | |
| N | 32 | R | | H | | |
| N | 80 | P | | | R | |
| N | 203 | R | K | | K | K |
| N | 204 | G | R | | R | R |
| N | 205 | T | | I | | I |
| N | 212 | G | | C | | |
| N | 234 | M | | | | I |
| N | 235 | S | F | | | |
| E | 71 | P | | L | | |

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

| | | | | | | |
|---|---|---|---|---|---|---|
| ORF3a | 57 | Q | | H | | H |
| ORF3a | 131 | W | | L | | |
| ORF3a | 171 | S | | L | | |
| ORF3a | 253 | S | | | P | |
| ORF7a | 93 | V | | F | | |
| ORF8 | 27 | - | Stop | | | |
| ORF8 | 92 | E | | | K | |
| ORF8 | 121 | I | | L | | |
| nsp1 | 109 | P | | S | | |
| nsp2 | 85 | T | | I | | I |
| nsp2 | 339 | G | | S | | |
| nsp2 | 366 | S | | T | | |
| nsp2 | 427 | Q | | H | | |
| nsp2 | 563 | E | | D | | |
| nsp3 | 183 | T | I | | | |
| nsp3 | 186 | T | | | A | |
| nsp3 | 370 | S | | | L | |
| nsp3 | 778 | P | | | | S |
| nsp3 | 837 | K | | N | | |
| nsp3 | 890 | A | D | | | |
| nsp3 | 926 | C | | S | | |
| nsp3 | 977 | K | | | Q | |
| nsp3 | 1180 | T | | I | | |
| nsp3 | 1412 | I | T | | | |
| nsp3 | 1778 | N | | S | | |
| nsp4 | 395 | S | | | | T |
| nsp5 | 90 | K | | R | | |
| nsp5 | 193 | A | | V | | |
| nsp6 | 106 | S | Del | Del | Del | |
| nsp6 | 107 | G | Del | Del | Del | |
| nsp6 | 108 | F | Del | Del | Del | |
| nsp6 | 125 | L | | | | F |
| nsp6 | 135 | G | | S | | |
| nsp6 | 149 | V | | F | | |
| nsp6 | 167 | L | | | | F |
| nsp9 | 65 | I | | | | V |
| nsp10 | 105 | N | | K | | |
| nsp12 | 323 | P | L | L | L | L |
| nsp13 | 53 | P | | | | L |
| nsp13 | 209 | V | | | | F |
| nsp13 | 260 | D | | | | Y |
| nsp13 | 341 | E | | | D | |
| nsp13 | 588 | T | | I | | |
| nsp14 | 177 | L | | F | | |
| nsp14 | 326 | F | | | | L |
| nsp14 | 328 | V | | F | | |
| nsp15 | 91 | D | | | | Y |

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**Table S2**: Related to Figures 1, 2, 3 and 4. List of mutated peptides with respect to the ancestral Wuhan strain in the different variants studied.

| Protein | Aminoacid position (Start) | Aminoacid position (End) | SARS-CoV-2 strain | Sequence |
|---|---|---|---|---|
| nsp3 | 173 | 187 | B.1.1.7 | QDGSEDNQTTIIQTI |
| nsp3 | 178 | 192 | B.1.1.7 | DNQTTIIQTIVEVQP |
| nsp3 | 183 | 197 | B.1.1.7 | IIQTIVEVQPQLEME |
| nsp3 | 878 | 892 | B.1.1.7 | QDAYYRARAGEADNF |
| nsp3 | 883 | 897 | B.1.1.7 | RARAGEADNFCALIL |
| nsp3 | 888 | 902 | B.1.1.7 | EADNFCALILAYCNK |
| nsp3 | 1398 | 1412 | B.1.1.7 | NYLKSPNFSKLINIT |
| nsp3 | 1403 | 1417 | B.1.1.7 | PNFSKLINITIWFLL |
| nsp3 | 1408 | 1422 | B.1.1.7 | LINITIWFLLLSVCL |
| nsp6 | 92 | 106 | B.1.1.7 | MRIMTWLDMVDTSLK |
| nsp6 | 97 | 111 | B.1.1.7 | WLDMVDTSLKLKDCV |
| nsp6 | 102 | 116 | B.1.1.7 | DTSLKLKDCVMYASA |
| nsp6 | 107 | 121 | B.1.1.7 | KLKDCVMYASAVVLL |
| nsp12 | 309 | 323 | B.1.1.7 | HCANFNVLFSTVFPL |
| nsp12 | 314 | 328 | B.1.1.7 | NVLFSTVFPLTSFGP |
| nsp12 | 319 | 333 | B.1.1.7 | TVFPLTSFGPLVRKI |
