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

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Imbalance of Regulatory and Cytotoxic SARS-CoV-2-Reactive CD4+ T Cells in COVID-19

Authors
Benjamin J. Meckiff, Ciro Ramírez-Suásteegui, Vicente Fajardo, ..., Grégory Seumois, Christian H. Ottensmeier, Pandurangan Vijayanand

Correspondence
cho@soton.ac.uk (C.H.O.), vijay@lji.org (P.V.)

In Brief
Analyses of CD4+ T cells from 40 COVID-19 patients show that hospitalization is associated with increased cytotoxic follicular helper cells and cytotoxic T helper cells and a reduction in regulatory T cells.
Imbalance of Regulatory and Cytotoxic SARS-CoV-2-Reactive CD4+ T Cells in COVID-19

Benjamin J. Meckiff,1,6 Ciro Ramírez-Suástegui,1,6 Vicente Fajardo,1,6 Serena J. Chee,2,6 Anthony Kusnadi,1 Hayley Simon,1 Simon Eschweiler,1 Alba Grifoni,1 Emanuela Pelosi,3 Daniela Weiskopf,1 Alessandro Sette,1,5 Ferhat Ay,1 Grégoire Seumois,1 Christian H. Ottensmeier,1,2,4,7,* and Pandurangan Vijayanand1,2,5,7,8,*

1La Jolla Institute for Immunology, La Jolla, CA 92037, USA
2Faculty of Medicine, University of Southampton, Southampton SO16 6YD, UK
3Southampton Specialist Virology Center, University Hospitals NHS Foundation Trust, Southampton SO16 6YD, UK
4Institute of Translational Medicine, Department of Molecular & Clinical Cancer Medicine, University of Liverpool, Liverpool L69 7ZX, UK
5Department of Medicine, University of California San Diego, La Jolla, CA 92037, USA
6These authors contributed equally
7Senior Author
8Lead Contact
*Correspondence: cho@soton.ac.uk (C.H.O.), vijay@lji.org (P.V.)

https://doi.org/10.1016/j.cell.2020.10.001

SUMMARY

The contribution of CD4+ T cells to protective or pathogenic immune responses to SARS-CoV-2 infection remains unknown. Here, we present single-cell transcriptomic analysis of >100,000 viral antigen-reactive CD4+ T cells from 40 COVID-19 patients. In hospitalized patients compared to non-hospitalized patients, we found increased proportions of cytotoxic follicular helper cells and cytotoxic T helper (T_H) cells (CD4-CTLs) responding to SARS-CoV-2 and reduced proportion of SARS-CoV-2-reactive regulatory T cells (T_REG). Importantly, in hospitalized COVID-19 patients, a strong cytotoxic T_H response was observed early in the illness, which correlated negatively with antibody levels to SARS-CoV-2 spike protein. Polyfunctional T_H1 and T_H17 cell subsets were underrepresented in the repertoire of SARS-CoV-2-reactive CD4+ T cells compared to influenza-reactive CD4+ T cells. Together, our analyses provide insights into the gene expression patterns of SARS-CoV-2-reactive CD4+ T cells in distinct disease severities.

INTRODUCTION

Coronavirus disease 2019 (COVID-19) is causing substantial mortality, morbidity, and societal disruption (Tay et al., 2020; Vabret et al., 2020), and effective vaccines and therapeutics may take several months or years to become available. A substantial number of patients become life-threateningly ill, and the mechanisms responsible for causing severe acute respiratory distress syndrome (SARS) in COVID-19 are not well understood. Therefore, there is an urgent need to understand the key players driving protective and pathogenic immune responses in COVID-19 (Vabret et al., 2020). This knowledge may help devise better therapeutics and vaccines for tackling the current pandemic. CD4+ T cells are key orchestrators of anti-viral immune responses, either by enhancing the effector functions of other immune cell types like cytotoxic CD8+ T cells, NK cells, and B cells or through direct killing of infected cells (Sallusto, 2016). Recent studies in patients with COVID-19 have verified the presence of CD4+ T cells that are reactive to SARS-CoV-2 (Braun et al., 2020; Thieme et al., 2020; Grifoni et al., 2020). However, the nature and types of CD4+ T cell subsets that respond to SARS-CoV-2 and the roles these subsets play in driving protective or pathogenic immune responses remain elusive. Here, we have analyzed single-cell transcriptomes of virus-reactive CD4+ T cells to determine associations with severity of COVID-19 illness and to compare the molecular properties of SARS-CoV-2-reactive CD4+ T cells to other common respiratory virus-reactive CD4+ T cells from healthy control subjects.

RESULTS

CD4+ T Cell Responses in COVID-19 Illness

To capture CD4+ T cells responding to SARS-CoV-2 in patients with COVID-19 illness, we employed the antigen-reactive T cell enrichment (ARTE) assay (Bacher et al., 2013, 2016, 2019; Schmiedel et al., 2018) that relies on in vitro stimulation of peripheral blood mononuclear cells (PBMCs) for 6 h with overlapping peptide pools targeting the immunogenic domains of the spike and membrane proteins of SARS-CoV-2 (see STAR Methods; Thieme et al., 2020). Following in vitro stimulation, SARS-CoV-2-reactive CD4+ memory T cells were isolated based on the expression of cell surface markers (CD154 and CD69) that reflect recent engagement of the T cell receptor (TCR) by cognate major
histocompatibility complex (MHC)-peptide complexes (Figure S1A). In the context of acute COVID-19 illness, CD4+ T cells expressing activation markers have been reported in the blood (Braun et al., 2020; Thevarajan et al., 2020); such CD4+ T cells, presumably activated in vivo by endogenous SARS-CoV-2 viral antigens, were also captured during the ARTE assay, thereby enabling us to study a comprehensive array of CD4+ T cell subsets responding to SARS-CoV-2. We sorted >300,000 SARS-CoV-2-reactive CD4+ T cells from >1.3 billion PBMCs isolated from a total of 40 patients with COVID-19 illness (22 hospitalized patients with severe illness, 9 of whom required intensive care unit [ICU] treatment, and 18 non-hospitalized subjects with relatively milder disease; Figures 1A and 1B and Tables S1A and S1B). In addition to expressing CD154 and CD69, sorted SARS-CoV-2-reactive CD4+ T cells co-expressed other activation-related cell surface markers like CD38, CD137 (4-1BB), CD279 (PD-1), and HLA-DR (Figures 1C and S1B and Table S1C).

Recent evidence from studies in non-exposed individuals (blood sample obtained pre-COVID-19 pandemic) indicates pre-existing SARS-CoV-2-reactive CD4+ T cells, possibly indicative of human coronavirus (HCoV) cross-reactivity. Such cells are observed in up to 50% of the subjects studied (Braun et al., 2020; Grifoni et al., 2020; Le Bert et al., 2020). To capture such SARS-CoV-2-reactive CD4+ T cells, likely to be coronavirus (CoV)-reactive, we screened healthy non-exposed subjects and isolated CD4+ T cells responding to SARS-CoV-2 peptide pools from 4 subjects with highest responder frequency (Figures 1A and S1C and Table S1D). Next, for defining the CD4+ T cell subsets and their properties that distinguish SARS-CoV-2-reactive cells from other common respiratory virus-reactive CD4+ T cells, we isolated CD4+ T cells responding to peptide pools specific to influenza hemagglutinin protein (FLU-reactive cells, see STAR Methods) from 8 additional healthy subjects who provided blood samples before and/or after influenza vaccination (Figures 1A, S1D, and S1E and Tables S1D and S1E). CD4+ T cells responding to peptide pools specific to other common respiratory viruses like human parainfluenza (HPIV) and human metapneumovirus (HMPV) were also isolated from healthy subjects (Figure S1C and Tables S1D and S1F). In total, we interrogated the transcriptome and TCR sequence of >100,000 viral-reactive CD4+ T cells from 53 subjects (Figures 1A, S2A, and S2B and Tables S2A–S2E).

SARS-CoV-2-Reactive CD4+ T Cells Are Enriched for TH1 Cells and CD4+ CTLs

Analysis of the single-cell transcriptomes of all viral-reactive CD4+ T cells from all subjects revealed 13 CD4+ T cell subsets that clustered distinctly, reflecting their unique transcriptional profiles (Figures 2A–2D and Table S2F). Strikingly, a number of clusters were dominated by cells reactive to particular viruses (Figures 2B and S2C). For example, the vast majority of cells in clusters 1 and 10 were FLU-reactive (>65%), whereas cells in clusters 5, 6, 7, and 12 mainly consisted of SARS-CoV-2-reactive CD4+ T cells (>70%) from COVID-19 patients (Figures 2B and S2C). Conversely, cells in clusters 2, 3, 4, 8, and 9 were not preferentially enriched for reactivity to any given virus (Figures 2B and S2C). These findings suggest that distinct viral infections generate CD4+ T cell subsets with distinct transcriptional programs, although the timing of survey (acute illness versus past infection) will also contribute to their cellular states. Our data highlight substantial heterogeneity in the nature of CD4+ T cells generated in response to different viral infections on the one hand and shared features on the other.

The clusters enriched for FLU-reactive CD4+ T cells (clusters 1 and 10) displayed features suggestive of polyfunctional T helper (Tfh) cells which have been associated with protective anti-viral immune responses (Seder et al., 2008). Such features include the expression of transcripts encoding for the cytokines linked to polyfunctionality such as IFN-γ, IL-2, and TNFα, and several other cytokines and chemokines like IL-3, CSF2, IL-23A, and CCL20 (Figures 2D, 2E, S2E, and S2F). SARS-CoV-2-reactive CD4+ T cells were underrepresented in these clusters (cluster 1 and 10, <2%) when compared to FLU-reactive cells (>70%) or HMPV- and HPIV-reactive cells (~5%–20%) (Figure S2C). Furthermore, SARS-CoV-2-reactive CD4+ T cells in cluster 1 expressed significantly lower levels of IFNG and IL2 transcripts when compared to FLU-reactive cells (Table S2G). Together, these data suggested a failure to generate robust polyfunctional Tfh1 cells in SARS-CoV-2 infection. A similar pattern was also observed in SARS-CoV-2-reactive CD4+ T cells from healthy non-exposed subjects (Figures 2B and S2C) but not for HPIV- or HMPV-reactive CD4+ T cells, suggesting the defect in generating polyfunctional Tfh1 cells may be a common feature for coronaviruses, although further studies specifically analyzing HCoV-reactive CD4+ T cells in healthy individuals will be required to verify this.

