Check for updates

Single-cell epigenomic landscape of peripheral immune cells reveals establishment of trained immunity in individuals convalescing from COVID-19

Maojun You [1,2,8], **Liang Chen**[3,4,5,7,8 ✉], **Dawei Zhang**[6,8], **Peng Zhao**[6,8], **Zhu Chen**[6], **En-Qiang Qin**[6], **Yanan Gao**[1,2], **Mark M. Davis** [3,4,5 ✉] **and Pengyuan Yang** [1,2 ✉]

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection often causes severe complications and even death. However, asymptomatic infection has also been reported, highlighting the difference in immune responses among individuals. Here we performed single-cell chromatin accessibility and T cell-receptor analyses of peripheral blood mononuclear cells collected from individuals convalescing from COVID-19 and healthy donors. Chromatin remodelling was observed in both innate and adaptive immune cells in the individuals convalescing from COVID-19. Compared with healthy donors, recovered individuals contained abundant TBET-enriched CD16$^+$ and IRF1-enriched CD14$^+$ monocytes with sequential trained and activated epigenomic states. The B-cell lineage in recovered individuals exhibited an accelerated developmental programme from immature B cells to antibody-producing plasma cells. Finally, an integrated analysis of single-cell T cell-receptor clonality with the chromatin accessibility landscape revealed the expansion of putative SARS-CoV-2-specific CD8$^+$ T cells with epigenomic profiles that promote the differentiation of effector or memory cells. Overall, our data suggest that immune cells of individuals convalescing from COVID-19 exhibit global remodelling of the chromatin accessibility landscape, indicative of the establishment of immunological memory.

The global coronavirus disease 2019 (COVID-19) pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has resulted in over 70 million recorded cases with a mortality rate of approximately 2.3%[1]. To date, the main strategy for ending the pandemic involves the development of effective vaccines, many of which are being evaluated[2–3]. Although most leading vaccine candidates can elicit virus-neutralizing antibodies, there is considerable uncertainty on how protective or durable these responses might be. Thus, a greater understanding of the innate and adaptive immune responses of individuals convalescing from COVID-19 is needed, particularly knowledge on whether immunological memory is established.

To date, studies on the immune responses to SARS-CoV-2 have mainly focused on the aspects of pathological inflammation[3–5]. According to single-cell transcriptome analyses, monocytes from the peripheral blood or the bronchoalveolar lavage fluid of COVID-19 patients are characterized by high levels of pro-inflammatory cytokines[5–8] such as IL-6, IL-1β and TNFα. In addition, most of the patients with COVID-19 develop lymphocytopenia—especially reduced numbers of CD4$^+$ and CD8$^+$ T cells—within two weeks of disease onset, contributing to increased serum IL-6 levels. Lymphocytopenia is also predictive of the disease severity[9–11], indicating that CD4$^+$ and CD8$^+$ T cells are critical for viral control and disease recovery. Moreover, single-cell transcriptomic

analyses showed an increased ratio of plasma cells in patients infected with COVID-19 and early convalescent individuals[7,12].

The epigenomic regulation of immune responses to primary SARS-CoV-2 infection remains unknown at present. Here we applied single-cell transposase-accessible chromatin with sequencing (scATAC-seq) to analyse chromatin remodelling in the peripheral immune cells of six matched and uninfected healthy donors (five donors collected in this study and an additional dataset from a public database[13]) and ten individuals convalescing from moderate or severe COVID-19 infection at 4–12 weeks following recovery. The approach of 10x-based scATAC-seq enables the production of high-quality single-cell accessible chromatin profiles at massive scale across all major immune-cell types in peripheral blood mononuclear cells (PBMCs). For the T-cell analysis, we developed a single-cell method that combines single-cell T cell-receptor (TCR) sequencing (scTCR-seq), fluorescence-activated cell sorting (FACS) with index sorting and ATAC-seq—termed TCR–FACS–index–ATAC sequencing (Ti-ATAC-seq)—thereby allowing us to obtain expression data for cell-surface markers, paired TCR sequences and the chromatin accessibility landscape for each cell analysed. This single-cell chromatin landscape approach facilitates a comprehensive understanding of how effector or memory cells are established in both the innate and adaptive immune responses of individuals convalescing from COVID-19.

[1]Key Laboratory of Infection and Immunity of CAS, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Beijing, China. [2]Chongqing International Institute for Immunology, Chongqing, China. [3]Department of Immunology and Microbiology, Stanford University, Stanford, CA, USA. [4]Institute for Immunity, Transplantation and Infection, Stanford University, Stanford, CA, USA. [5]Howard Hughes Medical Institute, Stanford University, Stanford, CA, USA. [6]The Fifth Medical Center, Chinese PLA General Hospital, Beijing, China. [7]Present address: School of Medicine, Shanghai University, Shanghai, China. [8]These authors contributed equally: Maojun You, Liang Chen, Dawei Zhang, Peng Zhao. ✉e-mail: lchen12@stanford.edu; mmdavis@stanford.edu; pyyang@ibp.ac.cn

Results

Overview of the peripheral immune-cell profiling in individuals convalescing from COVID-19. We mapped the transposase-accessible chromatin at the single-cell level using 10x chromium scATAC-seq (10x Genomics; Fig. 1a and Supplementary Table 1). We collected data from 72,318 PBMCs from eight individuals convalescing from COVID-19 as well as 24,997 PBMCs from five healthy donors and one additional dataset from a public database[13] as controls (Fig. 1b and Extended Data Fig. 1a). Using uniform manifold approximation and projection (UMAP), we analysed the distribution of the immune-cell populations based on the gene scores of canonical lineage markers, which demonstrate the aggregate accessibility of several enhancers linked to the indicated genes—including *CD14*, *CD16* (*FCGR3A*) and *CEBPB* for monocytes; *PAX5* and *MS4A1* for B cells; *CD8A* and *CD3G* for CD8+ T cells; *CD4* and *CD3G* for CD4+ T cells; and *KLRB1* for natural killer (NK) cells (Fig. 1c–e). The frequencies of monocyte-lineage and effector and memory CD8+ T cells increased in individuals convalescing from COVID-19, whereas those of the B-cell lineages decreased (Fig. 1f,g). We further analysed and clustered the cells of each lineage separately.

In addition, we explored Ti-ATAC-seq to further investigate special clonally expanded T-cell clones (Fig. 1a and Extended Data Fig. 2). We collected 2,880 αβ T cells from ten individuals convalescing from COVID-19, among which 1,574 quality-control-positive T cells had both paired TCRαβ sequences and ATAC-seq (Fig. 1h). Ti-ATAC-seq had similar performance to previously published microfluidic platform-based T-ATAC-seq[14] and other joint ATAC-RNA approaches[15,16] (Extended Data Fig. 2c).

Sequential trained and activated monocytes in COVID-19. We identified and annotated 13 subclusters of monocytes (Fig. 2a). Clusters of CD14+ monocytes (clusters 11, 12 and 13) showed a high *CD14* gene score. Clusters of CD16+ monocytes (clusters 3 and 4) were characterized by a high *CD16* gene score. CD1c+ conventional dendritic cells (clusters 1 and 8) were defined by their high *CD1c* gene score. CLEC4C+ plasmacytoid dendritic cells showed high gene accessibility in the *CLEC4C* locus. Clusters 9 and 10, with their predominantly high *KLF4* and *CD68* gene scores, probably represented monocyte-derived dendritic cells (Fig. 2b and Extended Data Fig. 3a,b).

We focused on CD14+ and CD16+ monocytes because the individuals convalescing from COVID-19 had a higher abundance of two clusters of CD14+ monocytes (clusters 11 and 12) and one cluster of CD16+ monocytes (cluster 4) than healthy individuals (Fig. 2c). We found and annotated sequential differentiation states from homeostasis to the mature immune inflammatory effector state in the COVID-19 group through differential *cis*- and *trans*-element analysis within the CD14+ and CD16+ clusters (Fig. 2a). For example, cluster 4 (CD16-activation) was associated with leukocyte activation and the regulation of cytokine production, and showed high levels of activity of transcription factors (TFs) involved in IFNγ-induced rapid signatures and myeloid differentiation, including TBET, JUN and FOSB[17–20], whereas cluster 3 (CD16-maturation) was associated with haemopoiesis and the regulation of the immune response (Fig. 2d and Extended Data Fig. 3c). Interestingly, cluster 4 had increased chromatin accessibility in the *CCL5*, *PRF1* and *GZMB* loci (Extended Data Fig. 3a), demonstrating that IFNγ drives the differentiation and activation of CD16+ monocytes in the hyper-inflammatory microenvironment of patients with COVID-19. For the CD14+ monocytes, the clusters 12 and 13 (CD14-activation and CD14-maturation, respectively) showed increased chromatin accessibility at inflammatory cytokine genes—including *IL1β*, *IL6*, *IL8*, *CCL2*, *CCL3* and *CCL7* (Extended Data Fig. 3a)—and enriched TFs that represent a mature state—including FOS, JUN and MAF (Fig. 2e–g)[19,21]. Meanwhile, cluster 11 (CD14-trained) was

characterized by the accessibility of genes involved in leukocyte activation (Extended Data Fig. 3c) and showed high activity of TFs involved in haematopoietic commitment and survival of monocytes, including HOXA9 and NR4A1 (Fig. 2e,f)[22,23]. In addition, NR4A1 TF activity, which was enriched in cluster 11 (Fig. 2e,f), was also enriched in the activation stage of the CD16+ monocyte (Fig. 2d).

Innate immune cells have recently been shown to display immunological memory after certain infections or vaccines, also termed trained immunity[24–26]. To understand the developmental dynamics of these monocytes, we constructed a lineage trajectory of CD14+ based on sequential differentiation states, which progressed from trained to activation and finally to maturation, and generated ordered single cells (termed as 'pseudotime') based on their epigenetic similarity (Fig. 3a). The dynamic *cis*-element and TF motif accessibility across the trajectory were consistent with the sequential differentiation states (Fig. 3b). For instance, the trained, activated and mature monocytes showed sequential chromatin accessibility in the *IL1β* and *CCL5* loci (Fig. 3b,c). The trained stage of the CD14+ lineage trajectory was accompanied by increased accessibilities of IRF1, IRF3 and IRF8 TF motifs (Fig. 3b,d,e), thus enhancing susceptibility to regulate rapid IFN-β induction and host inflammatory defences in human blood monocytes[20,27]. The activation and maturation stages of these trajectories shared accessibility at TFs involved in monocyte activation and maturation for the activator protein 1 (AP-1) factors FOS and JUN (Fig. 3b,d,e)[19,28,29].

To validate this, we re-challenged PBMCs with spike-nCoV pseudovirus or hepatitis B virus for 24 h in vitro, as monocytes predominantly showed high gene scores for *IL1β* and *IL6* (Fig. 1e). The PBMCs from individuals convalescing from COVID-19 secreted higher levels of IL-1β and IL-6 than those from the healthy donors (Fig. 3f), which is consistent with trained monocytes having enhanced sensitivity in response to different pathogens after an initial challenge[25]. There was no significant difference in the levels of cytokine secretion between the patients that were tested early (4 weeks) or late (10–12 weeks) following discharge (Fig. 3f).

Collectively, although systemic inflammation had returned to a normal and healthy level (Supplementary Table 1), the CD14+ and CD16+ monocytes maintained chromatin reprogramming that promotes trained immunity and thus enables a rapid inflammatory response against subsequent infection.

A facilitated B-cell developmental programme in individuals convalescing from COVID-19. We identified ten B-cell clusters: immature B cells (clusters 6 and 7) with a high *SDC1* score; naive B cells (clusters 4 and 5) with high *TCL1A*, *CD19* and *CD20* scores; memory B cells (clusters 1, 2 and 3) with high *CD27* and *CD38* scores; and plasma cells (clusters 8, 9 and 10) with a high *XBP1* score (Fig. 4a,b and Extended Data Fig. 4a).

Although the individuals convalescing from COVID-19 and healthy donors had comparable frequencies of B-cell subsets (Fig. 4c), the UMAP projection revealed preferential cell distribution in each B-cell subset between healthy donors and individuals convalescing from COVID-19 (Fig. 4d and Extended Data Fig. 4b). The transcriptional regulation of the B-cell developmental programme in healthy donors has been extensively studied[30,31]. However, the B-cell lineage trajectories based on accessible chromatin differed between the healthy donors and individuals recovering from COVID-19 (Extended Data Fig. 4b–d), suggesting that SARS-CoV-2 infection induced different B-cell developmental programmes.