| N | 1 | 15 | B.1.1.7 | MSLNGPQNQRNAPRI |
| N | 191 | 205 | B.1.1.7 | RNSSRNSTPGSSKRT |
| N | 196 | 210 | B.1.1.7 | NSTPGSSKRTSPARM |
| N | 201 | 215 | B.1.1.7 | SSKRTSPARMAGNGG |
| N | 221 | 235 | B.1.1.7 | LLLLDRLNQLESKMF |
| N | 226 | 240 | B.1.1.7 | RLNQLESKMFGKGQQ |
| N | 231 | 245 | B.1.1.7 | ESKMFGKGQQQQGQT |
| ORF8 | 41 | 55 | B.1.1.7 | FYSKWYIRVGAIKSA |
| ORF8 | 46 | 60 | B.1.1.7 | YIRVGAIKSAPLIEL |
| ORF8 | 51 | 65 | B.1.1.7 | AIKSAPLIELCVDEA |
| ORF8 | 61 | 75 | B.1.1.7 | CVDEAGSKSPIQCID |
| ORF8 | 66 | 80 | B.1.1.7 | GSKSPIQCIDIGNYT |
| ORF8 | 71 | 85 | B.1.1.7 | IQCIDIGNYTVSCLP |
| S | 56 | 70 | B.1.1.7 | LPFFSNVTWFHAISG |
| S | 61 | 75 | B.1.1.7 | NVTWFHAISGTNGTK |
| S | 66 | 80 | B.1.1.7 | HAISGTNGTKRFDNP |
| S | 131 | 145 | B.1.1.7 | CEFQFCNDPFLGVYH |
| S | 136 | 150 | B.1.1.7 | CNDPFLGVYHKNNKS |
| S | 141 | 155 | B.1.1.7 | LGVYHKNNKSWMESE |
| S | 491 | 505 | B.1.1.7 | PLQSYGFQPTYGVGY |
| S | 496 | 510 | B.1.1.7 | GFQPTYGVGYQPYRV |
| S | 501 | 515 | B.1.1.7 | YGVGYQPYRVVVLSF |
| S | 556 | 570 | B.1.1.7 | NKKFLPFQQFGRDID |
| S | 561 | 575 | B.1.1.7 | PFQQFGRDIDDTTDA |
| S | 566 | 580 | B.1.1.7 | GRDIDDTTDAVRDPQ |
| S | 601 | 615 | B.1.1.7 | GTNTSNQVAVLYQGV |
| S | 606 | 620 | B.1.1.7 | NQVAVLYQGVNCTEV |
| S | 611 | 625 | B.1.1.7 | LYQGVNCTEVPVAIH |
| S | 671 | 685 | B.1.1.7 | CASYQTQTNSHRRAR |
| S | 676 | 690 | B.1.1.7 | TQTNSHRRARSVASQ |
| S | 681 | 695 | B.1.1.7 | HRRARSVASQSIIAY |
| S | 706 | 720 | B.1.1.7 | AYSNNSIAIPINFTI |
| S | 711 | 725 | B.1.1.7 | SIAIPINFTISVTTE |
| S | 716 | 730 | B.1.1.7 | INFTISVTTEILPVS |
| S | 971 | 985 | B.1.1.7 | GAISSVLNDILARLD |
| S | 976 | 990 | B.1.1.7 | VLNDILARLDKVEAE |
| S | 981 | 995 | B.1.1.7 | LARLDKVEAEVQIDR |
| S | 1106 | 1120 | B.1.1.7 | QRNFYEPQIITTHNT |
| S | 1111 | 1125 | B.1.1.7 | EPQIITTHNTFVSGN |
| S | 1116 | 1130 | B.1.1.7 | TTHNTFVSGNCDVVI |
| N | 1 | 15 | B.1.351 | MSDNGPQNQRNASRI |
| N | 6 | 20 | B.1.351 | PQNQRNASRITFGGP |
| N | 11 | 25 | B.1.351 | NASRITFGGPSDSTG |
| N | 21 | 35 | B.1.351 | SDSTGSNQNGEHSGA |
| N | 26 | 40 | B.1.351 | SNQNGEHSGARSKQR |
| N | 31 | 45 | B.1.351 | EHSGARSKQRRPQGL |
| N | 191 | 205 | B.1.351 | RNSSRNSTPGSSRGI |
| N | 196 | 210 | B.1.351 | NSTPGSSRGISPARM |
| N | 201 | 215 | B.1.351 | SSRGISPARMACNGG |
| N | 206 | 220 | B.1.351 | SPARMACNGGDAALA |
| N | 211 | 225 | B.1.351 | ACNGGDAALALLLLD |
| ORF8 | 107 | 121 | B.1.351 | DFLEYHDVRVVLDFL |
| ORF7a | 81 | 95 | B.1.351 | SVSPKLFIRQEEFQE |
| ORF7a | 86 | 100 | B.1.351 | LFIRQEEFQELYSPI |
| ORF7a | 91 | 105 | B.1.351 | EEFQELYSPIFLIVA |
| M | 151 | 165 | B.1.351 | IAGHHLGRCDINDLP |
| M | 156 | 170 | B.1.351 | LGRCDINDLPKEITV |
| M | 161 | 175 | B.1.351 | INDLPKEITVATSRT |
| ORF3a | 46 | 60 | B.1.351 | LIVGVALLAVFHSAS |
| ORF3a | 51 | 65 | B.1.351 | ALLAVFHSASKIITL |
| ORF3a | 56 | 70 | B.1.351 | FHSASKIITLKKRWQ |
| ORF3a | 121 | 135 | B.1.351 | VRIIMRLWLCLKCRS |
