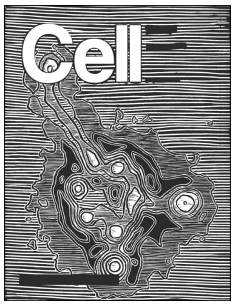
Journal Pre-proof



Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals

Alba Grifoni, Daniela Weiskopf, Sydney I. Ramirez, Jose Mateus, Jennifer M. Dan, Carolyn Rydyznski Moderbacher, Stephen A. Rawlings, Aaron Sutherland, Lakshmanane Premkumar, Ramesh S. Jadi, Daniel Marrama, Aravinda M. de Silva, April Frazier, Aaron Carlin, Jason A. Greenbaum, Bjoern Peters, Florian Krammer, Davey M. Smith, Shane Crotty, Alessandro Sette

PII: S0092-8674(20)30610-3

DOI: https://doi.org/10.1016/j.cell.2020.05.015

Reference: CELL 11420

To appear in: *Cell*

Received Date: 20 April 2020

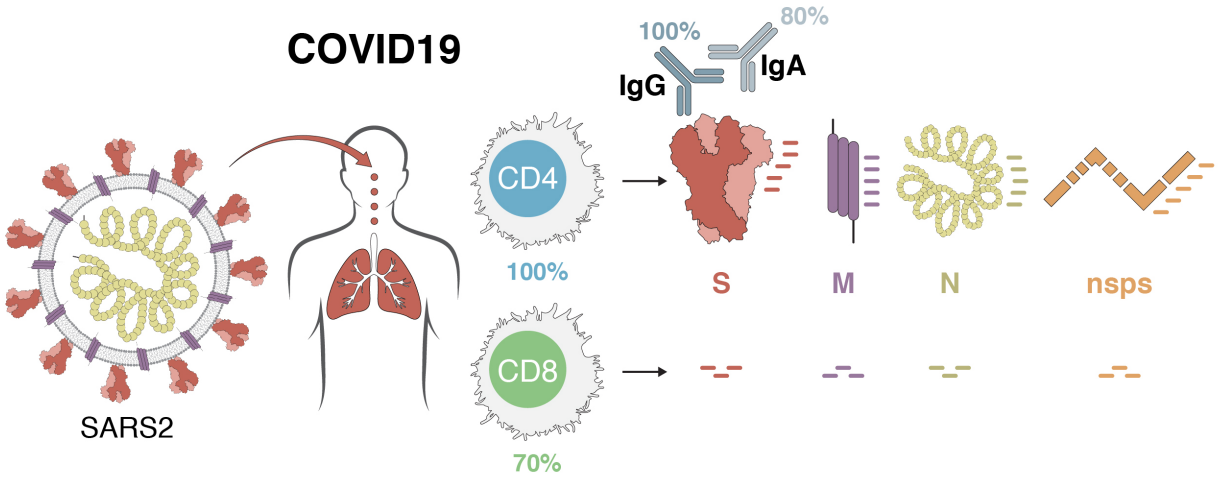
Revised Date: 4 May 2020

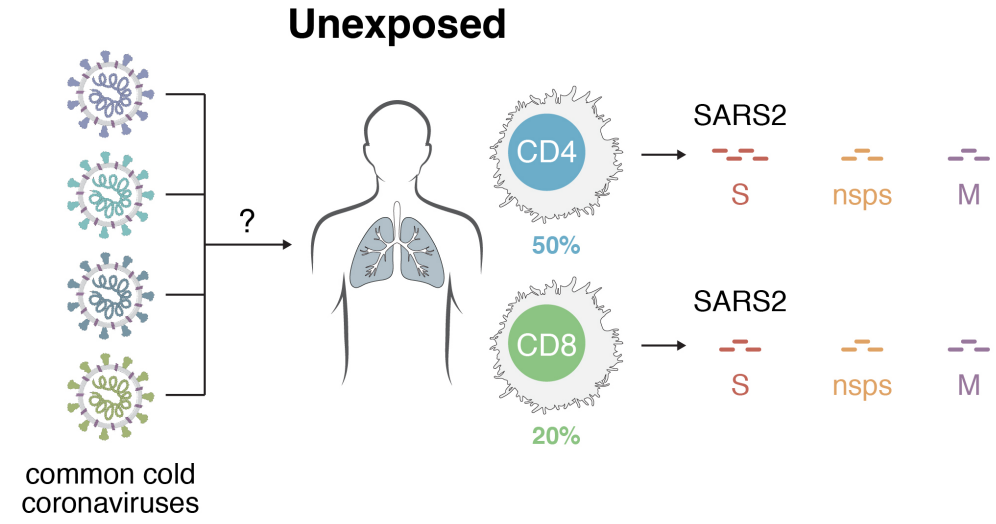
Accepted Date: 7 May 2020

Please cite this article as: 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., Marrama, D., de Silva, A.M., Frazier, A., Carlin, A., Greenbaum, J.A., Peters, B., Krammer, F., Smith, D.M., Crotty, S., Sette, A., Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals, *Cell* (2020), doi: https://doi.org/10.1016/j.cell.2020.05.015.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Elsevier Inc.


COVID19
100%
80%
IgG
IgA
CD4
100%
S
M
N
nsps
CD8
70%
SARS2


Unexposed
?
CD4
50%
SARS2
S
nsps
M
CD8
20%
SARS2
S
nsps
M
common cold
coronaviruses

Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals

AUTHORS
Alba Grifoni[1], Daniela Weiskopf[1], Sydney I. Ramirez[1,2], Jose Mateus[1], Jennifer M. Dan[1,2], Carolyn Rydyznski Moderbacher[1], Stephen A. Rawlings[2], Aaron Sutherland[1], Lakshmanane Premkumar[3], Ramesh S. Jadi[3], Daniel Marrama[1], Aravinda M. de Silva[3], April Frazier[1], Aaron Carlin[2], Jason A. Greenbaum[1], Bjoern Peters[1,2], Florian Krammer[4], Davey M. Smith[2], Shane Crotty[1,2]*, Alessandro Sette[1,2]*

AFFILIATIONS
[1] Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology (LJI), La Jolla, CA 92037, USA
[2] Department of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego (UCSD), La Jolla, CA 92037, USA
[3] Department of Microbiology and Immunology, University of North Carolina School of Medicine, USA.
[4] Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, New York
* indicates equal contributions
Correspondence: alex@lji.org (A.S.) and shane@lji.org (S.C.).
Lead Contact: alex@lji.org (A.S.)

SUMMARY
Understanding adaptive immunity to SARS-CoV-2 is important for vaccine development, interpreting coronavirus disease 2019 (COVID-19) pathogenesis, and calibration of pandemic control measures. Using HLA class I and II predicted peptide 'megapools', circulating SARS-CoV-2–specific $CD8^+$ and $CD4^+$ T cells were identified in ~70% and 100% of COVID-19 convalescent patients, respectively. $CD4^+$ T cell responses to spike, the main target of most vaccine efforts, were robust and correlated with the magnitude of the anti-SARS-CoV-2 IgG and IgA titers. The M, spike and N proteins each accounted for 11-27% of the total $CD4^+$ response, with additional responses commonly targeting nsp3, nsp4, ORF3a and ORF8, among others. For $CD8^+$ T cells, spike and M were recognized, with at least eight SARS-CoV-2 ORFs targeted. Importantly, we detected SARS-CoV-2–reactive $CD4^+$ T cells in ~40-60% of unexposed individuals, suggesting cross-reactive T cell recognition between circulating 'common cold' coronaviruses and SARS-CoV-2.

INTRODUCTION
COVID-19 is a world-wide emergency. The first cases occurred in December 2019 and now more than 240,000 deaths and 3,000,000 cases of SARS-CoV-2 infection have been reported globally as of May 1st, (Dong et al., 2020; Wu and McGoogan, 2020). Vaccines against SARS-CoV-2 are just beginning development (Amanat and Krammer, 2020; Thanh Le et al., 2020). An understanding of human T cell responses to SARS-CoV2 is lacking, due to the rapid emergence of the pandemic. There is an urgent need for foundational information about T cell responses to this virus.

The first steps for such an understanding are the ability to quantify the virus-specific $CD4^+$ and $CD8^+$ T cells. Such knowledge is of immediate relevance, as it will provide insights into immunity and pathogenesis of SARS-CoV-2 infection, and the same knowledge will assist vaccine design and evaluation of candidate vaccines. Estimations of immunity are also central to epidemiological model calibration of future social distancing pandemic control measures (Kissler et al., 2020). Such projections are dramatically different depending on whether SARS-CoV-2 infection creates substantial immunity, and whether any crossreactive immunity exists between SARS-CoV-2 and circulating seasonal 'common cold' human coronaviruses. Definition and assessment of human antigen-specific SARS-CoV2 T cell responses are best made with direct ex vivo T cell assays using broad-based epitope pools

and assays capable of detecting T cells of any cytokine polarization. Herein, we have completed such an assessment with blood samples from COVID-19 patients.

There is also great uncertainty about whether adaptive immune responses to SARS-CoV-2 are protective or pathogenic, or whether both scenarios can occur depending on timing, composition, or magnitude of the adaptive immune response. Hypotheses range the full gamut (Peeples, 2020), based on available clinical data from severe acute respiratory disease syndrome (SARS) or MERS (Alshukairi et al., 2018; Wong et al., 2004; Zhao et al., 2017) or animal model data with SARS in mice (Zhao et al., 2016; Zhao et al., 2010; Zhao et al., 2009), SARS in NHPs (Liu et al., 2019; Takano et al., 2008) or FIPV in cats (Vennema et al., 1990). Protective immunity, immunopathogenesis, and vaccine development for COVID-19 are each briefly discussed below, related to introducing the importance of defining T cell responses to SARS-CoV-2.

Based on data from SARS patients in 2003-2004 (caused by SARS-CoV, the most closely related human betacoronavirus to SARS-CoV-2), and based on the fact that most acute viral infections result in development of protective immunity (Sallusto et al., 2010), a likely possibility has been that substantial $CD4^+$ T cell, $CD8^+$ T cell, and neutralizing antibody responses develop to SARS-CoV-2 and all contribute to clearance of the acute infection; and, as a corollary, some of the T and B cells are retained long-term (i.e., multiple years) as immunological memory and protective immunity against SARS-CoV-2 infection (Guo et al., 2020b; Li et al., 2008). However, a contrarian viewpoint is also legitimate. While most acute infections result in the development of protective immunity, available data for human coronaviruses suggest the possibility that substantive adaptive immune responses can fail to occur (Choe et al., 2017; Okba et al., 2019; Zhao et al., 2017) and robust protective immunity can fail to develop (Callow et al., 1990). A failure to develop protective immunity could occur due to a T cell and/or antibody response of insufficient magnitude or durability, with the neutralizing antibody response being dependent on the $CD4^+$ T cell response (Crotty, 2019; Zhao et al., 2016). Thus, there is urgent need to understand the magnitude and composition of the human $CD4^+$ and $CD8^+$ T cell responses to SARS-CoV-2. If natural infection with SARS-CoV-2 elicits potent $CD4^+$ and $CD8^+$ T cell responses commonly associated with protective antiviral immunity, COVID-19 is a strong candidate for rapid vaccine development.

Immunopathogenesis in COVID-19 is a serious concern (Cao, 2020; Peeples, 2020). It is most likely that an early $CD4^+$ and $CD8^+$ T cell response against SARS-CoV-2 is protective, but an early response is difficult to generate because of efficient innate immune evasion mechanisms of SARS-CoV-2 in humans (Blanco-Melo et al., 2020). Immune evasion by SARS-CoV-2 is likely exacerbated by reduced myeloid cell antigen presenting cell (APC) function or availability in the elderly (Zhao et al., 2011). In such cases, it is conceivable that late T cell responses may instead amplify pathogenic inflammatory outcomes in the presence of sustained high viral loads in the lungs, by multiple hypothetical possible mechanisms (Guo et al., 2020a; Li et al., 2008; Liu et al., 2019). Critical (ICU) and fatal COVID-19 (and SARS) outcomes are associated with elevated levels of inflammatory cytokines and chemokines, including IL-6 (Giamarellos-Bourboulis et al., 2020; Wong et al., 2004; Zhou et al., 2020)

Vaccine development against acute viral infections classically focuses on vaccine-elicited recapitulation of the type of protective immune response elicited by natural infection. Such foundational knowledge is currently missing for COVID-19, including how the balance and the phenotypes of responding cells vary as a function of disease course and severity. Such knowledge can guide selection of vaccine strategies most likely to elicit protective immunity against SARS-CoV-2. Furthermore, knowledge of the T cell responses to COVID-19 can guide selection of appropriate immunological endpoints for COVID-19 candidate vaccine clinical trials, which are already starting.

Limited information is also available about which SARS-CoV-2 proteins are recognized by human T cell immune responses. In some infections, T cell responses are strongly biased towards certain viral proteins, and the targets can vary substantially between $CD4^+$ and $CD8^+$ T cells (Moutaftsi et al., 2010; Tian et al., 2019). Knowledge of SARS-CoV-2 proteins and epitopes recognized by human T cell responses is of immediate relevance, as it will allow for monitoring of COVID-19 immune responses in laboratories worldwide. Epitope knowledge will also assist candidate vaccine design and facilitate evaluation of vaccine candidate immunogenicity. Almost all of the current COVID-19 vaccine candidates are focused on the spike protein.

A final key issue to consider in the study of SARS-CoV-2 immunity is whether some degree of cross-reactive coronavirus immunity exists in a fraction of the human population, and whether this might influence susceptibility to COVID-19 disease. This issue is also relevant for vaccine development, as cross-reactive immunity could influence responsiveness to candidate vaccines (Andrews et al., 2015).

In sum, the ability to measure and understand the human $CD4^+$ and $CD8^+$ T cell responses to SARS-CoV-2 infection is a major knowledge gap currently impeding COVID-19 vaccine development, interpretation of COVID-19 disease pathogenesis, and calibration of future social distancing pandemic control measures.

