Signature of long-lived memory CD8$^+$ T cells in acute SARS-CoV-2 infection

https://doi.org/10.1038/s41586-021-04280-x

Received: 22 July 2021

Accepted: 24 November 2021

Published online: 7 December 2021

Open access

Check for updates

Sarah Adamo[1], Jan Michler[2], Yves Zurbuchen[1], Carlo Cervia[1], Patrick Taeschler[1], Miro E. Raeber[1], Simona Baghai Sain[2], Jakob Nilsson[1], Andreas E. Moor[2] & Onur Boyman[1,3]✉

Immunological memory is a hallmark of adaptive immunity and facilitates an accelerated and enhanced immune response upon re-infection with the same pathogen[1,2]. Since the outbreak of the ongoing COVID-19 pandemic, a key question has focused on which SARS-CoV-2-specific T cells stimulated during acute infection give rise to long-lived memory T cells[3]. Here, using spectral flow cytometry combined with cellular indexing of transcriptomes and T cell receptor sequencing, we longitudinally characterized individual SARS-CoV-2-specific CD8$^+$ T cells of patients with COVID-19 from acute infection to 1 year into recovery and found a distinct signature identifying long-lived memory CD8$^+$ T cells. SARS-CoV-2-specific memory CD8$^+$ T cells persisting 1 year after acute infection express CD45RA, IL-7 receptor-α and T cell factor 1, but they maintain low expression of CCR7, thus resembling CD45RA$^+$ effector memory T cells. Tracking individual clones of SARS-CoV-2-specific CD8$^+$ T cells, we reveal that an interferon signature marks clones that give rise to long-lived cells, whereas prolonged proliferation and mechanistic target of rapamycin signalling are associated with clonal disappearance from the blood. Collectively, we describe a transcriptional signature that marks long-lived, circulating human memory CD8$^+$ T cells following an acute viral infection.

The coronavirus disease 2019 (COVID-19) pandemic has taken an extraordinary toll on global health and economy, affecting billions of lives all over the world. The ongoing vaccination efforts appear to curtail the spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and prevent severe disease, even as new virus variants emerge[4,5]. Yet, prevailing questions concern whether and how exposure to SARS-CoV-2 by infection or immunization might result in long-term protective immunity.

On encountering their cognate antigen on antigen-presenting cells, antigen-specific CD8$^+$ T cells proliferate and differentiate into effector cells aimed at controlling the pathogen by killing virus-infected host cells. Following virus elimination, 90–95% of effector T cells undergo apoptosis, whereas some antigen-specific T cells survive to become long-lived memory T cells that are able to protect the host from re-infection with the same pathogen[2,6].

While antigen-specific effector T cell responses are generated during acute SARS-CoV-2 infection[7–12] and persist for several months[13–17], little is known about changes in memory phenotypes over time. Previous studies using live-attenuated virus vaccines in healthy donors[18–21] have described phenotypical trajectories of human antigen-specific T cell populations. However, it is unknown whether infection with a natural virus generates comparable memory T cell responses in humans, as infection route, viral load, inflammation and various host-related factors are likely to affect T cell responses and memory formation. Moreover, phenotypical and transcriptional trajectories at the single T cell receptor (TCR) level and the factors instructing individual effector T cell clones on their development to long-lived memory T cells have not been investigated in humans.

Phenotype of SARS-CoV-2$^+$ CD8$^+$ T cells

To assess the dynamics of antigen-specific T cells in COVID-19, we recruited 175 patients with real-time PCR (RT–PCR)-confirmed COVID-19, sampled during their symptomatic acute phase and followed up 6 months and 1 year after acute infection (Fig. 1a). We conducted human leukocyte antigen (HLA) typing on all patients and healthy controls and selected individuals carrying the *HLA-A*01:01*, *HLA-A*11:01* or *HLA-A*24:02* alleles for this study (n = 47 patients and n = 13 healthy controls; characteristics are included in Extended Data Table 1). In these individuals, SARS-CoV-2-specific CD8$^+$ T cells were detected by using HLA-A*01:01, HLA-A*11:01 and HLA-A*24:02 major histocompatibility complex class I (MHC-I) dextramers[12], hereafter termed CoV2-Dex (Fig. 1b, Extended Data Fig. 1a, b), and validated by using HLA-A*01:01 and HLA-A*11:01 MHC-I pentamers[22], hereafter termed CoV2-Pent (Extended Data Fig. 1c, d). Healthy controls were seronegative for SARS-CoV-2 spike-specific IgG and IgA (Extended Data Fig. 1e).

SARS-CoV-2-specific CD8$^+$ T cells were found in most patients carrying an *HLA-A*01:01* or *HLA-A*11:01* allele during acute infection and 6 months later (Fig. 1c). Moreover, we detected SARS-CoV-2-specific CD8$^+$ T cells 1 year after acute SARS-CoV-2 infection (Fig. 1b, c). Staining with HLA-A*24:02 CoV2-Dex (which carried a spike-derived peptide, QYIKWPWYI) showed much higher background staining in some healthy donors (Extended Data Fig. 2a), possibly due to

[1]Department of Immunology, University Hospital Zurich, Zurich, Switzerland. [2]Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland. [3]Faculty of Medicine, University of Zurich, Zurich, Switzerland. ✉e-mail: onur.boyman@uzh.ch

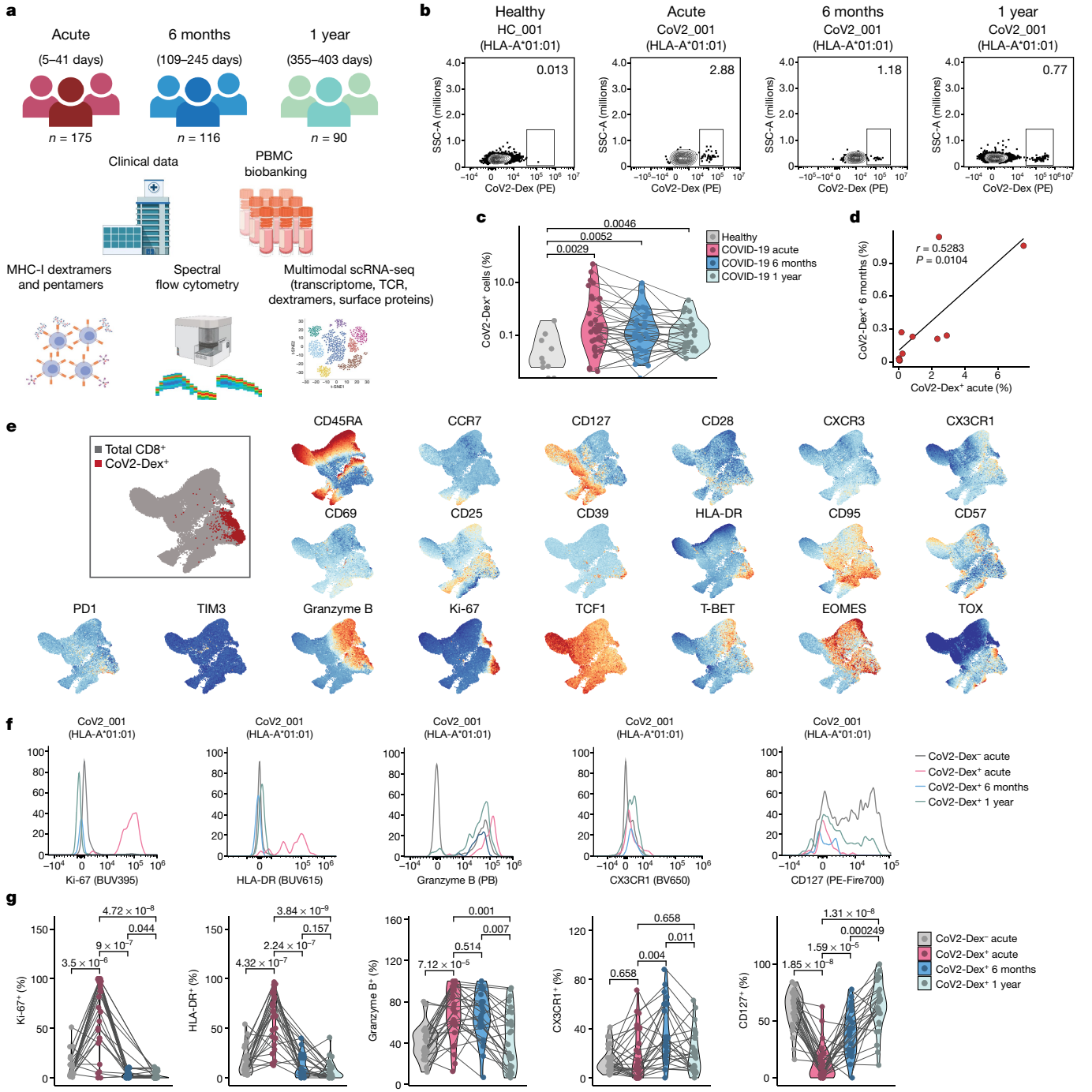


Fig. 1 | Characteristics of antigen-specific CD8⁺ T cells during acute and memory phases of SARS-CoV-2 infection. a, Overview of study design. PBMC, peripheral blood mononuclear cell. **b**, Representative plots of CoV2-Dex staining. PE, Phycoerythrin. Numbers in the plots indicate percentage of parent population. **c**, Frequency of CoV2-Dex⁺ cells in healthy donors and patients with COVID-19 during acute infection and 6 months and 1 year after infection. Each dot represents an independent donor at the indicated timepoint ($n = 10$ healthy, $n = 37$ acute, $n = 32$ 6 months, $n = 29$ 1 year after infection). *P* values are shown. **d**, Linear regression of frequency of CoV2-Dex⁺ cells 6 months after infection as a function of CoV2-Dex⁺ cell frequencies during acute infection ($n = 11$). The *P* value was calculated with *t*-statistic. **e**, Uniform manifold approximation and projection (UMAP) plots of marker expression for up to 2,000 CD8⁺ T cells from each sample collected during acute infection ($n = 37$) analysed by spectral flow cytometry. Regions with high marker expression appear in red. An overlay of CoV2-Dex⁺ cells (red) and total CD8⁺ T cells (grey) is shown in the top left. **f**, Representative histograms showing expression of selected markers on CoV2-Dex⁻ and CoV2-Dex⁺ cells. **g**, Frequency of Ki-67⁺, HLA-DR⁺, granzyme B⁺, CX3CR1⁺ and CD127⁺ cells in CoV2-Dex⁻ (grey) and CoV2-Dex⁺ cells during acute infection and 6 months and 1 year after infection. Analysis was conducted on paired samples from acute infection versus 6 months and/or 1 year after infection ($n = 28$ acute, $n = 24$ 6 months, $n = 29$ 1 year). The grey lines connect individual donors sampled at different timepoints. *P* values were calculated using a Wilcoxon–Mann–Whitney test in **c** and **g** and corrected for multiple comparisons in **g**. All tests were performed two-sided.

cross-reactivity. In individuals carrying *HLA-A*24:02* alleles, we did not observe increased frequencies of CoV2-Dex⁺ cells during acute infection compared to healthy donors (Extended Data Fig. 2b), contrary to patients with *HLA-A*01:01* or *HLA-A*11:01* alleles (Fig. 1c). This finding might indicate that T cells specific for this spike epitope did not undergo marked expansion during SARS-CoV-2 infection. We could

not determine whether HLA-A*24:02 CoV2-Dex$^+$ cells had an activated or proliferating phenotype due to low cell numbers. Furthermore, we noted a lower reactivity to HLA-A*11:01 dextramers than to HLA-A*01:01 dextramers (Extended Data Fig. 2c) during acute infection, which persisted 6 months after infection. These data suggest heterogeneity in effector and memory T cell responses based on HLA type and specific epitopes, although they need careful interpretation due to limited patient numbers.

In patients with an *HLA-A*01:01* or *HLA-A*11:01* allele, the frequency of CoV2-Dex$^+$ cells during acute infection correlated with the frequency of specific cells at the memory phase (Fig. 1d). In acute infection, flow cytometry analysis of CoV2-Dex$^+$ cells revealed a circumscript phenotype of activated cells, dominated by high abundance of Ki-67 and HLA-DR (Fig. 1e, f, Extended Data Fig. 2d, e). CoV2-Dex$^+$ cells also expressed granzyme B and tended to have higher expression of the terminal differentiation marker CX3CR1, whereas surface CD127 (also known as IL-7 receptor-α) was markedly downregulated (Fig. 1f, g). At the 6-month and 1-year timepoints, frequencies of Ki-67$^+$ and HLA-DR$^+$ CoV2-Dex$^+$ cells declined and the frequency of CD127$^+$ cells increased (Fig. 1f, g), indicating a transition from effector to memory state[6,23]. We did not observe phenotypical differences between HLA-A*01:01 and HLA-A*11:01 (Extended Data Fig. 2f).

Notably, disease severity seemed to positively correlate with the extent of CD8$^+$ T cell responses during acute infection, as well as frequencies of CoV2-Dex$^+$ cells 6 months and 1 year after infection, although expansion of CoV2-Dex$^+$ cells was also evident in patients with mild disease (Extended Data Fig. 3a). During acute infection, both proliferation and activation were only minimally affected by disease severity in the CoV2-Dex$^+$ compartment, whereas a relevant difference was observed in CoV2-Dex$^-$ cells (Extended Data Fig. 3b). This discrepancy could be due to higher bystander activation in severe disease or higher abundance of undetected SARS-CoV-2-reactive T cells. Patients with severe disease showed higher expression of granzyme B and CX3CR1 on CoV2-Dex$^+$ cells, possibly indicating a different T cell differentiation program during the acute phase of severe COVID-19. These differences were no longer evident 6 months and 1 year after infection (Extended Data Fig. 3c, d).