Other clusters that were relatively underrepresented for SARS-CoV-2-reactive CD4+ T cells included clusters 2 and 8, which were both enriched for Tfh17 signature genes, with cluster 2 highly enriched for cells expressing IL17A and IL17F transcripts, thus representing bona fide Tfh17 cells (Figures 2B–2F and S2C–S2E and Table S2F). Tfh17 cells have been associated with protective immune responses in certain models of viral infections (Acharya et al., 2016; Wang et al., 2011); however, in other contexts they have been shown to promote viral disease pathogenesis (Acharya et al., 2016; Ma et al., 2019). Therefore, the functional relevance of an impaired Tfh17 response in COVID-19 is not clear and requires further investigation.

Clusters that were evenly distributed across all viral-specific CD4+ T cells include clusters 3 and 4. Cluster 3 displayed a transcriptional profile consistent with enrichment of interferon (IFN)-response genes (IFIT3, IFI44L, ISG15, MX2, OAS1), and cluster 4 was enriched for CCR7, IL7R, and TCF7 transcripts, likely representing central memory CD4+ T cell subset (Figures 2B–2F and S2C–S2E and Table S2F). Cluster 12, which expressed high levels of transcripts linked to cell cycle genes MKI67 and CDK1, also contained a large proportion of SARS-CoV-2-reactive CD4+ T cells (Figures 2B–2D), indicative of actively proliferating cells responsive to SARS-CoV-2 antigens. Cluster 6, also dominated by SARS-CoV-2-reactive CD4+ T cells, was characterized by high levels of PRF1, GZMB, GZMH, GNLY, and NKG7 transcripts, which encode for molecules linked to cytotoxicity (Patil et al., 2018). Gene set enrichment analysis (GSEA) showed significant positive enrichment of signature genes for cytotoxicity in clusters...
Figure 1. CD4+ T Cell Responses in COVID-19 Illness
(A) Study overview.
(B) Representative FACS plots showing surface staining of CD154 (CD40L) and CD69 in memory CD4+ T cells stimulated for 6 h with SARS-CoV-2 peptide pools, post-enrichment (CD154-based), in 22 hospitalized and 18 non-hospitalized COVID-19 patients (left), and summary of numbers of cells sorted (right); data are mean ± SEM.
(C) Representative FACS plots (left) showing surface expression of CD137 (4-1BB) and HLA-DR in memory CD4+ T cells ex vivo (without in vitro stimulation) and in CD154+ CD69+ memory CD4+ T cells following stimulation, post-enrichment (CD154-based). (Right) Percentage of CD154+ CD69+ memory CD4+ T cells expressing CD137 (4-1BB) or HLA-DR in 17 hospitalized and 18 non-hospitalized COVID-19 patients; data are mean ± SEM. See also Figure S1 and Table S1.
Figure 2. SARS-CoV-2-Reactive CD4+ T Cells Are Enriched for T<sub>FH</sub> Cells and CD4-CTLs

(A) Single-cell transcriptomes of sorted CD154<sup>+</sup> CD69<sup>+</sup> memory CD4<sup>+</sup> T cells following 6 h stimulation with virus-specific peptide megapools are displayed by uniform manifold approximation and projection (UMAP). Seurat-based clustering of 102,230 cells colored based on cluster type.

(B) UMAPs showing memory CD4<sup>+</sup> T cells for individual virus-specific megapool stimulation conditions (left), and normalized proportions of each virus-reactive cells per cluster is shown (right).

(legend continued on next page)
Clusters 0, 5, and 7, which were colocalized in the uniform manifold approximation and projection (UMAP) plot, were dominated by SARS-CoV-2-reactive CD4+ T cells (Figures 2A and 2B). Cells in these clusters were uniformly enriched for transcripts encoding for cytokines, surface markers, and transcriptional coactivators associated with T follicular helper (TFH) cell function (CXCL13, IL21, CD200, BTLA, and POU2AF1) (Locci et al., 2013) (Figures 2B–2F and S2C–S2E and Table S2F). Independent GSEA showed significant positive enrichment of TFH signature genes in these clusters, confirming that cells in these clusters represent circulating TFH cells (Figure S2G and Table S2H). Bona fide TFH cells reside in the germinal center; however, TFH cells have been described in the blood where increased numbers have been reported during viral infections and following vaccinations (Bentebibel et al., 2013; Koutsakos et al., 2018; Smits et al., 2020). Thus, the increase in circulating SARS-CoV-2-reactive TFH subsets observed in patients with COVID-19 is consistent with published reports in acute infections. Overall, our single-cell transcriptomic analysis revealed substantial differences in the nature of CD4+ T cell responses to viral infections and highlight subsets that are specifically enriched or depleted in COVID-19 illness.

SARS-CoV-2-Reactive CD4+ T Cell Subsets Associated with Disease Severity

We next assessed if the proportions of SARS-CoV-2-reactive CD4+ T cells in any cluster were greater or lower in hospitalized COVID-19 patients when compared to non-hospitalized patients. Unsupervised clustering of patients, based on the proportions of SARS-CoV-2-reactive CD4+ T cells in different clusters, showed that patients with an increased proportion of TFH cells in cluster 0 clustered distinctly from those with increased proportions of TFH cells in cluster 5 or CD4-CTL cells (cluster 6) (Figure 3A). The total frequency of SARS-CoV-2-reactive CD4+ T cells with a TFH profile (cluster 0, 5, and 7) was not significantly different between hospitalized and non-hospitalized COVID-19 patients (Figure 3B). However, the relative proportion of TFH cells in cluster 5 was significantly greater in hospitalized patients (severe disease) compared to non-hospitalized patients (mild disease), and the inverse was observed for the proportion of TFH cells in cluster 0 (Figures 3C and S3A and Table S2B). This pattern was maintained irrespective of whether the patients’ samples were analyzed early (<3 weeks from symptom onset) or later (>3 weeks) in the course of illness (Figure S3B). Notably, the proportion of TFH cells in cluster 7 was not significantly different between hospitalized and non-hospitalized COVID-19 patients (Figure S3C).

To determine the transcriptional features that differentiated SARS-CoV-2-reactive TFH cells present in cluster 5 from those in cluster 0, we performed single-cell differential gene expression analysis (Figure S3D and Table S3A). Transcripts encoding for transcription factors zinc finger BED-type-containing 2 (ZBED2) and zinc finger and BTB domain-containing protein 32 (ZBTB32) were enriched in TFH cells in cluster 5 and were also expressed at significantly higher levels in hospitalized COVID-19 patients (Figures 3D and S3D and Tables S3A and S3B). ZBTB32, also known as PLZP, belongs to a broad-complex, tramtrack and bric-à-brac zinc finger (BTB-ZF) family of transcriptional repressors like PLZF, B-cell lymphoma 6 (BCL6), and T-helper-inducing POZ-Kruppel-like factor (ThPOK) and has been shown to play a role in impairing anti-viral immune responses by negatively regulating T cell proliferation, cytokine production, and development of long-term memory cells (Piazza et al., 2004; Shin et al., 2017). ZBED2, a novel zinc finger transcription factor without a mouse ortholog, has been linked to T cell dysfunction in the context of anti-tumor immune response (Li et al., 2019) and more recently shown to repress expression of IFN target genes (Somervile et al., 2020). In support of potential dysfunctional properties of the cells in the TFH cluster 5, we found increased expression of several transcripts encoding for molecules linked to inhibitory function, like TIGIT, LAG3, TIM3, and PD1 (Thommen and Schumacher, 2018), and to negative regulation of T cell activation and proliferation, like DUSP4 and CD70 (Huang et al., 2012; O’Neill et al., 2017) (Figures 3D and S3D and Table S3A).

Most strikingly, TFH cells in cluster 5 expressed high levels of cytotoxicity-associated transcripts (PRF1, GZMB) (Figures S3E, S3D, and S3E), reminiscent of the recently described cytotoxic TFH cells, which were shown to directly kill B cells and associated with the pathogenesis of recurrent tonsillitis in children (Dan et al., 2019). Of relevance, recent studies reported a striking loss of germinal center B cells in the thoracic lymph nodes and spleen of patients who died of SARS-CoV-2 infection (Kaneko et al., 2020), as well as slightly lower SARS-CoV-2 spike protein (S)-specific immunoglobulin M (IgM) antibodies in deceased COVID-19 patients (Atyeo et al., 2020). On the basis of these findings, we hypothesized that the cytotoxic TFH cells (cluster 5) observed in hospitalized COVID-19 patients may impair humoral (B cell) immune responses to SARS-CoV-2. To test this association, we assessed the correlation between the proportions of SARS-CoV-2-reactive TFH cell subsets and immunoglobulin G (IgG) antibody titers against the SARS-CoV-2 S1/S2 (S1 and S2 subunits), which was higher in hospitalized patients (Figures 3F, 3G, and S3G). Although the total frequency of SARS-CoV-2-reactive TFH cells (clusters 0, 5, and 7) showed a positive correlation with antibody levels in hospitalized COVID-19 patients, but...
not in non-hospitalized COVID-19 patients (Figure 3F), the relative proportions of cytotoxic $T_{FH}$ cells ($T_{FH}$ cells in cluster 5) showed a strong negative correlation with anti-S1/S2 antibody levels in hospitalized COVID-19 patients (Figure 3G and Table S3C). Conversely, the proportions of $T_{FH}$ cells in cluster 0 (non-cytotoxic) were positively correlated with antibody concentrations in hospitalized COVID-19 patients (Figure S3H). We noted that the magnitude of cytotoxic $T_{FH}$ response (cluster 5) also showed a significant negative correlation with the time interval between onset of illness and sample collection, suggesting that their association with antibody levels could be confounded by the timing of analysis of patients’ samples (Figure 3G and Table S3C). Furthermore, we did not observe this association between cytotoxic $T_{FH}$ cells and anti-S1/S2 antibody levels in non-hospitalized patients, which suggested that other mechanisms such as lower viral titers may explain the low levels of anti-S1/S2 antibodies in non-hospitalized patients. To further assess effects on B cell function, we analyzed B cells specific to SARS-CoV-2 spike protein (S1 and S2 subunits) from nine patients with varying proportion of cytotoxic $T_{FH}$ cells. Notably, in the hospitalized patients with high proportions of cytotoxic $T_{FH}$ cells (patients 08, 09, and 16), we observed a much smaller number of S1/S2-specific B cells compared to those with lower proportions of these cytotoxic $T_{FH}$ cells (Figure S3I). Future longitudinal studies that examine the kinetics of T and B cell responses to SARS-CoV-2 are likely to provide more definitive and time-resolved associations between cytotoxic $T_{FH}$ cell and antibody responses.