RESULTS

SARS-CoV-2 peptides and predicted class I and class II epitopes

We recently predicted SARS-CoV-2 T cell epitopes utilizing the Immune Epitope Database and Analysis Resource (IEDB) (Dhanda et al., 2019; Vita et al., 2019). Utilizing bioinformatic approaches, we identified specific peptides in SARS-CoV-2 with increased probability of being T cell targets (Grifoni et al., 2020). We previously developed the megapool (MP) approach to allow simultaneous testing of large numbers of epitopes. By this technique, numerous epitopes are solubilized, pooled and re-lyophilized to avoid cell toxicity problems (Carrasco Pro et al., 2015). These MPs have been used in human T cell studies of a number of indications, including allergies (Hinz et al., 2016), tuberculosis (Lindestam Arlehamn et al., 2016), tetanus, pertussis (Bancroft et al., 2016; da Silva Antunes et al., 2017) and dengue virus, for both $CD4^+$ and $CD8^+$ T cell epitopes (Grifoni et al., 2017; Weiskopf et al., 2015). Here, we generated MPs based on predicted SARS-CoV-2 epitopes. Specifically, one MP corresponds to 221 predicted HLA class II $CD4^+$ T cell epitopes (Grifoni et al., 2020) covering all proteins in the viral genome, apart from the spike (S) antigen (CD4_R MP). The prediction strategy utilized is geared to capture ~50% of the total response (Dhanda et al., 2018; Paul et al., 2015) and was designed and validated to predict dominant epitopes independently of ethnicity and HLA polymorphism. This approach takes advantage of the extensive cross-reactivity and repertoire overlap between different HLA class II loci and allelic variants to predict promiscuous epitopes, capable of binding many of the most common HLA class II prototypic specificities (Greenbaum et al., 2011; O'Sullivan et al., 1991; Sidney et al., 2010a, b; Southwood et al., 1998).

For the spike protein, to ensure that all T cell reactivity against this important antigen can be detected, we generated a separate MP covering the entire antigen with 253 15-mer peptides overlapping by 10-residues (MP_S, **Table S1**). As stated above, the MP used to probe the non-spike regions is expected to capture ~50% of the total response. The use of overlapping peptides spanning entire ORFs instead allows for a more complete characterization, but also requires more cells. This factor should be kept in mind in terms of comparison of the magnitude of the $CD4^+$ T cell responses to those pools.

In the case of CD8 epitopes, since the overlap between different HLA class I allelic variants and loci is more limited to specific groups of alleles, or supertypes (Sidney et al., 2008), we targeted a set of the 12 most prominent HLA class I A and B alleles, which together allow broad coverage (>85%) of the general population. Two class I MPs were synthesized based on epitope predictions for those 12 most common HLA A and B alleles (Grifoni et al., 2020), which collectively encompass 628 predicted HLA class I $CD8^+$ T cell epitopes from the entire SARS-CoV-2 proteome (CD8 MP-A and MP-B).

Immunological phenotypes of recovered COVID-19 patients

To test for the generation of SARS-CoV-2 $CD4^+$ and $CD8^+$ T cell responses following infection, we initially recruited 20 adult patients who had recovered from COVID-19 disease (**Table 1**). We also utilized PBMC and plasma samples from local healthy control donors collected in 2015-2018 (see STAR Methods). Blood samples were collected at 20-35 days post-symptoms onset from non-hospitalized COVID-19 patients who were no longer symptomatic. SARS-CoV-2 infection was determined by swab test viral PCR during the acute phase of the infection. Verification of SARS-CoV-2 exposure was attempted both by lateral flow serology and SARS-CoV-2 spike protein receptor binding domain (RBD) ELISA (Stadlbauer et al., 2020), using plasma from the convalescence stage blood

draw. Most patients were confirmed positive by lateral flow Ig tests (**Table 1**). All patients were confirmed COVID-19 cases by SARS-CoV-2 RBD ELISA (**Fig. 1, S1**). All cases were IgG positive; anti-RBD IgM and IgA was also detected in the large majority of cases (**Fig. 1, S1**).

We defined a 21-color flow cytometry panel of mononuclear leukocyte lineage and phenotypic markers (**Table S2**) to broadly assess the immunological cellular profile of recovered COVID-19 patients (**Fig. 1, S2**). The frequency of $CD3^+$ cells was slightly increased in recovered COVID-19 patients relative to non-exposed controls, while no significant differences overall were observed in the frequencies of $CD4^+$ or $CD8^+$ T cells between the two groups. Frequencies of $CD19^+$ cells were somewhat decreased, while no differences were observed in the frequencies of $CD3^-CD19^-$ cells or $CD14^+CD16^-$ monocytes (**Fig. 1, S2**). No evidence of general lymphopenia was observed in the convalescing patients, consistent with the literature. Next, we utilized the SARS-CoV-2 MPs to probe $CD4^+$ and $CD8^+$ T cell responses.

Identification and quantitation of SARS-CoV-2–specific $CD4^+$ T cell responses

We utilized T cell receptor– (TCR) dependent Activation Induced Marker (AIM) assays to identify and quantify SARS-CoV-2–specific $CD4^+$ T cells in recovered COVID-19 patients. Initial definition and assessment of human antigen-specific SARS-CoV-2 T cell responses are best made with direct ex vivo T cell assays using broad-based epitope pools, such as MPs, and assays capable of detecting T cells of unknown cytokine polarization and functional attributes. AIM assays are cytokine-independent assays to identify antigen-specific $CD4^+$ T cells (Havenar-Daughton et al., 2016; Reiss et al., 2017). AIM assays have been successfully used to identify virus-specific, vaccine-specific, or tuberculosis-specific $CD4^+$ T cells in a range of studies (Dan et al., 2019; Dan et al., 2016; Herati et al., 2017; Morou et al., 2019).

We stimulated PBMCs from 10 COVID-19 cases and 11 healthy controls (SARS-CoV-2 unexposed, collected in 2015-2018) with a spike MP (MP_S) and the class II MP covering the remainder of the SARS-CoV-2 orfeome ("non-spike", MP CD4_R). A CMV MP was used as a positive control, while DMSO was used as the negative control (**Fig. 2, S3**). SARS-CoV-2 spike-specific $CD4^+$ T cell responses ($OX40^+CD137^+$) were detected in 100% of COVID-19 cases ($p < 0.0001$ vs. unexposed donors spike MP, **Fig. 2A-B.** $p = 0.002$ vs. DMSO control, **Fig. 2C**). $CD4^+$ T cell responses to the remainder of the SARS-CoV-2 orfeome were also detected in 100% of COVID-19 cases ($p < 0.0079$ vs. unexposed donors non-spike MP, **Fig. 2A-B.** $p = 0.002$, non-spike vs. DMSO control, **Fig. 2C**). The magnitude of the SARS-CoV-2–specific $CD4^+$ T cell responses measured was similar to that of the CMV MP (**Fig. S3C**). The concordance between SARS-CoV-2–specific $CD4^+$ T cell measurements in independent experiments was high ($p < 0.0002$, **Fig S3D)**. To assess functionality and polarization of the SARS-CoV-2–specific $CD4^+$ T cell response, we measured cytokines secreted in response to MP stimulation. The SARS-CoV-2–specific $CD4^+$ T cells were functional, as the cells produced IL-2 in response to non-spike and spike MPs (**Fig. 2D**). Polarization of the cells appeared to be a classical T_H1 type, as substantial IFNγ was produced (**Fig. 2E**), while little to no IL-4, Il-5, IL-13, or IL-17α was expressed (**Fig. S3G-J**).

Thus, recovered COVID-19 patients consistently generated a substantial $CD4^+$ T cell response against SARS-CoV-2. Similar conclusions were reached using stimulation index as the metric (**Fig. S3E-F**). In terms of total $CD4^+$ T cell response per donor (**Fig. 2A**), on average ~50% of the detected response was directed against the spike protein, and ~50% was directed against the MP representing the remainder of the SARS-CoV-2 orfeome (**Fig. 2A**). This is of significance, since the SARS-CoV-2 spike protein is a key component of the vast majority of candidate COVID-19 vaccines under development. Of note, given the nature of the MP_R peptide predictions, the actual $CD4^+$ T cell response to be ascribed to non-spike ORFs was likely to be higher, addressed in further experiments below.

Identification and quantitation of SARS-CoV-2–specific $CD8^+$ T cell responses

To measure SARS-CoV-2–specific $CD8^+$ T cells in the recovered COVID-19 patients, we utilized two complementary methodologies, AIM assays and intracellular cytokine staining (ICS). The two SARS-CoV-2 class I MPs were used, CD8-A and CD8-B, with CMV MP and DMSO serving as positive and

negative controls, respectively (**Fig. 3, S4**). $CD8^+$ T cell responses were detected by AIM ($CD69^+CD137^+$) in 70% of COVID-19 cases ($p < 0.0011$ vs. unexposed donors "CD8 total", **Fig. 3A-B.** $p = 0.002$, CD8-A or CD8-B vs. DMSO control, **Fig. S4B**). MP CD8-A contains spike epitopes, among epitopes to other proteins. The magnitude of the SARS-CoV-2 reactive $CD8^+$ T cell responses measured by AIM was somewhat lower than the CMV MP (**Fig. S4C**). Similar conclusions were reached using stimulation index (**Fig. S3D-E**).

Independently, ICS assays detected $IFN\gamma^+$ SARS-CoV-2–specific $CD8^+$ T cells in the majority of COVID-19 cases (**Fig. 3C-D**). The majority of $IFN\gamma^+$ cells co-expressed granzyme B (**Fig. 3D-E**). A substantial fraction of the $IFN\gamma^+$ cells expressed TNF, but not IL-10 (**Fig. 3D**). Thus, the majority of recovered COVID-19 patients generated a $CD8^+$ T cell response against SARS-CoV-2.

Relationship between SARS-CoV-2–specific $CD4^+$ T cell responses and IgG and IgA titers

Most protective antibody responses are dependent on $CD4^+$ T cell help. Therefore, we assessed whether stronger SARS-CoV-2–specific $CD4^+$ T cell responses were associated with higher antibody titers in COVID-19 cases. Given that spike is the primary target of SARS neutralizing antibodies, we examined spike-specific $CD4^+$ T cells. Spike-specific $CD4^+$ T cell responses correlated well with the magnitude of the anti-spike RBD IgG titers ($R = 0.81$. $p < 0.0001$. **Fig. 4A.** Similar results were obtained using stimulation index, **Fig. S5A**). The non-spike SARS-CoV-2–specific $CD4^+$ T cell response did not correlate as well with anti-spike RBD IgG titers (**Fig. 4B, S5B**), consistent with a common requirement for intramolecular $CD4^+$ T cell help (Sette et al., 2008). Anti-spike IgA titers also correlated with spike-specific $CD4^+$ T cells ($p < 0.0002$, **Fig. S5**). Thus, COVID-19 patients make anti-spike RBD antibody responses commensurate with the magnitude of their spike-specific $CD4^+$ T cell response. We then assessed the relationship between the $CD4^+$ and $CD8^+$ T cell responses to SARS-CoV-2. SARS-CoV-2–specific $CD4^+$ and $CD8^+$ T cell responses were well correlated ($R=0.62$. $p = 0.0025$, **Fig. 4C and S5**). Thus, antibody, $CD4^+$ and $CD8^+$ T cell responses to SARS-CoV-2 were generally well correlated.

Pre-existing crossreactive coronavirus-specific T cells

While spike– and non-spike–specific $CD4^+$ T cell responses were detectable in all COVID-19 cases, cells were also detected in unexposed individuals (**Fig. 3A-B**). These responses were statistically significant for non-spike-specific $CD4^+$ T cell reactivity (non-spike, $p = 0.039$. Spike, $p = 0.067$. **Fig. 5A-B**). Non-spike–specific $CD4^+$ T cell responses were above the limit of detection in 50% of donors based on SI (**Fig. S3E**). All of the donors were recruited between 2015-2018, excluding any possibility of exposure to SARS-CoV-2. Four human coronaviruses are known causes of seasonal 'common cold' upper respiratory tract infections: HCoV-OC43, HCoV-HKU1, HCoV-NL63, and HCoV-229E. We tested the SARS-CoV-2 unexposed donors for seroreactivity to HCoV-OC43 and HCoV-NL63 as a representative betacoronavirus and alphacoronavirus, respectively. All donors were IgG seropositive to HCoV-OC43 and HCoV-NL63 RBD, to varying degrees (**Fig. 5C**), consistent with the endemic nature of these viruses (Gorse et al., 2010; Huang et al., 2020; Severance et al., 2008). We therefore examined whether these represented true pan-coronavirus T cells capable of recognizing SARS-CoV-2 epitopes.

SARS-CoV-2 ORF targets of $CD4^+$ and $CD8^+$ T cells

A most pressing, yet unresolved, set of issues in understanding SARS-CoV-2 immune responses is what antigens are targeted by $CD4^+$ and $CD8^+$ T cells, whether the corresponding antigens are the same or different, and how do they reflect the antigens currently considered for COVID-19 vaccine development. We synthesized sets of overlapping peptides spanning the entire sequence of SARS-CoV-2, and pooled them separately so that each pool would represent one antigen (with the exception of nsp3, for which two pools were made. **Table S1**).

In the case of $CD4^+$ T cell responses, no obvious pattern of antigen specificity was observed based on SARS-CoV-2 genome organization; however, coronaviruses increase protein synthesis of certain ORFs in infected cells via subgenomic RNAs. Accounting for the relative abundance of subgenomic RNAs (**Fig. 6A**) (Irigoyen et al., 2016; Snijder et al., 2003; Xie et al., 2020), the ORFs were re-ordered based on predicted protein abundance (**Fig. 6B**). A clear hierarchy of SARS-CoV-2–specific $CD4^+$ T cell targets was then apparent, with the majority of the $CD4^+$ T cell response in COVID-19

cases directed against highly expressed SARS-CoV-2 ORFs spike, M, and N. On average, these antigens accounted for 27%, 21% and 11% of the total $CD4^+$ T cell response, respectively. Most COVID-19 cases also had $CD4^+$ T cells specific for SARS-CoV-2 nsp3, nsp4 and ORF8 (**Fig. 6B**), on average each accounting for ~5% of the total $CD4^+$ T cell response (**Fig. 6C**). E, ORF6, hypothetical ORF10, and nsp1 are all small antigens (or potentially not expressed, in the case of ORF10) and were most likely predominantly unrecognized as a result. These results are somewhat unexpected, because data for other coronaviruses, from 27 different studies curated in the IEDB, reported that spike accounted for nearly two-thirds of reported $CD4^+$ T cell reactivity (**Table S3**). N accounted for most of the remaining epitopes in the published literature, although human N-specific $CD4^+$ T cell responses were not observed in one of the most comprehensive studies of human SARS-CoV-1 T cell responses (Li et al., 2008). Coronavirus M has not previously been described as a prominent target of $CD4^+$ T cell responses (**Table S3**). In sum, these results, fully scanning the SARS2 orfeome, demonstrate a pattern of robust and diverse SARS-CoV-2–specific $CD4^+$ T cell reactivity in convalescing COVID-19 cases that correlated largely with predicted viral protein abundance in infected cells.