Transcriptome of SARS-CoV-2$^+$ CD8$^+$ clones

To examine the transcriptional phenotype of individual SARS-CoV-2-specific CD8$^+$ T cells, we sorted CoV2-Dex$^+$CD8$^+$ T cells and CoV2-Dex$^-$CD8$^+$ T cells, mixed them at a 1:10 ratio, and performed single-cell RNA sequencing (scRNA-seq) on a subgroup of patients (n = 20 acute and n = 19 6-month timepoint). We classified sequenced cells as CoV2-Dex$^-$ or CoV2-Dex$^+$ based on their dCODE Dextramer unique molecular identifier counts (Methods, Extended Data Fig. 4a) and positivity for a single SARS-CoV-2 epitope (Extended Data Fig. 4b). Unbiased clustering revealed 12 distinct CD8$^+$ T cell clusters (Fig. 2a), none of which was dominated by a single patient (Extended Data Fig. 4c). Some clusters showed nearly complete segregation between the acute and memory phases (Extended Data Fig. 5a, b). In line with our flow cytometry data (Fig. 1e, Extended Data Fig. 5c), CoV2-Dex$^+$CD8$^+$ T cells showed a rather segregated transcriptional makeup during acute infection, whereas their transcriptional state was more heterogeneous 6 months after infection (Fig. 2b). Comparing the contribution of CoV2-Dex$^+$ cells to different clusters, we observed that clusters 1, 2 and 12 dominated the CoV2-Dex$^+$CD8$^+$ T cell response in the acute phase, whereas clusters 3, 6 and 11 became prominent in the recovery phase (Fig. 2c, Extended Data Fig. 5d). While clusters 1, 2 and 12 corresponded to cytotoxic, activated and proliferating cells, respectively, cluster 3 showed a signature marked by enrichment of NF-κB and Jun/Fos signalling, cluster 6 displayed an oxidative phosphorylation signature, and cluster 11 showed a dual signature marked by enrichment of interferon (IFN) response genes and genes encoding the effector cytokines IFNγ, tumour necrosis factor (TNF) and lymphotoxin-α (LTα) (Fig. 2d, e). Similarly, among genes with significantly higher expression in CoV2-Dex$^+$ cells from the acute phase versus the recovery phase, we found genes related to cytotoxicity (*GZMA*, *GZMK* and *PFN1*), activation (*HLA-DRA*, *CD38* and *PDCD5*) and proliferation (*MKI67*, *MCM7* and *NUDC1*), along with IFN response genes (*IFI6*, *MX1*, *IFI27L2* and *IFI44L*) (Extended Data Fig. 5e). *SELL* (which encodes CD62L) appeared to be enriched in cells retrieved during the acute phase rather than the recovery phase (Extended Data Fig. 5e).

To identify phenotypical trajectories in individual antigen-specific T cell clones, we performed TCR sequencing of CoV2-Dex$^+$ cells, which revealed several antigen-specific CD8$^+$ T cell clones for each epitope investigated (Fig. 2f). Clones were considered antigen-specific if any of the clonal cells were CoV2-Dex$^+$ (data available as Supplementary Dataset 1), and clones that were CoV2-Dex$^+$ in the acute phase were considered CoV2-Dex$^+$ independently of CoV2-Dex staining at six months after infection, and vice versa. The number of clones detected during convalescence was markedly lower than that detected during the acute phase of infection (Fig. 2f). In most cases, but not all, dominant clones in the acute phase corresponded to the largest clones found in the recovery phase (Fig. 2g). The phenotypical changes in acute infection versus the recovery phase on the general CoV2-Dex$^+$ population were reflected in individual T cell clones. Thus, analysis of individual CoV2-Dex$^+$ clones showed multiple clones containing cells from clusters 1, 2 or 12 during acute infection and cells from clusters 3, 6 and 11 during recovery (Fig. 2h, Extended Data Fig. 6). To better compare gene expression in acute infection versus recovery across all clones, we compiled an 'acute gene signature' comprising *NKG7*, *PRF1*, *GZMB*, *CENPU*, *CENPF* and *MKI67*, and a 'recovery gene signature' comprising *TNF*, *IFIT2*, *IFIT3*, *MT-CO1*, *MT-CO2* and *MT-ATP6*. We observed a significant decrease in acute gene signature transcripts in individual T cell clones from the acute phase to the recovery phase, which was paralleled by an increase in the recovery gene signature (Fig. 2i). Accordingly, individual T cell clones showed a decrease in *MKI67* and *HLA-DRB5* expression between the acute phase and the recovery phase (Fig. 2j).

Memory paths of SARS-CoV-2$^+$CD8$^+$ cells

To better understand the phenotypical memory trajectories of antigen-specific CD8$^+$ T cells following a naturally occurring acute virus infection, we followed CoV2-Dex$^+$ cells longitudinally, at both the population level and the clonal level. In the acute phase, CoV2-Dex$^+$ cells showed mostly an effector/effector memory ($T_{effector}/T_{EM}$) phenotype, whereas frequencies of naive (T_{naive}) cells were lower in CoV2-Dex$^+$ than in CoV2-Dex$^-$ CD8$^+$ T cells (Fig. 3a, b, Extended Data Fig. 7a).

These data were confirmed in CoV2-Pent$^+$ cells (Extended Data Fig. 7b, c). At 6 months and 1 year after infection, we observed a progressive switch from a $T_{effector}/T_{EM}$ phenotype to a terminally differentiated T_{EM} cell re-expressing CD45RA (T_{EMRA}) phenotype; thus, 1 year after infection, most CoV2-Dex$^+$ cells were of a T_{EMRA} phenotype (Fig. 3c, Extended Data Fig. 7d, e). Furthermore, we observed progressive enrichment in stem cell memory T (T_{SCM}) cells, particularly at the 1-year timepoint (Fig. 3c, Extended Data Fig. 7e). Of note, the increase in T_{SCM} cell enrichment was accompanied by an enrichment in T_{naive} cells 1 year after infection (Fig. 3c), possibly indicating that memory cells can reacquire CD45RA and CCR7 also in the absence of CD95 expression. We did not observe differences in memory phenotypes based on HLA, except for a tendency towards more central memory (T_{CM}) cells in HLA-A*11:01 1 year after infection (Extended Data Fig. 7f–h). Conversely, memory phenotypes were strongly influenced by disease severity (Extended Data Fig. 7i–k). Patients with severe disease had fewer CoV2-Dex$^+$ T_{naive} cells 6 months and 1 year after infection, tended to have fewer T_{SCM} cells and showed predominantly CoV2-Dex$^+$ T_{EMRA} cells 1 year after infection. When examining individual T cell clones, we observed an increase in CD45RA expression and a concomitant decrease in CCR7 protein expression determined by TotalSeq from acute infection to 6 months after infection (Fig. 3d), thus confirming an enrichment of a T_{EMRA} phenotype also on a single TCR level.

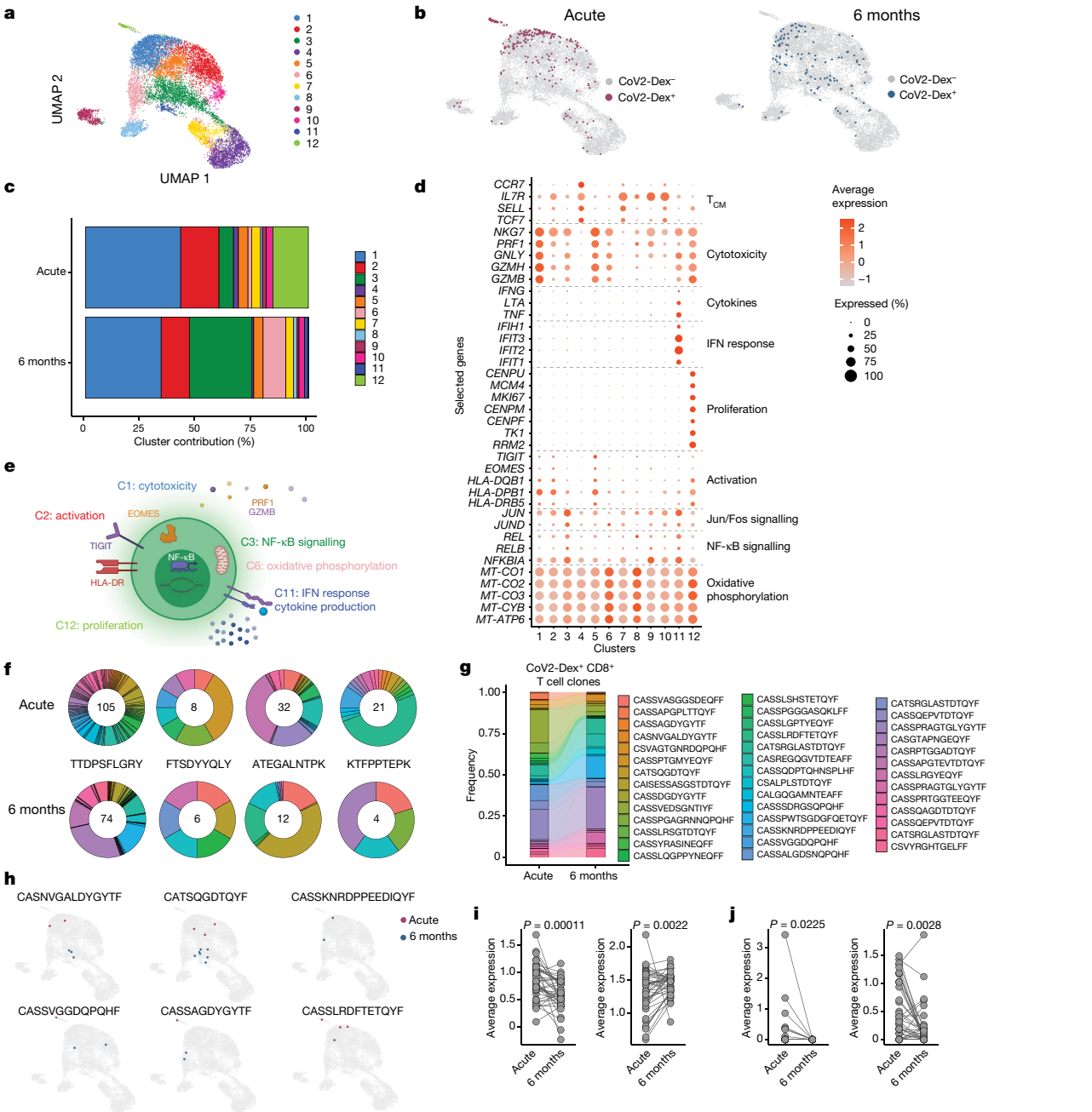


Fig. 2 | Transcriptional makeup of SARS-CoV-2-specific CD8⁺ T cell clones. **a**, Single-cell transcriptomes of CD8⁺ T cells displayed by UMAP. Seurat-based clustering of 14,853 cells, coloured based on cluster ID. **b**, UMAP as in **a**; CoV2-Dex⁺ cells from the acute infection (red) and 6 months after infection (blue) are highlighted. **c**, Cluster composition of CoV2-Dex⁺CD8⁺ T cells in acute infection versus 6 months after infection. **d**, Average expression (colour scale) and the percentage of expressing cells (size scale) of selected genes in indicated clusters. **e**, Schematic summary of the main clusters differentially represented in acute infection and 6 months after infection. **f**, Clonotype distribution in CoV2-Dex⁺ T cell clones (at least one CoV2-Dex⁺ cell per clone) for each of the four epitopes assessed. The number of T cell clones specific for the indicated epitopes at acute infection (top) and 6 months after infection (bottom) is provided within the circle. **g**, Alluvial plot showing relative representation of single clones present during acute infection and 6 months after infection ($n = 41$). **h**, UMAP as in **a**; cells from individual CoV2-Dex⁺CD8⁺ T cell clones in acute infection (red) and 6 months after infection (blue) are highlighted. **i**, Gene signature scores of individual CD8⁺ T cell clones in acute infection (acute gene signature, left) versus 6 months after infection (recovery gene signature, right) ($n = 41$). **j**, Expression of *MKI67* (left) and *HLA-DRB5* (right) in individual CD8⁺ T cell clones in acute infection versus 6 months after infection ($n = 41$). *P* values were calculated using a Wilcoxon signed-rank test in **i** and **j**.

Subsequently, we assessed whether the $T_{effector}/T_{EM}$ and T_{EMRA} phenotypes were associated with specific T cell markers, suggesting distinct differentiation states. Indeed, in CoV2-Dex⁺ cells, we observed several differences between the $T_{effector}/T_{EM}$ and T_{EMRA} populations. CoV2-Dex⁺ T_{EM} cells showed higher expression of Ki-67 and HLA-DR, whereas they had lower abundance of CX3CR1 already during acute infection (Fig. 3e). Notably, we observed the same phenotypical differences between CoV2-Dex⁺ T_{EM} and CoV2-Dex⁺ T_{EMRA} cells 6 months and 1 year after infection (Fig. 3f, g).

As T cell phenotypes are driven by specific transcription factors, we assessed the expression of T cell factor 1 (TCF1), T-box expressed in T cells (T-BET), eomesodermin (EOMES) and thymocyte selection-associated