To investigate this, we reconstructed the lineage trajectory of B cells from the healthy donors and individuals convalescing from COVID-19 (Fig. 4d). We observed *cis*-elements near known regulators of every stage of B-cell development—such as *TCL1A*, *CXCR4*, *MEF2C*, *BHLHE41*, *BAFF*, *PLCG2*, *PAX5*, *EBF1* and *RORA*
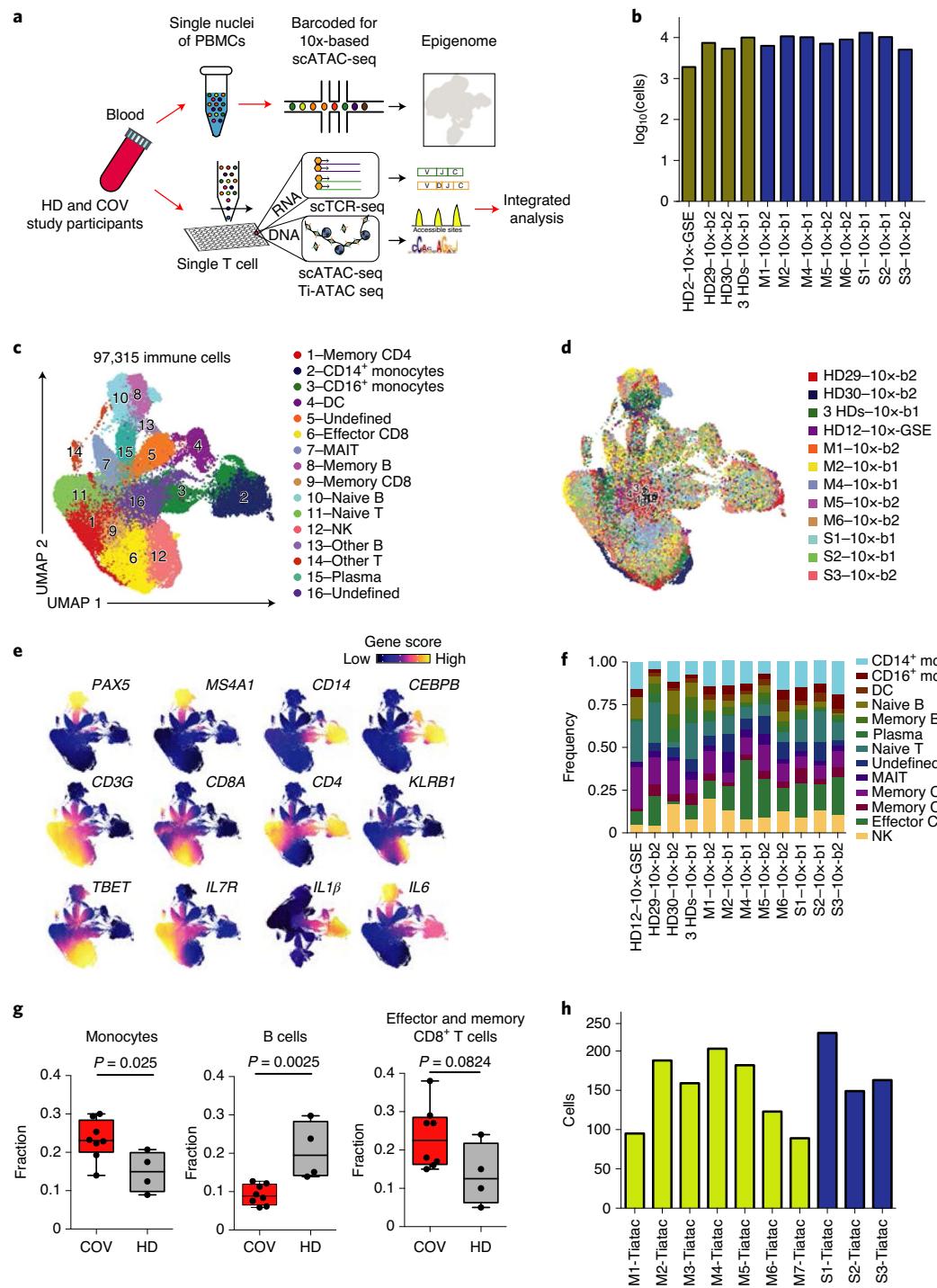
Fig. 1 | Overview of the immune-cell epigenomic landscape of the blood of individuals convalescing from COVID-19. a, Outline of the two approaches used for scATAC-seq: 10x-based scATAC-seq (top) and Ti-ATAC-seq (bottom). **b**, Cell numbers ($\log_{10}$-transformed) in each sample for 10x-based scATAC-seq. **c**, UMAP plot showing 97,315 10x-based scATAC-seq profiles of immune cells in peripheral blood—including B, NK and T cells, monocytes and other clusters. The dots indicate individual cells and the cell-type identity is indicated by colour. **d**, UMAP plot showing the sample origin. **e**, Canonical markers overlaid on UMAP embedding, including *CD14* and *CEBPB* for monocytes; *PAX5* and *MS4A1* for B cells; *CD3G*, *CD8A* and *CD4* for T cells; and *KLRB1* for NK cells. The UMAP plot is coloured based on the log-transformed normalized gene scores, which exhibited the accessibility of peaks linked to the indicated genes. The gene scores were calculated as $\log_2$(normalized count $+1$). **f**, Cell-type frequencies in each sample, determined using 10x-based scATAC-seq. The colours indicate the cell type. **g**, Differences in the proportion of monocyte and B-cell lineages as well as effector and memory CD8+ T cells in the samples from individuals convalescing from COVID-19 (COV; $n=8$ samples collected from eight individuals, one sample per individual) and healthy donors (HD; $n=4$ samples, three samples collected from five HD in this study (including a pooled sample from three HD) and an additional sample from published data (GSE139369) as indicated in **b** and Extended Data Fig. 1a). A two-sided unpaired Student's *t*-test was performed to determine the *P* values. The boxplots denote the median with the quartile range (25–75%), and the length of whiskers represents 1.5× the interquartile range (IQR). **h**, Number of T cells in each sample that had both scTCR-seq and ATAC-seq data.
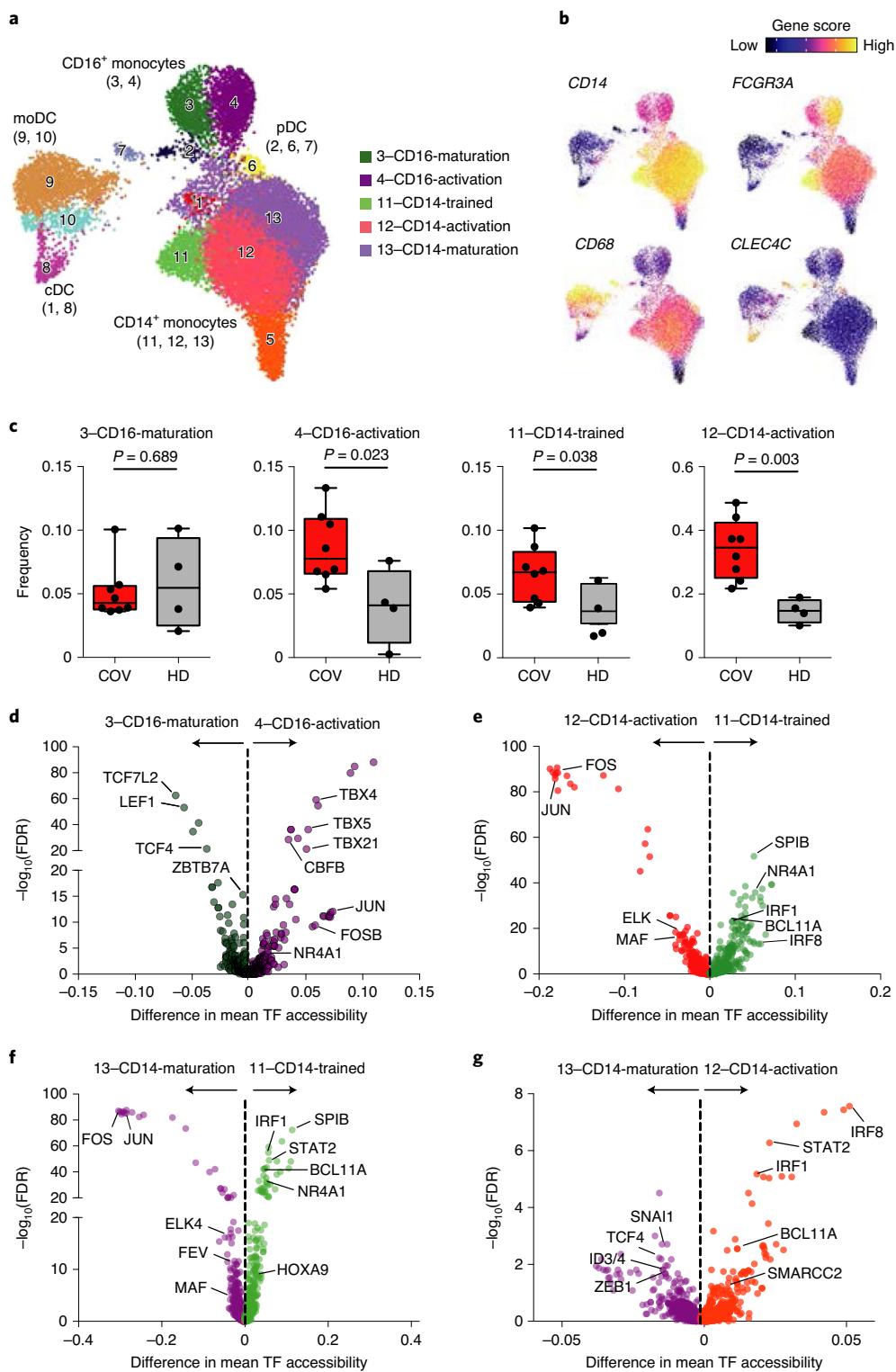
Fig. 2 | Epigenomic signatures of trained and activated monocytes in individuals convalescing from COVID-19. a, Subclustering UMAP of all monocytes (see Fig. 1c). The five clusters of CD14+ and CD16+ monocytes indicated in the legend were annotated according to sequential differentiation states; cDC, conventional dendritic cells; pDC, plasmacytoid dendritic cells and moDC, monocyte-derived dendritic cells. **b**, Surface markers overlaid on UMAP embedding coloured according to the log-transformed normalized gene scores. **c**, Differences in the proportions of clusters 3, 4, 11 and 12 in the samples of individuals convalescing from COVID-19 (COV; $n = 8$ samples collected from eight individuals, one sample per individual) and healthy donors (HD; $n = 4$ samples, three samples collected from five HD in this study (including a pooled sample from three HD) and an additional sample from published data (GSE139369)). A two-sided unpaired Student's t-test was performed to determine the P values. The boxplots denote the median with the quartile range (25–75%), and the length of whiskers represents 1.5× the IQR. **d–g**, Volcano plots showing the differential TF motif accessibility using the mean TF motif accessibility in the chromVAR TF bias-corrected deviation in clusters 3 and 4 (**d**), 12 and 11 (**e**), 13 and 11 (**f**), and 13 and 12 (**g**). The P values were calculated using two-sided pairwise Wilcoxon test and the false discovery rate (FDR) was corrected using the Benjamini–Hochberg procedure.
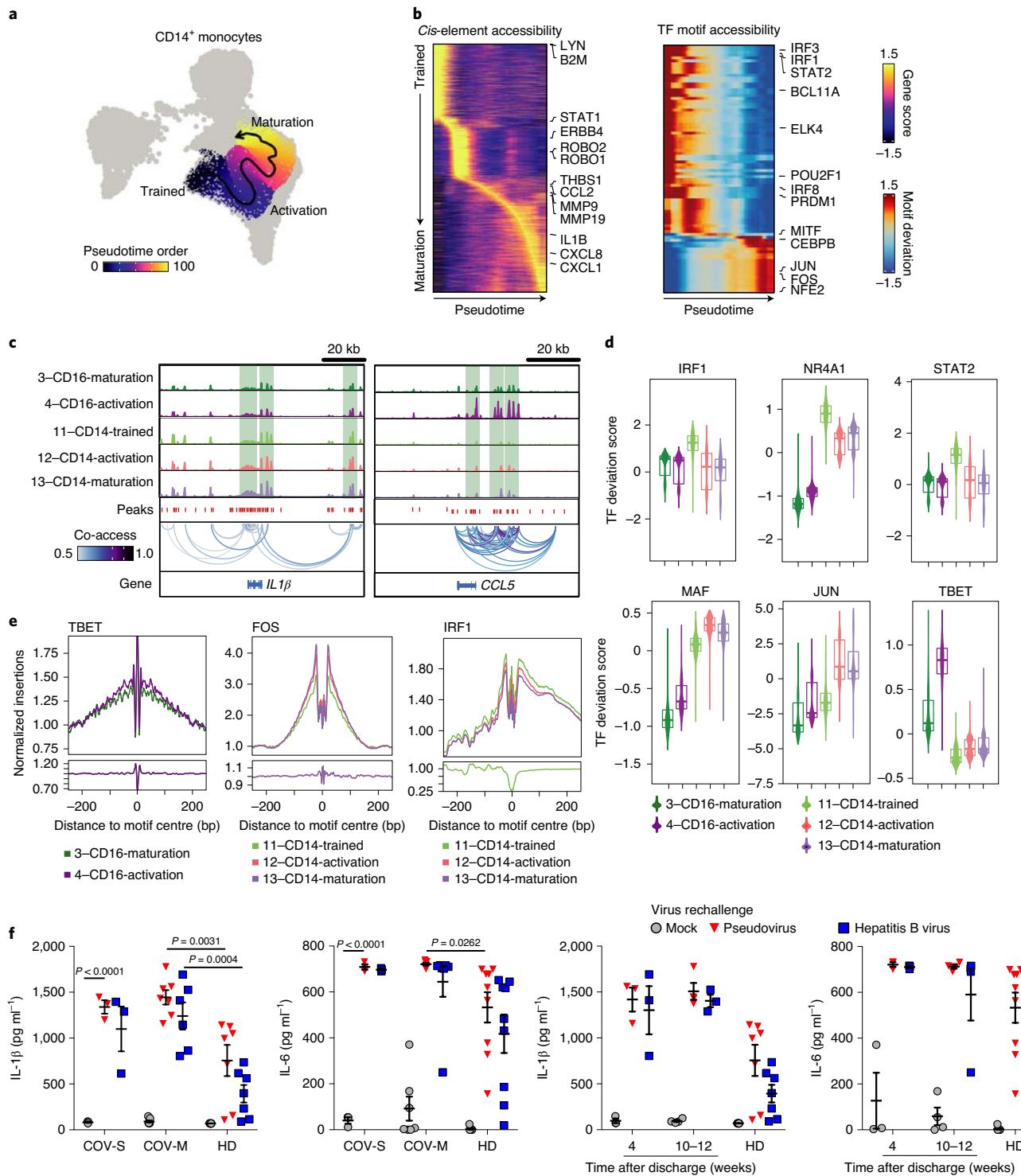
Fig. 3 | Epigenomic differentiation trajectory of CD14⁺ and CD16⁺ monocytes. a, UMAP showing the lineage trajectory of CD14⁺ monocytes ordered based on trained, activation and maturation states. Pseudotime values were overlaid on the UMAP embedding; the smoothed line and arrow represent the visualization of the trajectory path from the spline fit. **b**, Heatmaps of the ordered *cis*-element accessibility (left) and TF motif accessibility (right) across pseudotime in the CD14⁺ monocytes (see Fig. 3a). The *cis*-element and TF motif accessibilities are indicated by the gene score and chromVAR TF-motif bias-corrected deviation, respectively. **c**, Aggregated single-cell genome tracks for the indicated clusters at the *IL1β* (left) and *CCL5* (right) loci with peak co-accessibility (Co-access). The Co-access is indicated by the inferred peak-to-gene links for distal regulatory elements. Green shading indicates differential peaks within clusters. **d**, Violin plots showing the ChromVAR TF-motif bias-corrected deviation scores of the indicated TF regulators across clusters 3 ($n=1,003$ cells), 4 ($n=1,472$ cells), 11 ($n=988$ cells), 12 ($n=5,719$ cells) and 13 ($n=4,537$ cells). The boxplots denote the medians and the quartile range (25–75%), and the length of whiskers represents 1.5× the IQR. **e**, TF footprints of the TBET motif in clusters 3 and 4 (left) as well as the FOS (middle) and IRF1 (right) motifs in clusters 11, 12 and 13. The Tn5 insertion bias track is shown. **f**, Levels of IL-1β and IL-6 secreted by virus-exposed PBMCs (COV-S, $n=3$ samples from three individuals that recovered from severe COVID-19; COV-M, $n=7$ samples from seven individuals that recovered from moderate COVID-19; HD, $n=10$ samples from ten healthy donors). The error bars indicate the mean ± s.e.m. A two-sided unpaired Student's *t*-test was performed to determine the *P* values.
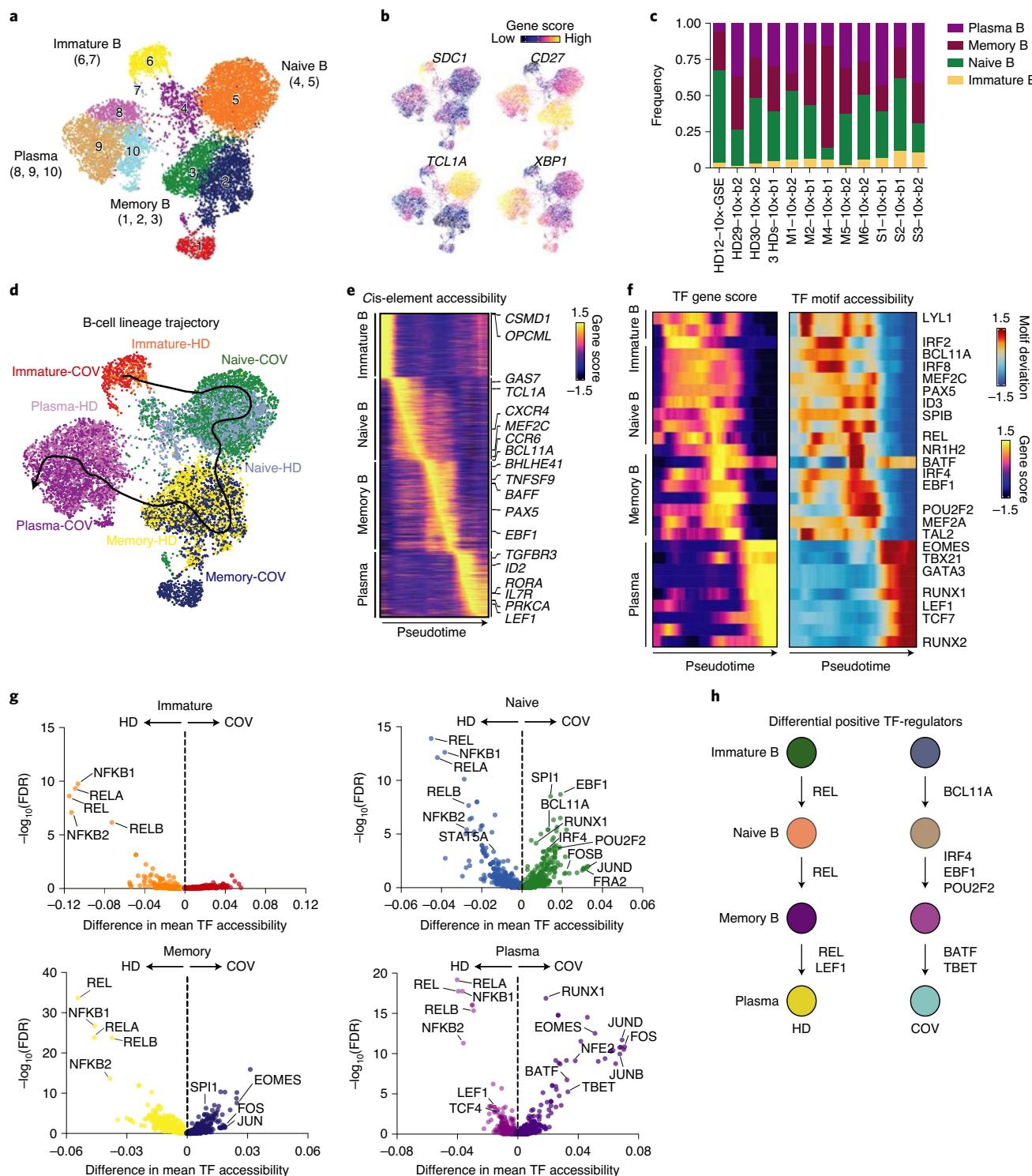
Fig. 4 | Accelerators facilitated the B-cell developmental programme in individuals convalescing from COVID-19. a, Subclustering UMAP plot of all B cells (see Fig. 1c). **b**, Gene markers overlaid on UMAP coloured according to the log-transformed normalized gene scores for *SDC1*, *TCL1A*, *CD27* and *XBP1*. **c**, Relative frequencies of the B-cell subclusters in the different samples. **d**, Integrative lineage trajectory of B-cell states in healthy donors (HD) and individuals convalescing from COVID-19 (COV). The smoothed line and arrow represent the visualization of the trajectory path across different states (immature, naive, memory and plasma), and the sample origins are denoted by colour. **e**, Heatmaps showing the ordered gene-score trajectory across pseudotime for B-cell differentiation. **f**, Heatmaps showing the positive TF regulators obtained from the integration of ordered TF gene scores (right) with ordered TF motif accessibility (left) across pseudotime for B-cell differentiation. Positive TF regulators are TF motifs that show high bias-corrected chromVAR TF-motif deviations that also exhibit similarly dynamic gene scores across differentiation states. **g**, Volcano plots demonstrating the differential TF motif accessibility using the mean TF motif accessibility in the chromVAR TF bias-corrected deviation between COV and HD individuals in the indicated B-cell states. The *P* values were calculated using a two-sided pairwise Wilcoxon test and the false discovery rate (FDR) was corrected using the Benjamini–Hochberg procedure. **h**, Schematic of differential positive regulatory TFs driving B-cell differentiation of immature B cells to plasma cells in HD and COV individuals.

