

ARTICLE

Open Access

Immune cell profiling of COVID-19 patients in the recovery stage by single-cell sequencing

Wen Wen¹, Wenru Su², Hao Tang^{3,4}, Wenqing Le⁵, Xiaopeng Zhang⁶, Yingfeng Zheng², Xiuxing Liu², Lihui Xie², Jianmin Li⁶, Jinguo Ye², Liwei Dong¹, Xiuliang Cui¹, Yushan Miao³, Depeng Wang⁷, Jiantao Dong⁸, Chuanle Xiao², Wei Chen⁶ and Hongyang Wang^{1,9,10}

Abstract

COVID-19, caused by SARS-CoV-2, has recently affected over 1,200,000 people and killed more than 60,000. The key immune cell subsets change and their states during the course of COVID-19 remain unclear. We sought to comprehensively characterize the transcriptional changes in peripheral blood mononuclear cells during the recovery stage of COVID-19 by single-cell RNA sequencing technique. It was found that T cells decreased remarkably, whereas monocytes increased in patients in the early recovery stage (ERS) of COVID-19. There was an increased ratio of classical CD14⁺⁺ monocytes with high inflammatory gene expression as well as a greater abundance of CD14⁺⁺IL1β⁺ monocytes in the ERS. CD4⁺ T cells and CD8⁺ T cells decreased significantly and expressed high levels of inflammatory genes in the ERS. Among the B cells, the plasma cells increased remarkably, whereas the naïve B cells decreased. Several novel B cell-receptor (BCR) changes were identified, such as IGHV3-23 and IGHV3-7, and isotypes (IGHV3-15, IGHV3-30, and IGKV3-11) previously used for virus vaccine development were confirmed. The strongest pairing frequencies, IGHV3-23-IGHJ4, indicated a monoclonal state associated with SARS-CoV-2 specificity, which had not been reported yet. Furthermore, integrated analysis predicted that IL-1β and M-CSF may be novel candidate target genes for inflammatory storm and that TNFSF13, IL-18, IL-2, and IL-4 may be beneficial for the recovery of COVID-19 patients. Our study provides the first evidence of an inflammatory immune signature in the ERS, suggesting COVID-19 patients are still vulnerable after hospital discharge. Identification of novel BCR signaling may lead to the development of vaccines and antibodies for the treatment of COVID-19.

Introduction

COVID-19, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has spread in many countries^{1–3}. As of April 6, 2020, SARS-CoV-2 has affected over 1,200,000 people and killed more than 60,000 of those affected in more than 160 countries. Following its global spread, the World Health

Organization declared it a public health emergency of international concern⁴. COVID-19 shows symptoms of fever, dry cough, fatigue, diarrhea, conjunctivitis, and pneumonia. Some patients develop severe pneumonia, acute respiratory distress syndrome (ARDS), or multiple organ failure^{5–7}. Although scientists and clinicians worldwide have made great efforts to produce vaccines and explored antiviral drugs^{8,9}, there is still no specific medicine and highly effective clinical treatment for COVID-19^{10,11}.

Immune system dysregulation, such as lymphopenia and inflammatory cytokine storm, have been observed and are believed to be associated with the severity of pathogenic coronavirus infections, such as severe acute respiratory syndrome coronavirus (SARS-CoV) and

Correspondence: Chuanle Xiao (xiaochuanle@126.com) or Wei Chen (cw0226@foxmail.com) or Hongyang Wang (hywangk@vip.sina.com)
¹National Center for Liver Cancer, Second Military Medical University, 200438 Shanghai, China
²State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, 510060 Guangzhou, China
Full list of author information is available at the end of the article
These authors contributed equally: Wen Wen, Wenru Su, Hao Tang, Wenqing Le, Xiaopeng Zhang, Yingfeng Zheng

© The Author(s) 2020



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Middle East respiratory syndrome coronavirus (MERS-CoV) infections^{12,13}. With regard to COVID-19, recent studies also found decreases in lymphocyte numbers and increases in serum inflammatory cytokine levels in peripheral blood^{5,14}. However, the manner in which key immune cell subsets change and their states during COVID-19 have remained largely unclear. Thus, defining these key cellular subsets and their states in COVID-19 is a crucial step in obtaining critical insights into the immune clearance mechanism and developing new therapeutic strategies for COVID-19.