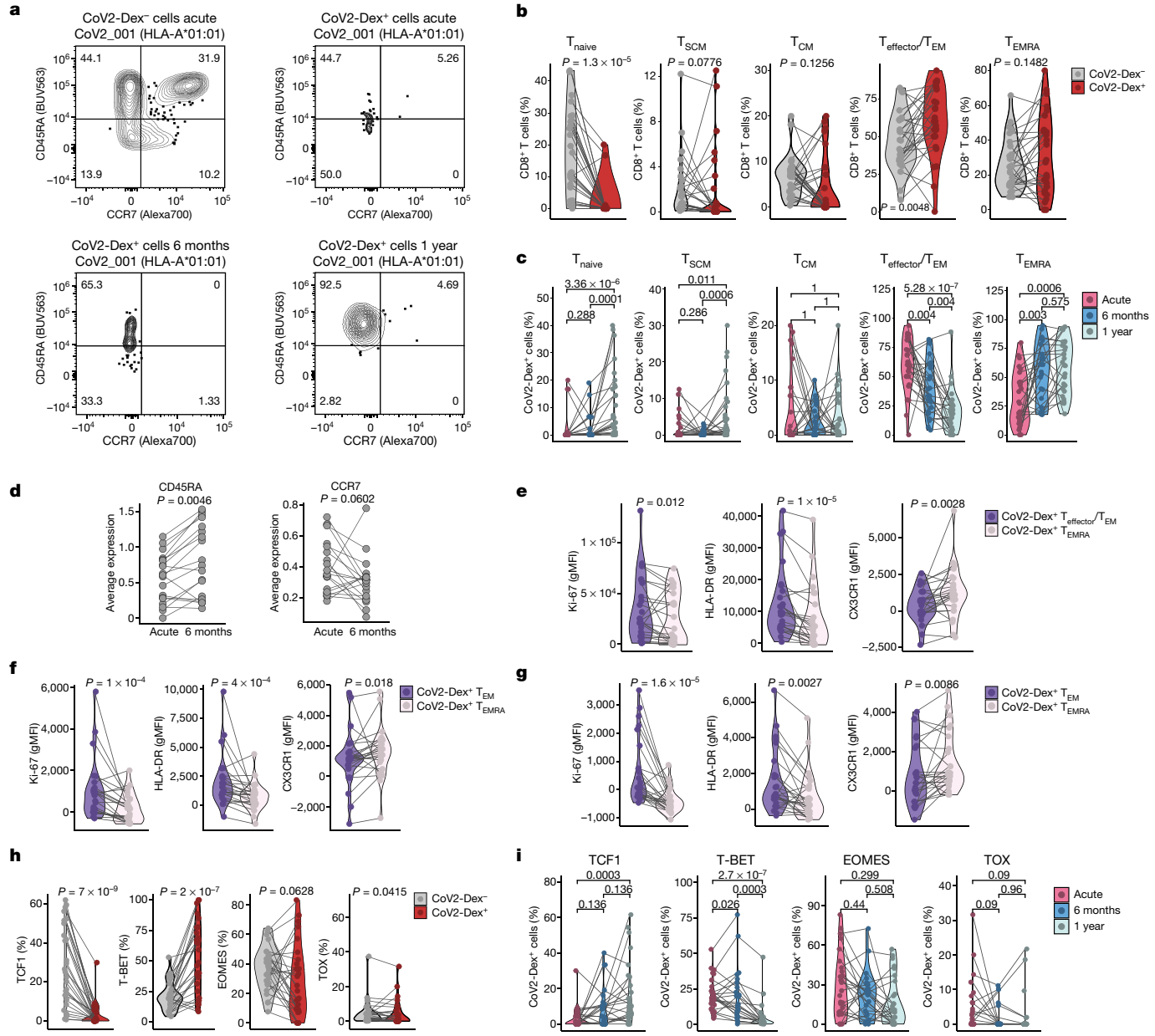


Fig. 3 | Transition of antigen-specific CD8⁺ T cells to TCF1⁺CD45RA⁺ effector memory cells at 1 year. a, Representative plots of CD45RA and CCR7 staining on CoV2-Dex⁻ and CoV2-Dex⁺ cells during acute infection and 6 months and 1 year after infection. Numbers in the plots indicate percentage of parent population. **b**, Percentages of T_{naive}, T_{SCM}, T_{CM}, $T_{effector}/T_{EM}$ and T_{EMRA} cells in CoV2-Dex⁻ and CoV2-Dex⁺ cells during acute infection ($n = 28$). **c**, Percentages of T_{naive}, T_{SCM}, T_{CM}, $T_{effector}/T_{EM}$ and T_{EMRA} CoV2-Dex⁺ cells in acute infection and 6 months and 1 year after infection ($n = 28$ acute, $n = 24$ 6 months, $n = 29$ 1 year). The grey lines connect individual donors sampled at different timepoints. *P* values are also shown. **d**, Expression of CD45RA (left) and CCR7 (right) determined by TotalSeq in individual CD8⁺ T cell clones in acute infection versus 6 months after infection ($n = 41$). **e**–**g**, Geometric mean fluorescence intensity (gMFI) of selected markers on $T_{effector}/T_{EM}$ and T_{EMRA} CoV2-Dex⁺ cells in acute infection (**e**), and 6 months (**f**) and 1 year (**g**) after infection. Phenotypes were evaluated only in patients with more than 5 $T_{effector}/T_{EM}$ and T_{EMRA} CoV2-Dex⁺ cells per sample ($n = 24$ acute, $n = 24$ 6 months, $n = 26$ 1 year). **h**, Percentages of TCF1⁺, T-BET⁺, EOMES⁺ and TOX⁺ CoV2-Dex⁻ and CoV2-Dex⁺ cells during acute infection ($n = 28$). **i**, Percentages of TCF1⁺, T-BET⁺, EOMES⁺ and TOX⁺ CoV2-Dex⁺ cells in acute infection, and 6 months and 1 year after infection ($n = 28$ acute, $n = 24$ 6 months, $n = 29$ 1 year). *P* values were calculated using a Wilcoxon signed-rank test in **b**, **d**–**g**, and a Wilcoxon–Mann–Whitney test with a correction for multiple comparisons using the Holm method in **c** and **i**. All tests were performed two-sided.

high-mobility group box (TOX), which are transcription factors known to have important roles in T cell differentiation[24–27]. CoV2-Dex⁺ cells downregulated TCF1 expression during the acute phase, which was progressively restored at subsequent timepoints (Fig. 3h, i, Extended Data Fig. 8a–c). Conversely, the expression of T-BET was increased in the acute phase (Fig. 3h, Extended Data Fig. 8a) and progressively decreased 6 months and 1 year after infection (Extended Data Fig. 8b, c). A difference in T-BET expression between $T_{effector}/T_{EM}$ and T_{EMRA} CoV2-Dex⁺ cells was not evident, except for a tendency 1 year after infection. However, T_{EMRA} CoV2-Dex⁺ cells expressed lower levels of TCF1 and TOX in the memory phase and lower levels of EOMES at all timepoints (Extended Data Fig. 8d).

Signatures of CD8⁺ memory precursors

Next, we sought to identify the factors present during acute infection that instruct T cell clones towards a memory fate. We compared clones detected in the peripheral blood in both the acute and the convalescent phases (termed persistent) to those that were only seen in

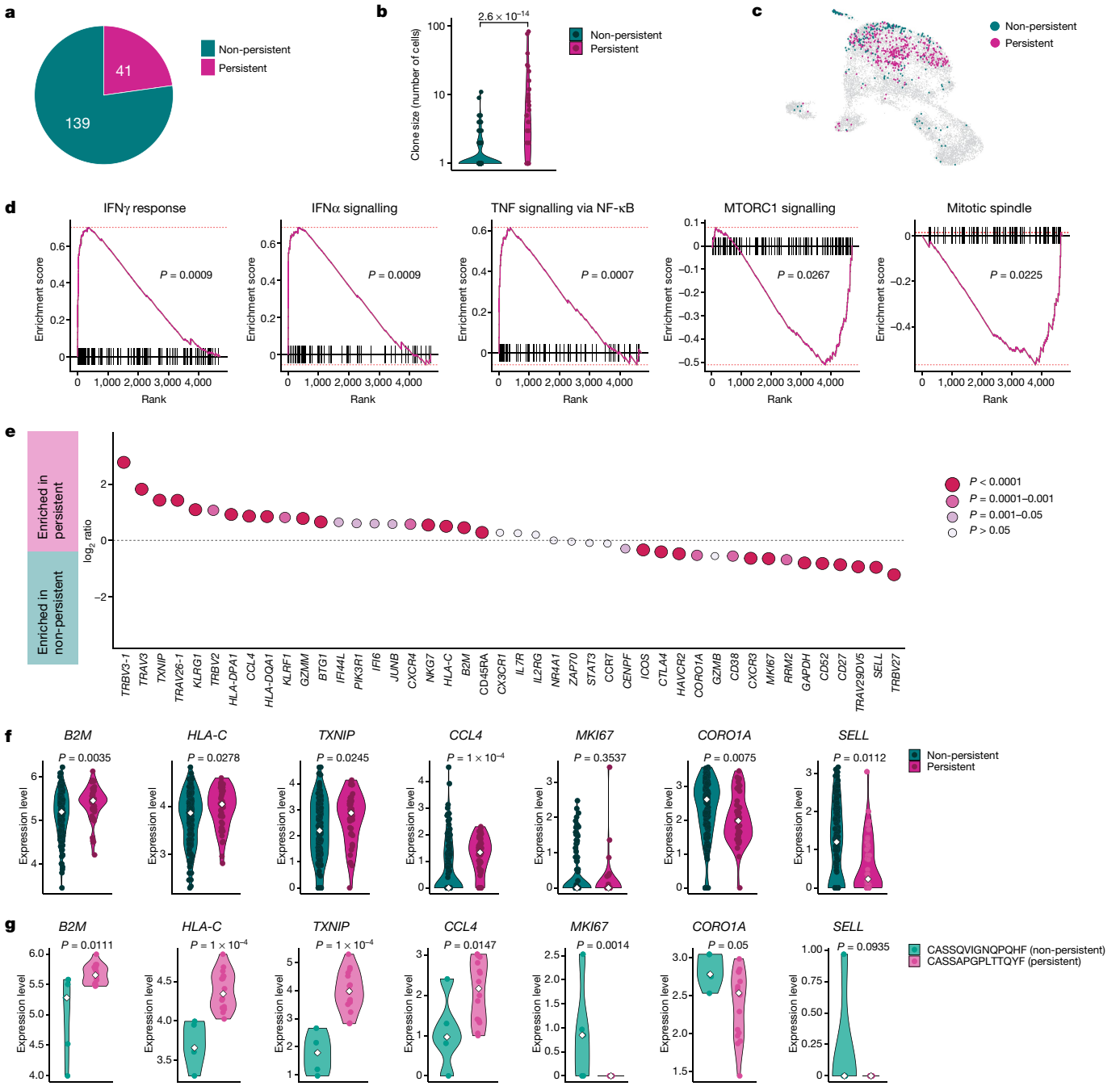


Fig. 4 | Transcriptional signature of antigen-specific CD8⁺ T cell clones persisting at 6 months. a, Proportion of CoV2-Dex⁺ CD8⁺ T cell clones present during acute infection that were also detectable 6 months after infection. **b**, Clone size of persisting versus non-persisting CoV2-Dex⁺ CD8⁺ T cell clones ($n = 41$ persistent, $n = 139$ non-persistent). **c**, UMAP plot of persistent (red) versus non-persistent (green) CoV2-Dex⁺CD8⁺ T cell clones detected during acute infection. **d**, Gene set enrichment analysis showing enrichment of genes associated with cytokine signalling in persistent clones and mTOR signalling and proliferation in non-persistent CoV2-Dex⁺ T cell clones. Red dashed lines indicate minimal and maximal cumulative enrichment values. *P* value calculation was performed as detailed in the Method section. **e**, Expression of selected genes and CCR7 and CD45RA protein determined by Totalseq for persistent versus non-persistent CoV2-Dex⁺ T cell clones. *P* values were calculated using a Wilcoxon signed-rank test; a Bonferroni correction was applied for multiple comparisons. **f**, Expression level of selected genes in persistent versus non-persistent individual T cell clones; each dot represents one clone. **g**, Expression level of selected genes in cells from a single non-persistent clone compared to cells from a single persistent T cell clone; each dot represents one cell ($n = 5$ CASSQVIGNQPQHF, $n = 16$ CASSAPGPLTTQYF). In **f**, **g**, the white diamonds indicate median expression. For **b**, **f**, **g**, *P* values were calculated using a Wilcoxon–Mann–Whitney test. All tests were performed two-sided.

the acute phase and became undetectable in the convalescent phase (non-persistent) (Fig. 4a). Not all clones detected at 6 months after infection were present in the acute phase, probably reflecting a limitation of detection (Extended Data Fig. 9a). Generally, clone size correlated positively with persistence (Fig. 4b). Cells of persistent clones showed a different transcriptional makeup in the acute phase when compared to cells of non-persistent clones (Fig. 4c), which also resulted in a different distribution in the previously identified CD8⁺ T cell clusters (Extended Data Fig. 9b). This effect was robustly seen in different clones and was not due to a few hyper-expanded clones (Extended Data

Fig. 9c). Gene set enrichment analysis revealed distinct signatures in persistent versus non-persistent clones. Genes involved in IFNγ and IFNα responses and TNF signalling were enriched in cells from persistent clones, whereas mechanistic target of rapamycin (mTOR) signalling and mitosis-related genes were enriched in cells from non-persistent clones (Fig. 4d). By comparing differentially regulated genes between cells from persistent and non-persistent clones, we observed genes associated with activation (*HLA-DQA1* and *HLA-DPA1*), terminal differentiation (*KLRG1*) and cytotoxicity (*GZMM* and *NKG7*), as well as certain IFN-induced (*B2M* and *HLA-C*) and TNF-induced (*CCL4*) genes to be enriched in persisters, along with CD45RA protein expression determined by TotalSeq (Fig. 4e). Conversely, cells from non-persistent clones showed higher expression of CTLA4, TIM3 (encoded by *HAVCR2*), Ki-67 (encoded by *MKI67*) and the mTOR-induced gene *CORO1A* (Fig. 4e). The same differences in gene expression could be observed at the clonal level; thus, genes were upregulated or downregulated accordingly in non-persistent individual clones compared to persistent clones (Fig. 4f), as exemplified in two selected clones of a patient (Fig. 4g). Non-persisters showed higher expression of *SELL* during the acute phase (Fig. 4e, f). We also observed differential TCR-Vβ usage between persistent and non-persistent clones (Fig. 4e).

Discussion

In this study, we address outstanding questions related to $CD8^+$ T cell memory upon acute SARS-CoV-2 infection in humans by longitudinally following individual memory $CD8^+$ T cell clones. Phenotypically, we find a transition of $CD8^+$ T cells from $T_{effector}/T_{EM}$ cells to T_{EMRA} cells with progressive enrichment of $TCF1^+$ cells, which is paralleled by a modest enrichment in T_{SCM} cells. While two previous papers have reported a high prevalence of T_{EMRA} cells among SARS-CoV-2-specific $CD8^+$ T cells[13,28], our study provides a description of progressive enrichment in this specific phenotype during the memory phase, at both the clonal level and the population level, revealing that $CD8^+$ T_{EMRA} cells might constitute the main circulating memory subset following an acute viral infection in humans.

Our data provide a different and more detailed view of individual antigen-specific human memory $CD8^+$ T cells than the one observed in the tetramer-positive memory $CD8^+$ T cell population in individuals vaccinated against yellow fever virus[19], where the prevalent subset observed was T_{SCM}. As that study dealt with human memory $CD8^+$ T cells examined several years after vaccination, T_{EMRA} and T_{SCM} cells could be part of the same phenotypical trajectory, with progressive enrichment in T_{SCM} cells over time due to differentiation or competitive advantage. We favour an alternative hypothesis based on the phenotypical differences between memory cells in mild and severe COVID-19 that we observed, suggesting that other factors—such as antigen availability, type of antigen-presenting cells stimulated and cytokine milieu—might influence the type of memory formed, with increased T_{EMRA} cell differentiation upon severe disease versus prevalent T_{SCM} cell formation in mild disease and upon inoculation with live-attenuated virus vaccines.

We also observed enrichment of CD62L expression in $CoV2\text{-}Dex^+$ cells during the acute phase rather than the recovery phase and in non-persistent clones disappearing from the circulation rather than in clones giving rise to circulating memory T cells. Whereas CD62L expression is a hallmark of recirculating T_{CM} cells[29], which are the prevalent memory cells in lymph and secondary lymphoid organs[30], CD62L is not typically associated with the T_{EMRA} phenotype[31]. This might explain our findings, as CD62L expression appears not to be part of the memory differentiation trajectory observed[20]. Whether a fraction of T cell clones detected in peripheral blood during the acute phase contributes to different memory pools in secondary lymphoid organs cannot be excluded in our present study.

Understanding how the immune system maintains the balance between effector response and memory formation could provide insights on why some infections result in robust and long-lasting T cell memory, whereas others fail to do so. Our study helps to unravel the complexity of these processes by finding a transcriptional signature at the level of T cell clones that correlates with the acquisition of long-lived, circulating memory T cells. We find that a strongly proliferative phenotype is associated with clonal contraction and disappearance. Furthermore mTOR signalling, probably stimulated by TCR engagement, appears to instruct the fate of short-lived effector cells, similar to previous results in mice[32] and in vitro studies[33]. Conversely, cytokine signalling marks cells destined to become long-lived, circulating memory cells, in agreement with previous studies showing the importance of type I IFN for memory generation[34]. As we sampled SARS-CoV-2-specific T cells from the peripheral blood, we can only infer persistence of $CD8^+$ T cell clones in this compartment. Importantly, memory phenotypes and the factors instructing their differentiation might vary in different immunological contexts, such as the lymph node or tissue. Similarly, as we could examine specific cells only from donors with *HLA-A*01:01* and *HLA-A*11:01* alleles and, to some extent, donors with an *HLA-A*24:02* allele, further studies will be needed to compare our findings in other HLA types.