(Fig. 4e and Extended Data Fig. 5a)[30–34]—suggesting that the integrated B-cell lineage trajectory was a well-defined developmental programme that could be used to compare regulatory mechanisms between the healthy donors and individuals convalescing from COVID-19. Next, through integrated chromVAR TF deviations with similarly dynamic gene scores across differentiation states, we identified positive TF regulators with sequential activities of BCL11A, IRF8, PAX5, REL, BATF, IRF4, EBF1, POU2F2, TBET and LEF1 that promote B-cell commitment, differentiation, maintenance and class-switch recombination[30,35–39], thus resolving the integrated timing of TF activity for comparison (Fig. 4f and Extended Data Fig. 5b–d).

We then measured the TF deviation scores and variations to identify the differential positive TF regulators of four B-cell subsets of the healthy donors and individuals convalescing from COVID-19 (Fig. 4g). Interestingly, NF-kB subunits, including REL, RELA and RELB, which are involved in germinal centre B-cell maintenance and homeostasis[40,41], were enriched in four subsets of B cells of the healthy donors (Fig. 4g and Extended Data Fig. 5d). Meanwhile, AP-1 factors, including FOS and JUN, which are involved in the B-cell receptor signalling pathway and indicate B-cell differentiation and activation[42], were enriched in naive, memory and plasma B cells of the individuals convalescing from COVID-19. Transcription factors—including SPI1, EBF1, IRF4 and POU2F2—that are vital for B-cell survival, differentiation and receptor signalling responses were enriched in the naive B cells of the individuals convalescing from COVID-19[36,38,43–46]. In comparison to the memory and plasma B cells of healthy donors that showed a high REL deviation score, the plasma cells and a subset of memory B cells of the COVID-19 group showed high activity of TFs that are involved in class-switch recombination and promote specialized immune function in class-specific IgG+ memory B cells, including TBET and BATF[39,47] (Fig. 4g and Extended Data Fig. 5d). We also detected specific IgG antibodies of the spike protein of SARS-CoV-2 in the blood samples of the individuals convalescing from COVID-19 (Extended Data Fig. 5e). Collectively, according to the results of integrated timing analysis of TF activity, differential positive TF regulators promoted B-cell maintenance and homeostasis in healthy donors, whereas they facilitated B-cell activation, differentiation and IgG class-switch recombination in the individuals convalescing from COVID-19 (Fig. 4h).

CD8+ T cell-fate decisions in individuals convalescing from COVID-19. We first sub-grouped NK and T cells into 14 subclusters (Fig. 5a and Extended Data Fig. 6a–c). The CD8+ T-cell states consisted of naive, intermediate, memory and effector T cells as well as mucosal-associated invariant T cells (MAIT), whereas the CD4+ T-cell states were composed of naive, central memory (TCM) and effector memory (TEM) T cells. In contrast to healthy donors enriched in naive CD8+ T cells, the individuals convalescing from

COVID-19 showed increased numbers of effector and memory CD8+ T cells (Figs. 1h and 5b). Thus, clusters 6 and 8 may represent SARS-CoV-2-induced effector and memory CD8+ T cells, respectively.

We next orchestrated the effector and memory CD8+ T-cell trajectories to explore CD8+ T cell-fate decisions. The effector and memory trajectories almost overlapped at the early state and diverged into two distinct branches, that were predominantly comprised of cells from the individuals convalescing from COVID-19 (Fig. 5c). The analysis of cis-elements near effector genes revealed distinct regulatory patterns of accessibility across the pseudotime of the effector and memory CD8+ T cells (Fig. 5d,e). For example, the cis-element accessibility in the promoter and distal enhancers in the GZMB and IFNG loci gradually increased from the naive to the intermediate and then to the memory and effector states (Fig. 5f), demonstrating that the expression of effector genes was accurately regulated by a distinct cis-element network in the different states.

The results of TF-activity analysis demonstrated the shared and unique TF programmes across the pseudotime of effector and memory CD8+ trajectories in the individuals convalescing from COVID-19 (Fig. 5d,e). For instance, in the first stage of trajectories from the naive to intermediate state, effector and memory CD8+ T cells shared accessibility at the AP-1 factors FOS and JUNB, consistent with their roles in T-cell differentiation and activation[14,31]; however, effector CD8+ T cells also showed accessibility at NFKB1/2 and RUNX3, whereas memory CD8+ T cells were characterized by accessibility of BATF. IRF4 activity gradually increased across the effector CD8+ T-cell trajectory but was depleted in that of memory CD8+ T cells (Fig. 5g and Extended Data Fig. 6d), which is consistent with their role in limiting the development of memory T cells[48]. In contrast, BATF motifs showed accessibility earlier in the trajectory of memory CD8+ T cells than that of effector cells (Fig. 5d,e). Similarly, the second stage of trajectories from the intermediate state to either the effector or memory state identified the pivotal roles of EOMES and TBET motifs in each pathway (Fig. 5d,e,g). Effector CD8+ T cells were also characterized by accessibility of SREBF1/2 (Fig. 5d,g), which is involved in the metabolic reprogramming of effector T cells during extensive clonal expansion[49]. Memory CD8+ T-cell commitment was accompanied by the accessibility of motifs involved in promoting memory T-cell trafficking and maintenance of Krüppel-like factor 2/13 (KLF2/13; Fig. 5e,g)[50–53]. Notably, compared with healthy donors, we also observed gradual increased gene accessibility of GZMB, IFNG and TBET in the intermediate, memory and effector CD8+ T cells of individuals convalescing from COVID-19 (Fig. 5h,i).

Furthermore, we performed TCR clonality analysis using the paired TCRαβ sequences generated from Ti-ATAC-seq. The clonal expansion rates of CD8+ T cells were significantly higher than those of CD4+ T cells of individuals convalescing from COVID-19

Fig. 5 | Single-cell epigenomic profiles of T cells in individuals convalescing from COVID-19. a, Subclustering UMAP of NK and T cells (see Fig. 1c). For the CD8+ T cells, naive cells had high *TCF7* and *CCR7* gene scores; effector cells had high *CD8A*, *TBX21* (*TBET*) and *IFNG* gene scores; memory cells had high *CD8A*, *TBX21* and *KLF2/13* gene scores but low effector gene scores; and MAIT cells had high *SLC4A10* gene scores. For the CD4+ T cells, naive cells had high *TCF7* and *CCR7* gene scores; TCM cells had high *CD4*, *CCR7*, *AQP3* and *SELL* gene scores; and TEM cells had high *CD4* and effector gene scores but low CCR7 gene scores. The NK cells had high *NCAM*, *KLRC1*, *KLRD1* and *FCGR3A* gene scores; and NKT cells had high *NCAM* and *CD3E* gene scores. **b,** Cell-type frequencies in the different samples with the cluster identities indicated. **c,** Lineage trajectory of CD8+ T-cell states, which included naive, intermediate, effector and memory CD8+ T cells. Pseudotime values overlapped on the UMAP. **d,e,** Pseudotime heatmap showing the ordered cis-element accessibility (left) and TF motif accessibility (right) in effector (**d**) and memory (**e**) CD8+ T-cell lineage trajectories (see **c**). **f,** Aggregated single-cell genome tracks for the indicated clusters and states of CD8+ T cells at the *GZMB* (top) and *IFNG* (bottom) loci with peak co-accessibility (Co-access). The regions shaded in blue indicate differential peaks within clusters. **g,** Ridge plots showing TF deviation scores across different clusters and states of CD8+ T cells. **h,i,** Gene scores of *IFNG* (left) and *GZMB* (right; **h**), as well as the gene score (left) and TF deviation score of TBET (right; **i**) in the indicated cell types. COV, individuals convalescing from COVID-19; HD, healthy donors; COV-CD8-effector, n=12,353 cells; COV-CD8-memory, n=1,313 cells; COV-CD8-intermediate, n=2,737 cells; COV-CD8-naive, n=1,858 cells; HD-CD8-effector, n=2,052 cells; HD-CD8-memory, n=279 cells; HD-CD8-intermediate, n=1,408 cells; HD-CD8-naive, n=1,272 cells. The boxplots denote the medians and the quartile range (25–75%), and the length of whiskers represents 1.5× the IQR.

(Fig. 6a,b). Interestingly, CD8+ T cells were strongly dominated by one or two T-cell clones in each patient and the clonal expansion rate of the largest clone ranged from 2.6% to 41.1% of the total CD8+ T cells (Fig. 6a). These results suggest that these putative SARS-CoV-2-specific CD8+ T-cell clones, rather than CD4+ T cells, play a critical role in viral control and long-term immune protection. We next integrated TCR clonality with the single-cell epigenomic profiles generated from Ti-ATAC-seq. Notably, on comparison with unexpanded CD8+ T-cell clones, the two largest clonally expanded CD8+ T-cell clones were particularly enriched in TBET and EOMES TFs

(Fig. 6c). This is consistent with our 10x-based scATAC-seq data showing that these two TFs are important for the development of effector and memory CD8+ T cells in the individuals convalescing from COVID-19 (Fig. 6d).