Next, to characterize upstream regulators that may induce the differentiation and maintenance of the cytotoxic $T_{FH}$ cells, we performed Ingenuity Pathway analysis (IPA) of the transcripts increased in SARS-CoV-2-reactive $T_{FH}$ cells in cluster 5 (cytotoxic) when compared to those in cluster 0 (Tables S3D and S3E). Surprisingly, we found that type 1 and 2 IFNs emerged as the top upstream activators of genes enriched in the cytotoxic $T_{FH}$ cluster (Figure S3J and Tables S3D and S3E). GSEA confirmed that IFN response signatures were also significantly enriched in the cytotoxic $T_{FH}$ cluster (cluster 5) (Figure S3K). Single-cell trajectory analysis showed that a large fraction of cytotoxic $T_{FH}$ cells (cluster 5) followed a separate trajectory from cluster 0 cells (Figure 3H), and cells in this track were enriched for the IFN response signature. In addition, we found that transcripts encoding perforin (PRF1) and the transcription factor ZBED2 were also enriched in the cytotoxic $T_{FH}$ cell trajectory, which suggested the hypothesis that ZBED2 may contribute to the differentiation or function of cytotoxic $T_{FH}$ cells, although further studies will be needed to verify this.

**Massive Clonal Expansion of CD4-CTLs**

While T cells with cytotoxic function are thought to predominantly consist of conventional MHC class I-restricted CD8$^+$ T cells, MHC class II-restricted CD4$^+$ T cells with cytotoxic potential (CD4-CTLs) have also been reported in several viral infections in humans and are associated with better clinical outcomes (Cheroutre and Husain, 2013; Juno et al., 2017; Meckiff et al., 2019; Weiskopf et al., 2015a). Paradoxically, in SARS-CoV-2 infection, we find that cells in the CD4-CTL clusters (Figure 4A; cluster 6 and 9) were present at higher frequencies in some hospitalized COVID-19 patients compared to non-hospitalized patients, potentially contributing to disease severity, although we observed substantial heterogeneity in responses among patients (Figures 4B and 3A and Table S2B).

Interrogation of the transcripts enriched in the CD4-CTL subsets pointed to several interesting molecules and transcription factors that are likely to play an important role in their maintenance and effector function. These include molecules like CD72 and GPR18 that are known to enhance T cell proliferation and maintenance of mucosal T cell subsets, respectively (Jiang et al., 2017; Wang et al., 2014) (Figures 4C and S4A). Additional examples include transcription factors HOPX and ZEB2 (Figures 4C and S4A) that have been shown to positively regulate effector differentiation, function, persistence, and survival of T cells (Albrecht et al., 2010; Omlusik et al., 2015). Besides cytotoxicity-associated transcripts, the CD4-CTL subsets (clusters 6 and 9) and cytotoxic $T_{FH}$ cells (cluster 5) were highly enriched for transcripts encoding for a number of chemokines like CCL3 (also known as macrophage inflammatory protein [MIP]-1x), CCL4 (MIP-1β), and CCL5 (Figures 4D and S2F); these chemokines

Figure 3. SARS-CoV-2-Reactice CD4$^+$ T Cell Subsets Associated with Disease Severity

(A) Unsupervised clustering of COVID-19 patients based on the proportions of SARS-CoV-2-reactive CD4$^+$ T cells in different clusters following 6 h peptide stimulation. Clusters with fewer than 5% of the total dataset are not depicted. Gender and hospitalization status per patient are indicated by different color schemes above the heatmap.

(B) Percentage of $T_{FH}$ cells (clusters 0, 5, and 7) in the total SARS-CoV-2-reactive CD4$^+$ T cell pool for non-hospitalized and hospitalized COVID-19 patients; dots indicate data from a single subject. Data are mean ± SEM; significance for comparisons was computed using Mann-Whitney U test; ns, non-significant P value.

(C) Proportion of clusters 5 and 0 cells in SARS-CoV-2-reactive $T_{FH}$ cells (clusters 0, 5, and 7) in non-hospitalized and hospitalized COVID-19 patients. Data are mean ± SEM; significance for comparisons was computed using Mann-Whitney U test; ***p < 0.0001.

(D) Violin plots showing normalized expression level (log2(CPM+1)) of $T_{FH}$ cells (clusters 0, 5, and 7) in non-hospitalized and hospitalized COVID-19 patients. Numbers indicate percentage of cells expressing indicated transcript. Plots below show average expression and percent of cells expressing selected transcripts in indicated clusters.

(E) Scatterplot displaying normalized co-expression level (log2(CPM+1)) between PRF1 and GZMB transcripts in SARS-CoV-2-reactive cells in clusters 5 (left) and 0 (right). Numbers indicate percentage of cells in each quadrant.

(F) Correlation between percentage of SARS-CoV-2-reactive CD4$^+$ $T_{FH}$ cells and S1/S2 antibody titers in 15 non-hospitalized (left) and 20 hospitalized (right) COVID-19 patients. Correlation coefficient $r$ and the related P value were computed using Spearman correlation; *p < 0.05.

(G) Correlation between percentage of SARS-CoV-2-reactive CD4$^+$ $T_{FH}$ cells form cluster 5 as a frequency of total CD4$^+$ $T_{FH}$ and S1/S2 antibody titers (left two plots) and interval between symptom onset and blood draw (right two plots) in 15 non-hospitalized and 20 hospitalized (left) COVID-19 patients. Correlation coefficient $r$ and the related P value were computed using Spearman correlation; ***p < 0.001; ***p < 0.0001; ns, non-significant P value.

(H) Single-cell trajectory analysis of cells in cluster 5 and 0 showing pseudotime, expression of indicated genes, and IFN response signature score.

See also Figure S3 and Table S3.
Figure 4. SARS-CoV-2-Reactive CD4-CTLs and Single-Cell TCR Sequence Analysis
(A) UMAPs showing Seurat-normalized expression level of PRF1, GZMB, GNLY, and NKG7 transcripts in each virus-reactive cell.
(B) Percentage of CD4-CTLs (clusters 6 and 9) in the total SARS-CoV-2-reactive CD4+ T cell pool for non-hospitalized and hospitalized COVID-19 patients; dots indicate data from a single subject. Data are mean ± SEM; significance for comparisons was computed using Mann-Whitney U test; ns, non-significant P value.
(C) Violin plots showing normalized expression level (log2(CPM+1)) of transcription factors HOPX and ZEB2 and effector molecules CD72, GPR18, and SLAMF7 transcripts in virus-reactive cells from designated clusters (6 and 9) compared to an aggregation of remaining cells (Rest).
(D) UMAPs showing Seurat-normalized expression of CCL3, CCL4, CCL5, XCL1, and XCL2 transcripts in each virus-reactive cell.
(E) UMAP showing TCR clone size (log2, color scale) of SARS-CoV-2-reactive cells from COVID-19 patients (6 h stimulation condition).
(F) Histogram bar graph (top) displaying single-cell TCR sequence analysis of SARS-CoV-2-reactive cells. Each bar shows the number of TCRs shared between cells from individual clusters (rows, connected by lines). Connected lines (bottom) indicates what clusters are sharing TCRs. Clusters 6 (green), 9 (blue), and 11 (pink), i.e., CD4-CTLs, are highlighted.

(legend continued on next page)
play an important role in the recruitment of myeloid cells (neutrophils, monocytes, macrophages), NK cells, and T cells expressing C-C type chemokine receptors (CCR1, CCR3, and CCR5) (Hughes and Nibbs, 2018). The CD4-CTL subset in cluster 6 and cytotoxic T
\( T_{\text{FH}} \) cells (cluster 5) also expressed high levels of transcripts encoding for chemokines XCL1 and XCL2 (Figures 4D, S4A, and S4C) that specifically recruit XCR1-expressing conventional type 1 dendritic cells (cDC1) to sites of immune responses where they play a key role in promoting the CD8+ T cell responses by antigen cross-presentation (Lei and Takahama, 2012). Overall, the transcriptomic features of SARS-CoV-2-reactive CD4-CTLs and cytotoxic T
\( T_{\text{FH}} \) cells suggest that they are likely to play an important role in orchestrating immune responses by recruiting innate immune cells to enhance CD8+ T cell responses, while also directly mediating cytotoxic death of MHC class II-expressing virally infected cells.

The recovery of paired TCR sequences from individual single cells enabled us to link transcriptome data to clonotype information and evaluate the clonal relationship between different CD4+ T cell subsets as well as determine the nature of subsets that display greatest clonal expansion (Tables S4A and S4B). In SARS-CoV-2 infection, hospitalized patients were characterized by large clonal expansion of the virus-reactive CD4+ T cells (mean of 55.8%); in contrast, in non-hospitalized patients, recovered TCRs were less clonally expanded (mean of 38.0%) (Figure S4D). Among SARS-CoV-2-reactive CD4+ T cells, CD4-CTL subsets (clusters 6 and 9) displayed the greatest clonal expansion (> 75% of cells were clonally expanded), indicating preferential expansion and persistence of CD4-CTLs in some patients with COVID-19 illness (Figure 4E and Tables S4A and S4B). Analysis of clonally expanded SARS-CoV-2-reactive CD4+ T cells from COVID-19 patients showed extensive sharing of TCRs between cells in clusters 6 and 9, as well as those in cluster 11 (Figure 4F), which, notably, was enriched for the expression of XCL1 and XCL2 transcripts and also for cytotoxicity-associated transcripts, albeit at lower levels compared to the established CD4-CTL clusters (Figures 4D and S4C and Table S2F). Thus, cells in cluster 11 are likely to be an intermediate transition population, a hypothesis supported by single-cell trajectory analysis that showed potential temporal connection and transcriptional similarity between these subsets (Figure 4G).