Collectively, our data demonstrate the formation of memory $CD8^+$ T cells to be dependent on a delicate balance between cytokine and TCR signalling during acute infection, which in turn influences outcomes of long-lived, circulating memory T cells in humans.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-04280-x.

1. Crotty, S. & Ahmed, R. Immunological memory in humans. *Semin. Immunol.* **16**, 197–203 (2004).
2. Sallusto, F., Lanzavecchia, A., Araki, K. & Ahmed, R. From vaccines to memory and back. *Immunity* **33**, 451–463 (2010).
3. Saad-roy, C. M. et al. Immune life history, vaccination, and the dynamics of SARS-CoV-2 over the next 5 years. *Science* **818**, 811–818 (2020).
4. Jalkanen, P. et al. COVID-19 mRNA vaccine induced antibody responses against three SARS-CoV-2 variants. *Nat. Commun.* **12**, 3991 (2021).
5. Chemaitelly, H. et al. mRNA-1273 COVID-19 vaccine effectiveness against the B.1.1.7 and B.1.351 variants and severe COVID-19 disease in Qatar. *Nat. Med.* **27**, 1614–1621 (2021).
6. Raeber, M. E., Zurbuchen, Y., Impellizzieri, D. & Boyman, O. The role of cytokines in T-cell memory in health and disease. *Immunol. Rev.* **283**, 176–193 (2018).
7. Weiskopf, D. et al. Phenotype and kinetics of SARS-CoV-2-specific T cells in COVID-19 patients with acute respiratory distress syndrome. *Sci. Immunol.* **5**, eabd2071 (2020).
8. Grifoni, A. et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* **181**, 1489-1501.e15 (2020).
9. Braun, J. et al. SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. *Nature* **587**, 270–274 (2020).
10. Le Bert, N. et al. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* **584**, 457–462 (2020).
11. Le Bert, N. et al. Highly functional virus-specific cellular immune response in asymptomatic SARS-CoV-2 infection. *J. Exp. Med.* **218**, e20202617 (2021).
12. Saini, S. K. et al. SARS-CoV-2 genome-wide T cell epitope mapping reveals immunodominance and substantial $CD8^+$ T cell activation in COVID-19 patients. *Sci. Immunol.* **6**, eabf7550 (2021).
13. Dan, J. M. et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* **371**, eabf4063 (2021).
14. Zuo, J. et al. Robust SARS-CoV-2-specific T cell immunity is maintained at 6 months following primary infection. *Nat. Immunol.* **22**, 620–626 (2021).
15. Bonifacius, A. et al. COVID-19 immune signatures reveal stable antiviral T cell function despite declining humoral responses. *Immunity* **54**, 340–354 (2021).
16. Hou, H. et al. Immunologic memory to SARS-CoV-2 in convalescent COVID-19 patients at one-year post-infection. *J. Allergy Clin. Immunol.* **148**, 1481–1492.e2 (2021).
17. Minervina, A. A. et al. Longitudinal high-throughput tcr repertoire profiling reveals the dynamics of T-cell memory formation after mild COVID-19 infection. *eLife* **10**, e63502 (2021).
18. Akondy, R. S. et al. The yellow fever virus vaccine induces a broad and polyfunctional human memory $CD8^+$ T cell response. *J. Immunol.* **183**, 7919–7930 (2009).
19. Akondy, R. S. et al. Origin and differentiation of human memory CD8 T cells after vaccination. *Nature* **552**, 362–367 (2017).
20. Mold, J. E. et al. Divergent clonal differentiation trajectories establish $CD8^+$ memory T cell heterogeneity during acute viral infections in humans. *Cell Rep.* **35**, 109174 (2021).
21. Graham, N. et al. Rapid induction and maintenance of virus-specific $CD8^+$ T_{EMRA} and $CD4^+$ T_{EM} cells following protective vaccination against dengue virus challenge in humans. *Front. Immunol.* **11**, 479 (2020).

22. Peng, Y. et al. Broad and strong memory CD4$^+$ and CD8$^+$ T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. *Nat. Immunol.* **21**, 1336–1345 (2020).
23. Surh, C. D. & Sprent, J. Homeostasis of naive and memory T cells. *Immunity* **29**, 848–862 (2008).
24. Escobar, G., Mangani, D. & Anderson, A. C. T cell factor 1: a master regulator of the T cell response in disease. *Sci. Immunol.* **5**, eabb9726 (2020).
25. Intlekofer, A. M. et al. Effector and memory CD8$^+$ T cell fate coupled by T-bet and eomesodermin. *Nat. Immunol.* **6**, 1236–1244 (2005).
26. Khan, O. et al. TOX transcriptionally and epigenetically programs CD8$^+$ T cell exhaustion. *Nature* **571**, 211–218 (2019).
27. Wieland, D. et al. TCF1$^+$ hepatitis C virus-specific CD8$^+$ T cells are maintained after cessation of chronic antigen stimulation. *Nat. Commun.* **8**, 15050 (2017).
28. Jung, J. H. et al. SARS-CoV-2-specific T cell memory is sustained in COVID-19 convalescent patients for 10 months with successful development of stem cell-like memory T cells. *Nat. Commun.* **12**, 4043 (2021).
29. Geginat, J., Lanzavecchia, A. & Sallusto, F. Proliferation and differentiation potential of human CD8$^+$ memory T-cell subsets in response to antigen or homeostatic cytokines. *Blood* **101**, 4260–4266 (2003).
30. Buggert, M. et al. The identity of human tissue-emigrant CD8$^+$ T cells. *Cell* **183**, 1946-1961. e15 (2020).
31. Sallusto, F., Lenig, D., Forster, R., Lipp, M. & Lanzavecchia, A. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature* **401**, 708–712 (1999).
32. Araki, K. et al. mTOR regulates memory CD8 T-cell differentiation. *Nature* **460**, 108–112 (2009).
33. Langenkamp, A. et al. T cell priming by dendritic cells: thresholds for proliferation, differentiation and death and intraclonal functional diversification. *Eur. J. Immunol.* **32**, 2046–2054 (2002).
34. Kolumam, G. A., Thomas, S., Thompson, L. J., Sprent, J. & Murali-Krishna, K. Type I interferons act directly on CD8 T cells to allow clonal expansion and memory formation in response to viral infection. *J. Exp. Med.* **202**, 637–650 (2005).

Methods

Human participants and patient characteristics

Following written informed consent, adult patients with symptomatic, RT–qPCR-confirmed SARS-CoV-2 infection were recruited in the Canton of Zurich, Switzerland, between 2 April and 24 September 2020. The study was approved by the Cantonal Ethics Committee of Zurich (BASEC 2016-01440). Patients (n = 175) donated peripheral blood at the time of inclusion into the study, and 116 and 90 patients donated peripheral blood approximately 6 months and 1 year after infection, respectively. Standardized clinical data were collected for all included patients and disease severity was assessed, as previously described for this cohort[35–37]. Peripheral blood mononuclear cells (PBMCs) and serum were bio-banked, as previously described[35–37]. Following HLA class I typing, patients carrying an *HLA-A*01:01*, *HLA-A*11:01* and/or an *HLA-A*24:02* allele with sufficient bio-banked samples at two different timepoints were selected for the study (n = 47). Thirteen healthy donors carrying *HLA-A*01:01*, *HLA-A*11:01* and/or an *HLA-A*24:02* allele were included for comparison. SARS-CoV-2-specific $CD8^+$ T cells were detected with MHC-I dextramers and pentamers in 42 and 12 patients, respectively.

IgA and IgG immunoassays

Spike S1-specific IgA and IgG antibodies were assessed with a commercial ELISA kit (SARS-CoV-2 IgA and IgG immunoassay, Euroimmun), as previously described[35]. OD ratios higher than 2.0 and 1.1 were considered positive for serum IgA and IgG, respectively.

Dextramer and pentamer staining

PBMCs (4×10^6) per patient were incubated with Human TruStain FcX blocking reagent (422302, BioLegend) for 10 min at 4 °C. After washing, cells were incubated with MHC-I dextramers (see Supplementary Table 1) in the presence of L-biotin and herring sperm DNA according to the manufacturer's instructions, for 10 min at room temperature. Two peptides presented on HLA-A*01:01 dextramers (FTSDYYQLY from ORF3a and TTDPSFLGRY from ORF1ab), two peptides presented on HLA-A*11:01 dextramers (ATEGALNTPK and KTFPPTEP from nucleocapsid protein) and one peptide presented on HLA-A*24:02 dextramers (QYIKWPWYI from spike protein) were included. For MHC-I pentamer staining, cells were incubated for 10 min at 37 °C with pentamers (Extended Data Table 1). One peptide presented on HLA-A*01:01 pentamers (FTSDYYQLY from ORF3a) and one peptide presented on HLA-A*11:01 pentamers (KTFPPTEP from nucleocapsid protein) were included. Frozen PBMCs were used throughout the study.

Spectral flow cytometry

After dextramer or pentamer staining, a concentrated surface staining antibody mix (Supplementary Tables 2, 3) was applied without washing and samples were incubated at room temperature for further 20 min. After four rounds of washing, cells were resuspended in a fixation permeabilization solution (eBioscience Foxp3/transcription factor staining buffer) and incubated for 60 min at room temperature. After washing, an antibody mix for intracellular staining (Supplementary Tables 2, 3) was added and cells were incubated for 30 min at room temperature. After washing, samples were acquired on a Cytek Aurora spectral flow cytometer using the SpectroFlo software. Quality control for the cytometer was performed daily. Data were analysed with FlowJo (version 10.7.1) and OMIQ (www.omiq.ai). Phenotypes were evaluated only in patients with more than 5 CoV2-Dex^+ cells per sample (n = 28 acute, n = 24 6 months after infection, n = 29 1 year after infection) or more than 5 CoV2-$Pent^+$ cells per sample (n = 7 acute, n = 9 6 months after infection). Correlation between frequency of CoV2-Dex^+ cells during acute infection and frequency of CoV2-Dex^+ cells 6 months after infection was assessed only for patients sampled at least 14 days after symptom onset (n = 11).

Sample sets of sorted and unsorted cells and healthy controls

A concentrated antibody mix containing TotalSeq antibodies (see Supplementary Table 4 for a complete list) was applied after dextramer staining without washing and cells were incubated at 4 °C for 30 min. After four rounds of washing, cells were resuspended in PBS with 2% FBS and 2 mM EDTA and sorted with a BD Aria cell sorter. For each patient, CoV2-Dex^- and CoV2-Dex^+ cells were sorted approximately in a 10:1 ratio. All CoV2-Dex^+ cells from each sample were sorted, the corresponding amount of CoV2-Dex^- cells was calculated and sorted in the same tube. Cells from ten patients at the same timepoint were pooled together, generating four individual sample sets in total: (1) patients CoV2_001–CoV2_010, acute; (2) patients CoV2_001–CoV2_010, 6 months after infection; (3) patients CoV2_011–CoV2_020, acute; and (4) patients CoV2_011–CoV2_20, 6 months after infection. Two additional sample sets were generated using 5,000 unsorted PBMCs from each patient's sample: (5) patients CoV2_001–CoV2_010, 6 months after infection unsorted; and (6) patients CoV2_011–CoV2_020, 6 months after infection unsorted. Finally, using PBMCs from four healthy donors, we generated sample set (7) by sorting and pooling 2,000 $CD8^+$ T cells from each healthy donor sample.

scRNA-seq library preparation and sequencing

Cells of sample sets 1–7 were analysed by scRNA-seq utilizing the 5′ Single Cell GEX and VDJ v1.1 platforms (10x Genomics). Each sample set was processed individually. Cell suspensions were pelleted, resuspended and loaded into the Chromium Chip following the manufacturer's instructions. Fourteen cycles of initial cDNA amplification were used for all sets and single-cell sequencing libraries for whole-transcriptome analysis (GEX), TCR profiling (VDJ), and combined cell-surface protein and dCODE Dextramers detection (ADT) were generated. Final libraries were quantified using a Qubit Fluorometer, pooled in a ratio of 5:1:1 (GEX:VDJ:ADT) and sequenced on a NovaSeq 6000 system with the following cycle configuration: read 1: 28 bp; index read 1: 10 bp; read 2: 101 bp.

Single-cell transcriptome analysis

Raw scRNA-seq FASTQ files were aligned to the human GRCh38 genome with Cell Ranger version 5.0.0 with default settings for the 'cellranger multi' pipeline (10x Genomics). The reference genome was downloaded from the 10x Genomics website (https://cf.10xgenomics.com/supp/cell-exp/refdata-gex-GRCh38-2020-A.tar.gz) and built as per official release notes (https://support.10xgenomics.com/single-cell-gene-expression/software/release-notes/build#GRCh38_2020A). Every sample set was analysed with the 'cellranger multi' pipeline, which allows to process together the paired GEX, ADT and VDJ libraries for each set. Downstream analysis was conducted in R version 4.1.0 with the package Seurat version 4.0.3 (ref. [38]). Cells with fewer than 200 or more than 2,500 detected genes and cells with more than 10% detected mitochondrial genes were excluded from the analysis.

To investigate possible patient biases, we demultiplexed cells from patient pools 1–6 based on genetic variants detected within the scRNA-seq reads. For this, we used the tool souporcell version 2 (ref. [39]). To cluster cells based on their patient-specific genetic variants, we merged sample sets 1, 2 and 5 (comprising sorted cells from both timepoints of patients CoV2_001–CoV2_010 and unsorted cells of the same patients) and sets 3, 4 and 6 (comprising cells from both timepoints of patients CoV2_011–CoV2_020 and unsorted cells of the same patients). Then, we executed the souporcell pipeline with option k = 10 (the number of clusters to be determined) for each of the two merged sample sets. This analysis allowed us to classify 88% of the cells passing the filtering steps from above into 20 genotype-driven 'patient' clusters.

After log normalization and variable feature calculation, independent datasets were integrated using Seurat's anchoring-based integration method. Data scaling, principal component analysis, clustering

and UMAP visualizations were performed on the integrated dataset using 15 principal components and a resolution of 0.5 for the shared nearest-neighbour clustering algorithm. To define distinct biological features of cell clusters, differential gene expression analyses were done on assay 'RNA' of the integrated dataset. FindAllMarkers was executed with logfc.threshold and min.pct cut-offs set to 0.25. For the analysis of clusters, FindMarkers was used with default settings for the comparison of persistent and non-persistent clones. For the differential expression analysis of manually selected genes and cell-surface proteins (CD45RA and CCR7), logfc.threshold and min.pct cut-offs were set to 0.