Here, we applied single-cell RNA sequencing (scRNA-seq) to comprehensively characterize the changes in peripheral blood mononuclear cells (PBMCs) from 10 COVID-19 patients. Our study depicted a high-resolution transcriptome landscape of blood immune cell subsets during the recovery stage of COVID-19. It revealed that, compared to that in the healthy controls (HCs), monocytes containing high inflammatory gene expression and $IL1\beta^+$ subsets predominated, whereas $CD4^+$ T cells decreased remarkably in patients in the early recovery stage of COVID-19. We found that T and B cell clones were highly expanded during the recovery stage in COVID-19 patients. Furthermore, several specific BCR changes in COVID-19 patients during the recovery stage may be helpful for vaccine and antibody production.

Results

Study design and analysis of single immune cell profiling in COVID-19 patients

To map the immune microenvironment of COVID-19 patients, we identified mirroring changes in the blood and pinpointed cell-specific alterations associated with disease severity and recovery; we then integrated scRNA-seq, single-cell paired BCR, and single-cell paired TCR analysis from a total of 10 COVID-19 patients in the early recovery stage (ERS) or late recovery stage (LRS) (70,858 PBMCs). We also collected scRNA-seq data (57,238 cells) from five healthy donors as controls (Fig. 1a and Supplementary Fig. S1a–c). This dataset passed stringent high-quality filtering. Single-cell suspensions of the scRNA-seq samples were converted to barcoded scRNA-seq libraries using 10x Genomics. Cell Ranger software (version 3.1.0) was used for the initial processing of the sequencing data.

Using t-distributed stochastic neighbor embedding (t-SNE), we analyzed the distribution of the three immune cell lineages, myeloid, NK and T, and B cells, based on the expression of canonical lineage markers and other genes specifically upregulated in each cluster (Fig. 1b, c). For marker genes, expression values in each cell positioned in a t-SNE are shown in Fig. 1d. We next clustered the cells of each lineage separately and identified a total of 20 immune cell clusters.

An overview of NK and T, B, and myeloid cells in the blood of convalescent patients with COVID-19

The immune cell compartment of patients who have recovered from COVID-19 infection comprised all major immune lineages. We analyzed 128,096 scRNA-seq profiles that passed quality control, including 36,442 myeloid cells, 64,247 NK and T cells, and 10,177 B cells from five HCs, five ERS, and five LRS patients. The sketchy clustering analysis landscape of each subject is presented in Supplementary Fig. S2a, and the merged image of each group is shown in Fig. 2a. We discovered that COVID-19 patients, including ERS and LRS, demonstrated a higher proportion of myeloid cells compared to the HCs, but with a lower proportion of NK and T cells (Fig. 2b, c). Interestingly, LRS patients had more B cells and NK and T cells, but less myeloid cells, than the ERS patients (Fig. 2b, c). Thus, these findings indicated that COVID-19 patients had decreased lymphocyte counts and increased counts of myeloid cells in peripheral blood.

To further understand the changes in the myeloid, NK and T, and B cells in COVID-19 patients, we conducted differential expression gene (DEG) analysis of the NK and T, B, and myeloid cells between the HCs and patients. The heatmaps are shown in Fig. 2d–f. Inflammatory cytokines and chemokines such as *IL1B*, *CCL3*, *IRF1*, *DUSP1*, *JUN*, and *FOS* were all expressed at high levels in patients, regardless of myeloid cells (Fig. 2d), NK and T cells (Fig. 2e), or B cells (Fig. 2f).

Collectively, our results demonstrated that myeloid cells increased, whereas NK and T cells decreased in the peripheral blood of COVID-19 patients and that the immune cell compositions differed between the patients in the ERS and LRS.

Myeloid cell subsets and their states in the blood of convalescent patients with COVID-19

To further understand the changes in the monocytes in patients in the early and late recovery stages of COVID-19, we conducted gene expression analysis and sub-clustered the myeloid cells into six transcriptionally distinct subsets using Uniform Manifold Approximation and Projection (UMAP). Classical $CD14^{++}$ monocytes (M1), non-classical $CD16^{++}$ (FCGR3A) $CD14^{-/+}$ monocytes (M2), intermediate $CD14^{++}$ $CD16^{+}$ monocytes (M3), $CD1C^{+}$ cDC2 (M4), $CLEC9A^{+}$ cDC1 (M5), and pDC ($CLEC4C^{+}CD123^{+}$) (M6) were present in the six distinct clusters (Fig. 3a, b). We found that the compartment of the monocyte subset differed remarkably among the HCs and COVID-19 patients (Fig. 3c). Among the myeloid cells, the ratio of classical $CD14^{++}$ monocytes (M1) higher in the ERS patients than in the HCs and was almost normal in the LRS patients (Fig. 3c).