When examining the non-exposed donors, the pattern of $CD4^+$ T cell targets changed. While S was still a relatively prominent target (23% of total, on average), there was no, or marginal, reactivity against SARS-CoV-2 N and M. Among donors with detectable $CD4^+$ T cells, a shift in reactivity was observed towards SARS-CoV-2 nsp14 (25%), nsp4 (15%) and nsp6 (14%) (**Fig 6B-C**). SARS-CoV-2–reactive $CD4^+$ T cells were detected in at least six different unexposed donors, demonstrating that the crossreactivity is relatively widely distributed (**Fig. S6A**).

Having scanned the full SARS-CoV-2 orfeome for $CD4^+$ T cell reactivity in multiple donors, it was possible to assess whether the epitope prediction MP approach successfully enriched for SARS-CoV-2 epitopes targeted by human $CD4^+$ T cells. When the total reactivity observed with the CD4_R MP was plotted versus the sum total of all antigen pools (excluding spike, given that spike predictions were not included in the CD4_R MP), a significant correlation was observed ($p < 0.0002$, **Fig S6C**). The single MP-R captured ~50% (44% +/- range 28-80%) of the non-spike response per COVID-19 donor, demonstrating the success of the prediction approach, which, as mentioned above, was devised to attempt to capture approximately 50% of the total response (Dhanda et al., 2018; Paul et al., 2015).

In the case of $CD8^+$ T cell responses, the data in the literature from other coronaviruses (57 different studies curated in the IEDB, **Table S3**) reported spike accounting for 50% and N accounting for 36% of the defined epitopes. In a large study of human SARS-CoV-1 responses, spike was reported as essentially the only target of $CD8^+$ T cell responses (Li et al., 2008), while in a study of MERS $CD8^+$ T cells, responses were noted for spike, N and a pool of M/E peptides (Zhao et al., 2017). Few epitopes have been reported from other coronavirus antigens (**Table S3**). Here, we scanned the full SARS-CoV-2 orfeome for $CD8^+$ T cell recognition. Our data indicate a somewhat different pattern of immunodominance for SARS-CoV-2 $CD8^+$ T cell reactivity (**Fig. 6D-E**), with spike protein accounting for ~26% of the reactivity, and N accounting for ~12%. Significant reactivity in COVID-19 recovered subjects was derived from other antigens, such as M (22%), nsp6 (15%), ORF8 (10%) and ORF3a (7%) (**Fig. 6D-E**). In unexposed donors, SARS-CoV-2–reactive $CD8^+$ T cells were detected in at least four different donors (**Fig. S7**), with less clear targeting of specific SARS-CoV-2 proteins than was observed for $CD4^+$ T cells, suggesting that coronavirus $CD8^+$ T cell crossreactivity exists but is less widespread than $CD4^+$ T cell crossreactivity.

DISCUSSION

There is a critical need for foundational knowledge about T cell responses to SARS-CoV-2. Here we report functional validation of predicted epitopes when arranged in epitope “megapools”, utilizing PBMCs derived from convalescing COVID-19 cases. The experiments also used protein-specific peptide pools to determine which SARS-CoV-2 proteins are the predominant targets of human SARS-CoV-2–specific $CD4^+$ and $CD8^+$ T cells generated during COVID-19 disease. Importantly, we utilized the exact same series of experimental techniques with blood samples from healthy control donors

(PBMCs collected in the 2015-2018 time frame), and substantial crossreactive coronavirus T cell memory was observed.

Our results demonstrate that the epitope MPs are reagents well suited to analyze and detect SARS-CoV-2–specific T cell responses with limited sample material. We also developed and tested peptide pools corresponding to each of the 25 proteins encoded in the SARS-CoV-2 genome. Data from both the epitope MPs and protein peptide pool experiments can be interpreted in the context of previously reported T cell response immunodominance patterns observed for other coronaviruses, particularly the SARS and MERS viruses, which have been studied in humans, HLA-transgenic mice, wild-type mice and other species. In the case of $CD4^+$ T cell responses, data for other coronaviruses found that spike accounted for nearly two-thirds of reported $CD4^+$ T cell reactivity, with N and M accounting for limited reactivity, and no reactivity in one large study of human SARS-CoV-1 responses (Li et al., 2008). Our SARS-CoV-2 data reveal that the pattern of immunodominance in COVID-19 is different. In particular, M, spike and N proteins were clearly co-dominant, each recognized by 100% of COVID-19 cases studied here. Significant $CD4^+$ T cell responses were also directed against nsp3, nsp4, ORF3s, ORF7a, nsp12 and ORF8. These data suggest that a candidate COVID-19 vaccine consisting only of SARS-CoV-2 spike would be capable of eliciting SARS-CoV-2–specific $CD4^+$ T cell responses of similar representation to that of natural COVID-19 disease, but the data also indicate that there are many potential $CD4^+$ T cell targets in SARS-CoV-2 and inclusion of additional SARS-CoV-2 structural antigens such as M and N would better mimic the natural SARS-CoV-2–specific $CD4^+$ T cell response observed in mild to moderate COVID-19 disease.

Regarding SARS-CoV-2 $CD8^+$ T cell responses, the pattern of immunodominance found here differed from the literature for other coronaviruses. However, stringent comparisons are not possible, as some earlier studies were not similarly comprehensive and did not utilize the same experimental strategy. The spike protein was a target of human SARS-CoV-2 $CD8^+$ T cell responses, but it is not dominant. SARS-CoV-2 M was just as strongly recognized, and significant reactivity was noted for other antigens, mostly nsp6, ORF3a, and N, which comprised nearly 50% of the total $CD8^+$ T cell response, on average. Thus, these data indicate that candidate COVID-19 vaccines endeavoring to elicit $CD8^+$ T cell responses against the spike protein will be eliciting a relatively narrow $CD8^+$ T cell response compared to the natural $CD8^+$ T cell response observed in mild to moderate COVID-19 disease. An optimal vaccine $CD8^+$ T cell response to SARS-CoV-2 might benefit from additional class I epitopes, such as the ones derived from the M, nsp6, ORF3a and/or N.

There have been concerns regarding vaccine enhancement of disease by certain candidate COVID-19 vaccine approaches, via antibody-dependent enhancement (ADE) or development of a T_H2 responses (Peeples, 2020). Herein, we saw predominant T_H1 responses in convalescing COVID-19 cases, with little to no T_H2 cytokines. Clearly more studies are required, but the data here appear to predominantly represent a classical T_H1 response to SARS-CoV-2.

While it was important to identify antigen-specific T cell responses in COVID-19 cases, it is also of great interest to understand whether crossreactive immunity exists between coronaviruses to any degree. A key step in developing that understanding is to examine antigen-specific $CD4^+$ and $CD8^+$ T cells in COVID-19 cases and in unexposed healthy controls, utilizing the exact same antigens and series of experimental techniques. $CD4^+$ T cell responses were detected in 40-60% of unexposed individuals. This may be reflective of some degree of crossreactive, preexisting immunity to SARS-CoV-2 in some, but not all, individuals. Whether this immunity is relevant in influencing clinical outcomes is unknown—and cannot be known without T cell measurements before and after SARS-CoV-2 infection of individuals—but it is tempting to speculate that the crossreactive $CD4^+$ T cells may be of value in protective immunity, based on SARS mouse models (Zhao et al., 2016). Clear identification of the crossreactive peptides, and their sequence homology relation to other coronaviruses, requires deconvolution of the positive peptide pools, which is not feasible with the cell numbers presently available, and time frame of the present study.

Regarding the value of crossreactive T cells, influenza (flu) immunology in relationship to pandemics may be instructive. In the context of the 2009 H1N1 influenza pandemic, preexisting T cell immunity existed in the adult population, which focused on the more conserved internal influenza viral proteins (Greenbaum et al., 2009). The presence of crossreactive T cells was found to correlate with

less severe disease (Sridhar et al., 2013; Wilkinson et al., 2012). The frequent availability of cross-reactive memory T cell responses might have been one factor contributing to the lesser severity of the H1N1 flu pandemic (Hancock et al., 2009). Cross-reactive immunity to influenza strains has been modeled to be a critical influencer of susceptibility to newly emerging, potentially pandemic, influenza strains (Gostic et al., 2016). Given the severity of the ongoing COVID-19 pandemic, it has been modeled that any degree of crossprotective coronavirus immunity in the population could have a very substantial impact on the overall course of the pandemic, and the dynamics of the epidemiology for years to come (Kissler et al., 2020).

Limitations and future directions.

Caveats of this study include the sample size and the focus on non-hospitalized COVID-19 cases. Sample size was limited by expediency. The focus on non-hospitalized cases of COVID-19 is a strength, in that these donors had uncomplicated disease of moderate duration, and thus it was encouraging that substantial $CD4^+$ T cell and antibody responses were detected in all cases, and $CD8^+$ T cell responses in the majority of cases. Complementing these data with MP T cell data from acute patients and patients with complicated disease course will also be of clear value, as will studies on the longevity of SARS-CoV-2 immunological memory. Additionally, lack of detailed information on common cold history or matched blood samples pre-exposure to SARS-CoV-2 prevents conclusions regarding the abundance of crossreactive coronavirus T cells before exposure to SARS-CoV-2 and any potential protective efficacy of such cells. Finally, full epitope mapping in the future will add important detailed resolution of the human coronavirus-specific T cell responses.

In sum, we measured SARS-CoV-2–specific $CD4^+$ and $CD8^+$ T cells responses in COVID-19 cases. Using multiple experimental approaches, SARS-CoV-2–specific $CD4^+$ T cell and antibody responses were observed in all COVID-19 cases, and $CD8^+$ T cell responses were observed in most. Importantly, pre-existing SARS-CoV-2–crossreactive T cell responses were observed in healthy donors, indicating some potential for pre-existing immunity in the human population. ORF mapping of T cell specificities revealed valuable targets for incorporation in candidate vaccine development, and revealed distinct specificity patterns between COVID-19 cases and unexposed healthy controls.

Journal Pre-proof

ACKNOWLEDGEMENTS

We would like to thank Cheryl Kim, Director of the LJI flow cytometry core facility for outstanding expertise. We thank Prof. Peter Kim, Abigail Powell, PhD, and colleagues (Stanford) for RBD protein synthesized from Prof. Florian Krammer (Mt. Sinai) constructs. J.M. was supported by Ph.D. student fellowships from the Departamento Administrativo de Ciencia, Tecnología e Innovación (COLCIENCIAS) and Pontificia Universidad Javeriana (Convocatoria 727 Doctorados Nacionales). This work was funded by the NIH NIAID under awards AI42742 (Cooperative Centers for Human Immunology) (S.C., A.S.), National Institutes of Health contract Nr. 75N9301900065 (A.S. and D.W.), and U19 AI118626 (A.S. and B.P.). The BD FACSymphony purchase was partially funded by the Bill and Melinda Gates Foundation and LJI Institutional Funds (S.C. and A.S.). This work was additionally supported in part by the Johnathan and Mary Tu Foundation (D.S.), the NIAID under K08 award AI135078 (J.D.), and UCSD T32s AI007036 and AI007384 Infectious Diseases Division (S.Ram., S.Raw.).

AUTHOR CONTRIBUTIONS

Conceptualization A.G, D.W, S.C and A.S.. Investigation A.G, D.W, J.M., C.M., J.M.D, D.M, L.P, R.S.J, A.S., D.W.. Formal Analysis A.G, D.W, C.M., J.M.D, J.M.. S.C.. Resources S.I.R, S.A.R, D.M.S, A.C., F.K., S.C., A.S.. Data curation J.A.G, B.P.. Writing S.C, A.S., A.G., D.W.. Supervision B.P, A. M d S., S.C and A.S.. Project Administration A.F. Funding Acquisition S.C., A.S, D.W., D.S., J.D..

Journal Pre-proof

FIGURE LEGENDS

Figure 1. SARS-CoV-2 IgM, IgA, and IgG responses of recovered COVID-19 patients.
(A-C) Plasma ELISA titers to SARS-CoV-2 spike RBD. (**A**) IgG. (**B**) IgM. (**C**) IgA. "Neg" = Unexposed donors from 2015-2018 (n = 20). "COVID" = Convalescing COVID-19 patients (n = 20). All data are shown as ELISA titers based on a standard. Dotted line indicates limit of detection. Geometric mean titers with geometric SDs are indicated.
(**D-H**) Immunophenotyping of mononuclear leukocytes. Frequency of (**D**) $CD3^+$ total T cells, (**E**) $CD4^+$ T cells ($CD4^+CD3^+$), (**F**) $CD8^+$ T cells ($CD8^+CD3^+$), (**G**) $CD19^+$ B cells ($CD19^+CD3^-$), (**H**) $CD3^-CD19^-$ cells, and (**I**) $CD14^+CD16^-$ monocytes ($CD3^-CD19^-CD56^-$) from the PBMCs of unexposed donors ("Neg", n=13) or convalescing COVID-19 patients ("COVID", n=14). Data were analyzed using Mann-Whitney with mean and standard deviation shown.
*p<0.05, ****p<0.0001. See also Figure S1 and S2.