For gene set enrichment analysis, the FindMarkers function from Seurat was first used for the differential expression of genes between cells belonging to persistent and non-persistent clones (using the default Wilcoxon rank-sum test, with options 'min.pct=0.1, logfc.threshold = -Inf', to account also for small expression changes, as long as the genes were expressed in at least 10% of cells of at least one group). The resulting 4,701 genes were pre-ranked in decreasing order by the negative logarithm of their *P* value, multiplied for the sign of their average log-fold change (in R, '-log(*P*_val)*sign(avg_log$_2$FC)'). Gene set enrichment analysis[40] was performed on this pre-ranked list using the R package FGSEA (https://github.com/ctlab/fgsea/)[41]. We used the FGSEA-simple procedure with 100,000 permutations and the hallmark gene sets for *Homo sapiens* from the Molecular Signatures Database (https://www.gsea-msigdb.org/gsea/msigdb/index.jsp, made accessible in R by the package msigdbr; https://github.com/cran/msigdbr) and set the seed value ('set.seed(42)' in R) before execution to make the results reproducible. For significance testing, the function fgsea::fgsea() was used, which performs a *P* value estimation based on an adaptive multi-level split Monte-Carlo scheme. A multiple hypothesis correction procedure was applied to get adjusted *P* values. The results were filtered for gene sets that were significantly enriched with adjusted $P < 0.1$.

TCR profiling

Paired chain TCR sequences were obtained through targeted amplification of full-length V(D)J segments during library preparation. Sequence assembly and clonotype calling was done through cellranger's immune profiling pipeline (cellranger multi). TCR profiling on filtered contig annotations was done using R package scRepertoire version 1.1.4 (ref. [42]), which assigns TCR nucleotide and amino acid sequences together with clonal frequency counts and a clonotype classification to each cell. The function combineTCR was executed with filterMulti=T to isolate the top two expressed chains in cell barcodes with multiple chains. Clonotypes were called based on the amino acid sequence of the CDR3 region of TCRα and TCRβ chains. For cells of which only one of the two chains could be identified, the available chain was used. Clone calling was done for each sample set independently before integration.

SARS-CoV-2 peptide-loaded dextramer binding of CD8$^+$ T cells

To identify SARS-CoV-2-specific CD8$^+$ T cells, we used dCODE Dextramers loaded with viral peptides presented on MHC-I molecules as described above. To assess unspecific binding, a negative control dextramer (peptide STEGGGLAY presented on HLA-A*01:01) and a general negative control dextramer were included. After analysis of the flow cytometry data, we noticed strong background staining of dextramer HLA-A*24:02 (peptide QYIKWPWYI) in samples of healthy donors, indicating unspecific dextramer binding. Thus, we excluded all sequencing counts from this dextramer in the downstream analysis. For other dextramers, cells were considered CoV2-Dex$^+$ when the unique molecular identifier (UMI) count of a CoV2-dextramer was more than ten and more than five times higher than the UMI count of the negative control in the same cell. Cells that were positive for more than one dextramer according to this classification (less than 0.2% of all cells with known TCR) were excluded from the analysis. A TCR clone was considered SARS-CoV-2-specific when at least one cell of the clone was CoV2-Dex$^+$.

Statistics

Wilcoxon–Mann–Whitney test was used for comparisons of two independent groups. Wilcoxon signed-rank test was used for paired testing. *P* values were adjusted for multiple comparisons with the Holm method. A linear regression model was used to quantify the relationship between variables. Significance was assessed by non-parametric methods unless otherwise specified. All tests were performed two sided. Analyses were performed with R (version 4.0.0 or 4.1.0).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The sequencing dataset generated in this study has been deposited at zenodo.org and is available at https://zenodo.org/record/5770747. Flow cytometry datasets are available from the corresponding author on reasonable request.

Code availability

The code generated during the current study is available at https://github.com/Moors-Code/SARS-CoV-2-Tcell-Boyman-collaboration.

35. Cervia, C. et al. Systemic and mucosal antibody responses specific to SARS-CoV-2 during mild versus severe COVID-19. *J. Allergy Clin. Immunol.* **147**, 545–557 (2021).
36. Chevrier, S. et al. A distinct innate immune signature marks progression from mild to severe COVID-19. *Cell Rep. Med.* **2**, 100166 (2021).
37. Adamo, S. et al. Profound dysregulation of T cell homeostasis and function in patients with severe COVID-19. *Allergy* **76**, 2866–2881 (2021).
38. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).
39. Heaton, H. et al. Souporcell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. *Nat. Methods* **17**, 615–620 (2020).
40. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
41. Korotkevich, G. et al. Fast gene set enrichment analysis. Preprint at https://doi.org/10.1101/060012 (2021).
42. Borcherding, N., Bormann, N. L. & Kraus, G. scRepertoire: an R-based toolkit for single-cell immune receptor analysis. *F1000Res.* **9**, 47 (2020).

Acknowledgements We thank S. Hasler for her assistance with patient recruitment and coordination; E. Baechli, A. Rudiger, M. Stüssi-Helbling, L. C. Huber and D. J. Schaer for their support in patient recruitment; and the members of the Boyman laboratory for helpful discussions. Graphical representations were generated with BioRender.com. This work was funded by the Swiss National Science Foundation (4078P0-198431 to O.B. and J.N.; and 310030-172978 and 310030-200669 to O.B.), the Clinical Research Priority Program of the University of Zurich for CRPP CYTIMM-Z (to O.B.), an Innovation grant of University Hospital Zurich (to O.B.), the Pandemic Fund of University of Zurich (to O.B.), and the Botnar Research Centre for Child Health (COVID-19 FTC to A.E.M.). S.A., C.C. and Y.Z. received Swiss Academy of Medical Sciences fellowships (323530-177975, 323530-191220 and 323530-191230, respectively), S.A. received a Forschungskredit Candoc grant from the University of Zurich (FK-20-022) and M.E.R. received a Young Talents in Clinical Research Project Grant by the Swiss Academy of Medical Sciences and G. & J. Bangerter-Rhyner Foundation (YTCR 08/20).

Author contributions S.A. designed and performed flow cytometry and scRNA-seq experiments, analysed and interpreted data. J.M. performed scRNA-seq experiments, analysed and interpreted data. Y.Z. contributed to flow cytometry and scRNA-seq experiments, patient recruitment and data collection. C.C. and P.T. contributed to patient recruitment and data collection. M.E.R. contributed to patient recruitment and clinical management. S.B.S. analysed scRNA-seq data. J.N. contributed to study design and patient recruitment. A.E.M. designed experiments and interpreted data. O.B. conceived the project, designed experiments and interpreted data. S.A. and O.B. wrote the manuscript with contribution by J.M. and S.B.S. All authors edited and approved the final draft of the article.

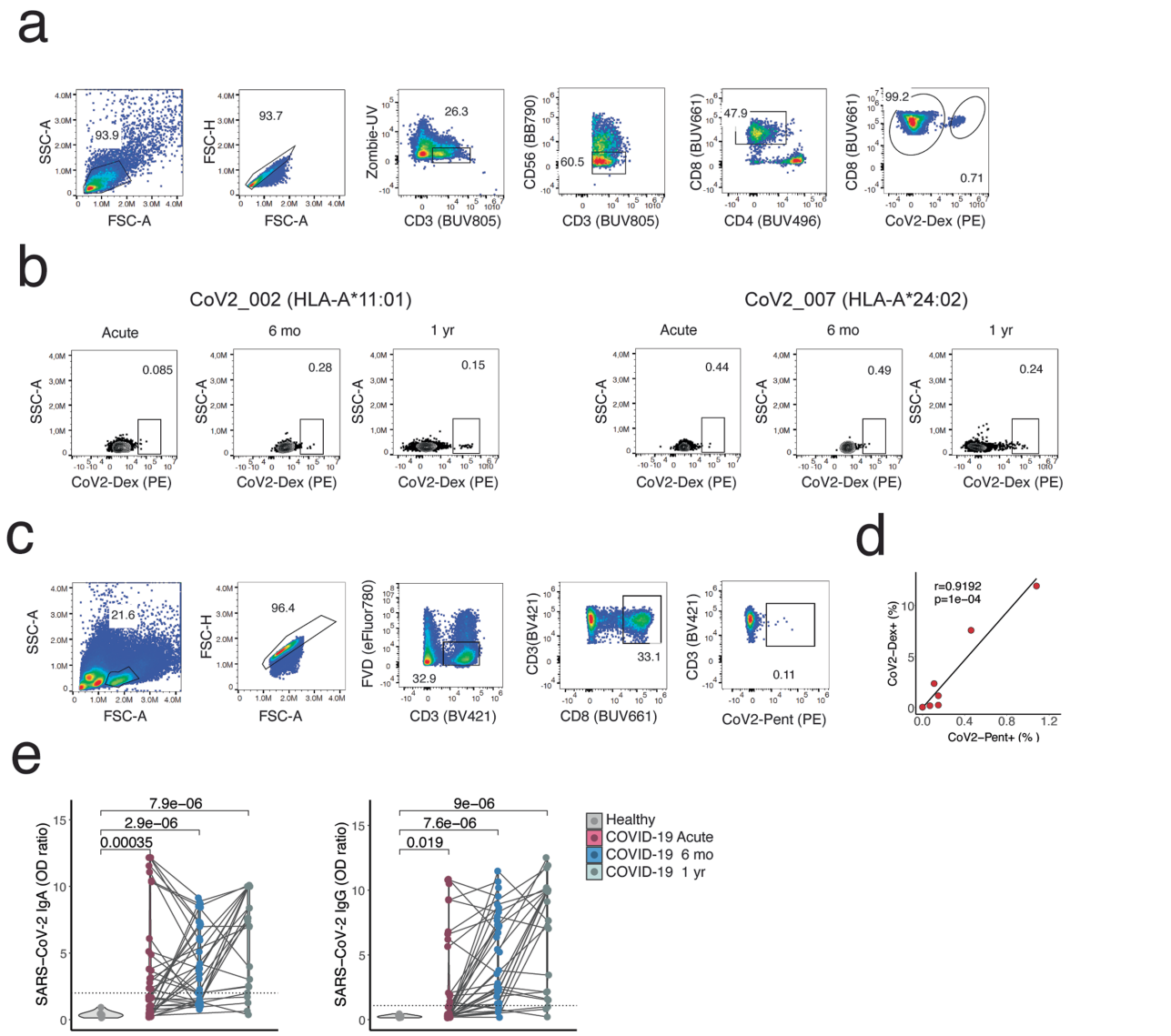
Competing interests The authors declare no competing interests.

Additional information

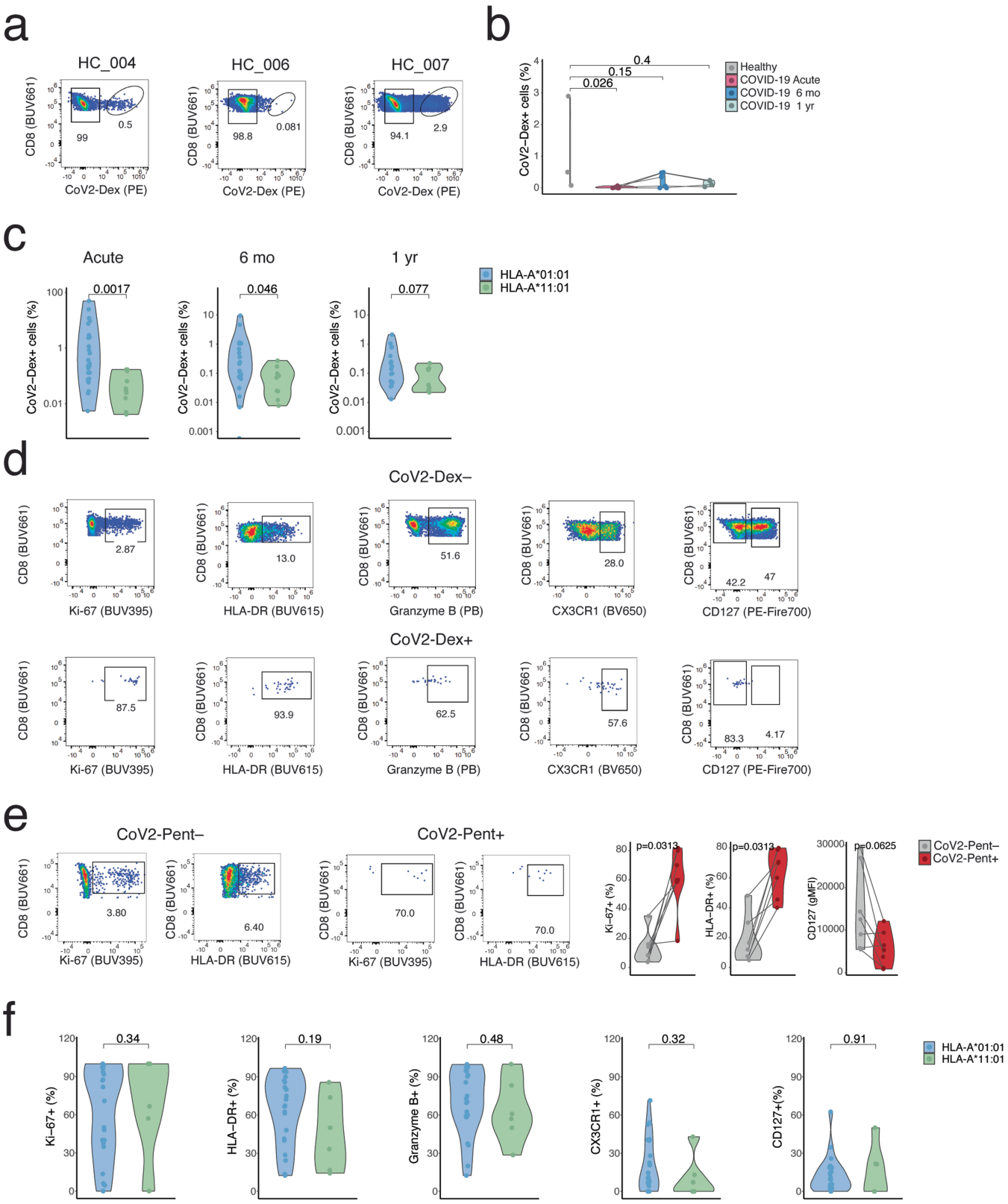
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-04280-x.

Correspondence and requests for materials should be addressed to Onur Boyman.