In summary, our in-depth analysis of the epigenomic landscape and single-cell TCR clonality in individuals convalescing from COVID-19 revealed pivotal roles for effector CD8+ T cells in the initial viral control and the formation of memory CD8+ T cells via global chromatin accessibility remodelling and accurate regulatory programmes.
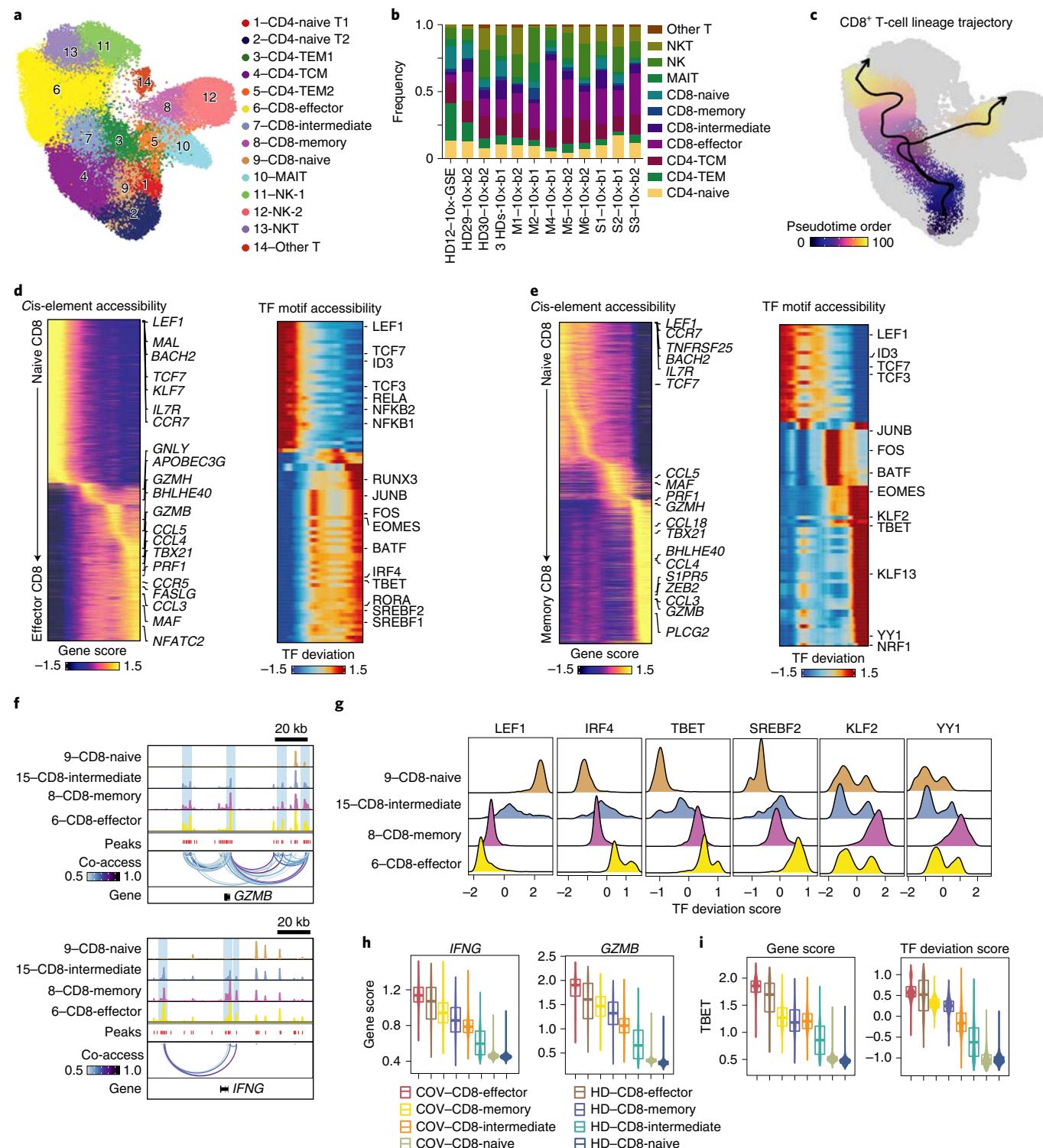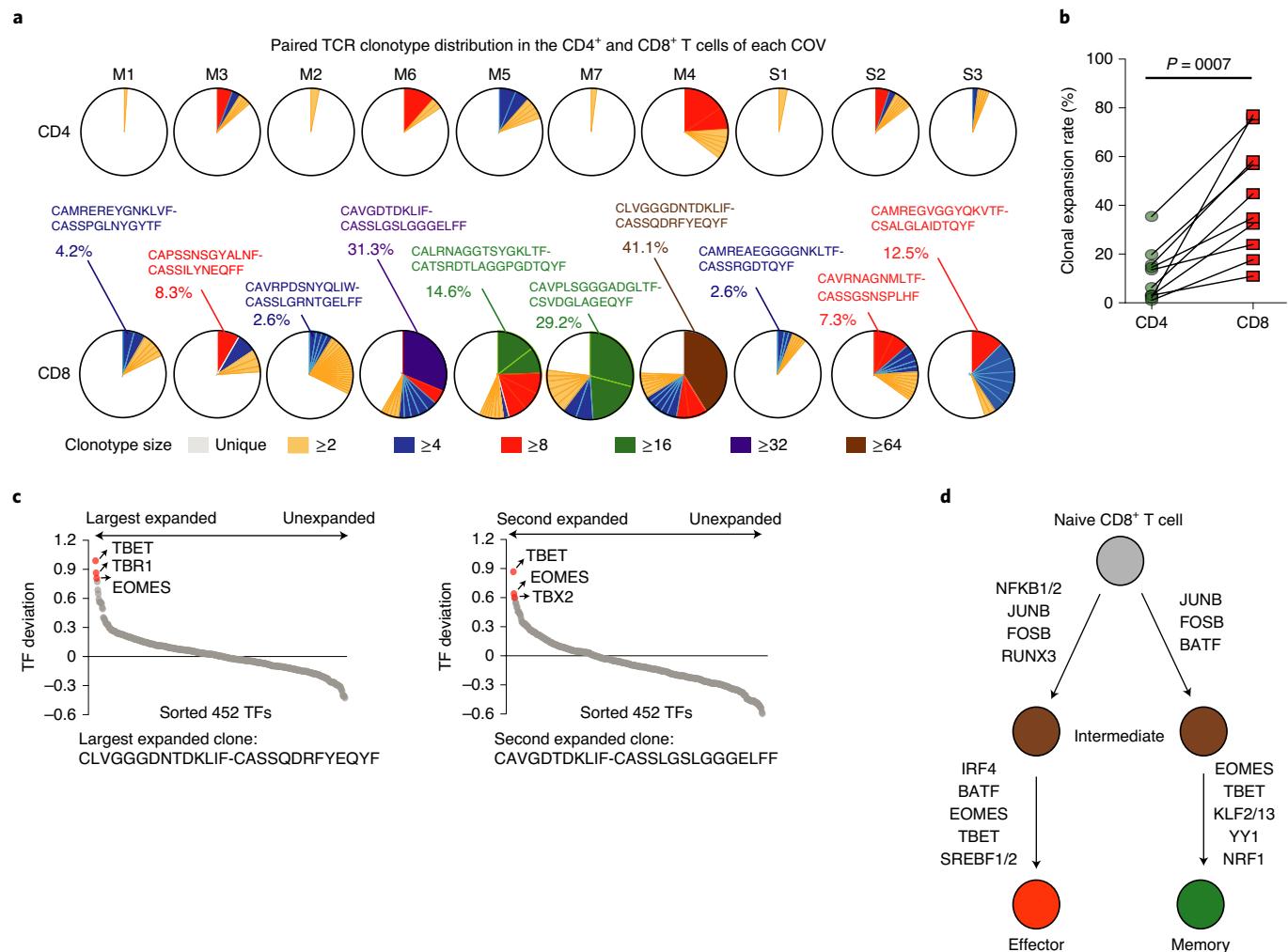
Fig. 6 | Single-cell epigenomic and TCR profiling of CD8+ T cells in individuals convalescing from COVID-19. a, Distribution of CD4+ (top) and CD8+ (bottom) T-cell TCR clonotypes, according to size, from ten individuals convalescing from COVID-19 (COV). The full-length paired TCR sequences were obtained from the Ti-ATAC-seq platform-based scTCR-seq. For the expanded TCRs with clonotype size >1, each slide indicates a unique clonotype. Single viable DAPI−CD3+TCRαβ+ T cells were gated and sorted, and the T-cell types were identified by FACS-indexed sorting. The CDR3α-CDR3β sequences and frequencies of the largest CD8+ T-cell clonotypes are shown. **b**, Clonal expansion rate of CD4+ and CD8+ T cells in ten individuals convalescing from COVID-19. Each data point indicates a single individual and the matched data of the CD4+ and CD8+ T cells are linked by distinct lines. A two-sided unpaired Student's *t*-test was performed to determine the *P* value. **c**, Ranked TF deviation enrichment values in the aggregated largest (left; TCR clone: CLVGGGDNTDKLIF-CASSQDRFYEQYF, *n* = 50 cells) and second-largest clonally expanded CD8+ T cells (right; TCR clone: CAVGDTDKLIF-CASSLGSLGGGELFF, *n* = 24 cells) versus aggregated unexpanded CD8+ T cells (*n* = 298 cells). The T-cell clonality and single-cell epigenomic profiles were obtained from the Ti-ATAC-seq platform. The TF enrichment values were calculated as the difference in the mean TF deviation between the two single-cell populations. **d**, Anti-SARS-CoV-2 CD8+ T-cell response. Schematic of the differential regulatory TFs driving CD8+ T-cell differentiation from the naive state to the effector-cell or memory state in individuals convalescing from COVID-19.

Discussion

SARS-CoV-2 causes severe pulmonary disease and complications with significant morbidity and mortality[1]. Our current understanding of the epigenomic regulatory mechanisms in the host immune response to SARS-CoV-2 infection and long-term immune protection is limited; thus, it is challenging to develop urgent therapeutics or assess the effect of different vaccine candidates.

Here we utilized high-throughput 10x-based scATAC-seq technology to measure the chromatin accessibility of all PBMCs, including monocytes and B, NK and T cells. We also used Ti-ATAC-seq to further investigate special clonally expanded T-cell clones with paired TCRαβ sequencing and responded chromatin accessibility from each T cell. Chromatin remodelling was significantly altered in almost all immune-cell compartments in the individuals

convalescing from COVID-19. Our epigenomic profiles revealed a sequential differentiation state from homeostasis to the mature immune inflammatory effector response in CD14+ and CD16+ monocytes of individuals convalescing from COVID-19, which featured trained, activation and maturation states. The trained and activation states of CD14+ and CD16+ monocytes were dominantly enriched in the individuals convalescing from COVID-19. Although immune memory is a well-known feature of the acquired immune system, the activation of the innate immune system can also heighten responsiveness to subsequent triggers, which may provide protection during the early stage of reinfection. In B-lineage cells, we found substantial differences in the TF regulators of each state between the individuals convalescing from COVID-19 and healthy donors. Furthermore, B-cell lineage trajectories revealed

an accelerated developmental programme from immune B cells to antibody-producing plasma cells in COVID-19. Integrated analysis of single-cell TCR clonality with the single-cell chromatin accessibility landscape showed the dramatic clonal expansion of CD8[+] T cells in individuals convalescing from COVID-19 as well as the bifurcation of cell-fate decisions for the putative SARS-CoV-2-specific clonally proliferating CD8[+] T cells. This is consistent with effector CD8[+] T cells in initial viral control and memory CD8[+] T cells in long-term immune protection.

Consistent with our data, the reported scRNA-seq profiles of PBMCs identified the abundant IL-1[+]CD14[++] monocytes and activated monocytes in individuals convalescing from COVID-19 (ref. [7]). However, our scATAC-seq profiles revealed further epigenetic changes in individuals convalescing from COVID-19, which can be attributable to the fact that epigenetic changes are more long-lived compared with transcriptional differences. Trained immunity caused by epigenetic and metabolic reprogramming is a 'double-edged sword'. Although it offers broad benefits for the host immune defence against pathogens, it identifies potentially detrimental outcomes in immune-mediated and chronic inflammatory diseases[24,25]. In the context of COVID-19, persistent infection with SARS-CoV-2 results in long-term chromatin accessibility reprogramming, which permits monocytes to remain in a 'trained' functional state. Well-controlled trained immunity is likely to protect against subsequent infection[25]. The effect of cytokine storms on trained immunity should be studied in the future.

Previous studies of immune responses to SARS-CoV-1 and Middle East respiratory syndrome coronavirus (MERS-CoV) can partially offer insights into SARS-CoV-2 memory immunity. Similar to SARS-CoV-2, SARS-CoV-1 and MERS-CoV induce strong inflammatory responses and associated lymphopenia[9]. Furthermore, long-term immune protection has been observed in patients that have recovered from SARS-CoV-1 and MERS-CoV and is attributed to virus-specific antibody responses and long-lived virus-specific memory T-cell responses[2,54,55], which is consistent with our observations on the chromatin reprogramming of B and CD8[+] T cells. Recent studies on memory immune protection against SARS-CoV-2 reinfection, including a rhesus macaque reinfection model and COVID-19 vaccines under phase I, II and III clinical trials, suggest that primary infection with SARS-CoV-2 and vaccines against SARS-CoV-2 can provide partial immune protection[56–59].

The mechanisms of memory immunity—including that in trained and activated monocytes, accelerated B cells and putative SARS-CoV-2-specific clonally expanded effector and memory CD8[+] T cells—to SARS-CoV-2 may be induced by cytokine storms, such as the high production of IFNγ in the serum. As we observed, IFNγ-induced TBET was significantly enriched in the activation state of trained monocytes, B-cell acceleration and CD8[+] T cell-fate decisions, which is consistent with its role in monocyte activation and maturation, IgG[+] memory B-cell development and effector and memory T-cell differentiation[17,20,24,31,39,60–62]. Therefore, the epigenomic regulation of innate and adaptive immune memory responses we demonstrated may not be specific for SARS-CoV-2 and could be elicited following other infections, such as SARS-CoV-1 and MERS. Future studies comparing epigenetic changes among convalescing individuals infected with SARS-CoV-2 and other viruses would be valuable.