We found that COVID-19 patients had a greater abundance of $CD14^{++}$ $IL1\beta^+$ monocytes and IFN-

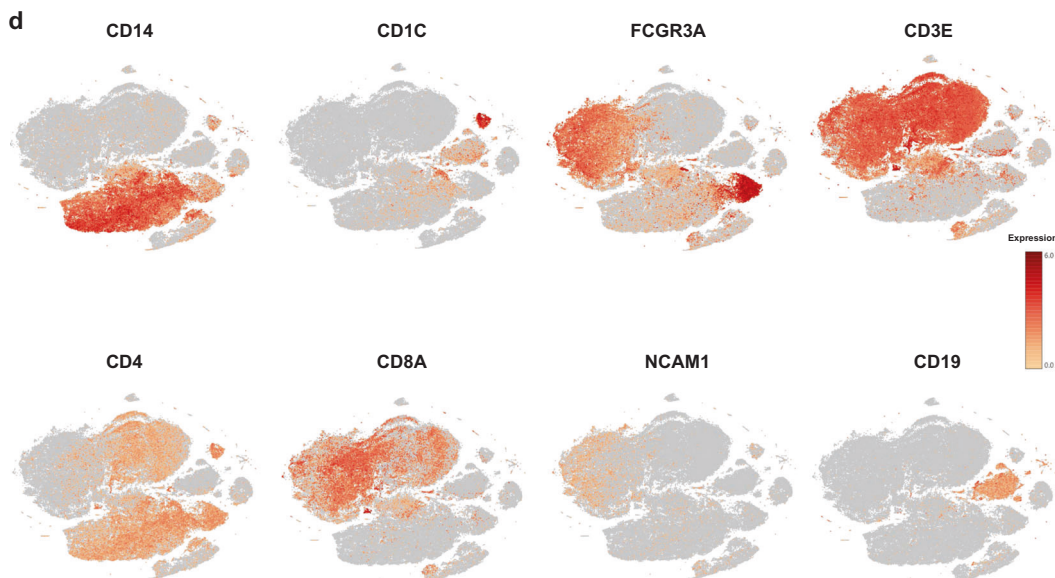
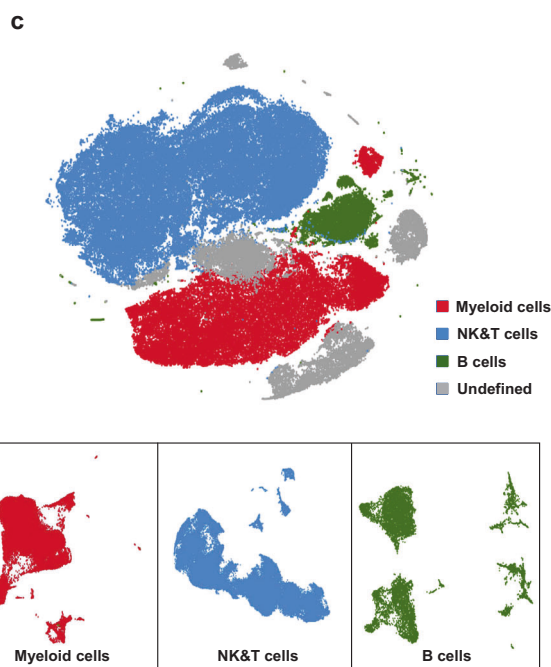
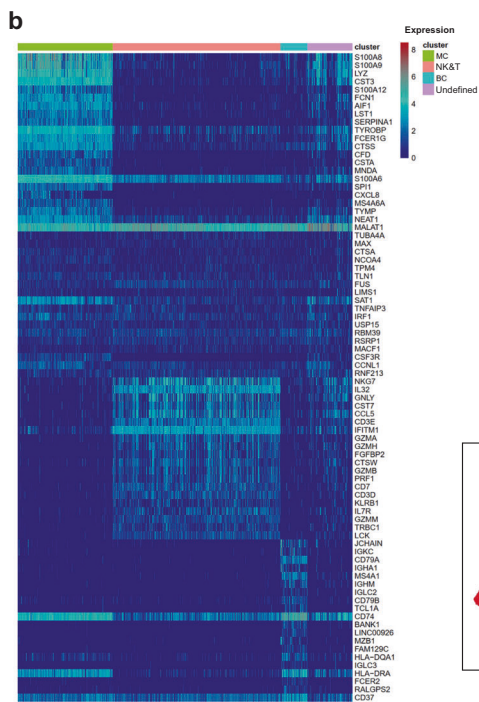
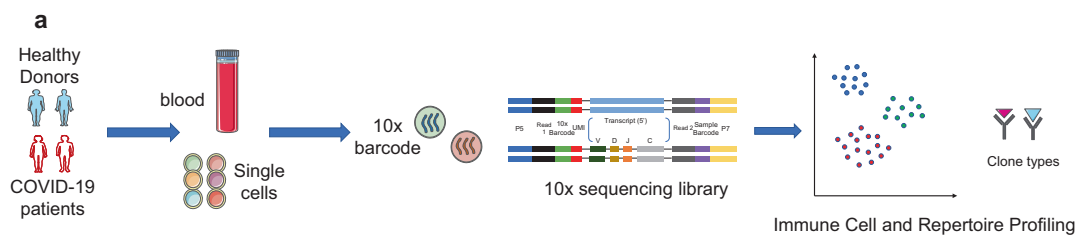