Figure 2. SARS-CoV-2–specific $CD4^+$ T cell responses of recovered COVID-19 patients.
(A) SARS-CoV-2–specific $CD4^+$ T cells measured as percentage of AIM^+ ($OX40^+CD137^+$) $CD4^+$ T cells after stimulation of PBMCs with peptide pools encompassing spike only ("Spike") MP or the CD4_R MP representing all the proteome without spike ("Non-spike"). Data were background subtracted against DMSO negative control and are shown with geometric mean and geometric standard deviation. Samples were from unexposed donors ("Unexposed", n=11) and recovered COVID-19 patients ("COVID-19", n=10).
(**B)** FACS plot examples, gated on total $CD4^+$ T cells.
(**C)** AIM^+ $CD4^+$ T cell reactivity in COVID-19 cases between the negative control (DMSO) and antigen-specific stimulations.
(**D-E**) Cytokine levels in the supernatant of PBMCs from COVID-19 donors after stimulation with peptide pools (Spike and Non-Spike) or the negative control (DMSO). (**D**) IL-2. (**E**) IFNγ.
Statistical comparisons across cohorts were performed with the Mann-Whitney test. Pairwise comparisons (C-E) were performed with the Wilcoxon test.
**p<0.01; ***p<0.001. See also Figure S3.

Figure 3. SARS-CoV-2–specific $CD8^+$ T cell responses by recovered COVID-19 patients.
(A) SARS-CoV-2–specific $CD8^+$ T cells measured as percentage of AIM^+ ($CD69^+CD137^+$) $CD8^+$ T cells after stimulation of PBMCs with class I MPs (CD8-A, CD8-B, and the combined data ("Total")). Data were background subtracted against DMSO negative control and are shown with geometric mean and geometric standard deviation. Samples were from unexposed donors ("Unexposed", n=11) and recovered COVID-19 patients ("COVID-19", n=10).
(**B)** FACS plot examples.
(C) Percentage of $CD8^+$ T cells producing IFNγ in response to SARS-CoV-2 MPs, or CMV MP, in PBMCs from COVID-19 and unexposed donors after background subtraction. Data are shown with geometric mean and geometric standard deviation.
(D) Functional profile of $IFN\gamma^+CD8^+$ T cells producing granzyme B (GzB), TNFα (TNF), or IL-10 in response to SARS-CoV-2 MPs. Mean and SD are shown.
(E) FACS plot examples of IFNγ and granzyme B co-expression.
Statistical comparisons across cohorts were performed with the Mann-Whitney test.
*p<0.05; **p<0.01.; ns not significant. See also Figure S4.

Figure 4. Correlations between SARS-CoV-2–specific $CD4^+$ T cells, antibodies, and $CD8^+$ T cells.
(A) Correlation between SARS-CoV-2 spike–specific $CD4^+$ T cells (%) and anti-spike RBD IgG.
(**B**) Correlation between SARS-CoV-2 non-spike–specific $CD4^+$ T cells (%) and anti-spike RBD IgG.
(**C**) Correlation between SARS-CoV-2–specific $CD4^+$ T cells and SARS-CoV-2–specific $CD8^+$ T cells. Total MP responses per donor were used in each case ("Non-spike" + "spike" (CD4_R + MP_S) for $CD4^+$ T cells, CD8_A + CD8_B for $CD8^+$ T cells).
Statistical comparisons were performed using Spearman correlation. See also Figure S5.

Figure 5. SARS-CoV-2 epitope reactivity in unexposed individuals.
(A) SARS-CoV-2–reactive $CD4^+$ T cells measured as percentage of AIM^+ ($OX40^+CD137^+$) $CD4^+$ T cells in unexposed (n=11) donors.
(**B)** FACS plot examples, gated on total $CD4^+$ T cells.
(**C**) Plasma IgG ELISAs for seroreactivity to RBD of HCoV-OC43 or HCoV-NL63.
Pairwise statistical comparisons (A) were performed with the Wilcoxon test. *p<0.05; ns not significant.

Figure 6. Protein immunodominance of SARS-CoV-2 specific $CD4^+$ and $CD8^+$ T cells in COVID-19 cases and unexposed donors.
(A) SARS-CoV-2 genome organization and predicted viral protein abundance in infected cells.
(**B)** SARS-CoV-2 antigen-specific $CD4^+$ T cells (AIM^+, $OX40^+CD137^+$) quantified by stimulation index, using a peptide pool for each viral protein (with two exceptions, see **Table S1**). COVID-19 cases (top, in blue. n=10) and unexposed donors (bottom, in white. n=10).
(**C)** Fraction of SARS-CoV-2 proteins recognized by $CD4^+$ T cells in COVID-19 cases (top) and unexposed donors (bottom).
(**D)** SARS-CoV-2 antigen-specific $CD4^+$ T cells (AIM^+, $OX40^+CD137^+$) quantified by stimulation index, using a peptide pool for each viral protein (with two exceptions, see **Table S1**). COVID-19 cases (top, in red. n=10) and unexposed donors (bottom, in grey. n=10).
(**E)** Fraction of SARS-CoV-2 proteins recognized by $CD8^+$ T cells in COVID-19 cases (top) and unexposed donors (bottom).
See also Figures S6, S7 and Table S6.

Journal Pre-proof

TABLES

Table 1. Participant characteristics

	Unexposed (n = 20)	COVID-19 (n = 20)
Age (years)	20-66 [Median = 31, IQR = 21]	20-64 [Median = 44, IQR = 9]
Gender		
Male (%)	35% (7/20)	45% (9/20)
Female (%)	65% (13/20)	55% (11/20)
Residency		
California (%)	95% (19/20)	100% (20/20)
USA, Non-California (%)	5% (1/20)	0% (0/20)
Sample Collection Date	March 2015-March 2018	March-April 2020
SARS-CoV-2 PCR Positivity	N/A	100% (16/16 tested)
Antibody Test Positivity[1]	N/A	90% (18/20)
Disease Severity[2]		
Mild	N/A	70% (14/20)
Moderate	N/A	20% (4/20)
Severe	N/A	10% (2/20)
Critical	N/A	0% (0/20)
Symptoms		
Cough	N/A	79% (15/19)
Fatigue	N/A	42% (8/19)
Fever	N/A	37% (7/19)
Anosmia	N/A	21% (4/19)
Dyspnea	N/A	16% (3/19)
Diarrhea	N/A	5% (1/19)
Days Post Symptom Onset at Collection	N/A	20-36 (18/20) [Median = 26, IQR = 7]
Past Medical History		
No known	N/A	65% (13/20)
Hyperlipidemia	N/A	15% (3/20)
Hypertension	N/A	10% (2/20)
Asthma	N/A	10% (2/20)
Known or suspected sick contact/exposure	N/A	75% (15/20)

[1] Commercial skin prick lateral flow assay.
[2] WHO criteria.

Journal Pre-proof

SUPPLEMENTAL MATERIALS

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. SARS-CoV-2 spike protein RBD Serology. Related to Figure 1.
(A-C) ELISA curves for (**A**) IgG, (**B**) IgM, and (**C**) IgA from 10 representative donors. Five COVID-19 cases and
(**D-F**) Area under the curve (AUC) SARS-CoV-2 spike protein RBD (**D**) IgG (**E**) IgM, and (**F**) IgA, ELISA quantitation, from the same donors and experiments shown in Figure 1. Geometric mean titers with geometric SDs are indicated. P values are two-tailed Mann-Whitney tests.

Figure S2. Phenotyping flow cytometry. Related to Figure 1.
Representative gating of $CD3^+$ T cells, $CD19^+$ B cells, $CD3^-CD19^-$ cells, $CD4^+$ T cells, $CD8^+$ T cells and $CD14^+$ monocytes from donor PBMCs is shown. Briefly, mononuclear cells were gated out of all events followed by subsequent singlet gating. Live cells are gated as Zombie UV^-. Cells were then gated as CD19-PE-$Cy5^+$, CD3-$buv395^+$ or $CD19^-CD3^-$ cells. T cells were further subdivided into either CD8-$buv805^+$ or CD4-$PerCPefluor710^+$ populations. $CD3^-CD19^-$ cells were defined as CD56-PE-$Dazzle^{bright}$ NK cells, $CD56^{dim}$CD-16$buv737^+$ NK cells or $CD56^-$ monocytes. Monocytes were further classified on differential expression of CD14-bv510 and CD16.

Figure S3. SARS-CoV-2–specific $CD4^+$ T cell responses of recovered COVID-19 patients. Related to Figure 2.
(A) Example flow cytometry gating strategy.
(**B**) FACS plot examples for controls. DMSO negative control, CMV positive control, PHA positive control.
(**C**) CMV-specific $CD4^+$ T cells as percentage of AIM^+ ($OX40^+CD137^+$) $CD4^+$ T cells after stimulation of PBMCs with CMV peptide pool. Data were background subtracted against DMSO negative control and are shown with geometric mean and geometric standard deviation. Samples were from unexposed donors (“Unexposed”, n=11) and recovered COVID-19 patients (“COVID-19”, n=10).
(**D**) Spearman correlation of SARS-CoV-2 spike–specific $CD4^+$ T cells (AIM^+ ($OX40^+CD137^+$) $CD4^+$ T cells, background subtracted) after stimulation with spike pool run on the same donors in two independent experiment series run on different dates. COVID-19 patient samples shown in blue. Unexposed donor samples shown in black.
(**E-F**) Stimulation index quantitation of AIM^+ ($OX40^+CD137^+$) $CD4^+$ T cells; the same samples as in Figure 2 and Fig S3C were analyzed.
(**G-H**) Cytokine levels in the supernatants of AIM assays after stimulation with (**G**) Spike MP (MP_S), or (**H**) CD4-R (“Non-spike”). Data are shown in comparison to the negative control (DMSO), per donor.
(**I-J**) Cytokine production by $CD4^+$ T cells in response to Non-spike (CD4-R MP) or Spike (MP_S) peptide pools (“CoV antigen (Ag)”) was confirmed by analyzing cytokine secretion from the subset of COVID-19 donors determined to have low or negative $CD8^+$ T cell responses (<0.1% by AIM) to the same peptide pool determined positive for SARS-CoV-2–specific $CD4^+$ T cells by AIM. (**I**) IL-2. (**J**) IFNγ. Statistical comparisons across cohorts were performed with the Mann-Whitney test, while paired sample comparisons were performed with the Wilcoxon test. **p<0.01; ***p<0.001. ns not significant.

Figure S4. SARS-CoV-2–specific $CD8^+$ T cell responses of recovered COVID-19 patients. Related to Figure 3.
(**A)** Flow cytometry gating strategy.
(**B)** SARS-CoV-2–specific $CD8^+$ T cells as determined by AIM^+ ($CD69^+CD137^+$) $CD8^+$ T cells. Response of PBMCs from COVID-19 cases between the negative control (DMSO) and antigen specific stimulation.
(**C)** CMV-specific $CD8^+$ T cells as percentage of AIM^+ ($CD69^+CD137^+$) $CD8^+$ T cells after stimulation of PBMCs with CMV peptide pool. Data were background subtracted against DMSO negative control and

are shown with geometric mean and geometric standard deviation. Samples were from unexposed donors ("Unexposed", n=11) and recovered COVID-19 patients ("COVID-19", n=10).
(**D-E)** Stimulation index quantitation of AIM^+ ($CD69^+CD137^+$) $CD8^+$ T cells; the same samples as in Figure 2 and Fig S4C were analyzed.
Statistical comparisons across cohorts were performed with the Mann-Whitney test, while paired sample comparisons were performed with the Wilcoxon test. **p<0.01; ***p<0.001. ns not significant.

Figure S5. Correlations between SARS-CoV-2–specific $CD4^+$ T cells, antibodies, and $CD8^+$ T cells. Related to Figure 4.
(A) Correlation between SARS-CoV-2 spike–specific $CD4^+$ T cells and anti-spike RBD IgG, using $CD4^+$ T cell stimulation index.
(B) Correlation between SARS-CoV-2 non-spike–specific $CD4^+$ T cells and anti-spike RBD IgG, using $CD4^+$ T cell stimulation index.
(**C**) Correlation between SARS-CoV-2 spike–specific $CD4^+$ T cells (%) and anti-spike RBD IgA.
(**D**) Correlation between SARS-CoV-2 spike–specific $CD4^+$ T cells (%) and anti-spike RBD IgA.
(**E**) Correlation between SARS-CoV-2–specific $CD4^+$ T cells and SARS-CoV-2–specific $CD8^+$ T cells, using stimulation index. Total MP responses per donor were used in each case ("Non-spike" + "spike" (CD4_R + MP_S) for $CD4^+$ T cells, CD8-A + CD8-B for $CD8^+$ T cells).
Statistical comparisons were performed using Spearman correlation.

Figure S6. Protein immunodominance of SARS-CoV-2 specific $CD4^+$ T cells in recovered COVID-19 patients and unexposed donors. **Related to Figure 6.**
(A) The same data as Figure 6B, but with each unexposed donor color coded.
(**B**) The same experiment as Figure 6B, but with SARS-CoV-2–specific $CD4^+$ T cells measured as percentage of AIM^+ ($OX40^+CD137^+$) $CD4^+$ T cells, after background subtraction. COVID-19 cases (top, in blue. n=10) and unexposed donors (bottom, in white. n=10).
(**C**) Correlation of SARS-CoV-2–specific $CD4^+$ T cells detected using the epitope prediction approach (CD4_R MP) compared against the sum total of all antigen pools of overlapping peptides (excluding spike), run with samples from the same donors in two different experiment series. Dotted line indicates 1:1 concordance. Statistical comparison was performed using Spearman correlation.

Figure S7. Protein immunodominance of SARS-CoV-2 specific $CD8^+$ T cells in recovered COVID-19 patients and unexposed donors. Related to Figure 6.
(A) The same data as Figure 6D, but with each unexposed donor color coded.
(**B**) The same experiment as Figure 6D, but with SARS-CoV-2–specific $CD8^+$ T cells measured as percentage of AIM^+ ($CD69^+CD137^+$) $CD8^+$ T cells, after background subtraction. COVID-19 cases (top, in red. n=10) and unexposed donors (bottom, in grey. n=10).

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

Healthy Unexposed Donors. Samples from healthy adult donors were obtained by the La Jolla Institute for Immunology (LJI) Clinical Core or provided by a commercial vendor (Carter Blood Care) for prior, unrelated studies between early 2015 and early 2018. These samples were considered to be from unexposed controls, given that SARS-CoV-2 emerged as a novel pathogen in late 2019, more than one year after the collection of any of these samples. These donors were considered healthy in that they had no known history of any significant systemic diseases, including, but not limited to, autoimmune disease, diabetes, kidney or liver disease, congestive heart failure, malignancy, coagulopathy, hepatitis B or C, or HIV. An overview of the characteristics of these unexposed donors is provided in **Table 1**.