Overall, our broad analysis of the epigenomic landscape and TCR profiling demonstrated that individuals convalescing from COVID-19 established immune memory formation and trained immunity via the global remodelling of the chromatin accessibility landscape. The stability of these changes over a longer period requires further study.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information,

References
1. WHO Coronavirus (COVID-19) Dashboard. https://covid19.who.int/ (World Health Organization, 2020).
2. Sariol, A. & Perlman, S. Lessons for COVID-19 immunity from other coronavirus infections. *Immunity* **53**, 248–263 (2020).
3. Vabret, N. et al. Immunology of COVID-19: current state of the science. *Immunity* **52**, 910–941 (2020).
4. Zhou, Z. et al. Heightened innate immune responses in the respiratory tract of COVID-19 patients. *Cell Host Microbe* **27**, 883–890 (2020).
5. Liao, M. et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020).
6. Wilk, A. J. et al. A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nat. Med.* **26**, 1070–1076 (2020).
7. Wen, W. et al. Immune cell profiling of COVID-19 patients in the recovery stage by single-cell sequencing. *Cell Discov.* **6**, 31 (2020).
8. Zheng, Y. et al. A human circulating immune cell landscape in aging and COVID-19. *Protein Cell* **11**, 740–770 (2020).
9. Lucas, C. et al. Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* **584**, 463–469 (2020).
10. Diao, B. et al. Reduction and functional exhaustion of T cells in patients with coronavirus disease 2019 (COVID-19). *Front. Immunol.* **11**, 827 (2020).
11. Del Valle, D. M. et al. An inflammatory cytokine signature predicts COVID-19 severity and survival. *Nat. Med.* **26**, 1636–1643 (2020).
12. Zhu, L. et al. Single-cell sequencing of peripheral mononuclear cells reveals distinct immune response landscapes of COVID-19 and influenza patients. *Immunity* **53**, 685–696 (2020).
13. Granja, J. M. et al. Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia. *Nat. Biotechnol.* **37**, 1458–1465 (2019).
14. Satpathy, A. T. et al. Transcript-indexed ATAC-seq for precision immune profiling. *Nat. Med.* **24**, 580–590 (2018).
15. Chen, S., Lake, B. B. & Zhang, K. High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. *Nat. Biotechnol.* **37**, 1452–1457 (2019).
16. Cao, J. et al. Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science* **361**, 1380–1385 (2018).
17. Lighvani, A. A. et al. T-bet is rapidly induced by interferon-γ in lymphoid and myeloid cells. *Proc. Natl Acad. Sci. USA* **98**, 15137–15142 (2001).
18. Yi, W. et al. T-bet-mediated Tim-3 expression dampens monocyte function during chronic hepatitis C virus infection. *Immunology* **150**, 301–311 (2017).
19. Friedman, A. D. Transcriptional control of granulocyte and monocyte development. *Oncogene* **26**, 6816–6828 (2007).
20. Langlais, D., Barreiro, L. B. & Gros, P. The macrophage IRF8/IRF1 regulome is required for protection against infections and is associated with chronic inflammation. *J. Exp. Med.* **213**, 585–603 (2016).
21. Kikuchi, K. et al. Macrophages switch their phenotype by regulating Maf expression during different phases of inflammation. *J. Immunol.* **201**, 635–651 (2018).
22. Hanna, R. N. et al. The transcription factor NR4A1 (Nur77) controls bone marrow differentiation and the survival of Ly6C[−] monocytes. *Nat. Immunol.* **12**, 778–785 (2011).
23. Ramos-Mejia, V. et al. HOXA9 promotes hematopoietic commitment of human embryonic stem cells. *Blood* **124**, 3065–3075 (2014).
24. Netea, M. G. et al. Defining trained immunity and its role in health and disease. *Nat. Rev. Immunol.* **20**, 375–388 (2020).
25. Netea, M. G. et al. Trained Immunity: a tool for reducing susceptibility to and the severity of SARS-CoV-2 infection. *Cell* **181**, 969–977 (2020).
26. Saeed, S. et al. Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. *Science* **345**, 1251086 (2014).
27. Li, P. et al. IRF8 and IRF3 cooperatively regulate rapid interferon-β induction in human blood monocytes. *Blood* **117**, 2847–2854 (2011).
28. Fontana, M. F. et al. JUNB is a key transcriptional modulator of macrophage activation. *J. Immunol.* **194**, 177–186 (2015).
29. Behmoaras, J. et al. *Jund* is a determinant of macrophage activation and is associated with glomerulonephritis susceptibility. *Nat. Genet.* **40**, 553–559 (2008).
30. Nutt, S. L. & Kee, B. L. The transcriptional regulation of B cell lineage commitment. *Immunity* **26**, 715–725 (2007).

31. Satpathy, A. T. et al. Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. *Nat. Biotechnol.* **37**, 925–936 (2019).

32. Stehling-Sun, S., Dade, J., Nutt, S. L., DeKoter, R. P. & Camargo, F. D. Regulation of lymphoid versus myeloid fate 'choice' by the transcription factor Mef2c. *Nat. Immunol.* **10**, 289–296 (2009).

33. Mackay, F. & Browning, J. L. BAFF: a fundamental survival factor for B cells. *Nat. Rev. Immunol.* **2**, 465–475 (2002).

34. Gutierrez, A. Jr. et al. LEF-1 is a prosurvival factor in chronic lymphocytic leukemia and is expressed in the preleukemic state of monoclonal B-cell lymphocytosis. *Blood* **116**, 2975–2983 (2010).

35. Ng, A. P. et al. An Erg-driven transcriptional program controls B cell lymphopoiesis. *Nat. Commun.* **11**, 3013 (2020).

36. Milanovic, M. et al. Differential requirements for the canonical NF-κB transcription factors c-REL and RELA during the generation and activation of mature B cells. *Immunol. Cell Biol.* **95**, 261–271 (2017).

37. Almaden, J. V. et al. B-cell survival and development controlled by the coordination of NF-κB family members RelB and cRel. *Blood* **127**, 1276–1286 (2016).

38. Vilagos, B. et al. Essential role of EBF1 in the generation and function of distinct mature B cell types. *J. Exp. Med.* **209**, 775–792 (2012).

39. Wang, N. S. et al. Divergent transcriptional programming of class-specific B cell memory by T-bet and RORα. *Nat. Immunol.* **13**, 604–611 (2012).

40. De Silva, N. S., Silva, K., Anderson, M. M., Bhagat, G. & Klein, U. Impairment of mature B cell maintenance upon combined deletion of the alternative NF-κB transcription factors RELB and NF-κB2 in B cells. *J. Immunol.* **196**, 2591–2601 (2016).

41. De Silva, N. S. et al. Transcription factors of the alternative NF-κB pathway are required for germinal center B-cell development. *Proc. Natl Acad. Sci. USA* **113**, 9063–9068 (2016).

42. Ohkubo, Y. et al. A role for c-fos/activator protein 1 in B lymphocyte terminal differentiation. *J. Immunol.* **174**, 7703–7710 (2005).

43. Heise, N. et al. Germinal center B cell maintenance and differentiation are controlled by distinct NF-κB transcription factor subunits. *J. Exp. Med.* **211**, 2103–2118 (2014).

44. Willis, S. N. et al. Environmental sensing by mature B cells is controlled by the transcription factors PU.1 and SpiB. *Nat. Commun.* **8**, 1426 (2017).

45. Xu, H. et al. Regulation of bifurcating B cell trajectories by mutual antagonism between transcription factors IRF4 and IRF8. *Nat. Immunol.* **16**, 1274–1281 (2015).

46. Ubieta, K. et al. Fra-2 regulates B cell development by enhancing IRF4 and Foxo1 transcription. *J. Exp. Med.* **214**, 2059–2071 (2017).

47. Ise, W. et al. The transcription factor BATF controls the global regulators of class-switch recombination in both B cells and T cells. *Nat. Immunol.* **12**, 536–543 (2011).

48. Man, K. et al. Transcription factor IRF4 promotes CD8+ T cell exhaustion and limits the development of memory-like T cells during chronic infection. *Immunity* **47**, 1129–1141 (2017).

49. Kidani, Y. et al. Sterol regulatory element-binding proteins are essential for the metabolic programming of effector T cells and adaptive immunity. *Nat. Immunol.* **14**, 489–499 (2013).

50. Weinreich, M. A. et al. KLF2 transcription-factor deficiency in T cells results in unrestrained cytokine production and upregulation of bystander chemokine receptors. *Immunity* **31**, 122–130 (2009).

51. Lai, D. et al. KLF13 sustains thymic memory-like CD8+ T cells in BALB/c mice by regulating IL-4-generating invariant natural killer T cells. *J. Exp. Med.* **208**, 1093–1103 (2011).

52. Das, H. et al. Kruppel-like factor 2 (KLF2) regulates proinflammatory activation of monocytes. *Proc. Natl Acad. Sci. USA* **103**, 6653–6658 (2006).

53. Bai, A., Hu, H., Yeung, M. & Chen, J. Kruppel-like factor 2 controls T cell trafficking by activating L-selectin (CD62L) and sphingosine-1-phosphate receptor 1 transcription. *J. Immunol.* **178**, 7632–7639 (2007).

54. Ng, O. W. et al. Memory T cell responses targeting the SARS coronavirus persist up to 11 years post-infection. *Vaccine* **34**, 2008–2014 (2016).

55. Tang, F. et al. Lack of peripheral memory B cell responses in recovered patients with severe acute respiratory syndrome: a six-year follow-up study. *J. Immunol.* **186**, 7264–7268 (2011).

56. Deng, W. et al. Primary exposure to SARS-CoV-2 protects against reinfection in rhesus macaques. *Science* **369**, 818–823 (2020).

57. Mulligan, M. J. et al. Phase I/II study of COVID-19 RNA vaccine BNT162b1 in adults. *Nature* **586**, 589–593 (2020).

58. Edwards, K. M. Vaccines targeting SARS-CoV-2 tested in humans. *Nat. Med.* **26**, 1336–1338 (2020).

59. Jackson, L. A. et al. An mRNA vaccine against SARS-CoV-2—preliminary report. *N. Engl. J. Med.* **383**, 1920–1931 (2020).

60. Intlekofer, A. M. et al. Effector and memory CD8+ T cell fate coupled by T-bet and eomesodermin. *Nat. Immunol.* **6**, 1236–1244 (2005).

61. Stone, S. L. et al. T-bet transcription factor promotes antibody-secreting cell differentiation by limiting the inflammatory effects of IFN-γ on B cells. *Immunity* **50**, 1172–1187 (2019).

62. Li, H., Wojciechowski, W., Dell'Agnola, C., Lopez, N. E. & Espinoza-Delgado, I. IFN-γ and T-bet expression in human dendritic cells from normal donors and cancer patients is controlled through mechanisms involving ERK-1/2-dependent and IL-12-independent pathways. *J. Immunol.* **177**, 3554–3563 (2006).

Methods

Subjects and specimen collection. We collected blood from five healthy donors, seven convalescing individuals who had recovered from moderate COVID-19 and three convalescing individuals who had recovered from severe COVID-19. They were enrolled at the Fifth Medical Center of PLA General Hospital in May 2020. We isolated PBMCs from the study participants using Ficoll solution according to the manufacturer's instructions. We defined individuals convalescing from COVID-19 as patients who had been diagnosed as infected with COVID-19 at admission by real-time RT–PCR, and had recovered and been discharged from the hospital at least one month before participating in the study. The individuals convalescing from COVID-19 were classified into mild and severe groups based on the Fifth Revised Trial Version of the Novel Coronavirus Pneumonia Diagnosis and Treatment Guidance. The healthy donors—three women and two men, 18–55 years old—were defined by negative test results in real-time RT–PCR analysis of the SARS-CoV-2 gene and IgG antibody detection. This study and study protocol was approved by the Ethics Committees of the Fifth Medical Center of Chinese PLA General Hospital, Beijing, China (2020005D) on healthy volunteers or individuals convalescing from COVID-19 in Medical Research and written informed consent was obtained from all of the participants. All of the participants received participant compensation. The clinical features of these individuals convalescing from COVID-19 and healthy donors are listed in Supplementary Table 1.

Single T-cell index sorting. Human PBMCs were thawed after storage in liquid nitrogen and allowed to recover overnight in R10 culture medium (RPMI1640 medium supplemented with 10% fetal calf serum, 25 mM HEPES, 1×non-essential amino acids, 50 μM β-mercaptoethanol and 1×penicillin and streptomycin). Cell surface staining was performed with 4,6-diamidino-2-phenylindole (DAPI) and the following antibodies: Alexa Fluor 700 anti-human CD3 (clone SK7, 1:100; BioLegend, cat. no. 344822), FITC anti-human TCRαβ (clone IP26, 1:50; BioLegend, cat. no. 306706), PerCP/Cyanine5.5 anti-human CD8 (clone SK1, 1:200; BioLegend, cat. no. 344710), PE/Dazzle 594 anti-human CD4 (clone A161A1, 1:200; BioLegend, cat. no. 357411), PE anti-human CD25 (clone BC96, 1:200; BioLegend, cat. no. 302606), Brilliant Violet 650 anti-human CD127 (clone A019D5, 1:200; BioLegend, cat. no. 351326), APC anti-human CD137 (4-1BB) (clone 4B4-1, 1:100; BioLegend, cat. no. 309810) and Brilliant Violet 605 anti-human CD279 (PD-1) (clone EH12.2H7, 1:100; BioLegend, cat. no. 329923). Single viable DAPI⁻CD3⁺TCRαβ⁺ T cells were immediately index sorted into 96-well plates for Ti-ATAC-seq analysis using a BD FACS Aria III system.

Preparation of 10x-based scATAC-seq libraries. After storage in liquid nitrogen, PBMCs were thawed, allowed to recover overnight in R10 culture medium and washed three times with PBS buffer to remove debris and aggregated cells. The cell viability of each sample exceeded 90%. For the healthy donor controls, we performed 10x-based scATAC-seq on PBMCs from five healthy donors, including a sample of mixed PBMCs from three healthy donors. The cells were lysed, the nuclei were isolated, the nuclei suspensions were washed and counted, and the transposition reaction and nuclei barcoding were performed according to the manufacturer's instructions. Approximately 5,000–10,000 nuclei were collected in each sample. The protocol for library preparation and the settings for the instrument and sequencing were provided by the manufacturer (Annoroad and Berry Genomics) and are available at https://support.10xgenomics.com/single-cell-atac. We also collected scATAC-seq data from a healthy donor control from a published dataset (GSE139369)[13].