| ORF3a | 126 | 140 | B.1.351 | RLWLCLKCRSKNPLL |
| ORF3a | 131 | 145 | B.1.351 | LKCRSKNPLLYDANY |
| ORF3a | 161 | 175 | B.1.351 | NSVTSSIVITLGDGT |
| ORF3a | 166 | 180 | B.1.351 | SIVITLGDGTTSPIS |
| ORF3a | 171 | 185 | B.1.351 | LGDGTTSPISEHDYQ |
| nsp1 | 96 | 110 | B.1.351 | QYGRSGETLGVLVSH |
| nsp1 | 101 | 115 | B.1.351 | GETLGVLVSHVGEIP |
| nsp1 | 106 | 120 | B.1.351 | VLVSHVGEIPVAYRK |
| nsp5 | 183 | 197 | B.1.351 | GPFVDRQTAQVAGTD |
| nsp5 | 188 | 202 | B.1.351 | RQTAQVAGTDTTITV |
| nsp5 | 193 | 207 | B.1.351 | VAGTDTTITVNVLAW |
| nsp10 | 93 | 107 | B.1.351 | KGKYVQIPTTCAKDP |
| nsp10 | 98 | 112 | B.1.351 | QIPTTCAKDPVGFTL |
| nsp10 | 103 | 117 | B.1.351 | CAKDPVGFTLKNTVC |
| nsp12 | 309 | 323 | B.1.351 | HCANFNVLFSTVFPL |
| nsp12 | 314 | 328 | B.1.351 | NVLFSTVFPLTSFGP |
| nsp12 | 319 | 333 | B.1.351 | TVFPLTSFGPLVRKI |

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

| | | | | |
|---|---|---|---|---|
| nsp14 | 316 | 330 | B.1.351 | VVKAALLADKFPVLH |
| nsp14 | 321 | 335 | B.1.351 | ADKFPVLHDIGNPKA |
| nsp14 | 326 | 340 | B.1.351 | FPVLHDIGNPKAIKC |
| S | 6 | 20 | B.1.351 | VLLPLVSSQCVNFTT |
| S | 11 | 25 | B.1.351 | VSSQCVNFTTRTQLP |
| S | 16 | 30 | B.1.351 | VNFTTRTQLPPAYTN |
| S | 66 | 80 | B.1.351 | HAIHVSGTNGTKRFA |
| S | 71 | 85 | B.1.351 | SGTNGTKRFANPVLP |
| S | 76 | 90 | B.1.351 | TKRFANPVLPFNDGV |
| S | 201 | 215 | B.1.351 | FKIYSKHTPINLVRH |
| S | 201 | 215 | B.1.351 | FKIYSKHTPINLVRG |
| S | 206 | 220 | B.1.351 | KHTPINLVRHLPQGF |
| S | 206 | 220 | B.1.351 | KHTPINLVRGLPQGF |
| S | 211 | 225 | B.1.351 | NLVRHLPQGFSALEP |
| S | 211 | 225 | B.1.351 | NLVRGLPQGFSALEP |
| E | 61 | 75 | B.1.351 | RVKNLNSSRVLDLLV |
| nsp2 | 71 | 85 | B.1.351 | LQTPFEIKLAKKFDI |
| nsp2 | 76 | 90 | B.1.351 | EIKLAKKFDIFNGEC |
| nsp2 | 81 | 95 | B.1.351 | KKFDIFNGECPNFVF |
| nsp2 | 326 | 340 | B.1.351 | CGNFKVTKGKAKKSA |
| nsp2 | 331 | 345 | B.1.351 | VTKGKAKKSAWNIGE |
| nsp2 | 336 | 350 | B.1.351 | AKKSAWNIGEQKSIL |
| nsp2 | 356 | 370 | B.1.351 | FASEAARVVRTIFSR |
| nsp2 | 361 | 375 | B.1.351 | ARVVRTIFSRTLETA |
| nsp2 | 366 | 380 | B.1.351 | TIFSRTLETAQNSVR |
| nsp2 | 416 | 430 | B.1.351 | VVMAYITGGVVHLTS |
| nsp2 | 421 | 435 | B.1.351 | ITGGVVHLTSQWLTN |
| nsp2 | 426 | 440 | B.1.351 | VHLTSQWLTNIFGTV |
| nsp2 | 551 | 565 | B.1.351 | MPLKAPKEIIFLDGE |
| nsp2 | 556 | 570 | B.1.351 | PKEIIFLDGETLPTE |
| nsp2 | 561 | 575 | B.1.351 | FLDGETLPTEVLTEE |
| nsp3 | 823 | 837 | B.1.351 | SFLGRYMSALNHTKN |
| nsp3 | 828 | 842 | B.1.351 | YMSALNHTKNWKYPQ |
| nsp3 | 833 | 847 | B.1.351 | NHTKNWKYPQVNGLT |
| nsp3 | 1168 | 1182 | B.1.351 | LHKPIVWHVNNAINK |
| nsp3 | 1173 | 1187 | B.1.351 | VWHVNNAINKATYKP |
| nsp3 | 1178 | 1192 | B.1.351 | NAINKATYKPNTWCI |
| nsp3 | 1768 | 1782 | B.1.351 | VAVKMFDAYVSTFSS |
| nsp3 | 1773 | 1787 | B.1.351 | FDAYVSTFSSTFNVP |
| nsp3 | 1778 | 1792 | B.1.351 | STFSSTFNVPMEKLK |