Initial reports in patients with acute COVID-19 have suggested that circulating T cells that express activation markers such as CD38, HLA-DR, and PD-1 ex vivo (without in vitro peptide stimulation) are enriched for SARS-CoV-2-reactive T cells (Braun et al., 2020; Thevarajan et al., 2020). However, a recent study indicated that bystander T cells reactive to other antigens (e.g., CMV and EBV) can also express these activation markers, likely to be non-specifically activated without TCR engagement (Se-kine et al., 2020). Thus, studies in active SARS-CoV-2 infection that just examine T cells expressing activation markers are not likely to reveal the full potential effector function of SARS-CoV-2-reactive T cells. To determine the specificity and molecular features of such T cells expressing activation markers ex vivo, we isolated CD38\textsuperscript{high} HLA-DR\textsuperscript{high} PD-1\textsuperscript{+} memory CD4+ T cells from hospitalized COVID-19 patients and performed single-cell transcriptome and TCR sequence analysis of >20,000 cells. CD4+ T cells expressing activation markers ex vivo clustered distinctly from the SARS-CoV-2-reactive CD4+ T cells, which were isolated following in vitro stimulation with SARS-CoV-2 peptides for 6 h (Figure S4E and Tables S2C–S2E, S4C, and S4D). The CD4+ T cells expressing activation markers ex vivo displayed reduced activation and T
\( T_{\text{FH}} \) signature scores and had lower expression of transcripts encoding effector cytokines (IFN-\( \gamma \), IL-2, TNF\( \alpha \)), activation markers (OX40), and T
\( T_{\text{FH}} \) associated genes (CD200, POU2AF1) (Figures S4F and S4G). Furthermore, by comparison of single-cell TCR sequences, we found that 33.8% of SARS-CoV-2-reactive CD4+ T cells shared clonotypes with CD4+ T cells expressing activation markers ex vivo, and 12.2% of CD4+ T cells expressing activation markers ex vivo shared their TCRs with SARS-CoV-2-reactive CD4+ T cells (Figure S4H and Tables S4E and S4F). Our findings indicate that using surface activation markers as a strategy to enrich for SARS-CoV-2-reactive T cells without SARS-CoV-2 peptide stimulation (ARTE assay) may not capture the full spectrum of SARS-CoV-2-reactive T cells, like T
\( T_{\text{FH}} \) biology and their cytokine profiles, although the transcriptomic features of such in vitro activated cells may be affected by antigen-presenting cells present in the cultures.

SARS-CoV-2-Reactive T
\( \text{REG} \) Cells Are Reduced in Hospitalized COVID-19 Patients

In order to capture SARS-CoV-2-reactive CD4+ T cells that may not upregulate the activation markers (CD154 and CD69) after 6 h of in vitro stimulation with SARS-CoV-2 peptide pools, we stimulated PMBCs from the same cultures for a total of 24 h (see STAR Methods) and captured cells based on co-expression of activation markers CD137 (4-1BB) and CD69, a strategy that allowed us to additionally capture antigen-specific regulatory T cells (T
\( \text{REG} \)) (Bacher et al., 2016) (Figures 5A and 5S). Our analysis of a total of 38,519 single-cell CD4+ T cell transcriptomes revealed 6 distinct clusters (Figures 5A–5C and Tables S5A–SSC). The T
\( T_{\text{FH}} \) subset (cluster D) was detectable at relatively lower frequencies in the 24 h condition, though they represented the major CD4+ T cell subsets in the 6 h stimulation condition (Figures 2A and 5A). Consistent with delayed kinetics of activation of central memory T (T
\( \text{CM} \)) cells, we identified a higher proportion of CD4+ T cells expressing transcripts linked to central memory cells (CCR7, IL7R, and TCF7) (cluster C) (Figures 2A, 5A, and 5C).

The largest cluster (cluster A) was characterized by high expression of FOXP3 transcripts, which encodes for the T
\( \text{REG} \) master transcription factor forkhead box P3 (FOXP3) (Rudensky, 2011) (Figures 5A–5D and Table SSC). Independent GSEA analysis showed significant positive enrichment of T
\( \text{REG} \) signature genes in this cluster, suggesting that cells in this cluster represented SARS-CoV-2-reactive T
\( \text{REG} \) cells (Figure S5B and Table S2F). Notably, the proportion of cells in the T
\( \text{REG} \) cluster was significantly lower in hospitalized COVID-19 patients compared to non-hospitalized patients (6 h stimulation condition) are shown.

See also Figure S4 and Table S4.
Figure 5. Analysis of SARS-CoV-2-Reactive CD4+ T Cells from 24 h Stimulation Condition

(A) Single-cell transcriptomes of sorted CD137+ CD69+ memory CD4+ T cells following 24 h stimulation with SARS-CoV-2-specific peptide megapools are displayed by UMAP. Seurat-based clustering of 38,519 cells colored based on cluster type.

(B) Heatmap showing expression of the most significantly enriched transcripts in each cluster (see Table S5C). Seurat marker gene analysis—comparison of cluster of interest versus all other cells—shown are top 200 transcripts with adjusted P value < 0.05, log2 fold change > 0.25, and > 10% difference in the percentage of cells expressing differentially expressed transcript between two groups compared.

(legend continued on next page)
to non-hospitalized patients (Figures 5D, 5E, and 5C and Tables SSA and SSB), suggesting a potential defect in the generation of immunosuppressive SARS-CoV-2-reactive TREG cells in hospitalized patients. Consistent with our data from 6 h stimulation condition, we found that cells in the CD4-CTL clusters (clusters B and F) were present at higher frequencies in some hospitalized COVID-19 patients (Figures 5E, 5F, and SSC and Tables SSA and SSB). They also showed the greatest clonal expansion compared to other clusters (Figures SSD an SSE and Table S4B), suggesting potential importance of the CD4-CTL subset in driving immune responses to SARS-CoV-2 infection.

Correlation analysis of the proportion of CD4-CTLs and TREG in our 24 h dataset revealed a significant negative correlation, which indicated that patients with an impaired TREG response to SARS-CoV-2 mounted a stronger CD4-CTL response (Figure 5G and Table SSD). A recent study in a murine model showed that cytotoxic TFH responses are curtailed by a subset of TREG cells called follicular regulatory T (TFR) cells (Xie et al., 2019). To determine if such association is observed in our datasets, we first quantified TFR cells based on the expression of IL1R2 (Eschweiler et al., 2020) from cells in the TREG Cluster A (Figure 5H). Independent GSEA confirmed that IL1R2-expressing cells were significantly enriched for follicular and TFR signature genes (Figure S5F), which indicated they represent TFR cells. Over 40% of the cells in the TREG cluster expressed IL1R2; this indicates that a strong circulating TFR response is generated in SARS-CoV-2 infection. Importantly, the proportion of TFR cells was significantly lower in hospitalized COVID-19 patients (Figure 5H) and showed a modest negative correlation with the proportion of cytotoxic TFH cells (Figure 5I and Table SSE). On the basis of these findings and the known function of these TREG subsets, we hypothesize that the magnitude of TREG and TFR responses to SARS-CoV-2 are likely to modulate cytotoxic CD4+ T and B cell responses in COVID-19 illness, although further studies are required to confirm this hypothesis.

**DISCUSSION**

There is an urgent need to better understand the molecular determinants of protective and pathogenic immune responses in COVID-19. Given the importance of CD4+ T cells in anti-viral immunity, studying this adaptive immune cell population is likely to provide insights into the nature of host responses observed in patients with COVID-19. Current studies on antigen-specific CD4+ T cells are limited to flow-cytometry-based phenotyping of SARS-CoV-2-responding cells using limited sets of markers (Braun et al., 2020; Grifoni et al., 2020; Thieme et al., 2020), which thus fail to comprehensively capture the breadth of CD4+ T cells that respond to SARS-CoV-2. Unbiased approaches employing single-cell RNA-seq assays can provide these insights; however, to our knowledge, single-cell studies to date have only examined total CD4+ T cells in blood or bronchoalveolar lavage specimens from patients with COVID-19 illness (Vabret et al., 2020). Due to the rarity of SARS-CoV-2-specific cells in the total CD4+ T cell populations, signals from these cells are likely to be masked by the relative abundance of other non-antigen-specific CD4+ T cells. Furthermore, despite a profusion of single-cell transcriptomic studies, the analysis of virus-specific or any antigen-specific T cells, as such in humans, has lagged behind, partly due to the challenges imposed by methods to isolate antigen-specific T cells in sufficient numbers. Here, we have overcame these issues and performed single-cell transcriptomic study of >100,000 virus-reactive CD4+ T cells, focused on SARS-CoV-2-reactive cytotoxic TFH cells from 40 COVID-19 patients with varying disease severity, and compared their molecular profile to CD4+ T cells reactive to other common respiratory viruses. We find remarkable heterogeneity in the nature of CD4+ T cell subsets that are reactive to SARS-CoV-2 and other respiratory viruses and across individual patients and with differing severity of COVID-19. Polyfunctional T**H1 cells, which are abundant among FLU-reactive CD4+ T cells and considered to be protective (Seder et al., 2008), were present in lower frequencies among SARS-CoV-2-reactive CD4+ T cells. Lower frequencies of T**H1**7 cells were also observed among SARS-CoV-2-reactive CD4+ T cells. In contrast, we find increased proportions of SARS-CoV-2-reactive cytotoxic TFH cells in hospitalized COVID-19 patients. Cytotoxic TFH cells can kill B cells and dampen germinal center responses (Xie et al., 2019), and to our knowledge this is the first description of circulating cytotoxic TFH cells in humans. Importantly, the magnitude of the cytotoxic TFH response to SARS-CoV-2 was stronger early in the course of illness and negatively correlated with antibody levels to SARS-
CoV-2 S. Recent reports have found that patients with fatal COVID-19 infections have abrogated germinal center B cell responses (Kaneko et al., 2020) and very slightly reduced levels of S-specific IgM antibodies (Atyeo et al., 2020), the mechanistic basis of which is not known. Our findings of strong cytotoxic T<sub>FH</sub> responses early in the illness may provide the link to defects in B cells responses in some patients with severe and fatal COVID-19 illness.

Another striking observation is the abundance in SARS-CoV-2-reactive CD4<sup>+</sup> T cells of CD4-CTLs that express high levels of transcripts encoding for multiple chemokines (XCL1, XCL2, CCL3, CCL4, and CCL5), particularly from some hospitalized COVID-19 patients. This suggests that the CD4-CTL responses in COVID-19 illness may be linked to pathogenesis, although further studies in animal models and large-scale association studies in COVID-19 patients are required to verify or refute this hypothesis. Notably, some hospitalized COVID-19 patients showed impaired T<sub>REG</sub> response to SARS-CoV-2, and such patients mounted a strong CD4-CTL response, raising another interesting association that warrants testing in larger studies.