The LJI Institutional Review Board approved the collection of these samples (LJI; VD-112). At the time of enrollment in the initial studies, all individual donors provided informed consent that their samples could be used for future studies, including this study.

Convalescent COVID-19 Donors. The Institutional Review Boards of the University of California, San Diego (UCSD; 200236X) and La Jolla Institute (LJI; VD-214) approved blood draw protocols for convalescent donors. All human subjects were assessed for capacity using a standardized and approved assessment. Subjects deemed to have capacity voluntarily gave informed consent prior to being enrolled in the study. Individuals did not receive compensation for their participation.

Study inclusion criteria included subjects over the age of 18 years, regardless of disease severity, race, ethnicity, gender, pregnancy or nursing status, who were willing and able to provide informed consent, or with a legal guardian or representative willing and able to provide informed consent when the participant could not personally do so. Study exclusion criteria included lack of willingness or ability to provide informed consent, or lack of an appropriate legal guardian or representative to provide informed consent.

Blood from convalescent donors was obtained at a UC San Diego Health clinic. Blood was collected in acid citrate dextrose (ACD) tubes and stored at room temperature prior to processing for PBMC isolation and plasma collection. A separate serum separator tube (SST) was collected from each donor. Samples were de-identified prior to analysis. Other efforts to maintain the confidentiality of participants included referring to specimens and other records via an assigned, coded identification number.

Prior to enrollment in the study, donors were asked to provide proof of positive testing for SARS-CoV-2, and screened for clinical history and/or epidemiological risk factors consistent with the World Health Organization (WHO) or Centers for Disease Control and Prevention (CDC) case definitions of COVID-19 or Persons Under Investigation (PUI) (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/surveillance-and-case-definitions, https://www.cdc.gov/coronavirus/2019-nCoV/hcp/clinical-criteria.html). Per CDC and WHO guidance, clinical features consistent with COVID-19 included subjective or measured fever, signs or symptoms of lower respiratory tract illness (e.g. cough or dyspnea). Epidemiologic risk factors

included close contact with a laboratory-confirmed case of SARS-CoV-2 within 14 days of symptom onset or a history of travel to an area with a high rate of COVID-19 cases within 14 days of symptom onset.

Disease severity was defined as mild, moderate, severe or critical based on a modified version of the WHO interim guidance, "Clinical management of severe acute respiratory infection when COVID-19 is suspected" (WHO Reference Number: WHO/2019-nCoV/clinical/2020.4). Mild disease was defined as an uncomplicated upper respiratory tract infection (URI) with potential non-specific symptoms (e.g. fatigue, fever, cough with or without sputum production, anorexia, malaise, myalgia, sore throat, dyspnea, nasal congestion, headache; rarely diarrhea, nausea and vomiting) that did not require hospitalization. Moderate disease was defined as the presence of lower respiratory tract disease or pneumonia without the need for supplemental oxygen, without signs of severe pneumonia, or a URI requiring hospitalization (including observation admission status). Severe disease was defined as severe lower respiratory tract infection or pneumonia with fever plus any one of the following: tachypnea (respiratory rate > 30 breaths per minute), respiratory distress, or oxygen saturation less than 93% on room air. Critical disease was defined as the need for ICU admission or the presence of acute respiratory distress syndrome (ARDS), sepsis, or septic shock, as defined in the WHO guidance document.

Convalescent donors were screened for symptoms prior to scheduling blood draws, and had to be symptom-free and approximately 3 weeks out from symptom onset at the time of the initial blood draw. Following enrollment, whole blood from convalescent donors was run on a colloidal-gold immunochromatographic 'lateral flow' assay to evaluate for prior exposure to SARS-CoV-2. This assay detects IgM or IgG antibodies directed against recombinant SARS-CoV-2 antigen labeled with a colloidal gold tracer (20/20 BioResponse CoronaCheck). Ninety percent of convalescent donors tested positive for IgM or IgG to SARS-CoV-2 by this assay (**Table 1**).

Convalescent donors were California residents, who were either referred to the study by a health care provider or self-referred. The majority (75%) of donors had a known sick contact with COVID-19 or suspected exposure to SARS-CoV-2 (**Table 1**). The most common symptoms reported were cough, fatigue, fever, anosmia, and dyspnea. Seventy percent of donors experienced mild illness. Donors were asked to self-report any known medical illnesses. Of note, 65% of these individuals had no known underlying medical illnesses.

METHOD DETAILS

Peptide Pools

Epitope MegaPool (MP) design and preparation. SARS-CoV-2 virus-specific CD4 and CD8 peptides were synthesized as crude material (A&A, San Diego, CA), resuspended in DMSO, pooled and sequentially lyophilized as previously reported (Carrasco Pro et al., 2015). SARS-CoV-2 epitopes were predicted using the protein sequences derived from the SARS-CoV-2 reference (GenBank: MN908947) and IEDB analysis-resource as previously described (Dhanda et al., 2019; Grifoni et al., 2020). Specifically, CD4 SARS-CoV-2 epitope prediction was carried out using a previously described approach in Tepitool resource in IEDB (Paul et al., 2015; Paul et al., 2016), to select peptides with median consensus percentile ≤ 20, similar to what was previously described, but removing the resulting spike glycoprotein epitopes from this prediction (CD4-R (remainder) "Non-spike" MP, n=221). This approach takes advantage of the extensive cross-reactivity and repertoire overlap between different HLA class II loci and allelic variants to predict promiscuous epitopes, capable of binding across the most common HLA class II prototypic specificities (Greenbaum et al., 2011; O'Sullivan et al., 1991; Sidney et al., 2010a, b; Southwood et al., 1998). The algorithm utilizes predictions for seven common HLA-DR alleles **(**DRB1*03:01, DRB1*07:01, DRB1*15:01, DRB3*01:01, DRB3*02:02, DRB4*01:01 and DRB5*01:01) empirically determined to allow coverage of diverse populations and for different pathogens and antigen systems (Dhanda et al., 2018; Paul et al., 2015).

To investigate in-depth spike-specific CD4 T cells, 15-mer peptides (overlapping by 10 amino acids) spanning the entire antigen have been synthesized and pooled separately (CD-4 S (spike) MP, n=253).

In the case of CD8 epitopes, since the overlap between different HLA class I allelic variants and loci is more limited to specific groups of alleles, or supertypes (Sidney et al., 2008), we targeted a set of

the 12 most prominent HLA class I A and B alleles (A*01:01, A*02:01, A*03:01, A*11:01, A*23:01, A*24:02, B*07:02, B*08:01, B*35:01, B*40:01, B*44:02, B*44:03), which have been shown to allow broad coverage of the general population. CD8 SARS-CoV-2 epitope prediction was performed as previously reported, using NetMHC pan EL 4.0 algorithm (Jurtz et al., 2017) for the top 12 more frequent HLA alleles and selecting the top 1 percentile predicted epitope per HLA allele clustered with nested/overlap reduction (Grifoni et al., 2020). The 628 predicted CD8 epitopes were split in two CD8 MPs containing 314 peptides each (CD8-A and CD8-B). The CMV MP is a pool of previously reported class I and class II epitopes (Carrasco Pro et al., 2015).

Protein peptide pools. In the case of the protein pools, peptides of 15 amino acid length overlapping by 10 spanning each entire protein sequence were tested in a single MP (6-253 peptides per pool). **Table S1** lists the number of peptides pooled for each of the viral proteins. Upon request we are prepared to make these MP available to the scientific community for use in a diverse set of investigations.

PBMC isolation

For all samples whole blood was collected in ACD tubes (COVID-19 donors) or heparin coated blood bag (healthy unexposed donors). Whole blood was then centrifuged for 15 min at 1850 rpm to separate the cellular fraction and plasma. The plasma was then carefully removed from the cell pellet and stored at -20C.

Peripheral blood mononuclear cells (PBMC) were isolated by density-gradient sedimentation using Ficoll-Paque (Lymphoprep, Nycomed Pharma, Oslo, Norway) as previously described (Weiskopf et al., 2013). Isolated PBMC were cryopreserved in cell recovery media containing 10% DMSO (Gibco), supplemented with 10% heat inactivated fetal bovine serum, depending on the processing laboratory, (FBS; Hyclone Laboratories, Logan UT) and stored in liquid nitrogen until used in the assays.

SARS-CoV-2 RBD ELISA

SARS-CoV-2 Receptor Binding Domain (RBD) protein was obtained courtesy of Florian Krammer and Peter Kim (Stadlbauer et al., 2020). Corning 96-well half-area plates (ThermoFisher 3690) were coated with 1µg/mL SARS-CoV-2 RBD overnight at 4°C. ELISA protocol generally followed that of the Krammer lab, which previously demonstrated specificity (Stadlbauer et al., 2020). Plates were blocked the next day with 3% milk (Skim Milk Powder ThermoFisher LP0031 by weight/volume) in Phosphate Buffered Saline (PBS) containing 0.05% Tween-20 (ThermoScientific J260605-AP) for 2 hours at room temperature. Plasma was then added to the plates and incubated for 1.5 hours at room temperature. Prior to plasma addition to the plates, plasma was heat inactivated at 56°C for 30-60 minutes. Plasma was diluted in 1% milk in 0.05% PBS-Tween 20 starting at a 1:3 dilution and diluting each sample at by 1:3. Plates were then washed 5 times with 0.05% PBS-Tween 20. Secondary antibodies were diluted in 1% milk in 0.05% Tween-20 and incubated for 1 hour. For IgG, anti-human IgG peroxidase antibody produced in goat (Sigma A6029) was used at a 1:5000 dilution. For IgM, anti-human IgM peroxidase antibody produced in goat (Sigma A6907) was used at a 1:10,000 dilution. For IgA, anti-human IgA horseradish peroxidase antibody (Hybridoma Reagent Laboratory HP6123-HRP) was used at a 1:1,000 dilution. Plates were washed 5 times with 0.05% PBS-Tween 20. Plates were developed with TMB Substrate Kit (ThermoScientific 34021) for 15 minutes at room temperature. The reaction was stopped with 2M sulfuric acid. Plates were read on a Spectramax Plate Reader at 450 nm using SoftMax Pro, and ODs were background subtracted. A positive control standard was created by pooling plasma from six convalescing COVID-19 patients. Positive control standard was run on each plate and was used to calculate titers (relative units) for all samples using non-linear regression interpolations, done to quantify the amount of anti-RBD IgG, anti-RBD IgM, and anti-RBD IgA present in each specimen. Titers were plotted for each specimen and compared to COVID-19 negative specimens. As a second analytical approach, Area under the curve was also calculated for each specimen to compare COVID-19 to negative specimens, using a baseline of 0.05 for peak calculations.

OC43 and NL63 coronavirus RBD ELISA

An in-house ELISA at UNC was performed by coating with recombinant S RBD antigens (SARS-CoV-2, SARS-CoV, OC43-CoV and NL63-CoV) in TBS for 1 h at 37 C. After blocking, we added 1:20 diluted serum and incubated at 37°C for 1 h. Antigen-specific antibodies (Ig) were measured at 405 nm by using alkaline phosphatase conjugated goat anti-human IgG, IgA and IgM Abs and 4-Nitrophenyl phosphate.

Flow Cytometry

Direct ex vivo PBMC immune cell phenotyping. For the surface stain, $1x10^6$ PBMCs were resuspended in 100 µl PBS with 2% FBS (FACS buffer) and stained with antibody cocktail for 1 hour at 4°C in the dark. Following surface staining, cells were washed twice with FACS buffer. Cells were then fixed/permeabilized for 40min at 4C in the dark using the eBioscience FoxP3 transcription factor buffer kit (ThermoFisher Scientific, Waltham, MA). Following fixation/permeabilization, cells were washed twice with 1x permeabilization buffer, resuspended in 100 µl permeabilization buffer and stained with intracellular/intranuclear antibodies for 1 hour at 4ºC in the dark. Samples were washed twice with 1x permeabilization buffer following staining. After the final wash, cells were resuspended in 200µl FACS buffer. All samples were acquired on a BD FACSymphony cell sorter (BD Biosciences, San Diego, CA). A list of antibodies used in this panel can be found in **Table S2.**

T cell stimulations. For all flow cytometry assays of stimulated T cells, cryopreserved cells were thawed by diluting them in 10 ml complete RPMI 1640 with 5% human AB serum (Gemini Bioproducts) in the presence of benzonase [20ul/10mL]. All samples were acquired on a ZE5 Cell analyzer (Bio-rad laboratories), and analyzed with FlowJo software (Tree Star, San Carlos, CA).

Activation induced cell marker assay. Cells were cultured for 24 hours in the presence of SARS-CoV-2 specific MPs [1 µg/ml] or 10 µg/mL PHA in 96-wells U bottom plates at $1x10^6$ PBMC per well. A stimulation with an equimolar amount of DMSO was performed as negative control, phytohemagglutinin (PHA, Roche, 1µg/ml) and stimulation with a combined CD4 and CD8 cytomegalovirus MP (CMV, 1µg/ml) were included as positive controls. Supernatants were harvested at 24 hours post-stimulation for multiplex detection of cytokines. Antibodies used in the AIM assay are listed in **Table S4.** AIM assays shown in Figure 2 and 3 and AIM assays shown in Figure 6 had five COVID-19 donors in common and nine Unexposed donors. Full raw data is listed in **Table S6.**

Intracellular cytokine staining assay. For the intracellular cytokine staining, PBMC were cultured in the presence of SARS-CoV-2 specific MPs [1 µg/ml] for 9 hours. Golgi-Plug containing brefeldin A (BD Biosciences, San Diego, CA) and monensin (Biolegend, San Diego, CA) were added 3 hours into the culture. Cells were then washed and surface stained for 30 minutes on ice, fixed with 1% of paraformaldehyde (Sigma-Aldrich, St. Louis, MO) and kept at 4^oC overnight. Antibodies used in the ICS assay are listed in **Table S5.** The gates applied for the identification of IFNγ, GzB, TNFα, or IL-10 production on the total population of $CD8^+$ T cells were defined according to the cells cultured with DMSO for each individual.