| nsp13 | 577 | 591 | B.1.351 | SDRDLYDKLQFISLE |
| nsp13 | 582 | 596 | B.1.351 | YDKLQFISLEIPRRN |
| nsp13 | 587 | 601 | B.1.351 | FISLEIPRRNVATLQ |
| nsp6 | 92 | 106 | B.1.351 | MRIMTWLDMVDTSLK |
| nsp6 | 97 | 111 | B.1.351 | WLDMVDTSLKLKDCV |
| nsp6 | 102 | 116 | B.1.351 | DTSLKLKDCVMYASA |
| nsp6 | 107 | 121 | B.1.351 | LKDCVMYASAVVLLI |
| nsp6 | 122 | 136 | B.1.351 | LLILMTARTVYDDSA |
| nsp6 | 127 | 141 | B.1.351 | TARTVYDDSARRVWT |
| nsp6 | 132 | 146 | B.1.351 | YDDSARRVWTLMNVL |
| nsp6 | 137 | 151 | B.1.351 | RRVWTLMNVLTLFYK |
| nsp6 | 142 | 156 | B.1.351 | LMNVLTLFYKVYYGN |
| nsp6 | 147 | 161 | B.1.351 | TLFYKVYYGNALDQA |
| S | 231 | 245 | B.1.351 | IGINITRFQTLHRSY |
| S | 236 | 250 | B.1.351 | TRFQTLHRSYLTPGD |
| S | 241 | 255 | B.1.351 | LHRSYLTPGDSSSGW |
| S | 406 | 420 | B.1.351 | EVRQIAPGQTGNIAD |
| S | 411 | 425 | B.1.351 | APGQTGNIADYNYKL |
| S | 416 | 430 | B.1.351 | GNIADYNYKLPDDFT |
| S | 476 | 490 | B.1.351 | EIYQAGSTPCNGVKG |
| S | 481 | 495 | B.1.351 | GSTPCNGVKGFNCYF |
| S | 486 | 500 | B.1.351 | NGVKGFNCYFPLQSY |
| S | 491 | 505 | B.1.351 | PLQSYGFQPTYGVGY |
| S | 496 | 510 | B.1.351 | GFQPTYGVGYQPYRV |
| S | 501 | 515 | B.1.351 | YGVGYQPYRVVVLSF |
| S | 606 | 620 | B.1.351 | GTNTSNQVAVLYQGV |
| S | 611 | 625 | B.1.351 | NQVAVLYQGVNCTEV |
| S | 616 | 630 | B.1.351 | LYQGVNCTEVPVAIH |
| S | 691 | 705 | B.1.351 | SIIAYTMSLGVENSV |
| S | 696 | 710 | B.1.351 | TMSLGVENSVAYSNN |
| S | 701 | 715 | B.1.351 | VENSVAYSNNSIAIP |
| N | 66 | 80 | P.1. | FPRGQGVPINTNSSR |
| N | 71 | 85 | P.1. | GVPINTNSSRDDQIG |
| N | 76 | 90 | P.1. | TNSSRDDQIGYYRRA |
| N | 191 | 205 | P.1. | RNSSRNSTPGSSKRT |
| N | 196 | 210 | P.1. | NSTPGSSKRTSPARM |
| N | 201 | 215 | P.1. | SSKRTSPARMAGNGG |
| ORF8 | 81 | 95 | P.1. | VSCLPFTINCQKPKL |
| ORF8 | 86 | 100 | P.1. | FTINCQKPKLGSLVV |
| ORF8 | 91 | 105 | P.1. | QKPKLGSLVVRCSFY |
| ORF3a | 241 | 255 | P.1. | EEHVQIHTIDGSPGV |
| ORF3a | 246 | 260 | P.1. | IHTIDGSPGVVNPVM |
| ORF3a | 251 | 265 | P.1. | GSPGVVNPVMEPIYD |
| nsp12 | 309 | 323 | P.1. | HCANFNVLFSTVFPL |
| nsp12 | 314 | 328 | P.1. | NVLFSTVFPLTSFGP |
| nsp12 | 319 | 333 | P.1. | TVFPLTSFGPLVRKI |
| nsp13 | 327 | 341 | P.1. | IDKCSRIIPARARVD |
| nsp13 | 332 | 346 | P.1. | RIIPARARVDCFDKF |
| nsp13 | 337 | 351 | P.1. | RARVDCFDKFKVNST |
| S | 6 | 20 | P.1. | VLLPLVSSQCVNFTN |
| S | 11 | 25 | P.1. | VSSQCVNFTNRTQLP |
| S | 16 | 30 | P.1. | VNFTNRTQLPSAYTN |
| S | 21 | 35 | P.1. | RTQLPSAYTNSFTRG |
| S | 26 | 40 | P.1. | SAYTNSFTRGVYYPD |
| S | 126 | 140 | P.1. | VVIKVCEFQFCNYPF |
| S | 131 | 145 | P.1. | CEFQFCNYPFLGVYY |
| S | 136 | 150 | P.1. | CNYPFLGVYYHKNNK |
| S | 176 | 190 | P.1. | LMDLEGKQGNFKNLS |
| S | 181 | 195 | P.1. | GKQGNFKNLSEFVFK |
| S | 186 | 200 | P.1. | FKNLSEFVFKNIDGY |
| S | 406 | 420 | P.1. | EVRQIAPGQTGTIAD |

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC BY-NC-ND 4.0 International license.