**Limitations and Future Directions**

The limitation of this study is the relatively small sample size considering the heterogeneity observed in the nature of CD4<sup>+</sup> T cell responses to SARS-CoV-2. Analysis of patients in the acute and convalescent phase of illness fails to discriminate effector and long-term memory CD4<sup>+</sup> T cell responses. Serial sampling of the same patients in the recovered phase is likely to provide insights into the nature and persistence of memory CD4<sup>+</sup> T cell responses to SARS-CoV-2.

Because the negative association between cytotoxic T<sub>FH</sub> cell responses and anti-spike antibody levels was not observed in non-hospitalized patients, the potential role of cytotoxic T<sub>FH</sub> cells in antibody responses cannot be generalized. Furthermore, the higher proportions of cytotoxic T<sub>FH</sub> cells in hospitalized patients may merely reflect higher viral titers and IFN production. Longitudinal studies are required to clarify the association between aberrant cytotoxic T<sub>FH</sub> responses and their impact on modulating the magnitude and duration of protective antibody responses to SARS-CoV-2. The role of CD4-CTLs in protective or pathogenic immune responses to SARS-CoV-2 needs to be clarified in pre-clinical models. Future studies in COVID-19 patients should also examine the relationships between the subsets of SARS-CoV-2-reactive CD4<sup>+</sup> T cells in the blood and those observed in the mucosal tissues where control of SARS-CoV-2 infection is critical.

**STAR METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and Code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - COVID-19 patients and samples
  - Healthy controls

**METHOD DETAILS**

- PBMC processing
- SARS-CoV-2 peptide pools
- SARS-CoV-2 antibody testing
- SARS-CoV-2 spike protein-specific B cell responses
- Epitope Megapool of peptide (MP) design
- Antigen-reactive T cell enrichment (ARTE) assay
- Cell isolation and single-cell RNA-seq assay (10x platform)
- Single-cell transcriptome analysis
- Single-cell differential gene expression analysis
- Gene Set Enrichment Analysis and Signature Module Scores
- Single-cell trajectory analysis
- T cell receptor (TCR) sequence analysis

**SUPPLEMENTAL INFORMATION**

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2020.10.001.

**ACKNOWLEDGMENTS**

We thank Luke Smith for patient recruitment and sample collection; Callum Dixon, Benjamin Johnson, Lydia Scarlett, and Silvia Austin for collection of clinical data; Céline Galloway, Oliver Wood, Katy McCarrn, and Lindsey Chudley for sample processing; and Sharon Gilchrist for the illustration. We thank the La Jolla Institute (LJI) Flow Cytometry Core for assisting with cell sorting and the LJI’s Clinical Studies Core for organizing sample collection. We thank Peter Friedmann and Anusha Preethi Ganesan for providing critical feedback on the manuscript. This work was funded by NIH grants U19AI142474, U19AI142742-05S1 (to P.V., A.S., and C.H.O.), U19AI118626 (to P.V., A.S., and G.S.), R01HL114093 (to P.V., A.S., and G.S.), R35-GM126938 (F.A), S10RR027366 (BD FACS Aria-II), and S10OD025052 (Illumina Novaseq6000); the William K. Bowes, Jr. Foundation (P.V.); and Whittaker Foundation (C.H.O.). Supported by the Wessex Clinical Research Network and National Institute for Health Research UK.

**AUTHOR CONTRIBUTIONS**

B.J.M., S.J.C., C.H.O., and P.V. conceived the work. B.J.M., S.J.C., C.H.O., and P.V. designed the study and wrote the manuscript. S.J.C. supervised patient identification, recruitment, sample collection, and processing. E.P. supervised the analysis of viral PCR and serology tests. A.G., D.W., and A.S. provided essential reagents for the isolation of viral-reactive CD4<sup>+</sup> T cells. B.J.M. and A.K. performed ARTE assay and FACS sorting. S.E. performed B cell analyses, and H.S. performed single-cell RNA-sequencing under the supervision of G.S., C.H.O., and P.V. performed bioinformatic analyses under the supervision of G.S., C.H.O., and P.V.

**DECLARATION OF INTERESTS**

The authors declare no competing interests.
Acharya, D., Wang, P., Paul, A.M., Dai, J., Gate, D., Lowery, J.E., Stokic, D.S., Leis, A.A., Flavell, R.A., Town, T., et al. (2016). Interleukin-17A Promotes CD8+ T Cell Cytotoxicity To Facilitate West Nile Virus Clearance. Journal of Virology 91, e01529–e16.

Albrecht, I., Niesner, U., Janke, M., Menning, A., Loddenkemper, C., Kühl, A.A., Lepenies, I., Lexcerpt, M.H., Westendorf, K., Hradilková, K., et al. (2010). Persistence of effector memory Th1 cells is regulated by Hopx. European Journal of Immunology 40, 2993–3006.

Atyeo, C., Fischinger, S., Zohar, T., Stein, M.D., Burke, J., Loos, C., McCulloch, D.J., Newman, K.L., Wolf, C., Yu, J., et al. (2020). Distinct Early Serological Signatures Track with Sars-Cov-2 Survival. Immunity 53, 524–532.e4.

Bacher, P., Schink, C., Teutschbein, J., Kriemeyer, O., Assenmacher, M., Brakhage, A.A., and Schefold, A. (2013). Antigen-reactive T cell enrichment for direct, high-resolution analysis of the human naive and memory T cell repertoire. J Immunol 190, 3967–3976.

Bacher, P., Heinrich, F., Stenvoo, U., Nienen, M., Vahldieck, M., Ivett, C., Vogt, K., Kollet, J., Babel, N., Sawitzki, B., et al. (2016). Regulatory T Cell Specificity Directs Tolerance Versus Allergy against Aeroantigens in Humans. Cell 167, 1067–1078.e16.

Bacher, P., Hohnstein, T., Beerbaum, E., Rocker, M., Blango, M.G., Kaufmann, S., Rohmel, J., Eschenhagen, P., Grehn, C., Seidel, K., et al. (2019). Human Anti-Fungal Th17 Immunity and Pathology Rely on Cross-Reactivity against Candida Albicans. Cell 176, 1340–1355.e15.

Bentebibel, S.E., Lopez, S., Obermoser, G., Schmitt, N., Mueller, C., Harrod, C., Flano, E., Mejas, A., Albrecht, R.A., Blankenship, D., et al. (2013). Induction of ICOS+CXCR3+CXCR5+ TH cells correlates with antibody responses to influenza vaccination. Sci Transl Med 5, 176ra32.

Braun, J., Loyal, L., Frentsch, M., Wendsch, D., Georg, P., Kurth, F., Hippensiel, S., Dingelely, D., Kruse, B., Fauchere, F., et al. (2020). Sars-Cov-2- Reactive T Cells in Healthy Donors and Patients with Covid-19. Nature. https://doi.org/10.1038/s41586-020-2598-9.

Carrasco Pro, S., Sidney, J., Paul, S., Lindestam Arlehamn, C., Weiskopf, D., Peters, B., and Sette, A. (2015). Automatic Generation of Validated Specific T-cell Epitope Candidates. Curr Protoc Immunol. https://doi.org/10.1002/0471142235.im0809fs01.18.19.1–18.19.24.
Piazza, F., Costoya, J.A., Merghoub, T., Hobbs, R.M., and Pandolfi, P.P. (2004). Disruption of PLZP in mice leads to increased T-lymphocyte proliferation, cytokine production, and altered hematopoietic stem cell homeostasis. Mol Cell Biol 24, 10456–10469.

Rudensky, A.Y. (2011). Regulatory T cells and Foxp3. Immunol Rev 241, 260–268.

Sallusto, F. (2016). Heterogeneity of Human CD4(+) T Cells Against Microbes. Annu Rev Immunol 34, 317–334.

Schmiedel, B.J., Singh, D., Madrigal, A., Valdivino-Gonzalez, A.G., White, B.M., Zapardiel-Gonzalo, J., Ha, B., Altay, G., Greenbaum, J.A., Mcvicker, G., Seumois, G., Rao, A., Kronenberg, M., Peters, B., and Vijayanand, P. (2018). Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. Cell 175, 1701–1715.e16.

Seder, R.A., Darrah, P.A., and Roederer, M. (2008). T-cell quality in memory and protection: implications for vaccine design. Nat Rev Immunol 8, 247–258.

Somerville, T.D.D., Xu, Y., Wu, X.S., Maia-Silva, D., Hur, S.K., de Almeida, L.M.N., Preall, J.B., Koo, P.K., and Vakoc, C.R. (2020). ZBED2 is an antagonist of interferon regulatory factor 1 and modifies cell identity in pancreatic cancer. Proc Natl Acad Sci USA 117, 11471–11482.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell 177, 1888–1902 E21.

Tay, M.Z., Poh, C.M., Rénia, L., MacAny, P.A., and Ng, L.F.P. (2020). The trinity of COVID-19: immunity, inflammation and intervention. Nat Rev Immunol 20, 363–374.

Thevarajan, I., Nguyen, T.H.O., Koutsakos, M., Druce, J., Caly, L., van de Sandt, C.E., Jia, X., Nicholson, S., Catton, M., Cowie, B., et al. (2020). Breadth of concomitant immune responses prior to patient recovery: a case report of non-severe COVID-19. Nat Med 26, 453–455.

Thieme, C., Arnf, M., Paniskaki, K., Blázquez Navarro, A., Doevelaar, A., Seibert, F.S., Hoeister, B., Konik, M.J., Brenner, T., and Tempfer, C. (2020). The Sars-Cov-2 T-Cell Immunity Is Directed against the Spike, Membrane, and Nucleocapsid Protein and Associated with Covid 19 Severity. Ssrn. https://doi.org/10.2139/Ssrn.3606763.

Thommen, D.S., and Schumacher, T.N. (2018). T Cell Dysfunction in Cancer. Cancer Cell 33, 547–562.

Trappnell, C., Ciacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 32, 381–386.

Vabret, N., Britton, G.J., Gruber, C., Hegde, S., Kim, J., Kuksin, M., Levantovsky, R., Malle, L., Moreira, A., Park, M.D., et al.; Sinai Immunology Review Proj ect (2020). Immunology of COVID-19: Current State of the Science. Immunity 52, 910–941.

Wang, X., Chan, C.C., Yang, M., Deng, J., Poon, V.K., Leung, V.H., Ko, K.H., Zhou, J., Yuen, K.Y., Zheng, B.J., and Lu, L. (2011). A critical role of IL-17 in modulating the B-cell response during H5N1 influenza virus infection. Cell Mol Immunol 8, 462–468.