Cytokine bead assays

Supernatants were collected from 24-hour stimulation cultures of the AIM assays and stored in 96 well plates at -20 °C. Cytokines in cell culture supernatants of the same samples used for AIM were quantified using a human Th cytokine panel (13-plex) kit (LEGENDplex, Biolegend) according to the manufacturer's instruction. Supernatants were mixed with beads coated with capture antibodies specific for IL-5, IL-13, IL-2, IL-6, IL-9, IL-10, IFNγ, TNFα, IL-17a, IL-17F, IL-4, IL-21 and IL-22 and incubated on a 96 well filter plate for 2 hours. Beads were washed and incubated with biotin-labelled detection antibodies for 1 hour, followed by a final incubation with streptavidin-PE. Beads were analyzed by flow cytometry using a FACS Canto cytometer. Analysis was performed using the LEGENDplex analysis software v8.0, which distinguishes between the 13 different analytes on basis of bead size and internal dye.

Identification of coronavirus epitopes and associated literature references

To identify coronavirus epitopes and associated references, the IEDB was searched (on April 16, 2020) utilizing the following queries. A first query was run to identify references associated with class I restricted CD8 epitopes, which utilized the criteria settings "Antigen": Organism = Coronavirus (taxonomy ID 11118); "Assay": Positive assays only; "Assay": T cell assay; "MHC restriction" = MHC Class II; no parameters were defined for "Host" or "Disease." This query identified 57 references, which are listed and displayed under the "References" tab on the results page.

A second query was run to identify references associated with class II restricted CD4 epitopes which utilized the criteria settings "Antigen": Organism = Coronavirus (taxonomy ID 11118); "Assay": Positive assays only; "Assay": T cell assay; "MHC restriction" = MHC Class II; no parameters were defined for "Host" or "Disease." This query identified 27 references, which are listed and displayed under the "References" tab on the results page.

A third query was run to specifically capture epitopes and map them back to the antigen of origin using the setting; "Antigen": Organism = Coronavirus (taxonomy ID 11118); "Assay": Positive assays only; "Assay": T cell assay; no parameters were defined for "MHC restriction," "Host" or "Disease." Results were exported as csv files, and then examined in Excel to tabulate the number of CD4 and CD8 epitopes recognized in humans, mice, transgenic mice and other hosts associated with each respective antigen.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data and statistical analyses were done in FlowJo 10 and GraphPad Prism 8.4, unless otherwise stated. The statistical details of the experiments are provided in the respective figure legends. Data plotted in linear scale were expressed as Mean + Standard Deviation (SD). Data plotted in logarithmic scales were expressed as Geometric Mean + Geometric Standard Deviation (SD). Correlation analyses were performed using Spearman, while Mann-Whitney or Wilcoxon tests were applied for unpaired or paired comparisons, respectively. Details pertaining to significance are also noted in the respective legends. T cell data have been calculated as background subtracted data or stimulation index. Background subtracted data were derived by subtracting the percentage of AIM^+ cells after SARS-CoV-2 stimulation from the DMSO stimulation. Stimulation Index was calculated instead by dividing the percentage of AIM+ cells after SARS-CoV-2 stimulation with the percentage of AIM^+ cells derived from DMSO stimulation. If the AIM^+ cells percentage after DMSO stimulation was equal to 0, the minimum value across each cohort was used. When two stimuli were combined together, the percentage of AIM^+ cells after SARS-CoV-2 stimulation was combined and either subtracted twice or divided by twice the value of the percentage of AIM+ cells derived from DMSO stimulation. Additional data analysis techniques are described in the Methods sections above.

SUPPLEMENTAL EXCEL TABLE LEGEND

Table S6. Raw data for Figures 2,3 and 6. Related to Figures 2,3 and 6.

DECLARATION OF INTEREST

The authors declare no competing interests.

REFERENCES

Alshukairi, A.N., Zheng, J., Zhao, J., Nehdi, A., Baharoon, S.A., Layqah, L., Bokhari, A., Al Johani, S.M., Samman, N., Boudjelal, M.*, et al.* (2018). High Prevalence of MERS-CoV Infection in Camel Workers in Saudi Arabia. mBio *9*, e01985-01918.

Amanat, F., and Krammer, F. (2020). SARS-CoV-2 Vaccines: Status Report. Immunity *52*, 583-589.

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.

Bancroft, T., Dillon, M.B., da Silva Antunes, R., Paul, S., Peters, B., Crotty, S., Lindestam Arlehamn, C.S., and Sette, A. (2016). Th1 versus Th2 T cell polarization by whole-cell and acellular childhood pertussis vaccines persists upon re-immunization in adolescence and adulthood. Cell Immunol *304-305*, 35-43.

Blanco-Melo, D., Nilsson-Payant, B.E., Liu, W.-C., Møller, R., Panis, M., Sachs, D., Albrecht, R.A., and tenOever, B.R. (2020). SARS-CoV-2 launches a unique transcriptional signature from in vitro, ex vivo, and in vivo systems. bioRxiv, 2020.2003.2024.004655.

Callow, K.A., Parry, H.F., Sergeant, M., and Tyrrell, D.A. (1990). The time course of the immune response to experimental coronavirus infection of man. Epidemiol Infect *105*, 435-446.

Cao, X. (2020). COVID-19: immunopathology and its implications for therapy. Nat Rev Immunol *20*, 269-270.

Carrasco Pro, S., Sidney, J., Paul, S., Lindestam Arlehamn, C., Weiskopf, D., Peters, B., and Sette, A. (2015). Automatic Generation of Validated Specific Epitope Sets. J Immunol Res *2015*, 763461.

Choe, P.G., Perera, R., Park, W.B., Song, K.H., Bang, J.H., Kim, E.S., Kim, H.B., Ko, L.W.R., Park, S.W., Kim, N.J.*, et al.* (2017). MERS-CoV Antibody Responses 1 Year after Symptom Onset, South Korea, 2015. Emerg Infect Dis *23*, 1079-1084.

Crotty, S. (2019). T Follicular Helper Cell Biology: A Decade of Discovery and Diseases. Immunity *50*, 1132-1148.

da Silva Antunes, R., Paul, S., Sidney, J., Weiskopf, D., Dan, J.M., Phillips, E., Mallal, S., Crotty, S., Sette, A., and Lindestam Arlehamn, C.S. (2017). Definition of Human Epitopes Recognized in Tetanus Toxoid and Development of an Assay Strategy to Detect Ex Vivo Tetanus CD4+ T Cell Responses. PLoS One *12*, e0169086.

Dan, J.M., Havenar-Daughton, C., Kendric, K., Al-Kolla, R., Kaushik, K., Rosales, S.L., Anderson, E.L., LaRock, C.N., Vijayanand, P., Seumois, G.*, et al.* (2019). Recurrent group A Streptococcus tonsillitis is an immunosusceptibility disease involving antibody deficiency and aberrant TFH cells. Sci Transl Med *11*, eaau3776.

Dan, J.M., Lindestam Arlehamn, C.S., Weiskopf, D., da Silva Antunes, R., Havenar-Daughton, C., Reiss, S.M., Brigger, M., Bothwell, M., Sette, A., and Crotty, S. (2016). A Cytokine-Independent Approach To Identify Antigen-Specific Human Germinal Center T Follicular Helper Cells and Rare Antigen-Specific CD4+ T Cells in Blood. J Immunol *197*, 983-993.

Dhanda, S.K., Karosiene, E., Edwards, L., Grifoni, A., Paul, S., Andreatta, M., Weiskopf, D., Sidney, J., Nielsen, M., Peters, B.*, et al.* (2018). Predicting HLA CD4 Immunogenicity in Human Populations. Front Immunol *9*, 1369.

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.

Dong, E., Du, H., and Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis.

Giamarellos-Bourboulis, E.J., Netea, M.G., Rovina, N., Akinosoglou, K., Antoniadou, A., Antonakos, N., Damoraki, G., Gkavogianni, T., Adami, M.E., Katsaounou, P.*, et al.* (2020). Complex Immune Dysregulation in COVID-19 Patients with Severe Respiratory Failure. Cell Host Microbe, pii: S1931-3128(1920)30236-30235.

Gorse, G.J., Patel, G.B., Vitale, J.N., and Connor, T.Z. (2010). Prevalence of Antibodies to Four Human Coronaviruses Is Lower in Nasal Secretions than in Serum. Clinical and Vaccine Immunology *17*, 1875.

Gostic, K.M., Ambrose, M., Worobey, M., and Lloyd-Smith, J.O. (2016). Potent protection against H5N1 and H7N9 influenza via childhood hemagglutinin imprinting. Science *354*, 722-726.

Greenbaum, J., Sidney, J., Chung, J., Brander, C., Peters, B., and Sette, A. (2011). Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. Immunogenetics *63*, 325-335.

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., Angelo, M.A., Lopez, B., O'Rourke, P.H., Sidney, J., Cerpas, C., Balmaseda, A., Silveira, C.G.T., Maestri, A., Costa, P.R.*, et al.* (2017). Global Assessment of Dengue Virus-Specific CD4+ T Cell Responses in Dengue-Endemic Areas. Front Immunol *8*, 1309.

Grifoni, A., Sidney, J., Zhang, Y., Scheuermann, R.H., Peters, B., and Sette, A. (2020). A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. Cell Host Microbe *27*, 671-680 e672.

Guo, T., Fan, Y., Chen, M., Wu, X., Zhang, L., He, T., Wang, H., Wan, J., Wang, X., and Lu, Z. (2020a). Cardiovascular Implications of Fatal Outcomes of Patients With Coronavirus Disease 2019 (COVID-19). JAMA Cardiol, doi: 10.1001/jamacardio.2020.1017. [Epub ahead of print].

Guo, X., Guo, Z., Duan, C., chen, Z., Wang, G., Lu, Y., Li, M., and Lu, J. (2020b). Long-Term Persistence of IgG Antibodies in SARS-CoV Infected Healthcare Workers. medRxiv, 2020.2002.2012.20021386.

Hancock, K., Veguilla, V., Lu, X., Zhong, W., Butler, E.N., Sun, H., Liu, F., Dong, L., DeVos, J.R., Gargiullo, P.M.*, et al.* (2009). Cross-reactive antibody responses to the 2009 pandemic H1N1 influenza virus. N Engl J Med *361*, 1945-1952.

Havenar-Daughton, C., Reiss, S.M., Carnathan, D.G., Wu, J.E., Kendric, K., Torrents de la Pena, A., Kasturi, S.P., Dan, J.M., Bothwell, M., Sanders, R.W.*, et al.* (2016). Cytokine-Independent Detection of Antigen-Specific Germinal Center T Follicular Helper Cells in Immunized Nonhuman Primates Using a Live Cell Activation-Induced Marker Technique. J Immunol *197*, 994-1002.

Herati, R.S., Muselman, A., Vella, L., Bengsch, B., Parkhouse, K., Del Alcazar, D., Kotzin, J., Doyle, S.A., Tebas, P., Hensley, S.E.*, et al.* (2017). Successive annual influenza vaccination induces a recurrent oligoclonotypic memory response in circulating T follicular helper cells. Sci Immunol *2*, eaag2152.

Hinz, D., Seumois, G., Gholami, A.M., Greenbaum, J.A., Lane, J., White, B., Broide, D.H., Schulten, V., Sidney, J., Bakhru, P.*, et al.* (2016). Lack of allergy to timothy grass pollen is not a passive phenomenon but associated with the allergen-specific modulation of immune reactivity. Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology *46*, 705-719.

Huang, A.T., Garcia-Carreras, B., Hitchings, M.D.T., Yang, B., Katzelnick, L., Rattigan, S.M., Borgert, B., Moreno, C., Solomon, B.D., Rodriguez-Barraquer, I.*, et al.* (2020). A systematic review of antibody mediated immunity to coronaviruses: antibody kinetics, correlates of protection, and association of antibody responses with severity of disease. medRxiv, 2020.2004.2014.20065771.

Irigoyen, N., Firth, A.E., Jones, J.D., Chung, B.Y., Siddell, S.G., and Brierley, I. (2016). High-Resolution Analysis of Coronavirus Gene Expression by RNA Sequencing and Ribosome Profiling. PLoS Pathog *12*, e1005473.

Jurtz, V., Paul, S., Andreatta, M., Marcatili, P., Peters, B., and Nielsen, M. (2017). NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. J Immunol *199*, 3360-3368.

Kissler, S.M., Tedijanto, C., Goldstein, E., Grad, Y.H., and Lipsitch, M. (2020). Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science, eabb5793.

Li, C.K., Wu, H., Yan, H., Ma, S., Wang, L., Zhang, M., Tang, X., Temperton, N.J., Weiss, R.A., Brenchley, J.M.*, et al.* (2008). T cell responses to whole SARS coronavirus in humans. J Immunol *181*, 5490-5500.

Lindestam Arlehamn, C.S., McKinney, D.M., Carpenter, C., Paul, S., Rozot, V., Makgotlho, E., Gregg, Y., van Rooyen, M., Ernst, J.D., Hatherill, M.*, et al.* (2016). A Quantitative Analysis of Complexity of Human Pathogen-Specific CD4 T Cell Responses in Healthy M. tuberculosis Infected South Africans. PLoS Pathog *12*, e1005760.

Liu, L., Wei, Q., Lin, Q., Fang, J., Wang, H., Kwok, H., Tang, H., Nishiura, K., Peng, J., Tan, Z.*, et al.* (2019). Anti-spike IgG causes severe acute lung injury by skewing macrophage responses during acute SARS-CoV infection. JCI Insight *4*, 10.1172/jci.insight.123158.

Morou, A., Brunet-Ratnasingham, E., Dube, M., Charlebois, R., Mercier, E., Darko, S., Brassard, N., Nganou-Makamdop, K., Arumugam, S., Gendron-Lepage, G.*, et al.* (2019). Altered differentiation is central to HIV-specific CD4(+) T cell dysfunction in progressive disease. Nat Immunol *20*, 1059-1070.