| | | | | |
|---|---|---|---|---|
| S | 471 | 485 | P.1. | ENQAGSTPCNGVKG |
| S | 476 | 490 | P.1. | GSTPCNGVKGFNCYF |
| S | 481 | 495 | P.1. | NGVKGFNCYFPLQSY |
| S | 491 | 505 | P.1. | PLQSYGFQPTYGVGY |
| S | 496 | 510 | P.1. | GFQPTYGVGYQPYRV |
| S | 501 | 515 | P.1. | YGVGYQPYRVVVLSF |
| S | 601 | 615 | P.1. | GTNTSNQVAVLYQGV |
| S | 606 | 620 | P.1. | NQVAVLYQGVNCTEV |
| S | 611 | 625 | P.1. | LYQGVNCTEVPVAIH |
| S | 641 | 655 | P.1. | NVFQTRAGCLIGAEY |
| S | 646 | 660 | P.1. | RAGCLIGAEYVNNSY |
| S | 651 | 665 | P.1. | IGAEYVNNSYECDIP |
| S | 1016 | 1030 | P.1. | AEIRASANLAAIKMS |
| S | 1021 | 1035 | P.1. | SANLAAIKMSECVLG |
| S | 1026 | 1040 | P.1. | AIKMSECVLGQSKRV |
| S | 1166 | 1180 | P.1. | LGDISGINASFVNIQ |
| S | 1171 | 1185 | P.1. | GINASFVNIQKEIDR |
| S | 1176 | 1190 | P.1. | FVNIQKEIDRLNEVA |
| nsp3 | 173 | 187 | P.1. | QDGSEDNQTTTIQAI |
| nsp3 | 178 | 192 | P.1. | DNQTTTIQAIVEVQP |
| nsp3 | 183 | 197 | P.1. | TIQAIVEVQPQLEME |
| nsp3 | 358 | 372 | P.1. | AVFDKNLYDKLVLSF |
| nsp3 | 363 | 377 | P.1. | NLYDKLVLSFLEMKS |
| nsp3 | 368 | 382 | P.1. | LVLSFLEMKSEKQVE |
| nsp3 | 963 | 977 | P.1. | KGVQIPCTCGKQATQ |
| nsp3 | 968 | 982 | P.1. | PCTCGKQATQYLVQQ |
| nsp3 | 973 | 987 | P.1. | KQATQYLVQQESPFV |
| nsp6 | 92 | 106 | P.1. | MRIMTWLDMVDTSLK |
| nsp6 | 97 | 111 | P.1. | WLDMVDTSLKLKDCV |
| nsp6 | 102 | 116 | P.1. | DTSLKLKDCVMYASA |
| nsp6 | 107 | 121 | P.1. | LKDCVMYASAVVLLI |
| N | 191 | 205 | CAL.20C | RNSSRNSTPGSSKRI |
| N | 196 | 210 | CAL.20C | NSTPGSSKRISPARM |
| N | 201 | 215 | CAL.20C | SSKRISPARMAGNGG |
| N | 221 | 235 | CAL.20C | LLLLDRLNQLESKIS |
| N | 226 | 240 | CAL.20C | RLNQLESKISGKGQQ |
| N | 231 | 245 | CAL.20C | ESKISGKGQQQQGQT |
| nsp2 | 71 | 85 | CAL.20C | LQTPFEIKLAKKFDI |
| nsp2 | 76 | 90 | CAL.20C | EIKLAKKFDIFNGEC |
| nsp2 | 81 | 95 | CAL.20C | KKFDIFNGECPNFVF |
| nsp3 | 768 | 782 | CAL.20C | MSMTYGQQFGSTYLD |
| nsp3 | 773 | 787 | CAL.20C | GQQFGSTYLDGADVT |
| nsp3 | 778 | 792 | CAL.20C | STYLDGADVTKIKPH |
| nsp4 | 383 | 397 | CAL.20C | ICISTKHFYWFFTNY |
| nsp4 | 388 | 402 | CAL.20C | KHFYWFFTNYLKRRV |
| nsp4 | 393 | 407 | CAL.20C | FFTNYLKRRVVFNGV |
| nsp6 | 112 | 126 | CAL.20C | DCVMYASAVVLLIFM |
| nsp6 | 117 | 131 | CAL.20C | ASAVVLLIFMTARTV |
| nsp6 | 122 | 136 | CAL.20C | LLIFMTARTVYDDGA |
| nsp6 | 157 | 171 | CAL.20C | ALDQAISMWAFIISV |
| nsp6 | 162 | 176 | CAL.20C | ISMWAFIISVTSNYS |
| nsp6 | 167 | 181 | CAL.20C | FIISVTSNYSGVVTT |
| nsp9 | 51 | 65 | CAL.20C | LKWARFPKSDGTGTV |
| nsp9 | 56 | 70 | CAL.20C | FPKSDGTGTVYTELE |
| nsp9 | 61 | 75 | CAL.20C | GTGTVYTELEPPCRF |
| nsp12 | 309 | 323 | CAL.20C | HCANFNVLFSTVFPL |
| nsp12 | 314 | 328 | CAL.20C | NVLFSTVFPLTSFGP |