Peer review information *Nature* thanks Katherine Kedzierska and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

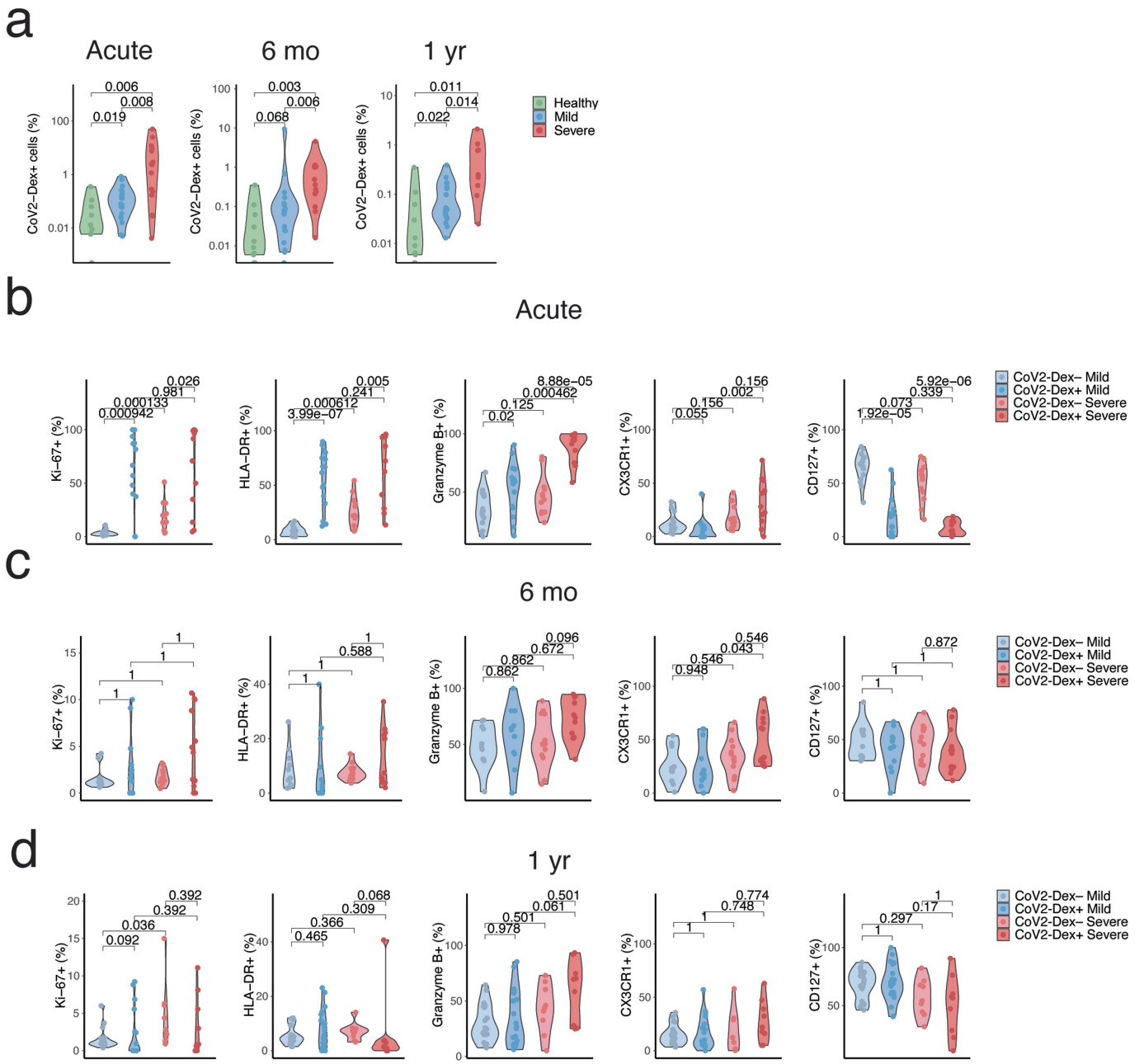


Extended Data Fig. 1 | Gating strategy for antigen-specific CD8$^+$ T cells and SARS-CoV2-specific antibodies in healthy donors and COVID-19 patients. **a**, Gating strategy for CoV2-Dex$^+$ cells. **b**, Representative plot of CoV2-Dex staining for HLA-A*11:01 and HLA-A*24:02. **c**, Gating strategy for CoV2-Pent$^+$ cells. **d**, Linear regression of frequency of CoV2-Dex$^+$ cells as a function of frequency of CoV2-Pent$^+$ cells ($n = 7$). The p-value was calculated with t-statistic. **e**, SARS-CoV2-specific IgA (left) and IgG (right) of healthy donors compared to COVID-19 patients during acute infection, six months and one year after infection ($n = 13$ healthy , $n = 46$ acute, $n = 41$ six months, $n = 30$ one year). Dashed lines represent diagnostic cut-off values, i.e., 2.0 and 1.1 for IgA and IgG, respectively. P-values were calculated with a Mann-Whitney-Wilcoxon test. All tests were performed two-sided.

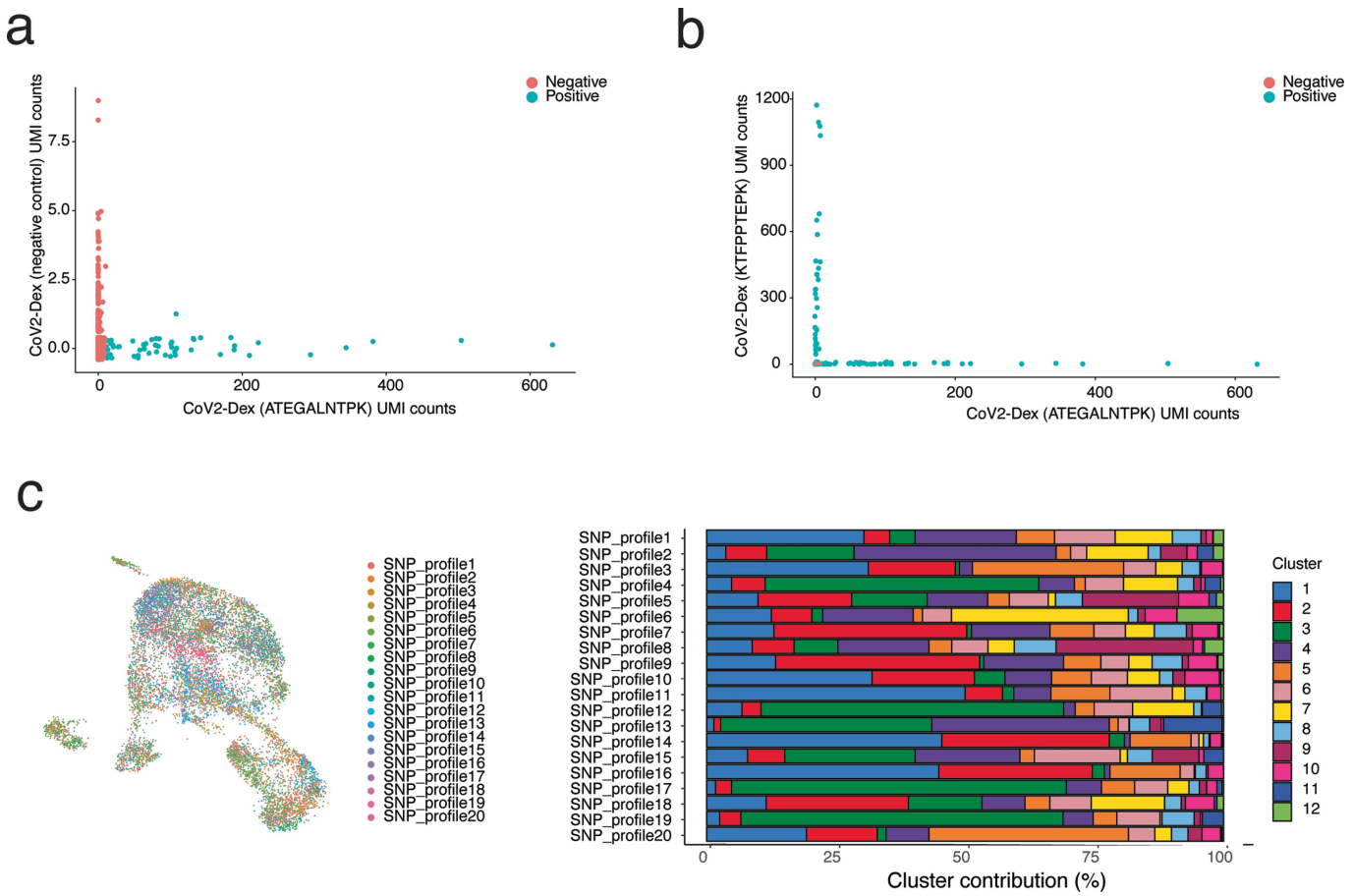


Extended Data Fig. 2 | See next page for caption.

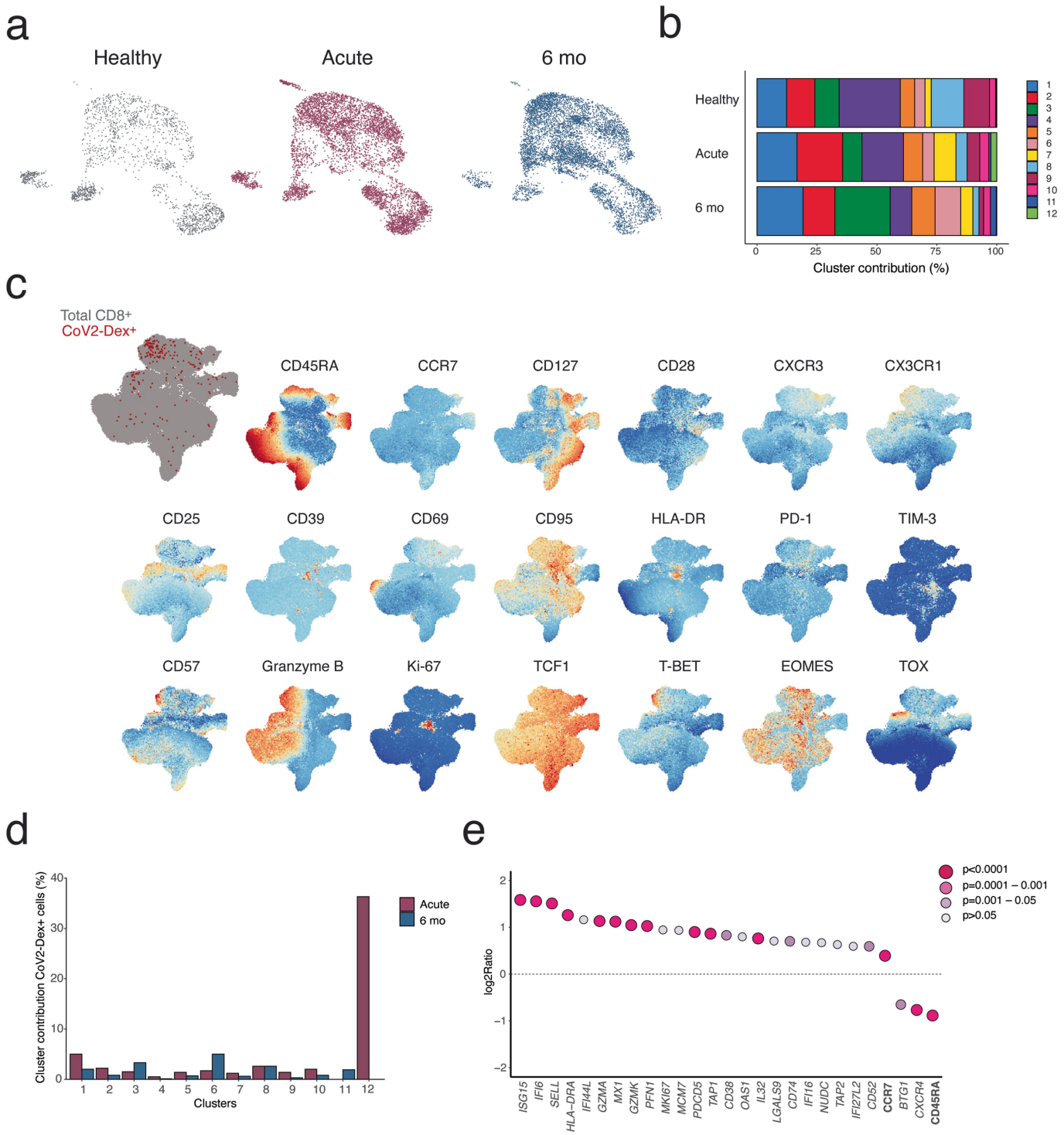
Extended Data Fig. 2 | HLA-A*24:02 dextramer staining and CoV2-Dex$^+$ cell frequency and phenotype based on HLA. a, Representative plots of CoV2-Dex (HLA-A*24:02) staining in healthy donors. **b**, Percentage of CoV2-Dex$^+$ cells in HLA-A*24:02 healthy donors and COVID-19 patients in acute infection and six months and one year after infection. Dots represent independent donors at each timepoint (n = 3 healthy, n = 6 acute, n = 5 six months, n = 3 one year). Gray lines connect individual donors sampled at different timepoints. **c**, Percentage of CoV2-Dex$^+$ cells in HLA-A*01:01 and HLA-A*11:01 in acute infection and six months and one year after infection (HLA-A*01:01 n = 26 and HLA-A*11:01 n = 10 acute, HLA-A*01:01 n = 21 and HLA-A*11:01 n = 9 six months, HLA-A*01:01 n = 20 and HLA-A*11:01 n = 9 one year). **d**, Gating strategy for phenotypical analysis of CoV2-Dex$^+$ compared to CoV2-Dex$^-$ cells. **e**, Gating strategy for phenotypical analysis of CoV2-Pent$^+$ compared to CoV2-Pent$^-$ cells and frequency of Ki-67$^+$, HLA-DR$^+$ and CD127 levels in CoV2-Pent$^-$ (gray) and CoV2-Pent$^+$ (red) cells in acute COVID-19 (n = 7). **f**, Frequency of Ki-67$^+$, HLA-DR$^+$, granzyme B$^+$, CX3CR1$^+$ and CD127$^+$ cells in CoV2-Dex$^+$ cells in patients with an HLA-A*01:01 versus HLA-A*11:01 allele (HLA-A*01:01 n = 22 and HLA-A*11:01 n = 6) in the acute phase. **b**, **c**, and **f**, P-values were calculated with a Mann-Whitney-Wilcoxon test. **e**, P-values were calculated with a Wilcoxon signed-rank test. All tests were performed two-sided.



Extended Data Fig. 3 | Frequency and phenotype of antigen-specific CD8$^+$ T cells relative to disease severity. a, Frequency of CoV2-Dex$^+$ CD8$^+$ T cells in healthy controls and patients with mild and severe disease during acute COVID-19 six months and one year after infection ($n = 10$ healthy; acute: $n = 21$ mild, $n = 15$ severe; six months: $n = 18$ mild, $n = 16$ severe; one year: $n = 19$ mild, $n = 10$ severe). **b**–**d**, Frequency of Ki-67$^+$, HLA-DR$^+$, granzyme B$^+$, CX3CR1$^+$ and CD127$^+$ in CoV-2Dex$^-$ and CoV2-Dex$^+$ cells in patients with mild versus severe disease (**b**) during acute infection, (**c**) six months and (**d**) one year after infection (acute: $n = 16$ mild, $n = 12$ severe; six months: $n = 11$ mild, $n = 16$ severe; one year: $n = 19$ mild, $n = 10$ severe). P-values were calculated with a Mann-Whitney-Wilcoxon test and adjusted for multiple comparisons using the Holm method. All tests were performed two-sided.