Moutaftsi, M., Tscharke, D.C., Vaughan, K., Koelle, D.M., Stern, L., Calvo-Calle, M., Ennis, F., Terajima, M., Sutter, G., Crotty, S.*, et al.* (2010). Uncovering the interplay between CD8, CD4 and antibody responses to complex pathogens. Future Microbiol *5*, 221-239.

O'Sullivan, D., Arrhenius, T., Sidney, J., Del Guercio, M.F., Albertson, M., Wall, M., Oseroff, C., Southwood, S., Colon, S.M., Gaeta, F.C.*, et al.* (1991). On the interaction of promiscuous antigenic peptides with different DR alleles. Identification of common structural motifs. J Immunol *147*, 2663-2669.

Okba, N.M.A., Raj, V.S., Widjaja, I., GeurtsvanKessel, C.H., de Bruin, E., Chandler, F.D., Park, W.B., Kim, N.J., Farag, E., Al-Hajri, M.*, et al.* (2019). Sensitive and Specific Detection of Low-Level Antibody Responses in Mild Middle East Respiratory Syndrome Coronavirus Infections. Emerg Infect Dis *25*, 1868-1877.

Paul, S., Lindestam Arlehamn, C.S., Scriba, T.J., Dillon, M.B., Oseroff, C., Hinz, D., McKinney, D.M., Carrasco Pro, S., Sidney, J., Peters, B.*, et al.* (2015). Development and validation of a broad scheme for prediction of HLA class II restricted T cell epitopes. J Immunol Methods *422*, 28-34.

Paul, S., Sidney, J., Sette, A., and Peters, B. (2016). TepiTool: A Pipeline for Computational Prediction of T Cell Epitope Candidates. Current protocols in immunology *114*, 18 19 11-18 19 24.

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.

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.

Sallusto, F., Lanzavecchia, A., Araki, K., and Ahmed, R. (2010). From vaccines to memory and back. Immunity *33*, 451-463.

Sette, A., Moutaftsi, M., Moyron-Quiroz, J., McCausland, M.M., Davies, D.H., Johnston, R.J., Peters, B., Rafii-El-Idrissi Benhnia, M., Hoffmann, J., Su, H.P*., et al.* (2008). Selective CD4+ T cell help for antibody responses to a large viral pathogen: deterministic linkage of specificities. Immunity *28*, 847-858.

Severance, E.G., Bossis, I., Dickerson, F.B., Stallings, C.R., Origoni, A.E., Sullens, A., Yolken, R.H., and Viscidi, R.P. (2008). Development of a nucleocapsid-based human coronavirus immunoassay and estimates of individuals exposed to coronavirus in a U.S. metropolitan population. Clin Vaccine Immunol *15*, 1805-1810.

Sidney, J., Peters, B., Frahm, N., Brander, C., and Sette, A. (2008). HLA class I supertypes: a revised and updated classification. BMC Immunol *9*, 1.

Sidney, J., Steen, A., Moore, C., Ngo, S., Chung, J., Peters, B., and Sette, A. (2010a). Divergent motifs but overlapping binding repertoires of six HLA-DQ molecules frequently expressed in the worldwide human population. J Immunol *185*, 4189-4198.

Sidney, J., Steen, A., Moore, C., Ngo, S., Chung, J., Peters, B., and Sette, A. (2010b). Five HLA-DP molecules frequently expressed in the worldwide human population share a common HLA supertypic binding specificity. J Immunol *184*, 2492-2503.

Snijder, E.J., Bredenbeek, P.J., Dobbe, J.C., Thiel, V., Ziebuhr, J., Poon, L.L., Guan, Y., Rozanov, M., Spaan, W.J., and Gorbalenya, A.E. (2003). Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J Mol Biol *331*, 991-1004.

Southwood, S., Sidney, J., Kondo, A., del Guercio, M., Appella, E., Hoffman, S., Kubo, R., Chesnut, R., Grey, H., and Sette, A. (1998). Several common HLA-DR types share largely overlapping peptide binding repertoires. J Immunol *160*, 3363 - 3373.

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.

Takano, T., Kawakami, C., Yamada, S., Satoh, R., and Hohdatsu, T. (2008). Antibody-dependent enhancement occurs upon re-infection with the identical serotype virus in feline infectious peritonitis virus infection. J Vet Med Sci *70*, 1315-1321.

Thanh Le, T., Andreadakis, Z., Kumar, A., Gomez Roman, R., Tollefsen, S., Saville, M., and Mayhew, S. (2020). The COVID-19 vaccine development landscape. Nat Rev Drug Discov, doi: 10.1038/d41573-41020-00073-41575. [Epub ahead of print].

Tian, Y., Grifoni, A., Sette, A., and Weiskopf, D. (2019). Human T Cell Response to Dengue Virus Infection. Front Immunol *10*, 2125.

Vennema, H., de Groot, R.J., Harbour, D.A., Dalderup, M., Gruffydd-Jones, T., Horzinek, M.C., and Spaan, W.J. (1990). Early death after feline infectious peritonitis virus challenge due to recombinant vaccinia virus immunization. J Virol *64*, 1407-1409.

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.

Weiskopf, D., Angelo, M.A., de Azeredo, E.L., Sidney, J., Greenbaum, J.A., Fernando, A.N., Broadwater, A., Kolla, R.V., De Silva, A.D., de Silva, A.M*., et al.* (2013). Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. Proc Natl Acad Sci U S A *110*, E2046-2053.

Weiskopf, D., Cerpas, C., Angelo, M.A., Bangs, D.J., Sidney, J., Paul, S., Peters, B., Sanches, F.P., Silvera, C.G., Costa, P.R.*, et al.* (2015). Human CD8+ T-Cell Responses Against the 4 Dengue Virus Serotypes Are Associated With Distinct Patterns of Protein Targets. J Infect Dis *212*, 1743-1751.

Wilkinson, T.M., Li, C.K.F., Chui, C.S.C., Huang, A.K.Y., 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. Nature Medicine *18*, 274-280.

Wong, C.K., Lam, C.W., Wu, A.K., Ip, W.K., Lee, N.L., Chan, I.H., Lit, L.C., Hui, D.S., Chan, M.H., Chung, S.S.*, et al.* (2004). Plasma inflammatory cytokines and chemokines in severe acute respiratory syndrome. Clin Exp Immunol *136*, 95-103.

Wu, Z., and McGoogan, J.M. (2020). Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. JAMA *323*, 1239-1242.

Xie, X., Muruato, A., Lokugamage, K.G., Narayanan, K., Zhang, X., Zou, J., Liu, J., Schindewolf, C., Bopp, N.E., Aguilar, P.V.*, et al.* (2020). An Infectious cDNA Clone of SARS-CoV-2. Cell Host Microbe, https://doi.org/10.1016/j.chom.2020.1004.1004 in press.

Zhao, J., Alshukairi, A.N., Baharoon, S.A., Ahmed, W.A., Bokhari, A.A., Nehdi, A.M., Layqah, L.A., Alghamdi, M.G., Al Gethamy, M.M., Dada, A.M.*, et al.* (2017). Recovery from the Middle East respiratory syndrome is associated with antibody and T-cell responses. Sci Immunol *2*, eaan5393.

Zhao, J., Zhao, J., Legge, K., and Perlman, S. (2011). Age-related increases in PGD(2) expression impair respiratory DC migration, resulting in diminished T cell responses upon respiratory virus infection in mice. J Clin Invest *121*, 4921-4930.

Zhao, J., Zhao, J., Mangalam, A.K., Channappanavar, R., Fett, C., Meyerholz, D.K., Agnihothram, S., Baric, R.S., David, C.S., and Perlman, S. (2016). Airway Memory CD4(+) T Cells Mediate Protective Immunity against Emerging Respiratory Coronaviruses. Immunity *44*, 1379-1391.

Zhao, J., Zhao, J., and Perlman, S. (2010). T cell responses are required for protection from clinical disease and for virus clearance in severe acute respiratory syndrome coronavirus-infected mice. J Virol *84*, 9318-9325.

Zhao, J., Zhao, J., Van Rooijen, N., and Perlman, S. (2009). Evasion by stealth: inefficient immune activation underlies poor T cell response and severe disease in SARS-CoV-infected mice. PLoS Pathog *5*, e1000636.

Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X.*, et al.* (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet *395*, 1054-1062.

Journal Pre-proof

- Measuring immunity to SARS-CoV-2 is key for understanding COVID19 and vaccine development
- Epitope pools detect $CD4^+$ and $CD8^+$ T cells in 100 and 70% of convalescent COVID patients
- T cell responses are focused not only on spike but also on M, N and other ORFs
- T cell reactivity to SARS-CoV-2 epitopes is also detected in non-exposed individuals

An analysis of immune cell responses to SARS-CoV-2 from recovered patients identifies the regions of the virus that is targeted and also reveals cross-reactivity with other common circulating coronaviruses

Journal Pre-proof

Figure 1

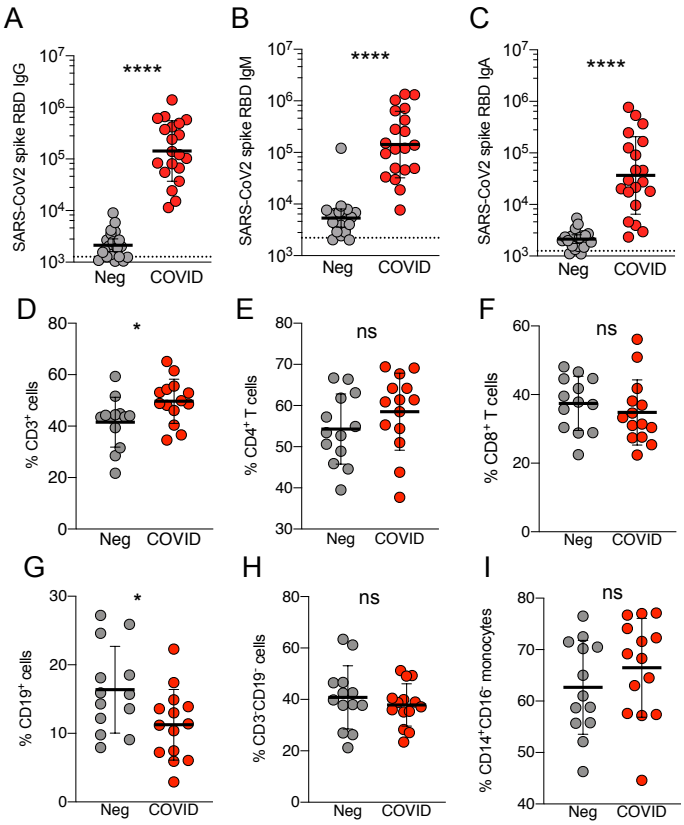

A
B
C
SARS-CoV2 spike RBD IgG
SARS-CoV2 spike RBD IgM
SARS-CoV2 spike RBD IgA
****
Neg
COVID
D
E
F
% CD3+ cells
% CD4+ T cells
% CD8+ T cells
*
ns
G
H
I
% CD19+ cells
% CD3-CD19- cells
% CD14+CD16- monocytes

Figure 2

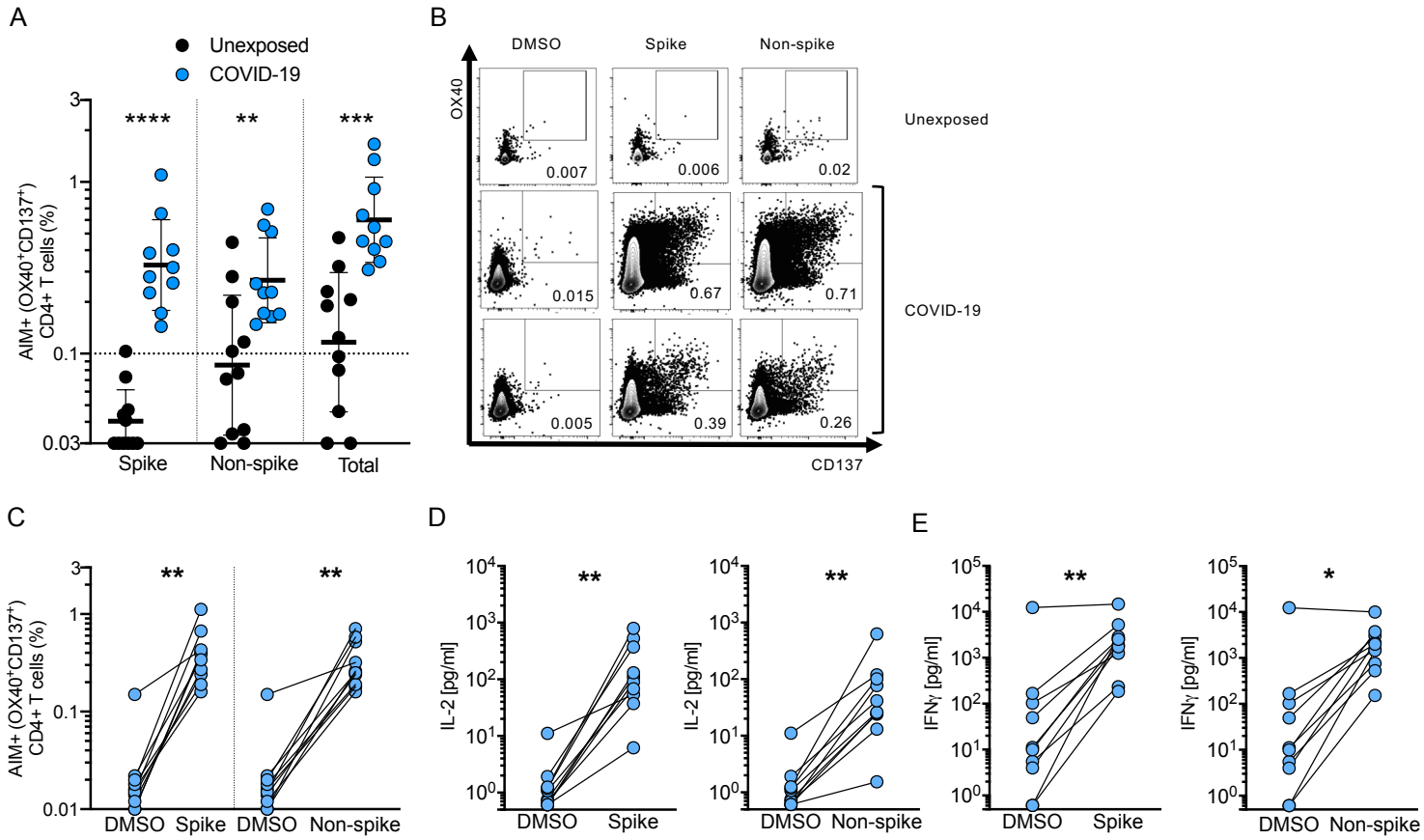

A
Unexposed
COVID-19
AIM+ (OX40+CD137+) CD4+ T cells (%)
****
**
***
Spike
Non-spike
Total
B
DMSO
Spike
Non-spike
OX40
CD137
Unexposed
COVID-19
0.007
0.006
0.02
0.015
0.67
0.71
0.005
0.39
0.26
C
AIM+ (OX40+CD137+) CD4+ T cells (%)
DMSO Spike DMSO Non-spike
D
IL-2 [pg/ml]
DMSO Spike
DMSO Non-spike
E
IFNγ [pg/ml]
DMSO Spike
DMSO Non-spike