| nsp12 | 319 | 333 | CAL.20C | TVFPLTSFGPLVRKI |
| nsp13 | 42 | 56 | CAL.20C | VLSVNPYVCNAPGCD |
| nsp13 | 47 | 61 | CAL.20C | PYVCNAPGCDVTDVT |
| nsp13 | 52 | 66 | CAL.20C | APGCDVTDVTQLYLG |
| nsp13 | 197 | 211 | CAL.20C | EYTFEKGDYGDAFVY |
| nsp13 | 202 | 216 | CAL.20C | KGDYGDAFVYRGTTT |
| nsp13 | 207 | 221 | CAL.20C | DAFVYRGTTTYKLNV |
| nsp13 | 247 | 261 | CAL.20C | VRITGLYPTLNISYE |
| nsp13 | 252 | 266 | CAL.20C | LYPTLNISYEFSSNV |
| nsp13 | 257 | 271 | CAL.20C | NISYEFSSNVANYQK |
| nsp14 | 316 | 330 | CAL.20C | VVKAALLADKLPVLH |
| nsp14 | 321 | 335 | CAL.20C | LLADKLPVLHDIGNP |
| nsp14 | 326 | 340 | CAL.20C | LPVLHDIGNPKAIKC |
| nsp15 | 79 | 93 | CAL.20C | IAANTVIWDYKRYAP |
| nsp15 | 84 | 98 | CAL.20C | VIWDYKRYAPAHIST |
| nsp15 | 89 | 103 | CAL.20C | KRYAPAHISTIGVCS |
| ORF3a | 46 | 60 | CAL.20C | LIVGVALLAVFHSAS |
| ORF3a | 51 | 65 | CAL.20C | ALLAVFHSASKIITL |
| ORF3a | 56 | 70 | CAL.20C | FHSASKIITLKKRWQ |
| S | 1 | 15 | CAL.20C | MFVFLVLLPLVSIQC |
| S | 6 | 20 | CAL.20C | VLLPLVSIQCVNLTT |
| S | 11 | 25 | CAL.20C | VSIQCVNLTTRTQLP |
| S | 141 | 155 | CAL.20C | LGVYYHKNNKSCMES |
| S | 146 | 160 | CAL.20C | HKNNKSCMESEFRVY |
| S | 151 | 165 | CAL.20C | SCMESEFRVYSSANN |
| S | 441 | 455 | CAL.20C | LDSKVGGNYNYRYRL |
| S | 446 | 460 | CAL.20C | GGNYNYRYRLFRKSN |
| S | 451 | 465 | CAL.20C | YRYRLFRKSNLKPFE |
| S | 601 | 615 | CAL.20C | GTNTSNQVAVLYQGV |
| S | 606 | 620 | CAL.20C | NQVAVLYQGVNCTEV |
| S | 611 | 625 | CAL.20C | LYQGVNCTEVPVAIH |
| S | 926 | 940 | CAL.20C | QFNSAIGKIQDSFSS |
| S | 931 | 945 | CAL.20C | IGKIQDSFSSTASAL |
| S | 936 | 950 | CAL.20C | DSFSSTASALGKLQD |
| S | 1181 | 1195 | CAL.20C | KEIDRLNEVANNLNE |
| S | 1186 | 1200 | CAL.20C | LNEVANNLNESLIDL |
| S | 1191 | 1205 | CAL.20C | NNLNESLIDLQELGK |

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433180; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**Table S3**. Related to Figure 4. Effect of mutations on CD8 epitope HLA class I binding capacity

| Origin | Ancestral reference sequence | Protein | Start | Mutation[a] | Mutated sequence[b] | HLA restriction | WT ($IC_{50}$ nM)[c] | Mutant ($IC_{50}$ nM) | Fold difference | Effect[d] |
|---|---|---|---|---|---|---|---|---|---|---|
| B.1.1.7 | HVSGTNGTK | S | 69 | HV69-70 del | HAISGTNGTK | A*68:01 | 55 | 44 | 0.8 | Neutral |
| B.1.1.7 | GVYYHKNNK | S | 142 | Y145 del | FLGVYHKNNK | A*03:01 | 28 | 1078 | 39 | Decrease |
| B.1.1.7 | YYHKNNKSW | S | 144 | Y145 del | VYHKNNKSW | A*24:02 | 117 | 308 | 2.6 | Decrease |
| B.1.1.7 | YGFQPTNGV | S | 495 | N501Y | YGFQPTYGV | B*51:01 | 3488 | 3541 | 1.0 | Neutral |