Wang, X., Sumida, H., and Cyster, J.G. (2014). GPR18 is required for a normal CD8(+x intestinal intraepithelial lymphocyte compartment. J Exp Med 211, 2351–2359.

Weiskopf, D., Bangs, D.J., Sidney, J., Kolla, R.V., De Silva, A.D., de Silva, A.M., Crotty, S., Peters, B., and Sette, A. (2015a). Dengue virus infection elicits highly polarized CX3CR1+ cytotoxic CD4(+) T cells associated with protective immunity. Proc Natl Acad Sci USA 112, E4256–E4263.

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. (2015b). 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.

Xie, M.M., Fang, S., Chen, Q., Liu, H., Wan, J., and Dent, A.L. (2019). Follicular regulatory T cells inhibit the development of granzyme B-expressing follicular helper T cells. JCI Insight, 4, E128076.
## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| **Antibodies**      |        |            |
| CD3                 | Biolegend | SK7; RRID: AB_10640737 |
| CD3                 | Biolegend | UCHT1; RRID: AB_314060 |
| CD4                 | Biolegend | OKT4; RRID: AB_2561866 |
| CD8a                | Biolegend | RPA-T8; RRID: AB_314134 |
| CD8b                | eBioscience | SIDI88BEE; RRID: AB_2762625 |
| CD14                | Biolegend | HCD14; RRID: AB_830893 |
| CD14                | Biolegend | M5E2; RRID: AB_493695 |
| CD19                | Biolegend | HIB19; RRID: AB_314248 |
| CD27                | Biolegend | M-T271; RRID: AB_2561825 |
| CD38                | Biolegend | HIT-2; RRID: AB_2072782 |
| CD38                | Biolegend | SK1; RRID: AB_2564510 |
| CD40 (blocking)     | Miltenyi Biotec | HB14; RRID: AB_10839704 |
| CD45                | BD Bioscience | HI30; RRID: AB_2744399 |
| CD45RA              | Biolegend | H100; RRID: AB_493763 |
| CD56                | Biolegend | HCD56; RRID: AB_10896424 |
| CD69                | Biolegend | FN50; RRID: AB_2563696 |
| CD137 (4-1BB)       | Biolegend | 4B4-1; RRID: AB_2566258 |
| CD137 (4-1BB)       | Miltenyi Biotec | REA765; RRID: AB_2654994 |
| CD138               | Biolegend | M15; RRID: AB_2562899 |
| CD154 (CD40L)       | Biolegend | 24-31; RRID: AB_314829 |
| CD154 (CD40L)       | Miltenyi Biotec | 5C8; RRID: AB_2751206 |
| CD197 (CCR7)        | BD Bioscience | 3D12; RRID: AB_2744306 |
| CD279 (PD-1)        | BD Bioscience | EH12.1; RRID: AB_2739399 |
| CD298 (i2M; TotalSeq-C) | Biolegend | LNH-94; 2M2; RRID: AB_2801031, AB_2801032, AB_2801033, AB_2820042, AB_2820043, AB_2820044, AB_2820045, AB_2820046 |
| HLA-DR              | BD Bioscience | G46-6; RRID: AB_2732846 |
| IgD                 | Biolegend | IA6-2; RRID: AB_2563269 |
| IgG                 | Biolegend | M1310G05; RRID: AB_2565788 |
| IgM                 | Biolegend | MHM-88; RRID: AB_2562916 |
| Ki-67               | BD Bioscience | B56; RRID: AB_2732007 |
| Nucleoprotein       | Sino Biological |            |
| S1/S2 Protein       | Sino Biological |            |
| **Biological Samples** |        |            |
| Cryopreserved PBMCs from hospitalized and non-hospitalized COVID-19 patients | Southampton University Hospital |            |
| Cryopreserved PBMCs from healthy non-exposed subjects | San Diego Blood Bank |            |
| Cryopreserved PBMCs from subjects before and/or after receiving flu vaccination | La Jolla Institute for Immunology |            |
| **Chemicals, Peptides, and Recombinant Proteins** |        |            |
| Peptivator SARS-CoV-2 Prot M (membrane glycoprotein) | Miltenyi Biotec | 130-126-703 |

(Continued on next page)
RESOURCE AVAILABILITY

Lead contact
Further information and requests for reagents may be directed to the lead contact, Pandurangan Vijayanand (vijay@lji.org).

Materials availability
SARS-CoV-2, Human Influenza (FLU), Parainfluenza (HPIV) and Metapneumovirus (HMPV) epitope pools utilized in this paper will be made available to the scientific community upon request and execution of a material transfer agreement (MTA) directed to Dr. Alessandro Sette (alex@lji.org). There might be restrictions on the availability of the peptide reagents due to cost and limited quantity.

Data and Code availability
Scripts are available in our repository on GitHub (https://github.com/vijaybioinfo/COVID19_2020). Sequencing data for this study has been deposited onto the Gene Expression Omnibus with the accession number GSE152522.
EXPERIMENTAL MODEL AND SUBJECT DETAILS

COVID-19 patients and samples
Ethical approval for this study from the Berkshire Research Ethics Committee 20/SC/0155 and the Ethics Committee of La Jolla Institute for Immunology (LJI) was in place. Written consent was obtained from all subjects. 22 hospitalized patients in a large teaching hospital in the south of England with SARS-CoV-2 infection, confirmed by reverse transcriptase polymerase chain reaction (RT-PCR) assay for detecting SARS-CoV-2, between April-May 2020 were recruited to the study. A further cohort of 18 participants consisting of healthcare workers who were not hospitalized with COVID-19 illness, confirmed based on RT-PCR assay or serological evidence of SARS-CoV-2 antibodies, were also recruited over the same period. All subjects provided up to 80 mL of blood for research studies. Clinical and demographic data were collected from patient records for hospitalized patients including comorbidities, blood results, drug intervention, radiological involvement, thrombotic events, microbiology, and virology results (Table S1A). The 22 hospitalized patients had a median age of 60 (33-82), 17 of these patients (77%) were men and this cohort consisted of 16 (73%) White British/White Other, 4 (18%) Indian, and 2 (9%) Black British patients. All hospitalized patients survived to discharge from hospital. All hospitalized patients were still symptomatic at time of blood collection, whereas some of the non-hospitalized patients (4/18) were symptom free (Table S1A). The 18 non-hospitalized participants had a median age of 39 (22-50), 8 (44%) of these participants were men and this cohort consisted of 15 (83%) White British/White Other, 2 (11%) Arab, and 1 (6%) Chinese participant. We noted that the median age of the non-hospitalized patients was lower than the hospitalized COVID-19 patients.

Healthy controls
To study HPIV, HMPV, and SARS-CoV-2-reactive CD4+ T cells from healthy non-exposed subjects (pre-COVID-19 pandemic), we utilized de-identified buffy coat samples from 5 healthy adult donors who donated blood at the San Diego Blood Bank before 2019, prior to the Covid-19 pandemic. Donors were considered to be in good health, free of cold or flu-like symptoms and with no history of Hepatitis B or Hepatitis C infection. The median age was 50 (32-71) and 4 of these patients (80%) were men. To study FLU-reactive cells, we obtained de-identified blood samples from 8 donors enrolled in the LJI Normal Blood Donor Program before and/or after (12 - 14 days) receiving the FLUCELVAX vaccine (September and October 2019). The median age was 37 (26-57) and 5 of these patients (63%) were women. Approval for the use of this material was obtained from the LJI Ethics Committee.

METHOD DETAILS

PBMC processing
Peripheral blood mononuclear cells (PBMCs) were isolated from up to 80ml of anti-coagulated blood by density centrifugation over Lymphoprep (Axis-Shield PoC AS, Oslo, Norway) and cryopreserved in 50% decomplemented human antibody serum, 40% complete RPMI 1640 medium and 10% DMSO.

SARS-CoV-2 peptide pools
Pools of lyophilized peptides covering the immunodominant sequence of the spike glycoprotein and the complete sequence of the membrane glycoprotein of SARS-CoV-2 (15-mer sequences with 11 amino acids overlap) were obtained from Miltenyi Biotec (Thieme et al., 2020) resuspended and stored according to the manufacturer’s instructions.

SARS-CoV-2 antibody testing
The LIAISON SARS-CoV-2 S1/S2 IgG (DiaSorin S.p.A., Saluggia, Italy) was utilized as per the manufacturer’s instructions to obtain quantitative antibody results from plasma samples via an indirect chemiluminescence immunoassay (CLIA) in a United Kingdom Accreditation Service (UKAS) diagnostic laboratory at University Hospital Southampton. Sample results were interpreted as positive (≥ 15 AU/mL), Equivocal (≥ 12.0 and < 15.0 AU/mL) and negative (< 12 AU/mL).

SARS-CoV-2 spike protein-specific B cell responses
To assess the level of SARS-CoV-2 S1/S2-specific B cells, cells were prepared in staining buffer (PBS with 2% FBS and 2 mM EDTA), FcγR blocked (clone 2.4G2, BD Biosciences), stained with indicated primary antibodies and biotinylated S1/S2 proteins (Sino Biological) for 30 min at 4°C; washed, and subsequently stained with streptavidin-BV421. Patients 10, 24 and 49 were analyzed on a different day with a lower intensity violet laser and required different gating.

Epitope Megapool of peptide (MP) design
The Human Parainfluenza (HPIV) and Metapneumovirus (HMPV) CD4+ T cell peptide megapools (MPs) were produced by sequential lyophilization of viral-specific epitopes as previously described (Carrasco Pro et al., 2015, Weiskopf et al., 2015b). Full lists of the viral protein sequences derived from the uniprot database and used for the HPIV and HMPV MP designs are available in Table S1F. T cell prediction was performed using TepiTool tool, available in identification epitope database analysis resources (IEDB-AR, LJI), applying the 7-allele prediction method and a median cutoff ≤ 20 (Dhanda et al., 2019, Paul et al., 2015, Paul et al., 2016). For the HA-influenza MP, we selected 177 experimentally defined epitopes, retrieved by querying the IEDB database (www.IEDB.org) on
07/12/19 with search parameters “positive assay only, No B cell assays, No MHC ligand assay, Host: Homo Sapiens and MHC restriction class II.” The list of epitopes was enriched with predicted peptides derived from the HA sequences of the vaccine strains available in 2017–2018 and 2019 (A/Michigan/45/2015(H1N1), B/Brisbane/60/2008, A/Hong Kong/4801/2014(H3N2), A/Michigan/45/2015(H1N1), A/Alaska/06/2016(H3N2), B/Iowa/06/2017, and B/Phuket/3073/2013). The resulting peptides were then clustered using the IEDB cluster 2.0 tool and the IEDB recommended method (cluster-break method) with a 70% cut off for sequence identity applied (Dhanda et al., 2019, Dhanda et al., 2018) (Table S1E). Peptides were synthesized as crude material (A&A, San Diego, CA), resuspended in DMSO, pooled according to each MP composition and finally sequentially lyophilized (Carrasco Pro et al., 2015).