Figure 3

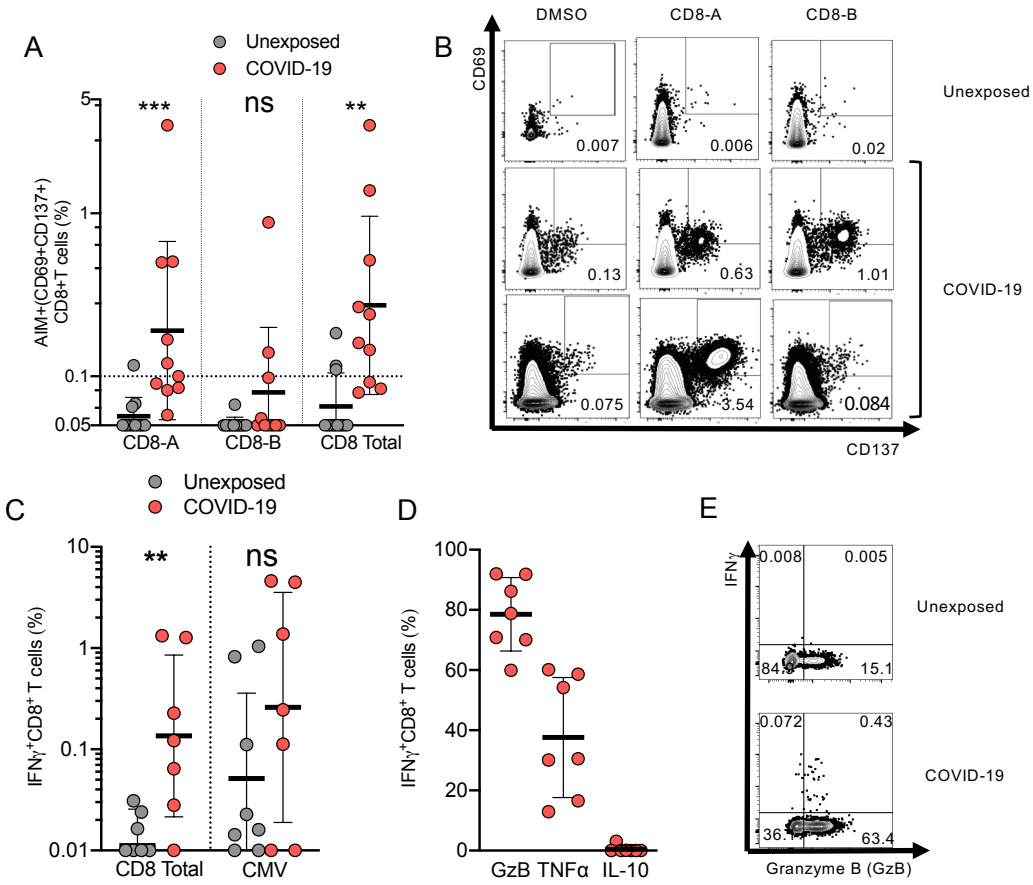

A
Unexposed
COVID-19
***
ns
**
AIM+(CD69+CD137+) CD8+T cells (%)
5
1
0.1
0.05
CD8-A
CD8-B
CD8 Total
B
DMSO
CD8-A
CD8-B
CD69
CD137
Unexposed
COVID-19
0.007
0.006
0.02
0.13
0.63
1.01
0.075
3.54
0.084
C
Unexposed
COVID-19
**
ns
IFNγ+CD8+ T cells (%)
10
1
0.1
0.01
CD8 Total
CMV
D
IFNγ+CD8+ T cells (%)
100
80
60
40
20
0
GzB
TNFα
IL-10
E
IFNγ
Granzyme B (GzB)
Unexposed
COVID-19
0.008
0.005
84.9
15.1
0.072
0.43
36.6
63.4

Figure 4

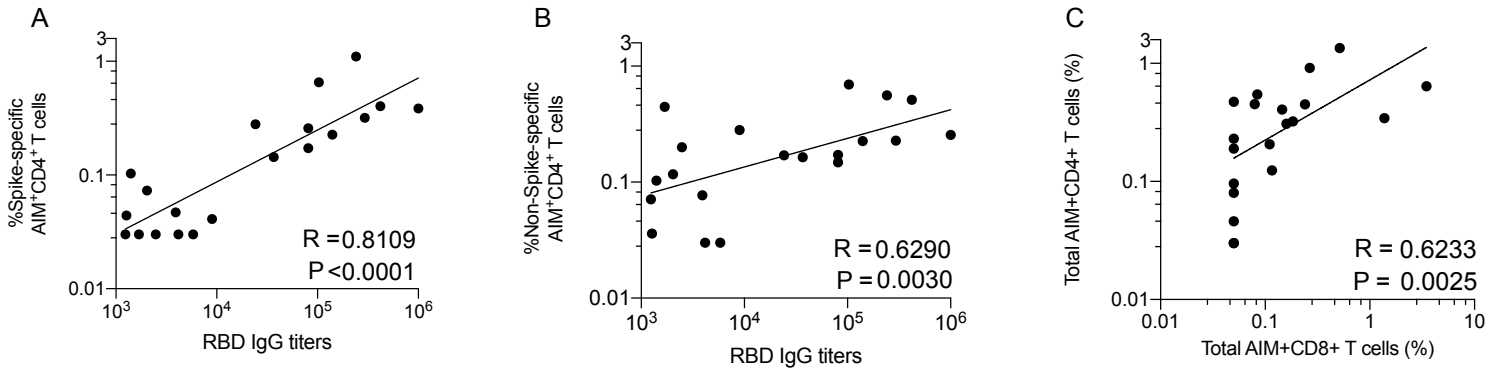

A
%Spike-specific AIM+CD4+ T cells
3
1
0.1
0.01
10^3
10^4
10^5
10^6
RBD IgG titers
R =0.8109
P <0.0001
B
%Non-Spike-specific AIM+CD4+ T cells
3
1
0.1
0.01
10^3
10^4
10^5
10^6
RBD IgG titers
R =0.6290
P =0.0030
C
Total AIM+CD4+ T cells (%)
3
1
0.1
0.01
0.01
0.1
1
10
Total AIM+CD8+ T cells (%)
R = 0.6233
P = 0.0025

Figure 5

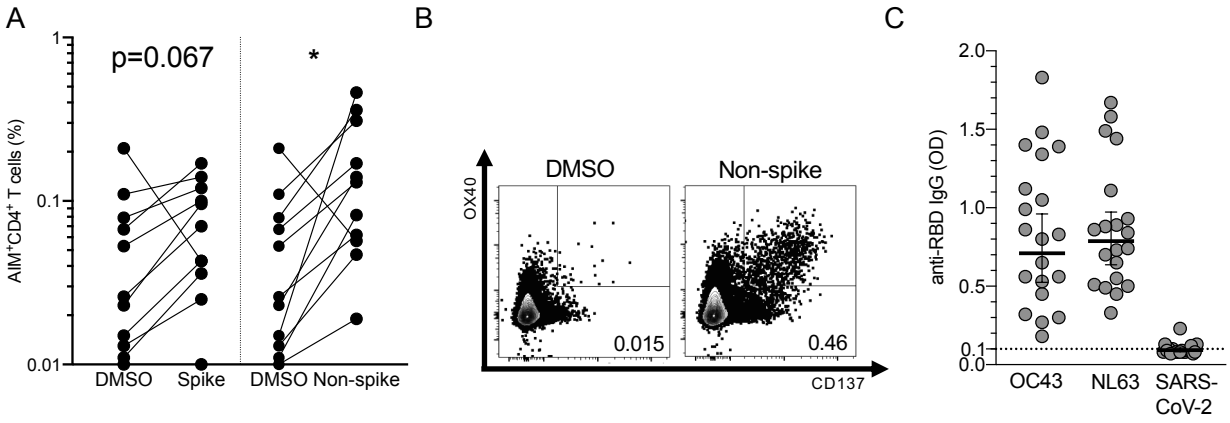
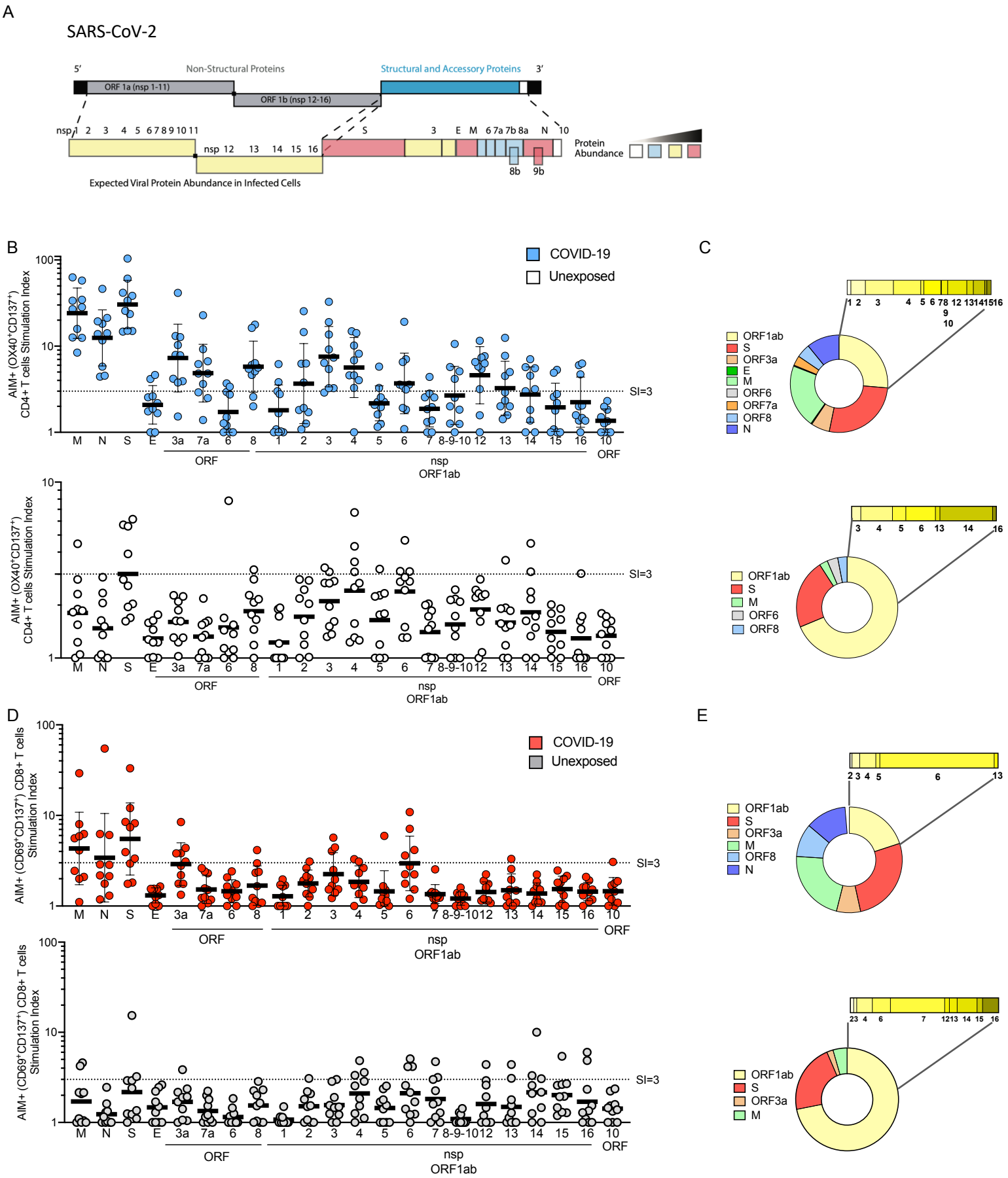

A
p=0.067
*
AIM+CD4+ T cells (%)
1
0.1
0.01
DMSO
Spike
DMSO
Non-spike
B
DMSO
Non-spike
OX40
0.015
0.46
CD137
C
anti-RBD IgG (OD)
2.0
1.5
1.0
0.5
0.1
0.0
OC43
NL63
SARS-CoV-2

Figure 6


A
SARS-CoV-2
5′
Non-Structural Proteins
Structural and Accessory Proteins
3′
ORF 1a (nsp 1-11)
ORF 1b (nsp 12-16)
nsp 1 2 3 4 5 6 7 8 9 10 11
nsp 12 13 14 15 16
S
3
E M 6 7a 7b 8a
N
10
8b
9b
Protein Abundance
Expected Viral Protein Abundance in Infected Cells
B
COVID-19
Unexposed
AIM+ (OX40+CD137+) CD4+ T cells Stimulation Index
SI=3
M N S E 3a 7a 6 8 1 2 3 4 5 6 7 8-9-10 12 13 14 15 16 10
ORF
nsp ORF1ab
C
ORF1ab
S
ORF3a
E
M
ORF6
ORF7a
ORF8
N
D
COVID-19
Unexposed
AIM+ (CD69+CD137+) CD8+ T cells Stimulation Index
SI=3
E
ORF1ab
S
ORF3a
M
ORF8
N