Preparation of single-cell Ti-ATAC-seq libraries. Single antibody-indexed cells were sorted into 96-well plates with 6 μl ATAC-RSB buffer (10 mM Tris–HCl at pH 7.5, 10 mM NaCl, 3 mM MgCl₂ and 0.8 U μl⁻¹ RNase inhibitor) per well. Next, 1.5 μl transposition reaction mix with 0.1 μl TTE Mix V5 (Vazyme Biotech), 50 mM TAPS–NaOH (pH 8.5), 25 mM MgCl₂, 50% dimethylformamide, 0.8 U μl⁻¹ RNase inhibitor and 0.2% NP40 was carefully added to each well. The transposition reaction was performed by incubating the mixture at 37 °C for 30 min and then stopped by the addition of 1.5 μl release buffer (125 mM EDTA and 10 mM Tris–HCl at pH 8.0), following which the reaction mixture was incubated at 50 °C for 30 min. Quenching buffer (1 μl; 187.5 mM MgCl₂ and 10 mM Tris–HCl at pH 8.0) was then added for MgCl₂ quenching. The *TRA* and *TRB* chains and transposed DNA fragments were reverse transcribed with 11 μl of the RT–PCR reaction mix (8.8 μl of 2×1 Step buffer, 1.5 μl of TRA and TRB RT primer mix, 0.2 μl of non-indexed custom Nextera ATAC-seq PCR primer mix (forward, 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3′ and reverse, 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3′) and 0.5 μl PrimeScript 1 Step enzyme mix (Takara Bio)) under the following RT–PCR conditions: 50 °C for 36 min; 95 °C for 15 min; and eight cycles at 94 °C for 30 s, 62 °C for 1 min and 72 °C for 1 min. The RT–PCR product was then immediately equally divided into two 96-well plates for amplification and the barcoding of transposed DNA fragments and *TRA* and *TRB* chains, which were performed as previously described.[63] Briefly, 1.25 μM of dual-index Nextera ATAC-seq PCR primers with 1×TruePrep Amplify reaction mix (Vazyme Biotech) was prepared, and ATAC-seq amplification and barcoding were performed under the following PCR conditions: 72 °C for 5 min, 98 °C for 30 s, and 15 cycles of 98 °C for 15 s and 63 °C for 30 s.

Finally, scTCR-seq and scATAC-seq libraries were respectively pooled and purified on a single Zymo DNA clean and concentrator 5 column. Fragment-length selection was performed (0.5×/1.5× for scATAC-seq and 300–400 bp for scTCR-seq). The scTCR-seq and scATAC-seq libraries were 150-bp paired-end sequenced on a HiSeq X Ten platform (Illumina) with the sequencing read length and dual indexing according to the manufacturer's instructions (Annoroad).

Data processing of scTCR-seq libraries. The scTCR-seq data were analysed as previously described[63,64]. Briefly, raw reads were unpacked and joined by shared regions and demultiplexed using a custom pipeline. The TCR V, D and J segments were analysed and assigned using MiXCR[65].

Primary data pre-processing of scATAC-seq libraries. We used a previously described workflow with minor modifications[14]. Briefly, a sample name was added to the header of each read, adaptor sequences were trimmed using Cutadapt, reads were mapped to the hg38 human genome with Bowtie2 and those aligned to the mitochondria were removed. The resulting cleaned single-cell files (./bam files) of each individual convalescing from COVID-19 were merged and sorted by read name using SAMtools. Finally, the resulting merged file of each individual convalescing from COVID-19 was converted to a fragment file, the Tn5 insertion site was adjusted and unique nuclear fragments were retained using BEDtools. We then calculated the TF deviation of cells using chromVAR[66].

For the pre-processing of the 10x-based scATAC-seq data, we used the 'cellranger-atac count' function (cellranger-atac, v1.2.0) to generate single-cell accessibility counts for each library (https://support.10xgenomics.com/single-cell-atac/software/pipelines/latest/using/count). Reads were aligned to the hg38 human genome. Similarly, we created an Arrow file for each resulting fragment file and included those single-cell libraries with at least 1,000 unique fragments and a transcription-start-site enrichment of eight using the ArchR package[67] in the R statistical environment (v3.5.1). Finally, we created an ArchRProject for downstream analysis by combining all of the Arrow files of the 10x-based scATAC-seq.

Dimensionality reduction and clustering analysis. We used the 'addIterativeLSI' function of ArchR to perform iterative latent semantic indexing[31,67]. We then used the harmony algorithm to correct for batch-effect differences[68] and added clusters using the 'addClusters' function. We ran UMAP using the 'addUMAP' function and plotted the results using the 'plotEmbedding' function in ArchR[67].

Identification of marker features and differential analysis for clusters. We created pseudo-bulk replicates using the 'addGroupCoverages' function for each cluster and called peaks using the 'findMacs2' function. We used the 'addDeviationsMatrix' function to compute per-cell deviations across all motif annotations to create the deviation matrix 'MotifMatrix'. Next, we used the 'getMarkerFeatures' function with 'bias' parameter to account for transcription-start-site enrichment and the number of unique fragments per cell to identify the markers of clusters, including peaks, genes (based on gene scores) and TF motifs (based on chromVAR deviations). We applied the 'getMarkers' function to obtain the marker list of each cluster, the 'addImputeWeights' function to impute the weights of marker features and the 'plotEmbedding' function to visualize the marker features. We performed pairwise comparisons of peaks, genes and TFs for the indicated two clusters using the getMarkerFeatures function and plotted an MA or volcano plot using the 'markerPlot' function in ArchR[67].

Plotting browser tracks of *cis*-element co-accessibility. Co-accessibility represents the correlation between accessibility peaks across many single cells and is useful for identifying cell type-specific peaks. We used the 'addCoAccessibility' function to examine co-accessibility in ArchR[67], which returned a loop track that represented the co-accessibility information using the 'getCoAccessibility' function. Finally, we plotted the genome browser tracks of peaks and co-accessibility using the 'plotBrowserTrack' function.

TF footprinting. We performed TF footprinting to precisely predict the binding location of a TF at a particular locus as previously described[31,67] using ArchR. The Tn5 bias signal was normalized using the 'Divide' strategy.

Trajectory construction. We performed cellular trajectory analyses as previously described using ArchR[67]. First, we defined the trajectory backbone of cell groups or clusters based on the cell differentiation states, such as the order of naive, intermediate and effector CD8⁺ cells in the effector CD8⁺ T-cell trajectory. We then created a trajectory using the 'addTrajectory' function and plotted the pseudotime values on UMAP embedding using the 'plotTrajectory' function. We next plotted pseudotime heatmaps of TFs, gene scores and peak accessibility using the 'plotTrajectoryHeatmap' function. To identify positive TF regulators, we performed an integrative analysis of gene scores and motif accessibility across pseudotime using the 'correlateTrajectories' and plotTrajectoryHeatmap functions.

Statistics and reproducibility. Statistical analyses were performed in GraphPad Prism (v6) or R (version 3.5.1). For the frequencies of the identified clusters and the virus rechallenge assay, a two-sided unpaired Student's *t*-test was performed

and the *P* values are reported. For the differential TF motif accessibility, differential *cis*-element accessibility and Gene Ontology analyses, multiple-test corrections were performed, the *P* values were calculated using a two-sided pairwise Wilcoxon test and the false discovery rate was corrected using the Benjamini–Hochberg procedure. No data were excluded from the analyses.

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

Data availability

Data were deposited in the Sequence Read Archive (SRA) under the accession number PRJNA718009. Data were also deposited in the Genome Sequence Archive for Human (GSA-Human) under the accession HRA000562. Source data are provided with this study. The minimum dataset is posted on GitHub (https://github.com/Yang-Chen-Lab-co/COVID19-scATAC-seq). The previously published data that were re-analysed here are available under the accession numbers GSE139369, GSE117089, GSE126074 and GSE107817. All other data supporting the findings of this study are available from the corresponding authors on reasonable request. Source data are provided with this paper.

Code availability

All of the R packages that were used are available online, as described in the Methods. Customized code used in this study has been posted on GitHub (https://github.com/Yang-Chen-Lab-co/COVID19-scATAC-seq).

References

63. Han, A., Glanville, J., Hansmann, L. & Davis, M. M. Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat. Biotechnol.* **32**, 684–692 (2014).
64. Glanville, J. et al. Identifying specificity groups in the T cell receptor repertoire. *Nature* **547**, 94–98 (2017).
65. Bolotin, D. A. et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods* **12**, 380–381 (2015).
66. Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat. Methods* **14**, 975–978 (2017).
67. Granja, J. M. et al. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat. Genet.* **53**, 403–411 (2021).
68. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

Author contributions

P.Y., M.M.D. and L.C. supervised the overall project design and execution. M.Y. and L.C. performed most of the experiments and analyses, and wrote the manuscript. D.Z. and P.Z. recruited the human participants with assistance from Z.C. and E.-Q.Q. Y.G. assisted with experiments. All authors read and gave comments to this manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41556-021-00690-1.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41556-021-00690-1.

Correspondence and requests for materials should be addressed to L.C., M.M.D. or P.Y.