For screening healthy non-exposed subjects (samples provided before the current pandemic) who cross-react to SARS-CoV-2, we screened 20 healthy non-exposed subjects using SARS-CoV-2 peptide CD4-R and CD4-S pools, as described (Grifoni et al., 2020).

**Antigen-reactive T cell enrichment (ARTE) assay**

Enrichment and FACS sorting of virus-reactive CD154⁺ CD4⁺ memory T cells following peptide pool stimulation was adapted from Bacher et al. 2016 (Bacher et al., 2016). Briefly, PBMCs from each donor, were thawed, washed, plated in 24-well culture plates at a concentration of 5 × 10⁶ cells/mL in 1 mL of serum-free TexMACS medium (Miltenyi Biotec) and left overnight (5% CO₂, 37 °C). Cells were stimulated by the addition of individual virus-specific peptide pools (1 μg/mL) for 6 h in the presence of a blocking CD40 antibody (1 μg/mL; Miltenyi Biotec). For subsequent MACS-based enrichment of CD154⁺, cells were sequentially stained with fluorescence-labeled surface antibodies (antibody list in Table S1G), Cell-hashtag TotalSeq-C antibody (0.5 μg/condition), and a biotin-conjugated CD154 antibody (clone 5C8; Miltenyi Biotec) followed by anti-biotin microbeads (Miltenyi Biotec). Labeled cells were added to MS columns (Miltenyi Biotec) and positively selected cells (CD154⁺) were eluted and used for FACS sorting of CD154⁺ memory CD4⁺ T cells. The flow-through from the column was collected and re-plated to harvest cells responding 24 h after peptide stimulation. Analogous to enrichment for CD154⁺, CD137-expressing CD4⁺ memory T cells were positively selected by staining with biotin-conjugated CD137 antibody (clone REA765; Miltenyi Biotec) followed by anti-biotin MicroBeads and applied to a new MS column. Following elution, enriched populations were immediately sorted using a FACSaria Fusion Cell Sorter (Becton Dickinson) based on dual expression of CD154 and CD69 for the 6 h stimulation condition, and CD137 and CD69 for the 24 h stimulation condition. The gating strategy used for sorting is shown in Figures S1A and S4B. All flow cytometry data were analyzed using FlowJo software (version 10).

**Cell isolation and single-cell RNA-seq assay (10x platform)**

For combined single-cell RNA-seq and TCR-seq assays (10x Genomics), a maximum of 60,000 virus-reactive memory CD4⁺ T cells from up to 8 donors were pooled by sorting into low retention 1.5 mL collection tubes, containing 500 μL of a 1:1 solution of PBS:FBS supplemented with recombinant RNase inhibitor (1:100, Takara). For healthy donors, when possible, equal numbers of cells were isolated from each donor and pooled according to equivalent molar concentrations and sequenced on Illumina NovaSeq6000 sequencing platform with the following read lengths: read 1 – 101 cycles; read 2 – 101 cycles; and i7 index - 8 cycles.

**Single-cell transcriptome analysis**

Reads from single-cell RNA-seq were aligned and collapsed into Unique Molecular Identifiers (UMI) counts using 10x Genomics’ Cell Ranger software (v3.1.0) and mapped to GRCh37 reference (v3.0.0) genome. HashTag UMI counts for each TotalSeq-C antibody capture library were generated with the Feature Barcoding Analysis pipeline from Cell Ranger. To demultiplex donors, UMI counts of cell barcodes were first obtained from the raw data output, and only cells with at least 100 UMI for the hashtag with the highest UMI counts were considered for donor assignment. Donor identities were inferred by MULTSeqDemux (autoThresh = TRUE and maxiter = 10) from Seurat (v3.1.5) using the UMI counts. Each cell barcode was assigned a donor ID, marked as a Doublet or having a Negative enrichment. Cells were re-classified as doublets if the ratio of UMI counts between the top 2 barcodes was less than 3. Cells labeled as Doublet or Negative were removed from downstream analyses. Raw 10x data were independently aggregated using Cell Ranger’s aggr function (v3.1.0). Donors P28 and P48 were not stained with hashtag antibodies and therefore did not contribute to any donor-specific data. The merged data was transferred to the R statistical environment for analysis using the package Seurat (v3.1.5) (Stuart et al., 2019). To further minimize doublets and to eliminate cells with low quality transcriptomes, cells expressing < 800 and > 4400
unique genes, < 1500 and > 20,000 total UMI content, and > 10% of mitochondrial UMIs were excluded. The summary statistics for all the single-cell transcriptome libraries are provided in Table S2C–E and indicate good quality data with no major differences in quality control metrics across multiple batches, where batches are groups of donors whose libraries were sequenced together (Figure S2A). This procedure was independently applied for data from CD4+ T cells stimulated for 0 and 6 h, 6 and 24 h.

For single-cell transcriptome analysis only genes expressed in at least 0.1% of the cells were included. The transcriptome data was then log-transformed and normalized (by a factor of 10,000) per cell, using default settings in Seurat software (Stuart et al., 2019). Variable genes with a mean UMI expression greater than 0.01 and explaining 25% of the total variance were selected using the Variance Stabilizing Transformation method, as described (Stuart et al., 2019). Transcriptomic data from each cell was then further scaled by regressing the number of UMI-detected and percentage of mitochondrial counts. For data from CD4+ T cells stimulated for 6 h, principal component analysis was performed using the variable genes, and based on the standard deviation of PCs in the “elbow plot,” the first 38 principal components (PCs) were selected for further analyses. Cells were clustered using the FindNeighbors and FindClusters functions in Seurat with a resolution of 0.6. The robustness of clustering was independently verified by other clustering methods and by modifying the number of PCs and variable genes utilized for clustering. Analysis of clustering patterns across multiple batches revealed no evidence of strong batch effects (Figure S2A). For data from CD4+ T cells stimulated for 24 h, the first 16 PCs were selected for further analyses. Cluster 6 (G) in the 24 h dataset was merged with cluster 0 (A) after being identified as TREG. For 0 and 6 h aggregation analysis, 30 PCs were taken. Finally, cells were clustered using the FindNeighbors and FindClusters functions in Seurat with a resolution of 0.6. The robustness of clustering was independently verified by other clustering methods and by modifying the number of PCs and variable genes utilized for clustering. Analysis of clustering patterns across multiple batches revealed no evidence of strong batch effects (Figure S2A). For data from CD4+ T cells stimulated for 24 h, the first 16 PCs were selected for further analyses. Cluster 6 (G) in the 24 h dataset was merged with cluster 0 (A) after being identified as TREG. For 0 and 6 h aggregation analysis, 30 PCs were taken. Finally, cells were clustered using the FindNeighbors and FindClusters functions in Seurat with a resolution of 0.6 and 0.2 for 6 and 0 h aggregation and 24 h, respectively. Further visualizations of exported normalized data such as UMAP or “violin” plots were generated using the Seurat package and custom R scripts. Violin shape represents the distribution of cell expressing transcript of interest (based on a Gaussian Kernel density estimation model) and are colored according to the percentage of cells expressing the transcript of interest.

**Single-cell differential gene expression analysis**

Pairwise single-cell differential gene expression analysis was performed using the MAST package in R (v1.8.2) (Finak et al., 2015) after conversion of data to log2 counts per million (log2(CPM + 1)). A gene was considered differentially expressed when Benjamini–Hochberg adjusted P-value was < 0.05 and a log2 fold change was more than 0.25. For finding cluster markers (transcripts enriched in a given cluster) the function FindAllMarkers from Seurat was used.

**Gene Set Enrichment Analysis and Signature Module Scores**

GSEA scores were calculated with the package fgsea in R using the signal-to-noise ratio (or the log2 fold change for cluster 5 versus cluster 0 comparison) as a metric. Gene sets were limited by minSize = 3 and maxSize = 500. Normalized enrichment scores were presented as GSEA plots. Signature module scores were calculated with AddModuleScore function, using default settings in Seurat. Briefly, for each cell, the score is defined by the mean of the signature gene list after the mean expression of an aggregate of control gene lists is subtracted. Control gene lists were sampled (same size as the signature list) from bins created based on the level of expression of the signature gene list. Gene lists used for analysis are provided in Table S2H.

**Single-cell trajectory analysis**

The “branched” trajectory was constructed using Monocle 3 (v0.2.1, default settings) (Trapnell et al., 2014) with the number of UMI, percentage of mitochondrial UMI as the model formula and including the highly variable genes from Seurat for consistency. After setting a single partition for all cells, the cell-trajectory was projected on the PCA and UMAP generated from Seurat analysis. The ‘root’ was selected by the get_earliest_principal_node function provided in the package. Monocle 3 alpha was used to analyze cluster 0 and 5 using the DDRTree algorithm for dimensional reduction after selecting the top 500 highly variable genes with Seurat.