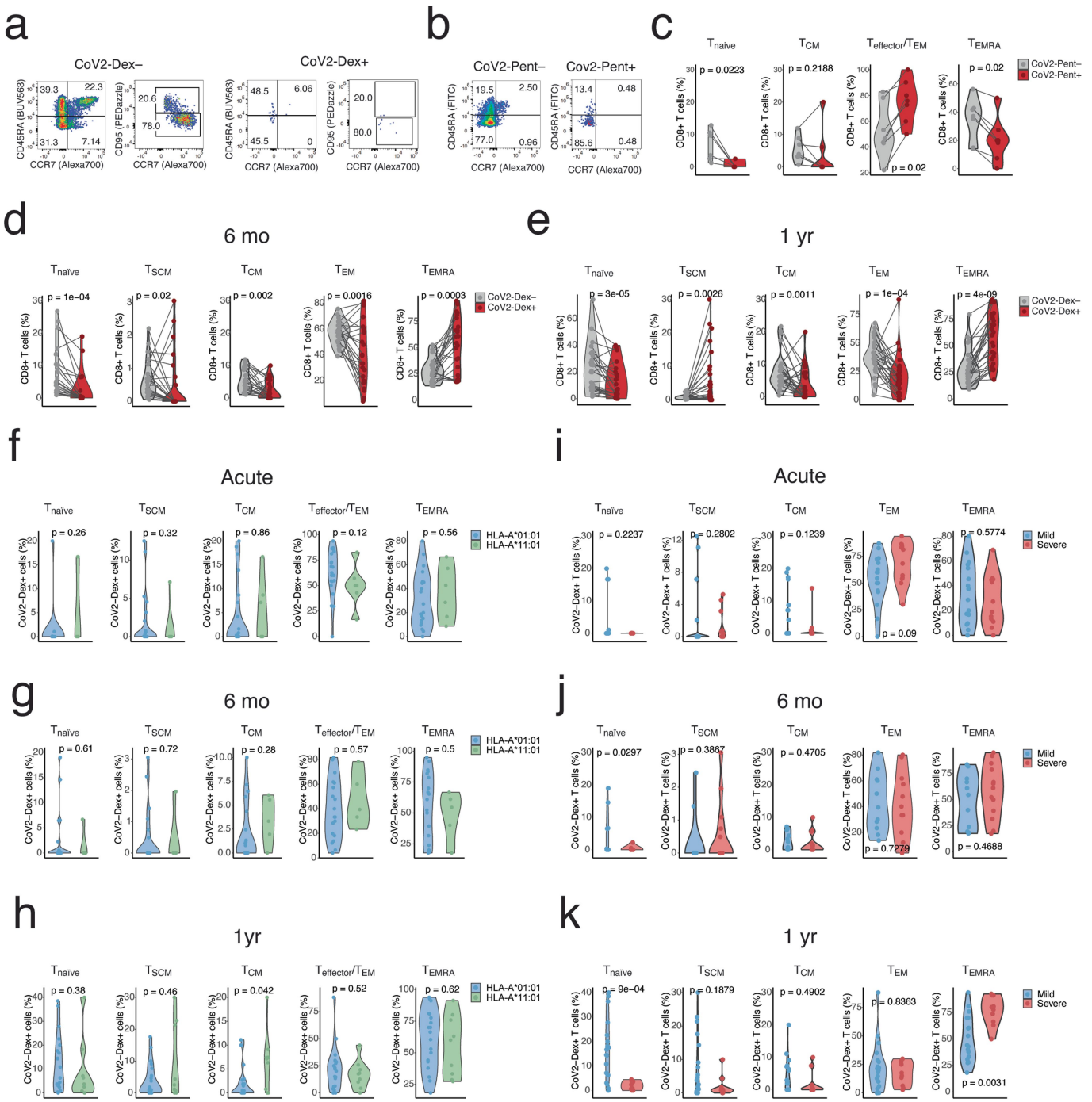


Extended Data Fig. 4 | Definition of CoV2-Dex$^+$ cells and single patient contribution to individual clusters. a, Unique molecular identifier (UMI) counts for CoV2-Dex HLA-A*11:01 (ATEGALNTPK) versus UMI counts for negative control dextramer; cells defined as CoV2-Dex HLA-A*11:01 (ATEGALNTPK)$^+$ are depicted in blue. **b**, (UMI) counts for CoV2-Dex HLA-A*11:01 (ATEGALNTPK) versus UMI counts for CoV2-Dex HLA-A*11:01 (KTFPPTEPK). **c**, Uniform manifold approximation and projection (UMAP) plot colored by patient ID (left) and cluster distribution for single patients (right).

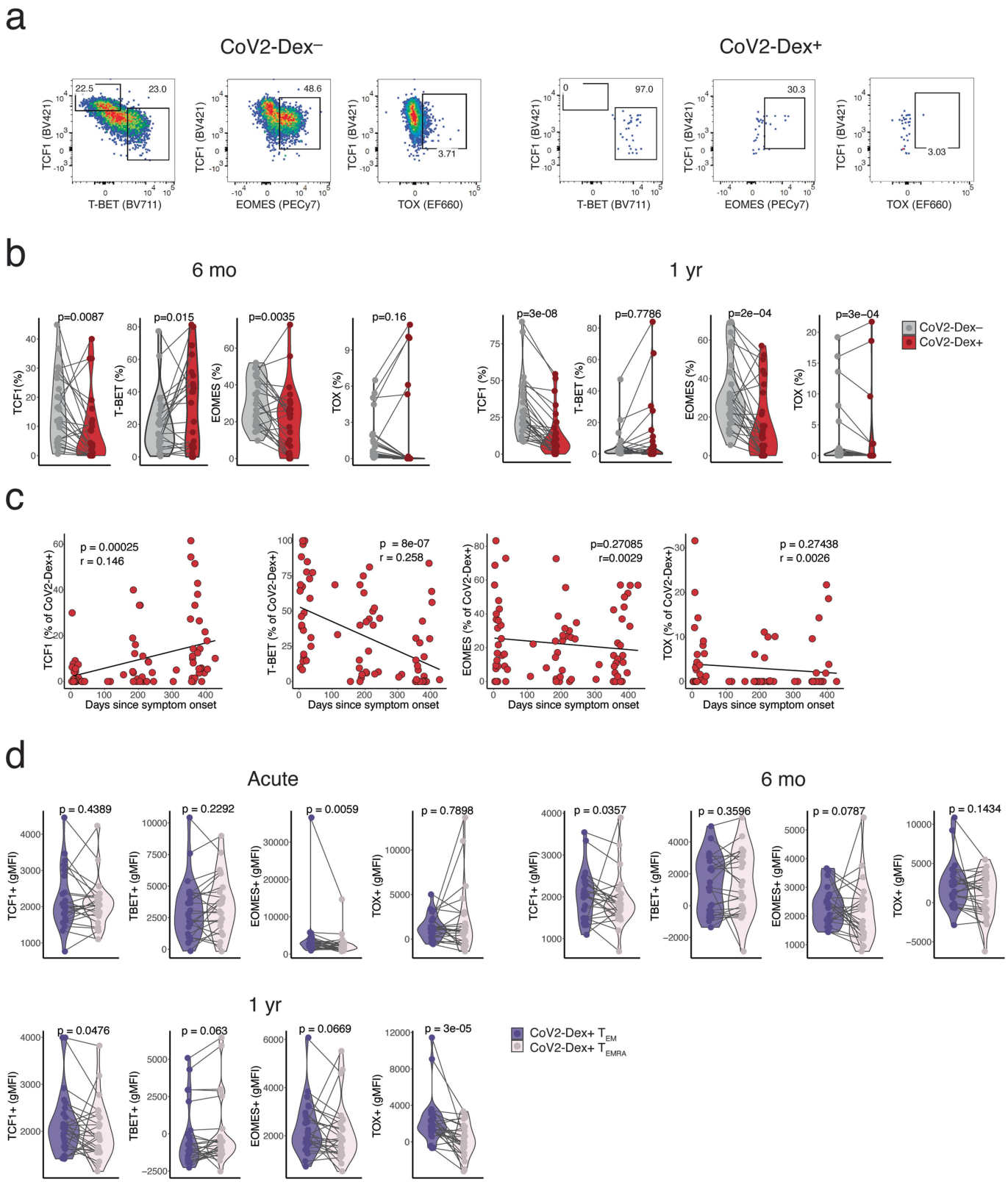


Extended Data Fig. 5 | Characteristics of antigen-specific CD8⁺ T cells six months after infection. a, UMAP plot of transcriptomes from CD8⁺ T cells of healthy individuals (gray), COVID-19 patients in acute infection (red), and six months after infection (blue). **b**, Cluster contribution to total CD8⁺ T cells of healthy subjects, acute infection, and six months after infection. **c**, UMAP plots of marker expression for up to 2,000 CD8⁺ T cells from each sample collected six months after infection analyzed by spectral flow cytometry. Regions with high expression of specific markers appear red. Overlay of CoV2-Dex⁺ cells (red) and total CD8⁺ T cells (gray) is shown on the upper left. **d**, Percentages of CoV2-Dex⁺ cells contributing to indicated clusters during acute infection or six months after infection. Percentages are calculated on total cells per cluster per timepoint. **e**, Expression of selected genes (italicized), as well as CCR7 and CD45RA proteins determined by Totalseq™, for CoV2-Dex⁺ cells detected in the acute versus recovery phase. P-values were calculated using a Mann-Whitney-Wilcoxon test and a Bonferroni correction was applied for multiple comparisons.

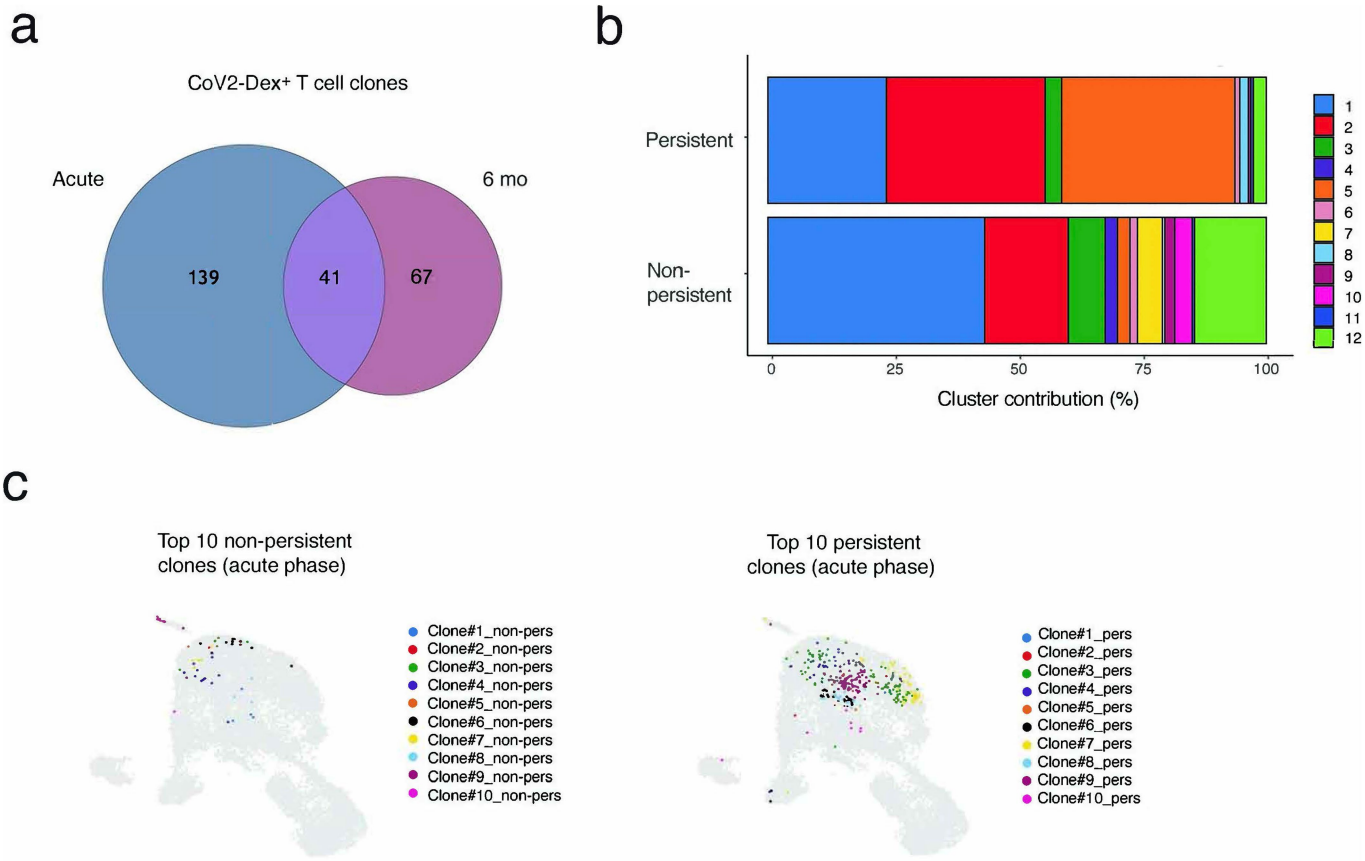
Extended Data Fig. 6 | Transcriptome of SARS-CoV-2-specific CD8$^+$ T cell clones. UMAP plots highlighting single-cell transcriptomes belonging to individual CoV2-Dex$^+$ T cell clones in acute infection (red) and six months after infection (blue).



Extended Data Fig. 7 | Memory phenotypes in CoV2-Dex$^+$ and CoV2-Pent$^+$ cells. a, **b**, Gating strategy for identifying naïve, stem cell memory (T_{SCM}), central memory (T_{CM}), effector/effector memory ($T_{effector}/T_{EM}$), and effector memory T cells re-expressing CD45RA (T_{EMRA}) in (a) CoV2-Dex$^-$ and CoV2-Dex$^+$ cells or (b) CoV2-Pent$^-$ and CoV2-Pent$^+$ cells. **c**, Percentages of naïve, T_{CM}, $T_{effector}/T_{EM}$, and T_{EMRA} cells among CoV2-Pent$^-$ (gray) and CoV2-Pent$^+$ (red) cells ($n = 7$). **d**, **e**, Percentages of naïve, T_{CM}, $T_{effector}/T_{EM}$, and T_{EMRA} cells among CoV2-Dex$^-$ (gray) and CoV2-Dex$^+$ (red) cells (**d**) six months and (**e**) one year after infection ($n = 24$ six months, $n = 29$ one year). **f**–**h**, Percentages of naïve, T_{SCM}, T_{CM}, $T_{effector}/T_{EM}$, and T_{EMRA} in CoV2-Dex$^+$ cells in patients with HLA-A*01:01 versus HLA-A*11:01 allele (**f**) during acute infection, (**g**) six months and (**h**) one year after infection (acute: $n = 22$ HLA-A*01:01 and $n = 6$ HLA-A*11:01; six months: $n = 19$ HLA-A*01:01 and $n = 5$ HLA-A*11:01; one year: $n = 20$ HLA-A*01:01 and $n = 9$ HLA-A*11:01). **i**–**k**, Percentages of naïve, T_{SCM}, T_{CM}, $T_{effector}/T_{EM}$, and T_{EMRA} in CoV2-Dex$^+$ cells in patients with mild versus severe disease (**i**) during acute infection, (**j**) six months and (**k**) one year after infection (acute: $n = 16$ mild, $n = 12$ severe; six months: $n = 11$ mild, $n = 16$ severe; one year: $n = 19$ mild, $n = 10$ severe). P-values were calculated with a Wilcoxon signed-rank test in **c**–**e** and with a Mann-Whitney-Wilcoxon test in **f**–**k**. All tests were performed two-sided.