| B.1.1.7 | YQDVNCTEV | S | 612 | D614G | YQGVNCTEV | A*02:06 | 18 | 57 | 3.2 | Decrease |
| B.1.1.7 | QTNSPRRAR | S | 677 | P681H | QTNSHRRAR | A*31:01 | 35 | 33 | 0.94 | Neutral |
| B.1.1.7 | SPRRARSV | S | 680 | P681H | SHRRARSV | B*08:01 | 429 | 2449 | 5.7 | Decrease |
| B.1.1.7 | NSIAIPTNF | S | 710 | T716I | NSIAIPINF | B*57:01 | 1335 | 968 | 0.73 | Neutral |
| B.1.1.7 | IAIPTNFTI | S | 712 | T716I | IAIPINFTI | B*51:01 | 209 | 189 | 0.90 | Neutral |
| B.1.1.7 | IAIPTNFTI | S | 712 | T716I | IAIPINFTI | B*53:01 | 396 | 266 | 0.67 | Neutral |
| B.1.1.7 | IPTNFTISV | S | 714 | T716I | IPINFTISV | B*07:02 | 188 | 168 | 0.89 | Neutral |
| B.1.1.7 | IPTNFTISV | S | 714 | T716I | IPINFTISV | B*51:01 | 156 | 94 | 0.60 | Neutral |
| B.1.1.7 | SVLNDILSR | S | 975 | S982A | SVLNDILAR | A*68:01 | 109 | 92 | 0.84 | Neutral |
| B.1.1.7 | KLINIIIWF | nsp3 | 1407 | I1412T | KLINITIWF | A*32:01 | 161 | 48 | 0.30 | Increase |
| B.1.1.7 | STVFPPTSF | nsp12 | 318 | P323L | STVFPLTSF | B*57:01 | 1583 | 637 | 0.40 | Increase |
| B.1.351 | RFDNPVLPF | S | 78 | D80A | RFANPVLPF | A*24:02 | 458 | 34 | 0.075 | Increase |
| B.1.351 | FDNPVLPFNDGVYF | S | 79 | D80A | FANPVLPFNDGVYF | B*35:01 | 65 | 65 | 1.0 | Neutral |
| B.1.351 | TPINLVRDL | S | 208 | D215G | TPINLVRGL | B*07:02 | 213 | 119 | 0.56 | Neutral |
| B.1.351 | TPINLVRDL | S | 208 | D215H | TPINLVRHL | B*07:02 | 213 | 199 | 0.93 | Neutral |
| B.1.351 | QIAPGQTGK | S | 409 | K417N | QIAPGQTGN | A*68:01 | 137 | 27998 | 204 | Decrease |
| B.1.351 | YGFQPTNGV | S | 495 | N501Y | YGFQPTYGV | B*51:01 | 3488 | 3541 | 1.0 | Neutral |
| B.1.351 | YQDVNCTEV | S | 612 | D614G | YQGVNCTEV | A*02:06 | 18 | 57 | 3.2 | Decrease |
| B.1.351 | YTMSLGAENSVAY | S | 695 | A701V | YTMSLGVENSVAY | A*26:01 | 184 | 253 | 1.4 | Neutral |
| B.1.351 | LGAENSVAY | S | 699 | A701V | LGVENSVAY | B*35:01 | 19 | 21 | 1.1 | Neutral |
| B.1.351 | GPQNQRNAPRITF | N | 5 | K17N | GPQNQRNASRITF | B*07:02 | 640 | 696 | 1.1 | Neutral |
| B.1.351 | QSASKIITL | ORF3a | 57 | Q57H | HSASKIITL | B*08:01 | 1788 | 573 | 0.32 | Increase |
| B.1.351 | MSALNHTKK | nsp3 | 829 | K837N | MSALNHTKN | A*30:01 | 102 | 7035 | 69 | Decrease |
| B.1.351 | MSALNHTKKW | nsp3 | 829 | K837N | MSALNHTKNW | B*57:01 | 16 | 14 | 0.88 | Neutral |
| B.1.351 | SALNHTKKW | nsp3 | 830 | K837N | SALNHTKNW | B*57:01 | 111 | 93 | 0.84 | Neutral |
| B.1.351 | STVFPPTSF | nsp12 | 318 | P323L | STVFPLTSF | B*57:01 | 1583 | 637 | 0.40 | Increase |
| P.1. | LPPAYTNSF | S | 24 | P26S | LPSAYTNSF | B*07:02 | 294 | 51 | 0.17 | Increase |
| P.1. | LPPAYTNSF | S | 24 | P26S | LPSAYTNSF | B*35:01 | 44 | 4.1 | 0.093 | Increase |
| P.1. | LPPAYTNSF | S | 24 | P26S | LPSAYTNSF | B*53:01 | 366 | 18 | 0.049 | Increase |