**T cell receptor (TCR) sequence analysis**

Reads from single-cell V(D)J TCR sequence enriched libraries (Table S2D) were processed with the vdj pipeline from Cell Ranger (v3.1.0 and human annotations reference GRCh38, v3.1.0, as recommended). In brief, the V(D)J transcripts were assembled and their annotations were obtained for each independent library. In order to perform combined analysis of single-cell transcriptome and TCR sequence from the same cells, V(D)J libraries were first aggregated using a custom script. Then cell barcode suffixes from these libraries were revised according to the order of their gene expression libraries. Unique clonotypes, as defined by 10x Genomics as a set of productive Complementarity-Determining Region 3 (CDR3) sequences, were identified across all library files and their frequency and proportion (clone statistics) were calculated based on the aggregation result considering only the cells present in the gene expression libraries. This procedure was independently applied for data from CD4+ T cells stimulated for 6 and 24 h. Based on the vdj aggregate files, barcodes captured by our gene expression data and previously filtered to keep only good-quality cells, were annotated with a specific clonotype ID alongside their clone size (number of cells with the same clonotypes in either one or both the TCR alpha and beta chains) and other statistics (Table S4A,B,E and F). Cells that share clonotype with more than 1 cell were called as clonally expanded (clone size >2). Clone size for each cell was visualized on UMAP, depicting only SARS-CoV-2-reactive CD4+ T cells. Sharing of clonotype between cells in different clusters was depicted using the tool UpSetR (Conway et al., 2017). Finally, in order to assess the sharing between the 0- and 6 h datasets, the same aggregation process was applied for all of the vdj libraries from these data and only SARS-CoV-2-reactive CD4+ T cells specifically isolated from matched patients between sets were considered.
QUANTIFICATION AND STATISTICAL ANALYSIS

Processing of data, applied methods and codes are described in the respective section in the STAR Methods. The number of subjects, samples, replicates analyzed, and the statistical test performed are indicated in the figure legends or STAR methods. Statistical analysis for comparison between two groups were assessed with Mann Whitney U test and correlation assessed with spearman test with using GraphPad Prism.
Figure S1. CD4+ T Cell Responses in COVID-19 Illness, Related to Figure 1

(A) Gating strategy to sort: lymphocytes size-scatter gate, single cells (Height versus Area forward scatter (FSC)), live, CD3+ CD4+ memory (CD45RA+ CCR7+ naive cells excluded) activated CD154+ CD69+ cells. Surface expression of activation markers was analyzed on memory CD4+ T cells.

(B) Representative FACS plots (left) showing surface expression of PD-1 and CD38 in memory CD4+ T cells ex vivo and in CD154+ CD69+ memory CD4+ T cells following 6 h of stimulation, post-enrichment (CD154-based). (Middle) Plots depicting percentage of CD154+ CD69+ memory CD4+ T cells expressing PD-1 or CD38 following stimulation and post-enrichment (CD154-based) in 17 hospitalized and 18 non-hospitalized COVID-19 patients. (Right) Plot showing the total number of sorted CD154+ CD69+ memory CD4+ T cells per million PBMCs; data are mean ± SEM.

(C) Representative FACS plots showing surface staining of CD154 and CD69 in memory CD4+ T cells stimulated for 6 h with individual virus megapools, pre-enrichment (top) and post-enrichment (CD154-based) (bottom) in healthy non-exposed subjects. (Right) Percentage of memory CD4+ T cells co-expressing CD154 and CD69 following stimulation with individual virus megapools (pre-enrichment); data are mean ± SEM.

(D) Representative FACS plots (left) showing surface staining of CD154 in memory CD4+ T cells stimulated with Influenza megapool, pre-enrichment in healthy subjects pre and/or post-vaccination. (Right) Percentage of memory CD4+ T cells expressing CD154 following stimulation with Influenza megapool (pre-enrichment); data are mean ± SEM.

(E) Representative FACS plots showing surface staining of CD154 in memory CD4+ T cells stimulated with Influenza megapool, post-enrichment (CD154-based), in healthy subjects pre and/or post-vaccination.
Figure S2. SARS-CoV-2-Reactive CD4+ T Cells Are Enriched for TFH Cells and CD4-CTLs, Related to Figure 2

(A) Number of genes recovered for each 10x library sequenced.

(B) Proportion of cells in each cluster for the 6 batches of donors.

(C) Donut charts show proportion of individual virus-reactive CD4+ T cells per cluster for different viruses. Notable clusters are highlighted.

(D) Violin plots showing enrichment patterns of TH17, IFN response, TFH, and CD4-CTLs gene signatures for each cluster. Color indicates mean signature score of cells within a cluster.

(E) Violin plots showing normalized expression level (log2(CPM+1)) of select TH1, TH17, IFN response, TFH, and CD4-CTL marker transcripts in designated clusters compared to an aggregation of remaining cells (Rest). Color indicates the percentage of cells expressing indicated transcript.

(F) Scatterplot displaying co-expression level (log2(CPM+1)) of IL2 and TNF transcripts in IFNG-expressing, virus-reactive memory CD4+ T cells in cluster 1. Numbers indicate percentage of cells in each quadrant.

(G) Gene set enrichment analysis (GSEA) for TH17, IFN response, cell cycling, TFH, and CD4-CTL signature genes in a given cluster compared to the rest of the cells; *p < 0.05; ***p < 0.01; ****p < 0.001.
Figure S3. SARS-CoV-2-Reactive CD4+ T Cell Subsets Associated with Disease Severity, Related to Figure 3

(A) Average frequency of cells per cluster from hospitalized and non-hospitalized COVID-19 patients. 

(B) Proportion of cluster 5 cells in SARS-CoV-2-reactive cytotoxic T_{fr} cells (cluster 0, 5, and 7) in non-hospitalized and hospitalized COVID-19 patients who provided blood samples under 21 days (left) and over 21 days (right) after onset of symptoms. Data are mean ± S.E.M; significance for comparisons was computed using Mann-Whitney U test; **p < 0.01; ***p < 0.001.

(C) Proportion of cluster 7 cells in SARS-CoV-2-reactive T_{fr} cells in non-hospitalized and hospitalized COVID-19 patients. Data are mean ± SEM. Significance for comparisons was computed using Mann-Whitney U test; ns identifies non-significant P value.

(D) Volcano plot showing differentially expressed genes between SARS-CoV-2-reactive CD4+ T cells in cluster 5 versus cluster 0.

(E) Violin plots showing expression level (log2(CPM+1)) of PRF1 and GZMB transcripts in cells from clusters 0, 5 and 7.

(F) Scatterplot displaying co-expression level (log2(CPM+1)) of PRF1 and GZMB transcripts in SARS-CoV-2-reactive cells present in cluster 7. Numbers indicate percentage of cells in each quadrant.

(G) Concentration of S1/S2 antibodies in the circulation of 22 hospitalized and 16 hospitalized non-hospitalized COVID-19 patients. Data are mean ± S.E.M; significance for comparisons was computed using Mann-Whitney U test; *p < 0.05.

(H) Correlation between percentage of SARS-CoV-2-reactive CD4+ T_{fr} cells form cluster 0 as a frequency of total CD4+ T_{fr} cells and S1/S2 antibody titers (left two plots) and interval between symptom onset and blood draw (right two plots) in 15 non-hospitalized and 20 hospitalized (left) COVID-19 patients. Correlation coefficient r and the related P value were computed using Spearman correlation; ***p < 0.001.

(I) FACS plots showing S1/S2-specific B cells in 9 COVID-19 patients. Patient ID and proportion of SARS-CoV-2-reactive T_{fr} cells is specified.

(J) Ingenuity pathway analysis (IPA) of genes with increased expression (adjusted p < 0.05 and log2 fold change > 1) between cells from cluster 5 versus cluster 0. Upstream regulatory network analysis of genes in IFN alpha pathway.

(K) GSEA for IFN response signature genes in cluster 5 versus cluster 0; ***p < 0.001.
Figure S4. Single-Cell TCR Sequence Analysis and Analysis of SARS-CoV-2-Reactive CD4+ T Cells from 24 h Stimulation and Ex Vivo Conditions, Related to Figure 4

(A) Average expression and percent expression of selected transcripts in indicated clusters.

(B) Violin plots showing normalized expression level (log2(CPM+1)) of CCL3, CCL4, CCL5, XCL1, and XCL2 transcripts in designated clusters (6 and 9) compared to an aggregation of remaining cells (Rest).

(legend continued on next page)
(C) Scatterplots displaying co-expression level ([log2(CPM+1)] of XCL1 and XCL2 transcripts in SARS-CoV-2-reactive cells present in designated clusters. Numbers indicate percentage of cells in each quadrant.

(D) Proportion of expanded SARS-CoV-2-reactive CD4+ T cells (clone size >2) in hospitalized and non-hospitalized COVID-19 patients (6 h stimulation condition). Data are mean ± S.E.M; significance for comparisons were computed using Mann-Whitney U test; *p < 0.05.

(E) Single-cell transcriptomes of memory CD4+ T cells expressing activation markers (CD38, HLA-DR, PD-1) ex vivo (0 h; blue) and sorted CD154+ CD69+ memory CD4+ T cells following 6 h stimulation with virus-specific peptide megapools (6 h; red) are displayed by UMAP. Seurat-based clustering of 122,292 cells.

(F) UMAP showing activation, Tfh, and CD4-CTL signature scores for each cell.

(G) Violin plots showing expression level ([log2(CPM+1)] of TNFRSF4, TNFRSF18, MIR155HG, CD200, IFNG, IL2, TNF, and POU2AF1 transcripts in 0- and 6 h time points.

(H) Number of cells from matched patients with shared (yellow) and unique (blue) TCRs between activation marker-positive cells sorted ex vivo (0 h) and 6 h peptide stimulated populations (left). Venn diagram illustrating the number of shared clones between activation marker-positive CD4+ T cells sorted ex vivo (0 h) and 6 h peptide stimulated populations.
Figure S5. Analysis of SARS-CoV-2-Reactive CD4+ T Cells from 24 h Stimulation Condition, Related to Figure 5

(A) Representative FACS plots showing surface staining of CD137 and CD69 in memory CD4+ T cells stimulated for 24 h with SARS-CoV-2 peptide pools, post-enrichment (CD137-based), in hospitalized and non-hospitalized COVID-19 patients (left). Summary of number of cells sorted in 14 hospitalized and 17 non-hospitalized COVID-19 patients (right); data are mean ± SEM.

(B) GSEA for TREG, cytotoxicity, TFH and TH17 signature genes in a given cluster compared to the rest of the cells; **p < 0.01; ***p < 0.001.

(legend continued on next page)
(C) Unsupervised clustering of 17 hospitalized and 13 non-hospitalized COVID-19 patients based on the proportions of SARS-CoV-2-reactive CD4+ T cells in different clusters following 24 h peptide stimulation. Clusters with fewer than 5% of the total dataset are not depicted. Hospitalization status (red versus green) and sex (pink versus blue) are indicated in the annotation rows immediately below the dendrogram.

(D) UMAP showing TCR clone size (log2, color scale) of SARS-CoV-2-reactive cells from COVID-19 patients (24 h stimulation condition).

(E) Proportion of clonally expanded (clone size >2) and non-expanded cells in each cluster (24 h stimulation condition).

(F) GSEA for TFH and TFR signature genes in IL1R2+ cells compared to IL1R2− cells in cluster A; *p < 0.05; ***p < 0.001.