Extended Data Fig. 8 | Gating strategy and pseudolongitudinal time course of transcription factor expression. a, Gating strategy TCF1$^+$, T-BET$^+$, EOMES$^+$, and TOX$^+$ cells in CoV2-Dex$^-$ and CoV2-Dex$^+$ cells. **b**, Percentages of TCF1$^+$, T-BET$^+$, EOMES$^+$, and TOX$^+$ cells in CoV2-Dex$^-$ and CoV2-Dex$^+$ cells (left) six months and (right) one year after infection ($n = 24$ six months, $n = 29$ one year). **c**, Linear regression of frequencies of TCF1$^+$, T-BET$^+$, EOMES$^+$, and TOX$^+$ cells in CoV2-Dex$^+$ cells as a function of time since symptom onset ($n = 81$). P-values were calculated with t-statistic. **d**, Expression of transcription factors in $T_{effector}/T_{EM}$ and T_{EMRA} in CoV2-Dex$^+$ cells in acute infection and six months and one year after infection ($n = 24$ acute, $n = 24$ six months, $n = 26$ one year). P-values were calculated using a Wilcoxon signed-rank test in **b** and **d**. All tests were performed two-sided.



Extended Data Fig. 9 | Clonality and phenotype of persistent versus non-persistent CoV2-Dex$^+$ T cell clones. a, Venn diagram showing overlapping clones during acute infection and at six months after infection. **b**, Cluster composition of persistent versus non-persistent CD8$^+$ T cell clones. **c**, UMAP plots showing top ten persistent CoV2-Dex$^+$ CD8$^+$ T cell clones and top ten non-persistent CoV2-Dex$^+$ CD8$^+$ T cell clones.

Extended Data Table 1 | Clinical and demographic characteristics of healthy subjects and COVID-19 patients

Disease severity[a]	**Healthy (*n* = 13)**	**Mild cases (*n* = 25)**		**Severe cases (*n* = 22)**			
Disease grade[b]		Mild illness	Pneumonia	Severe pneumonia	Mild ARDS	Moderate ARDS	Severe ARDS
Disease grade – no.	–	23	2	9	2	5	6
Demographic characteristics							
Age (median (IQR) [yrs])	29.0 (27.5–37)	29.0 (24.5–38.5)		62.5 (53.5–79.3)**			
Gender (m/f)	5/8	13/12		13/9			
Time since symptom onset (median (IQR) [days]) – acute infection	–	7.0 (4.3–10.8)		15.0 (8–33)*			
Time since symptom onset (median (IQR) [days]) – 6 months	–	197 (185.8–205.0)		213 (170.5–227)			
Time since symptom onset (median (IQR) [days]) – 1 year	–	375 (360.5–382.8)		367 (355–397)			
Maximal level of care							
Outpatient – no. (%)	–	23 (92.0)		–			
Inpatient – no. (%)	–	2 (8.0)		22 (100)*			
– ward – no. (%)	–	2 (8.0)		11 (50.0)			
– ICU – no. (%)	–	–		11 (50.0)			
Outcome							
Released / recovered	–	25 (100)		22 (100)			
Deceased	–	0 (0)		0 (0)			
Systemic immunosuppression – no. (%)	0 (0)	0 (0)		3 (13.6)			
COVID-19 treatment at first sampling							
Hydroxychloroquine – no. (%)	–	0 (0)		9 (40.9)			
Remdesivir – no. (%)	–	1 (4.0)		4 (18.2)			
Glucocorticoids – no. (%)[c]	–	0 (0)		4 (18.2)			
Tocilizumab – no. (%)	–	0 (0)		2 (9.1)			

Mann-Whitney-Wilcoxon test was used to test for differences between continuous variables adjusted for multiple testing using the Holm method. * Indicates significance (P-value threshold <0.05) compared to mild COVID-19. ** Indicates significance (P-value threshold <0.05) in comparison to healthy individuals and mild COVID-19. Categorical variables were compared using Chi-square test.

[a] Maximum disease severity recorded for patients prospectively followed until recovery. Mild illness and pneumonia were considered mild COVID-19 disease, whereas severe pneumonia and any severity of ARDS were considered severe COVID-19 disease.

[b] COVID-19 severity according to WHO guidelines, maximum grade recorded for patients prospectively followed until recovery.

[c] Glucocorticoids initiated as part of COVID-19 treatment.

Abbreviations: ARDS, acute respiratory distress syndrome; ICU, intensive care unit; IQR, interquartile range.

nature portfolio

Corresponding author(s): Onur Boyman

Last updated by author(s): Nov 15, 2021

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Cytek SprectroFlo was used for acquisition of flow cytometry data. The cell Ranger pipeline (version 5.0.0) was used for acquisition of scRNAseq data.
Data analysis	FlowJo (version 10.7.0) was used for manual gating of flow cytometry data. OMIQ (Sept 2021) was used for visualization. R (version 4.0.4) was used for visualization and testing of flow cytometry data. Seurat version (4.0.37) and R (version 4.1.0) was used for analysis of scRNAseq data. Demultiplexing was conducted with Souporcell (version 2). TCR profiling on filtered contig annotations was done using R package scRepertoire (version 1.1.4). Code is available at https://github.com/TheMoorLab.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The sequencing dataset generated during the current study is available under https://doi.org/10.5281/zenodo.5119633.
Flow cytometry datasets are available from the corresponding author on reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was based on the availability of samples rather than on a pre-defined sample size calculation.
Data exclusions	Samples stained with HLA-A*24:02 dextramers were excluded from further analysis due to a strong background staining on cells from unexposed individuals. CoV2-Dex+ clones that were double positive for different dextramers were excluded from the analysis. Phenotypes were only analyzed when ≥5 cells were available.
Replication	Samples from each patient were analyzed once due to limited availability. All phenotypical and functional differences between patient groups were identified in multiple patient samples.
Randomization	In this observational study randomization was not applicable.
Blinding	Flow cytometry samples were stained in 4 batches on 4 consecutive days and in one additional fifth batch for revision. While performing the experiments the investigators were blinded to disease severity and timepoint of collection (acute, six months, one year) of the sample.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Antibodies used for flow cytometry are listed in Supplementary Tables 1-3.
Validation	All antibodies used in this study are commercially available antibodies validated for flow cytometry analysis and/or sequencing. CD4 Pacific Blue Biolegend 344620.Reactivity: Human, Application: FC, quality tested CD56 BV510 Biolegend 318339. Reactivity: Human, African Green, Baboon, Cynomolgus, Rhesus. Application: FC, quality tested CD3 BV785 Biolegend 300472. Reactivity: Human cross-reactivity chimpanzee, Application: FC, quality tested CD8 AF488 Biolegend 344716 Reactivity: Human, Cross-Reactivity: African Green, Chimpanzee, Cynomolgus, Pigtailed Macaque, Rhesus, Sooty Mangabey, Application: FC, quality tested

CD39 APC Biolegend 328209, Reactivity: Human, Rhesus. Application: FC, quality tested
Fixable viability dye EF 780 Invitrogen 1 65-0865-14 Application: FC,
CD45RA TotalSeqTM Biolegend 304163 Reactivity: Human cross-reactivity chimpanzee, Proteogenomics, quality tested
CCR7 TotalSeqTM Biolegend 353251 Reactivity: Human, African Green, Baboon, Cynomolgus, Rhesus, Proteogenomics, quality tested
Ki-67 BUV 395 BD 564071, Reactivity: Human (QC Testing), Application: Intracellular staining (flow cytometry) (Routinely Tested)
Zombie UV UV450 Biolegend 423107, Application: FC, ICFC
CD4 BUV496 BD 564652, Reactivity: Human (QC Testing), Application: Flow cytometry (Routinely Tested)
CD45RA BUV563 BD 565703, Reactivity: Human (QC Testing), Application: Flow cytometry (Routinely Tested)
HLA-DR BUV615 BD 751142, Reactivity: Human (QC Testing), Application: Flow cytometry (Routinely Tested)
CD8 BUV661 BD 741683, Reactivity: Human (Tested in Development), Application: Flow cytometry (Qualified)
CD28 BUV737 BD 564438, Reactivity: Human (QC Testing), Application: Flow cytometry (Routinely Tested)
CD3 BUV 805 BD 612893, Reactivity: Human (QC Testing), Rhesus, Cynomolgus, Baboon (Tested in Development), Flow cytometry (Routinely Tested)
TCF-7/TCF-1 BV421 BD 566692 Reactivity: Human (QC Testing), Rhesus, Cynomolgus, Baboon (Tested in Development), Flow cytometry (Routinely Tested)
Granzyme B Pacific blue Biolegend 515408, Reactivity: Human, Mouse, Cross-Reactivity: Rat, Application: FC, ICFC
CXCR3 BV510 Biolegend 353725 , Reactivity: Human, African Green, Baboon, Cynomolgus, Rhesus, Application: FC, ICFC
PD-1 BV605 Biolegend 329923, Reactivity: Human, African Green, Baboon, Cynomolgus, Rhesus, Application: FC, ICFC
CX3CR1 BV650 Biolegend 341625, Reactivity: Human, African Green, Baboon, Chimpanzee, Common Marmoset, Cynomolgus, Rhesus, Squirrel Monkey, Application: FC - Quality tested
TBET BV711 Biolegend 644819 Reactivity:Human, Mouse, Application: ICFC - Quality tested
CD39 BV785 Biolegend 328239, Reactivity: Human, Rhesus, Application: FC - Quality tested
CD69 FITC Biolegend 310904, Reactivity: Human, African Green, Baboon, Chimpanzee, Cynomolgus, Pigtailed Macaque, Rhesus, Application: FC - Quality tested
CD57 PerCP Cy5.5 Biolegend 359621, Reactivity: Human, Application: FC - Quality tested
CD56 BB790-P BD 624296 (custom), Reactivity: Human (QC Testing), Application: Flow cytometry- Quality tested
CD95 PE Dazzle BD 562395 Reactivity: Human (QC Testing), Rhesus, Cynomolgus, Baboon (Tested in Development), Application: Flow cytometry- Quality tested
CD25 PECy5 Biolegend 302608 Reactivity: Human, Baboon, Chimpanzee, Pigtailed Macaque, Rhesus, Application: FC- Quality tested
CD127 PE-Fire 700 Biolegend 351365 Reactivity: Human, African Green, Baboon, Cynomolgus, Rhesus, Application: FC- Quality tested
EOMES PECy7 Invitrogen 25-4877-41 Reactivity: Human, Pig, Application: Flow cytometry
TOX EF 660 Invitrogen 50-6502-80 Reactivity: Human, mouse, Application: Flow cytometry
CCR7 AF700 BD 561143 Reactivity: Human (QC Testing), Application: Flow cytometry- Quality tested
TIM-3 APC fire Biolegend 345043 Reactivity: Application: FC- Quality tested

Human research participants

Policy information about studies involving human research participants

Population characteristics	Clinical characteristics of this cohort have been previously published: Cervia, C. et al. Systemic and mucosal antibody responses specific to SARS-CoV-2 during mild versus severe COVID-19. J. Allergy Clin. Immunol. 147: 545-557.e9 (2021). Chevrier, S. et al. A distinct innate immune signature marks progression from mild to severe COVID-19. Cell Reports Med. 2, 100166 (2021). Adamo, S. et al. Profound dysregulation of T cell homeostasis and function in patients with severe COVID-19. Allergy 1–16 (2021) doi:10.1111/all.14866. Age, median (IQR) [yrs]: Healthy: 29.0 (27.5–37), Mild: 29.00 (24.50–38.50), Severe: 62.50 (53.5–79.3) Gender m/f: Healthy: 5/8, Mild: 13/12, Severe: 13/9 Genotypes, frequency no. (%): HLA A01:01: Healthy: 5 (38%), Mild: 14 (56%), Severe: 14 (64%) HLA A11:01: Healthy: 5 (38%), Mild: 8 (32%), Severe: 8 (36%) HLA A24:02: Healthy: 3 (24%), Mild: 3 (12%), Severe: 3 (14%) Diagnosis no. (%): Hypertension – Mild: 0, Severe: 1 (5%) Diabetes – Mild: 0, Severe: 0 Heart disease – Mild: 0, Severe: 1 (5%) Lung disease – Mild: 0, Severe: 1 (5%) Malignancy – Mild: 0, Severe: 1 (5%) Kidney disease – Mild: 0, Severe: 1 (5%) Cerebro-vascular disease – Mild: 0, Severe: 1 (5%) Vascular thrombosis – Mild: 0, Severe: 1 (5%) Treatment no. (%): Hydroxychloroquine – Mild: 0 Severe: 9 (41%) Remdesivir – Mild: 1 (4%) Severe: 4 (19%) Glucocorticoids – Mild: –0 Severe: 4 (19%) Tocilizumab – Mild: 0 Severe: 2 (9%)
Recruitment	Patients 18 years and older with symptomatic, RT-qPCR-confirmed SARS-CoV-2 infection were recruited at four different

Recruitment	hospitals in Zurich, Switzerland. Both hospitalized patients and outpatients were recruited into the study and all participants gave written informed consent. Patients had to be competent at the time of consent. All out- and inpatients were recruited at the time of positive PCR test in each of the recruiting centers, if they were able to give informed consent. This introduced a potential bias as patients who were mechanically ventilated at the time of positive test could not be recruited.
Ethics oversight	The study was approved by the Cantonal Ethics Committee of Zurich (BASEC 2016-01440).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Whole blood was centrifuged and plasma was removed. The remaining blood was diluted with PBS and layered into a SepMate tube (STEMCELL, catalog number 85460) filled with Lymphodex solution (Inno-Train Diagnostik GmBH, catalog number 002041500). After centrifugation, peripheral blood mononuclear cells (PBMCs) were collected and washed with PBS, resuspended in FBS with 10% DMSO and frozen.
Instrument	Cytek Aurora
Software	FlowJo (version 10.7.0) was used for gating and visualization. OMIQ was used for visualization of UMAPs.
Cell population abundance	The post sort ratio of specific to non-specific cells was checked by acquisition on a Cytek Aurora immediately after the sort for one representative donor and was approximately 1:10.
Gating strategy	Lymphocytes were gated from ungated samples (FSC-A/SSC-A), single cells were gated from lymphocytes (FSC-H, FSC-A). Live CD3+ cells were gated from single cells (CD3, live/dead). T cells were gated from CD3+ cells (CD3, CD56). CD8+ T cells were gated from CD3+ T cells. For complete gating strategy see Extended Data Figures.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.