| P.1. | QIAPGQTGK | S | 409 | K417T | QIAPGQTGT | A*68:01 | 137 | 20478 | 149 | Decrease |
| P.1. | YGFQPTNGV | S | 495 | N501Y | YGFQPTYGV | B*51:01 | 3488 | 3541 | 1.0 | Neutral |
| P.1. | YQDVNCTEV | S | 612 | D614G | YQGVNCTEV | A*02:06 | 18 | 57 | 3.2 | Decrease |
| P.1. | AEHVNNSY | S | 653 | H655Y | AEYVNNSY | B*44:02 | 1038 | 904 | 0.87 | Neutral |
| P.1. | AEHVNNSY | S | 653 | H655Y | AEYVNNSY | B*44:03 | 1020 | 577 | 0.57 | Neutral |
| P.1. | RASANLAATK | S | 1019 | L1027I | RASANLAAIK | A*03:01 | 85 | 99 | 1.2 | Neutral |
| P.1. | NASVVNIQK | S | 1173 | V1176F | NASFVNIQK | A*68:01 | 13 | 6.2 | 0.48 | Increase |
| P.1. | NTNSSPDDQIGYY | N | 75 | P80R | NTNSSRDDQIGYY | A*01:01 | 44 | 44 | 1.0 | Neutral |
| P.1. | SPDDQIGYY | N | 79 | P80R | SRDDQIGYY | B*35:01 | 101 | 17781 | 175 | Decrease |
| P.1. | LYDKLVSSF | nsp3 | 364 | S370L | LYDKLVLSF | A*24:02 | 77 | 70 | 0.91 | Neutral |
| P.1. | STVFPPTSF | nsp12 | 318 | P323L | STVFPLTSF | B*57:01 | 1583 | 637 | 0.40 | Increase |
| CAL.20C | LPLVSSQCV | S | 8 | S13I | LPLVSIQCV | B*51:01 | 402 | 272 | 0.68 | Neutral |
| CAL.20C | YYHKNNKSW | S | 144 | W152C | YYHKNNKSC | A*24:02 | 117 | 11134 | 95 | Decrease |
| CAL.20C | SWMESEFRVY | S | 151 | W152C | SCMESEFRVY | A*29:02 | 49 | 980 | 20 | Decrease |
| CAL.20C | KVGGNYNYLY | S | 444 | L452R | KVGGNYNYRY | A*29:02 | 101 | 505 | 5.0 | Decrease |
| CAL.20C | VGGNYNYLY | S | 445 | L452R | VGGNYNYRY | A*29:02 | 94 | 519 | 5.5 | Decrease |
| CAL.20C | NYNYLYRLF | S | 448 | L452R | NYNYRYRLF | A*24:02 | 21 | 108 | 5.1 | Decrease |
| CAL.20C | YNYLYRLFR | S | 449 | L452R | YNYRYRLFR | A*31:01 | 16 | 12 | 0.75 | Neutral |
| CAL.20C | YQDVNCTEV | S | 612 | D614G | YQGVNCTEV | A*02:06 | 18 | 57 | 3.2 | Decrease |
| CAL.20C | QSASKIITL | ORF3a | 57 | Q57H | HSASKIITL | B*08:01 | 1788 | 573 | 0.32 | Increase |
| CAL.20C | WFFSNYLKR | nsp4 | 392 | S395T | WFFTNYLKR | A*31:01 | 70 | 98 | 1.4 | Neutral |
| CAL.20C | STVFPPTSF | nsp12 | 318 | P323L | STVFPLTSF | B*57:01 | 1583 | 637 | 0.40 | Increase |

[a]Mutation noted as ancestral residue-position-variant residue. Del refers to deletion of the corresponding residue.

[b]For deletion mutants, the peptide sequence shown represents the variant encompassing the same region that has the highest predicted binding affinity for the corresponding restricting allele.

[c]Indicates predicted $IC_{50}$ for the corresponding reported restricting allele. Predictions were performed using the NetMHCpan BA 4.1 algorithm, hosted by the IEDB.

[d]Increase/decrease in affinity defined by a two-fold difference in predicted $IC_{50}$ nM.