Peer review information *Nature Cell Biology* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.
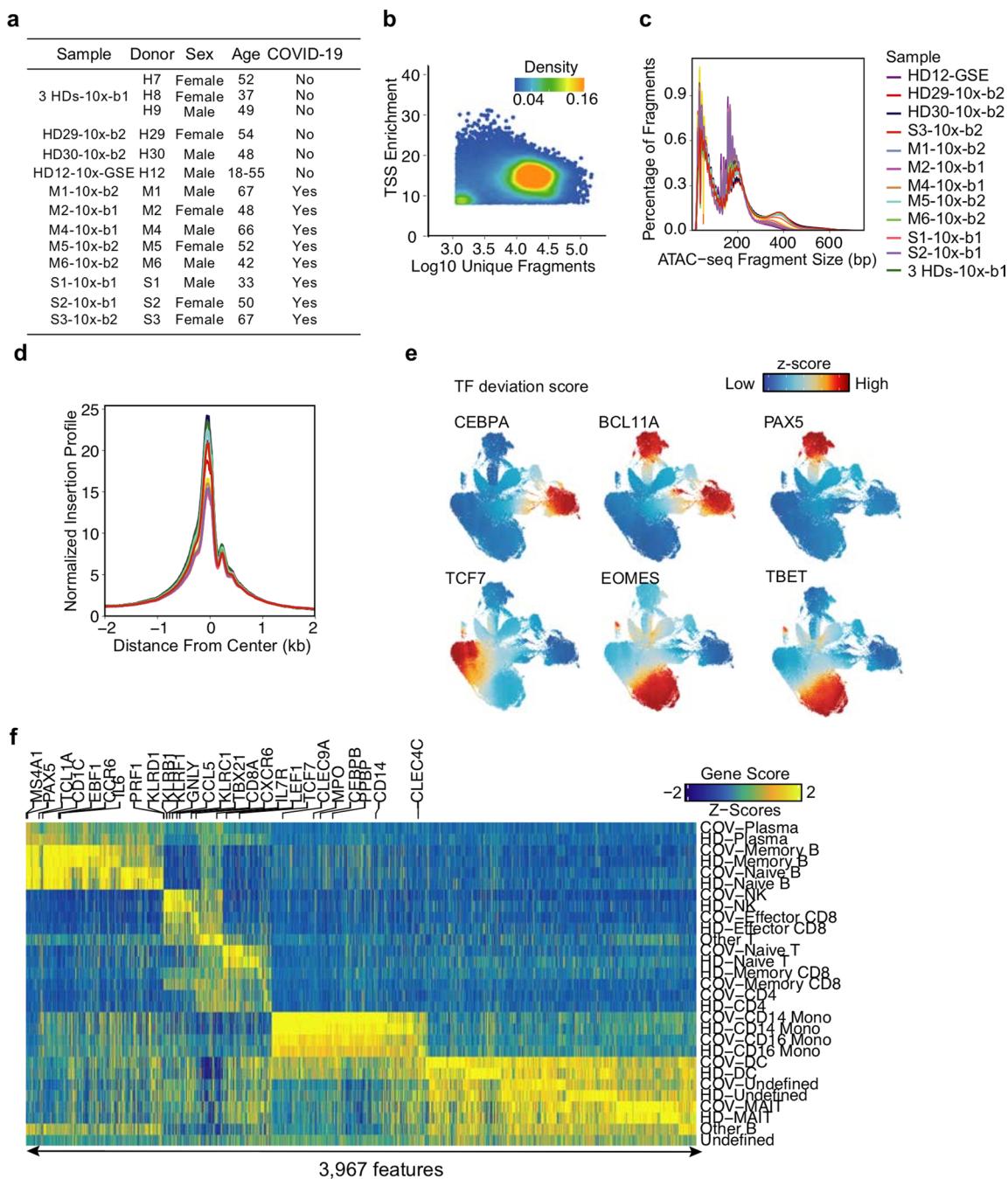
Extended Data Fig. 1 | Quality control for the 10x-based scATAC-seq datasets. a, Table showing the information of 10x scATAC-seq based samples.
b, the transcriptional start site (TSS) enrichment score compared with the number of unique ATAC-seq nuclear fragments in each single cell that passed.
c, Nucleosomal periodicity fragment lengths of aggregate single-cell profiles of 10x scATAC-seq based samples. **d**, the enrichments of normalized
Tn5 insertions around the TSSs of 10x scATAC-seq based samples. **e**, ChromVAR TF motif bias-corrected deviations overlaid on UMAP projection
of 10x-scATAC-seq-based single PBMCs as showed in Fig. 1c. **f**, Heatmap showing the 3,967 differentially accessible genes (FDR = < 0.05 and fold
change > =2) in various clusters of different sample origin (HD or COV) as indicated in Fig. 1c.
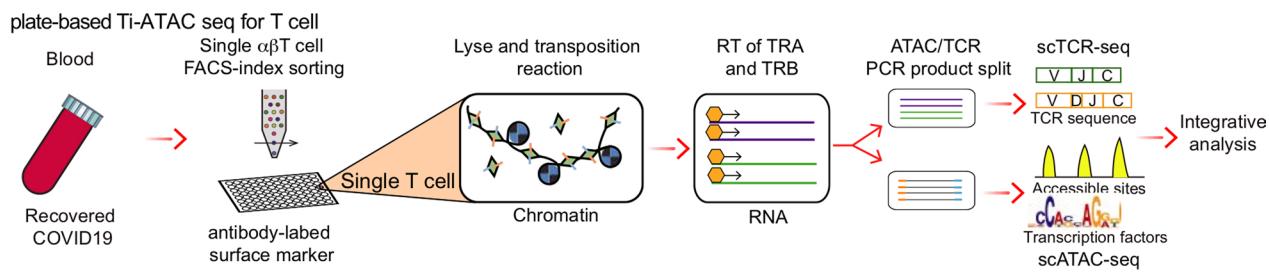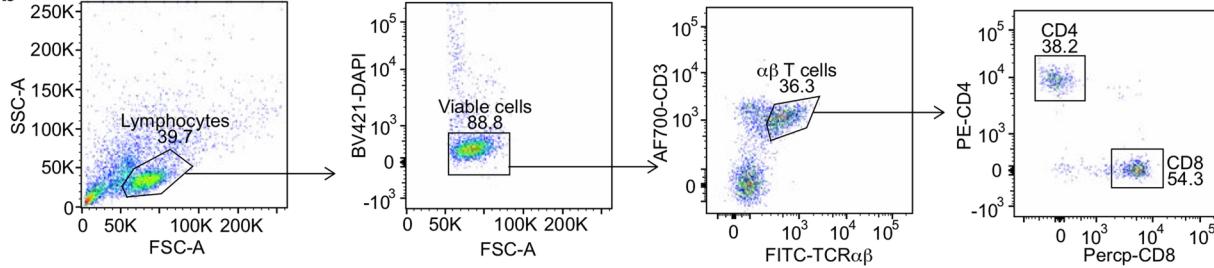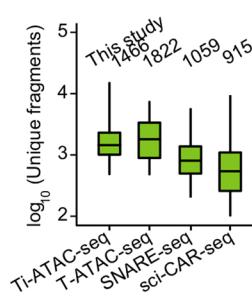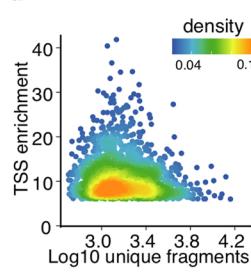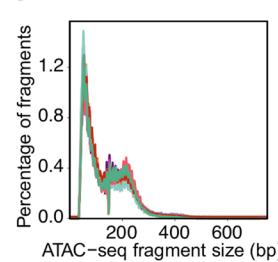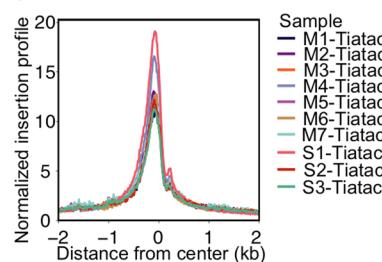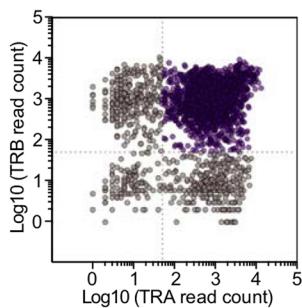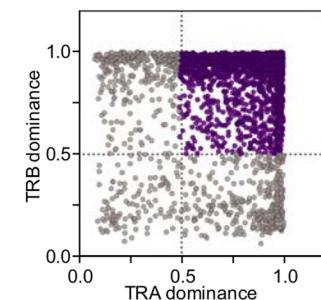
Extended Data Fig. 2 | Workflow and quality control for the Ti-ATAC-seq datasets. a, Workflow of plated-based single T cell Ti-ATAC-seq. Briefly, single DAPI⁻CD3⁺TCRαβ⁺ T cells were FACS index sorted into 96-well plates, transposed as single cells with the Tn5 transposase, reverse transcribed to TRA and TRB mRNA, and then the ATAC-seq product and TCR-seq product were split for further scATAC-seq and scTCR-seq analyses**. b,** Shown are the gating strategy to gate and sort single viable DAPI⁻CD3 + TCRαβ + T cells into 96-well plates. **c,** Number of unique ATAC-seq fragments for Ti-ATAC-seq (this study, n = 1,574 cells), T-ATAC-seq (n = 300 cells, GSE107817), SNARE-seq (n = 1048 cells, GSE126074), and sci-CAR-seq (n = 1000 cells, GSE117089). Boxplots show the medians and the quartile range (25% and 75%), and 1.5x interquartile ranges (IQRs) of the lengths of whiskers. **d,** TSS enrichment score compared with the number of unique ATAC-seq nuclear fragments in each single cell that passed. **e,** Nucleosomal periodicity fragment lengths of aggregate single-cell profiles of Ti-ATAC-seq based samples. **f,** the enrichments of normalized Tn5 insertions around the TSSs of Ti-ATAC-seq based samples. **g,** the read counts for TRA and TRB sequences of each single T cell TCR clone in Ti-ATAC-seq-based scTCR-seq data. **h,** TRA and TRB dominance of the top clone for each single T cell TCR clone in Ti-ATAC-seq-based scTCR-seq data. Each dot represents a dominant TCR clone of a single T cell.
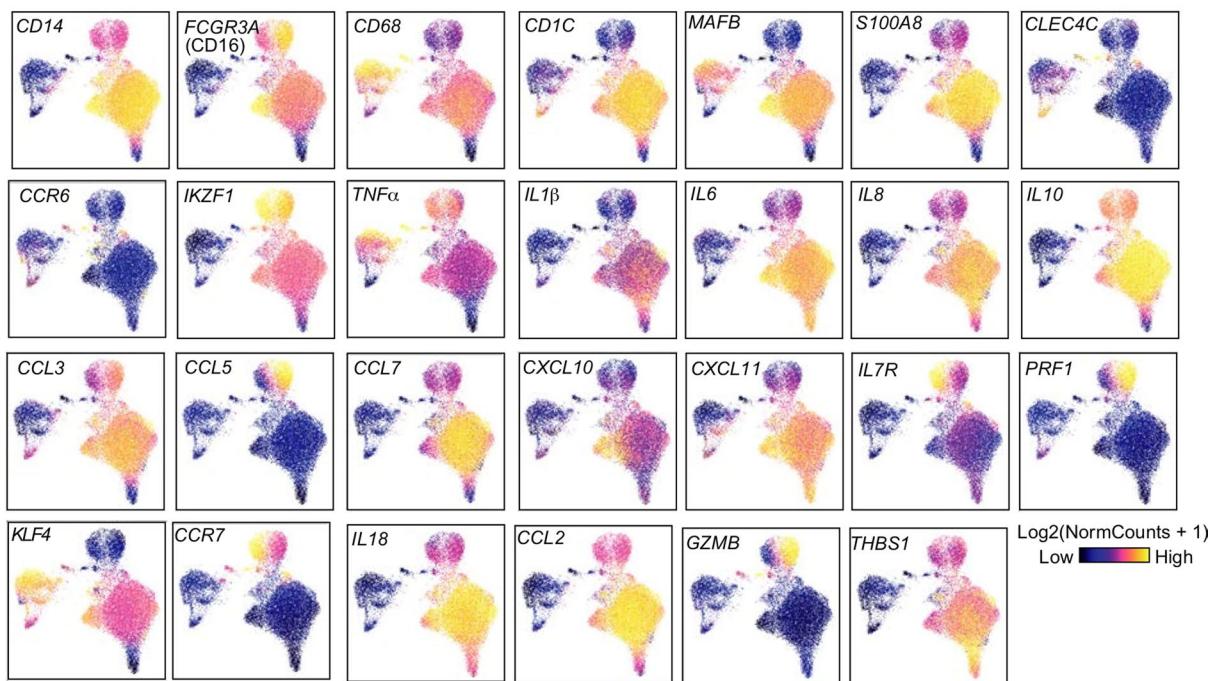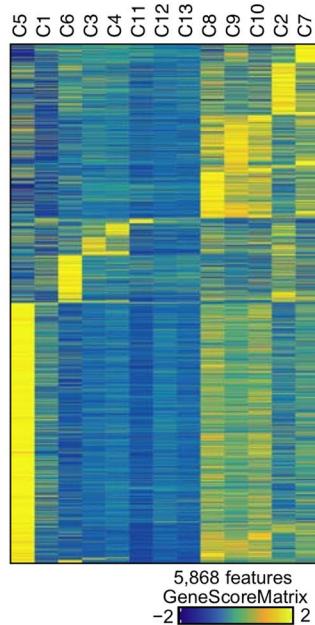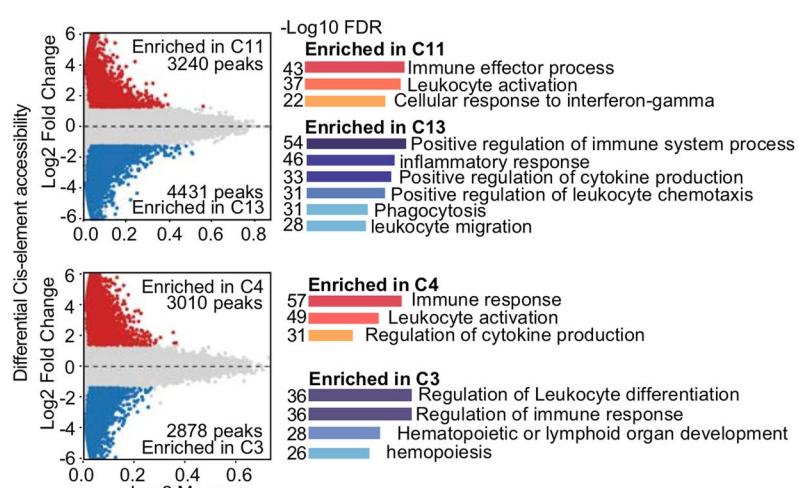
Extended Data Fig. 3 | Differentially accessible genes in various monocyte clusters. a, UMAP plot of monocytes as indicated in Fig. 2a coloured by log-normalized gene scores. **b**, Heatmap showing the 5,868 differentially accessible genes (FDR =< 0.05 and fold change > =2) in various monocyte clusters as indicated in Fig. 2a. **c**, Left: MA plots showing the differential ATAC-seq peaks (red or blue dot: log2 fold change > = 1 and false discovery rate (FDR) =< 0.05) for the indicated clusters including cluster C11 versus C13, cluster C4 versus C3. Right: Bar charts showing the top GO biological process pathways of differential peaks enriched in each cluster using GREAT analysis.
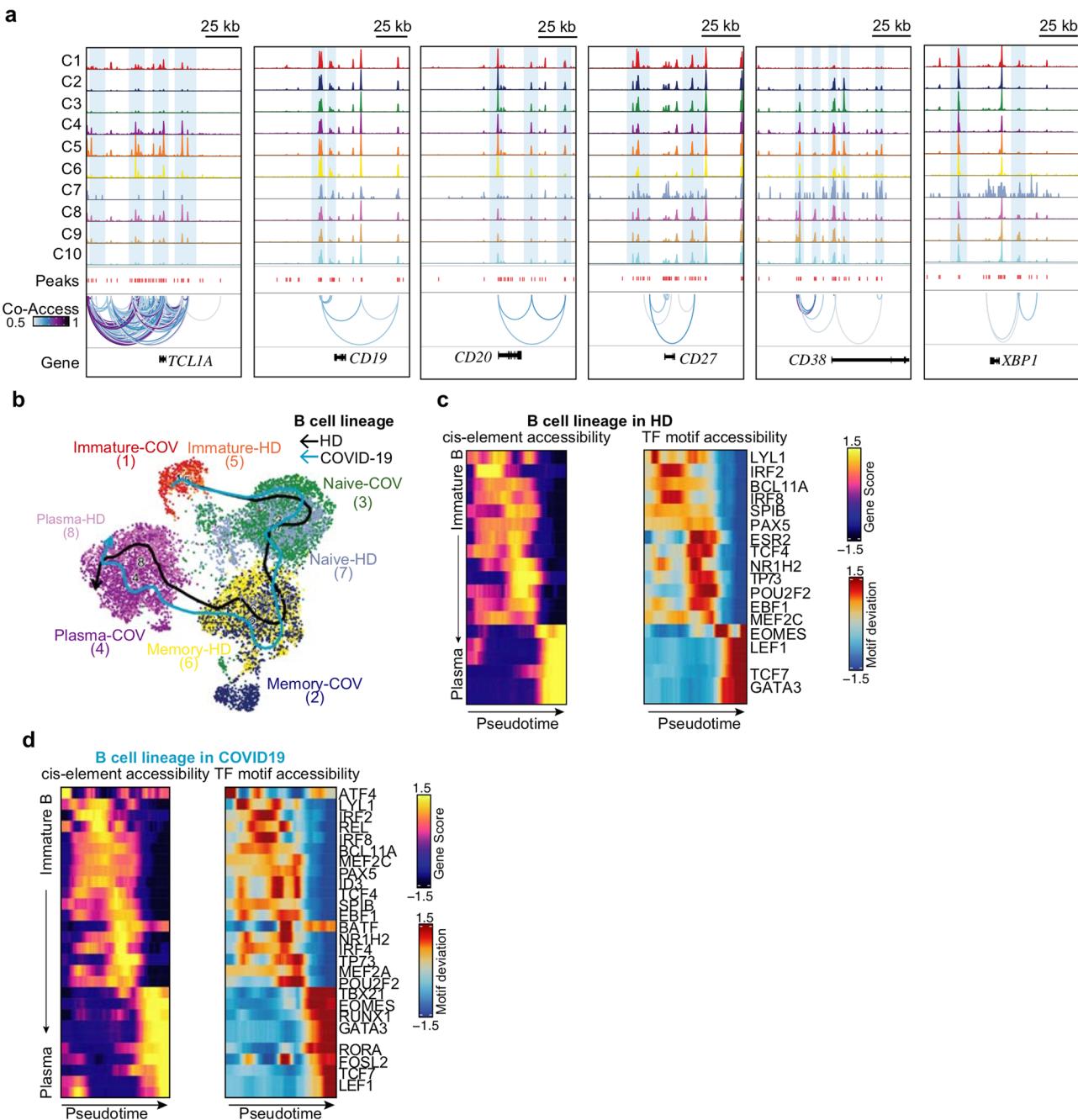
Extended Data Fig. 4 | Lineage trajectory of B cells of healthy donors and individuals convalescing from COVID-19. a, Aggregated single-cell genome tracks for clusters as indicated in Fig. 4a at the *TCL1A, CD19, CD20 (MS4A1), CD27, CD38* and *XBP1* gene loci with peak co-accessibility (Co-Access). **b**, Lineage trajectory of B-cell states in healthy donors and COVID-19 convalescent individuals as indicated in Fig. 4a, respectively. **c**, Heatmaps showing the positive TF regulators by the ordered TF gene scores (right) with ordered TF motif accessibility (left) across pseudotime for B-cell differentiation in healthy donors. Positive TF regulators are TF motifs that show high bias-corrected chromVAR TF motif deviations that also exhibit similarly dynamic gene scores across differentiation state. **d**, Heatmaps showing the positive TF regulators by the ordered TF gene scores (right) with ordered TF motif accessibility (left) across pseudotime for B-cell differentiation in COVID-19 convalescent individuals.
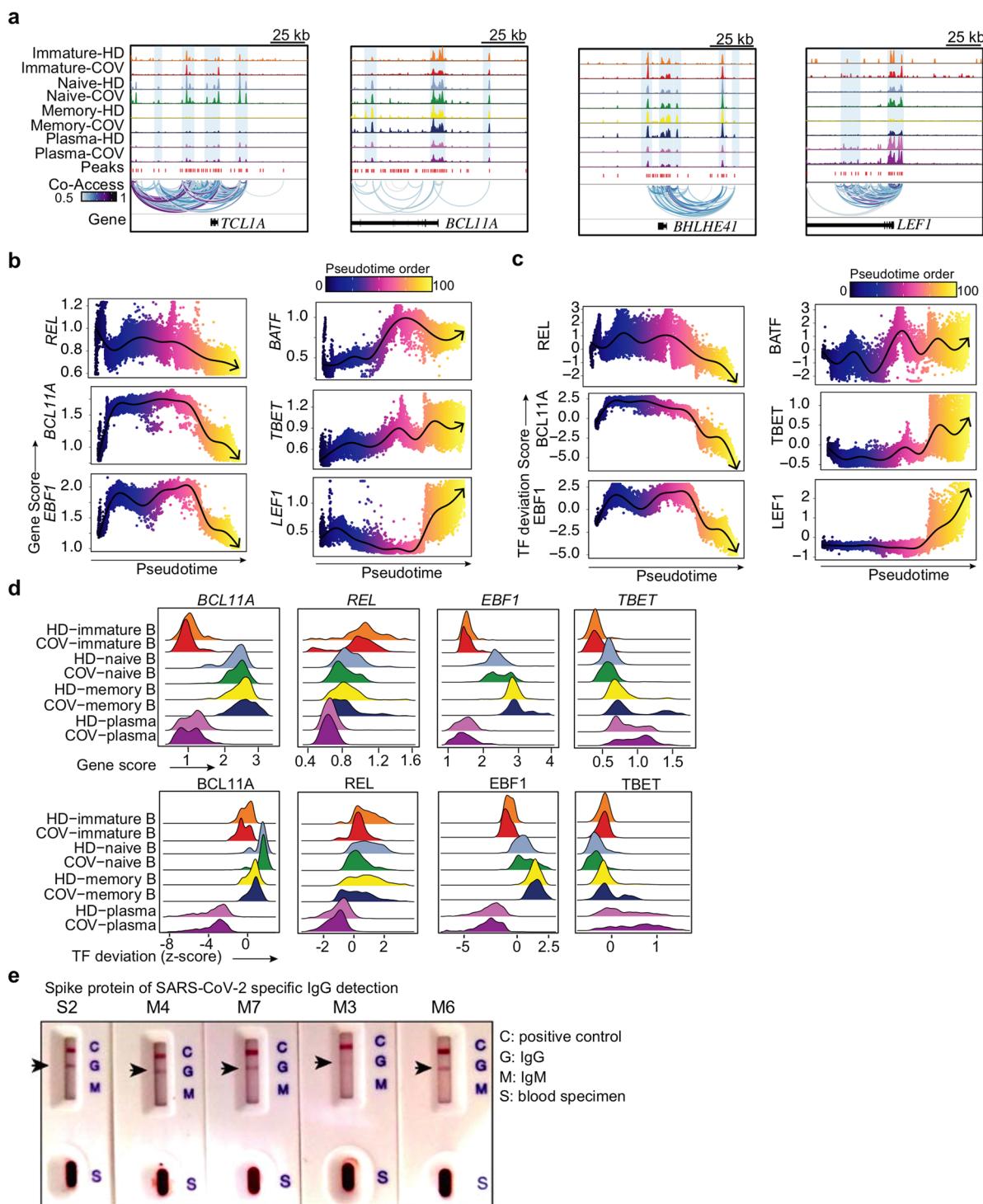
Extended Data Fig. 5 | Integrative lineage trajectory of B-cell states in healthy donors and individuals convalescing from COVID-19. a, Aggregated single-cell genome tracks for clusters as indicated in Fig. 4d at the *TCL1A, BCL11A, BHLHE41*, and *LEF1* gene loci with peak co-accessibility (Co-Access). **b**, Single-cell gene scores of the positive TF regulators as defined in Fig. 4f across pseudotime in the B-cell trajectory. Dots represent gene scores of an individual pseudotime-ordered scATAC-seq profiles. The smoothed line and arrow represent the visualization of the trajectory path from the spline fit. **c**, Single-cell bias-corrected chromVAR TF motif deviation scores of the positive TF regulators as defined in Fig. 4f across pseudotime in the B-cell trajectory. **d**, Ridge plots showing TF gene scores (top) and bias-corrected chromVAR TF motif deviation scores (bottom) of the indicated positive TF regulators across different states and samples origins as indicated in Fig. 4d. **e**, Shown are the SARS-CoV-2 virus specific IgG detection assays in the 5 indicated convalescent individuals.
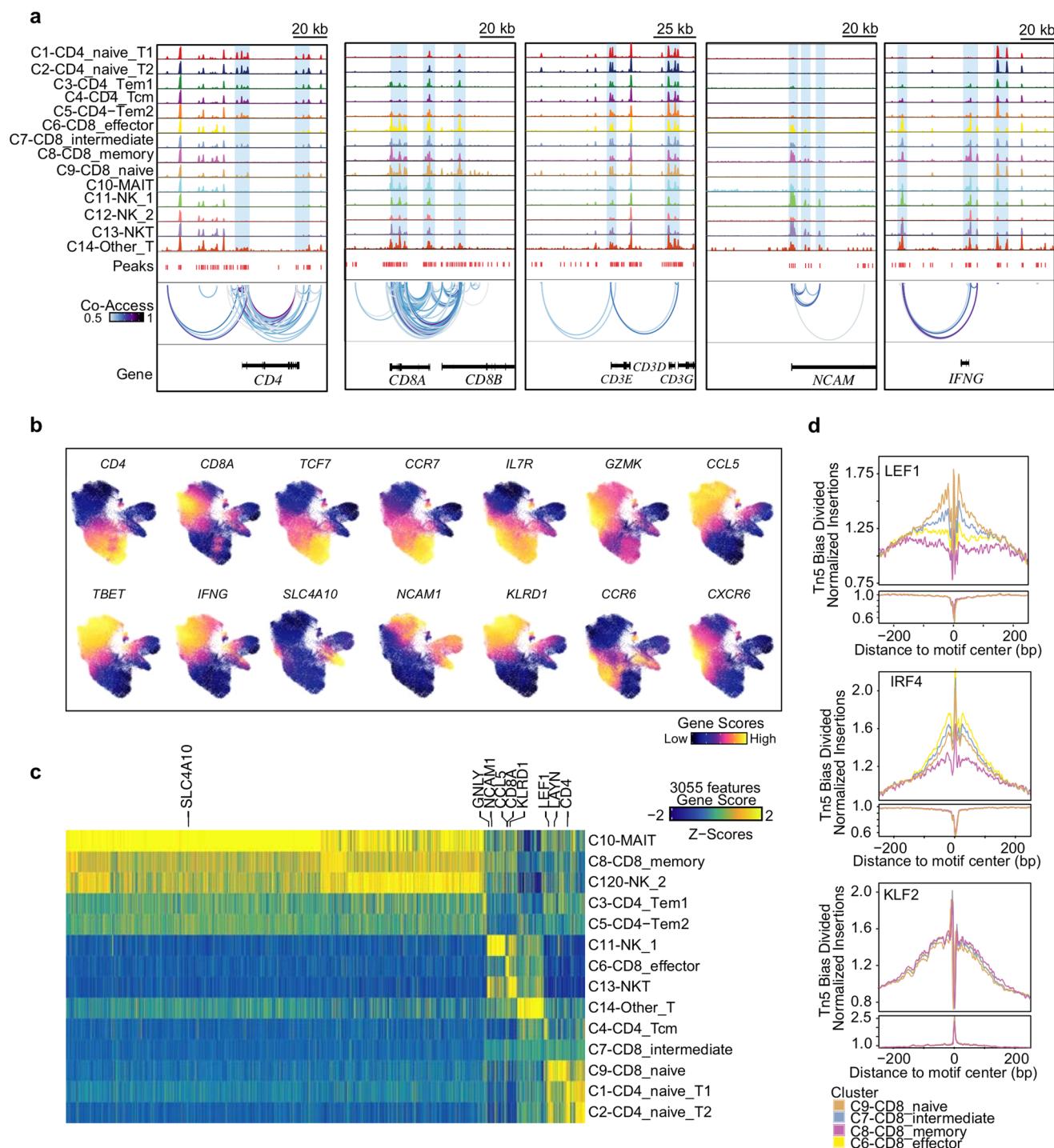
Extended Data Fig. 6 | Differentially accessible genes in various T- and NK-cell clusters. a, Aggregated single-cell genome tracks for clusters as indicated in Fig. 5a at the *CD4*, *CD8A*, *CD8B*, *CD3E*, *CD3D*, *CD3G*, *NCAM* and *IFNG* gene loci with peak co-accessibility (Co-Access). **b**, Canonical surface markers overlaid on UMAP embedding as indicated in Fig. 5a coloured by log-normalized gene scores for indicated genes. **c**, Heatmap showing the 3,055 differentially accessible genes (FDR = < 0.05 and fold change > =2) in various clusters as indicated in Fig. 5a. **d**, TF footprint of the LEF1, IRF4, and KLF4 motifs in the CD8+ T cell clusters as indicated in Fig. 5a. The Tn5 insertion bias track is shown below.

nature research

Corresponding author(s): Pengyuan Yang, Mark M. Davis, Liang Chen

Last updated by author(s): Apr 1, 2021

Reporting Summary

Nature Research 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 Research 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	cellranger-atac (v1.2.0). Code used in this study is posted on GitHub for main analysis (https://github.com/YouMaojun/COVID19-scATAC-seq)
Data analysis	Software used: R statistical environment (v3.5.1), bowtie2 (v2.3.4.1), cutadapt (v2.10), samtools (v1.7), GraphPad Prism (v6), bedtools (v2.27.1), MiXCR (v3.0.5), FlowJo V10. R package: ArchR (v0.9.5), ChromVAR (v1.8.0), Harmony (v0.99.9)

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 Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

Data were deposited in the Sequence Read Archive (SRA) under the accession PRJNA718009. Data were also deposited in the Genome Sequence Archive for Human (GSA-Human) under the accession HRA000562. Source data are provided with this study. The minimum dataset is posted on GitHub (https://github.com/YouMaojun/COVID19-scATAC-seq). There are no restrictions on data availability or use.
Previously published data were re-analyzed here are available under accession GSE139369,GSE117089,GSE126074, GSE107817.

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	For single cell analysis, the profiles obtained from over 100,000 single cells from 8 COVID-19 convalescent individuals and 6 matched healthy donors provide sufficient power. For functional studies, statistical significant can be obtained from the results, representing sufficient power using the given sample size.
Data exclusions	The single cell ATAC-seq profiles and TCR sequences obtained were subjected to quality control and low quality cells were excluded from analysis. Exclusion criteria established in previous work and were described in detailed in Methods section.
Replication	For human biological specimens, we collected data from 8 COVID-19 convalescent individuals and 6 healthy donors. For functional studies, we conducted at least 2 independent experiments. The results presented in manuscript were reliably reproduced.
Randomization	No randomization of human specimen was used. Human specimen were allocated into health or COVID-19 group. COVID-19 convalescent individuals were classified into mild and severe groups based on the Fifth Revised Trial Version of the Novel Coronavirus Pneumonia Diagnosis and Treatment Guidance.
Blinding	Blinding was not relevant to our study. Healthy donors were defined by negative results both in real-time RT-PCR of SARS-CoV-2 gene and IgG antibody detection, and COVID-19 convalescent individuals were predetermined by their known disease status.

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	Alexa Fluor® 700 anti-human CD3 (Clone SK7, 1:100, BioLegend, Cat #344822), FITC anti-human TCR α/β (Clone IP26, 1:50, BioLegend, Cat #306706), PerCP/Cyanine5.5 anti-human CD8 (Clone SK1, 1:200, BioLegend, Cat #344710), PE/Dazzle™ 594 anti-human CD4 (Clone A161A1, 1:200, BioLegend, Cat #357411), PE anti-human CD25 (Clone BC96, 1:200, BioLegend, Cat #302606), Brilliant Violet 650™ anti-human CD127 (Clone A019D5, 1:200, BioLegend, Cat #351326), APC anti-human CD137 (4-1BB) (Clone 4B4-1, 1:100, BioLegend, Cat #309810), Brilliant Violet 605 anti-human CD279 (PD-1) (Clone EH12.2H7, 1:100, BioLegend, Cat #329923)
Validation	All antibodies were validated by the manufacturer directly in human peripheral blood mononuclear cells.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Seven convalescent individuals who recovered from moderate COVID-19 and three convalescent individuals who recovered from severe COVID-19. These ten COVID-19 convalescent individuals who recovered and had been discharged from the hospital at lease for one month, included six women and four men, ranging from 22-69 years old. Healthy donors, included three women and three men, ranging from 18-55 years old, were defined by negative results both

	in real-time RT-PCR of SARS-CoV-2 gene and IgG antibody detection
Recruitment	No selective recruitment of healthy donors was performed. These ten COVID-19 convalescent individuals who recovered and had been discharged from the hospital for at lease for one month were consented and enrolled at the Fifth Medical Center of PLA General Hospital, regardless of age and gender. This study did not attempt to exclude healthy volunteers or COVID-19 convalescent individuals based on sex, race, or ethnicity.
Ethics oversight	This study was approved by the Ethics Committees of the Fifth Medical Center of Chinese PLA General Hospital, Beijing, China (2020005D) on healthy volunteers or COVID-19 convalescent individuals in Medical Research, and written informed consent was obtained from all participants.

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	The PBMCs from human blood were isolated using Ficoll solution according to the manufacturer's instructions. Human PBMCs were thawed after storage in liquid nitrogen and were allowed to recover in R10 culture medium overnight (RPMI1640 medium supplemented with 10% fetal calf serum (FCS), 25 mM HEPES, 1x nonessential amino acids, 50 uM b-mercaptoethanol, and 1x penicillin and streptomycin) overnight. Then, cells were collected and washed with FACS buffer for three times, and stained with the antibody cocktail for 25 min on ice. After staining, cells were washed with FACS buffer for three times. Single viable DAPI-CD3+TCRα/β+ T cells were immediately indexed sorted into 96-well plates for Ti-ATAC-seq.
Instrument	BD FACSAria III
Software	FlowJo V10 was used for data analysis
Cell population abundance	Single viable DAPI-CD3+TCR+ T cells were indexed sorted into 96-well plates via flow cytometor sorting
Gating strategy	single DAPI-, CD3+ and TCRα/β+ T cells were gated and used to identify α/β T cells from other cells

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