Fig. 1 (See legend on next page.)

(see figure on previous page)

Fig. 1 Study design and analysis of single immune cell profiling in COVID-19 patients. **a** Schematics of the experimental design for single-cell RNA (sc-RNA) sequencing. Peripheral blood mononuclear cells (PBMCs) were collected from COVID-19 patients and healthy controls (HCs) and then processed via sc-RNA, sc-BCR, and sc-TCR sequencing using the 10x-Based Genomics platform. **b** The heatmaps show differentially expressed genes (DEGs) upregulated in myeloid cells, NK and T cells, B cells, and other clusters of PBMCs. **c** t-distributed stochastic neighbor embedding (t-SNE) plot showing myeloid cells (red), NK and T cells (blue), B cells (green), and other clusters (gray) of PBMCs identified using integrated and classification analysis. **d** t-SNE projection of canonical markers, including *CD14*, *CD1C*, and *FCGR3A* for myeloid cells; *CD3E*, *CD4*, *CD8A*, and *NCAM1* for NK and T cells; and *CD19* for B cells as indicated in the legend.

activated monocytes than the HCs (Fig. 3d–f). Genes associated with $CD14^{++}$ inflammatory monocytes (M1) had high expression levels of inflammatory genes such as *IL1 β* , *JUN*, *FOS*, *JUNB*, and *KLF6*; chemokines, *CCL4*, *CXCR4*; and interferon-stimulated genes, *IFRD1*, *IRF1*, and *IFI6*. In contrast, anti-inflammatory genes associated with $CD14^{++}$ monocytes (M1) were downregulated in COVID-19 patients relative to that in the HCs (Fig. 3d, e). Notably, *IL1 β* expression values in a UMAP with simultaneous contrast indicated that *IL1 β* was upregulated in the ERS group and decreased in the LRS patients (Fig. 3f). This was also confirmed in the DC cluster of the ERS group compared to that of the HCs (Supplementary Fig. S3a, b). Next, we took the average of the inflammatory genes for each myeloid cell scRNA-seq subset in the COVID-19 patients versus that in the HCs (Supplementary Fig. S3c). These results demonstrated that cytokine activation drives the expansion of monocyte populations (especially $CD14^{++}$ inflammatory monocytes) in COVID-19-infected patients. To explore the biological significance of the transcriptional changes in the M1 cluster, we performed GO analysis with DEGs (Fig. 3g). We observed enrichment of the pathways related to cytokine signaling and inflammation activation, which were driven by the upregulation of *IFITM3* and *IFI6* and *IL1 β* , *JUN*, *FOS*, *JUNB*, and *KLF6* (Fig. 3g).

Collectively, these findings demonstrate that a dysregulated balance in the monocyte populations in ERS patients is manifested by substantially increased classical $CD14^{++}$ monocytes. Our results suggest that the classical $CD14^{++}$ monocytes increase in circulation to fuel inflammation during SARS-CoV-2-infection.

Characterization of T and NK cell responses in the blood of recovered COVID-19 patients

T and NK cells play critical roles in viral clearance during respiratory infections^{15,16}. Our clustering analysis sub-grouped T and NK lymphocytes into 10 subsets (Fig. 4a) based on canonical markers (Fig. 4b and Supplementary Fig. S4a). NK cells highly expressed *NCAM1*, *KLRF1*, *KLRC1*, and *KLRD1*; then, we sub-divided the NK cells into $CD56^{+}CD16^{-}$ NK cells (NK1), which expressed high levels of *CD56* and low levels of *CD16*; and $CD56^{-}CD16^{+}$ NK cells (NK2), which expressed high levels

of *CD16* and low levels of *CD56*. $CD4^{+}$ T cells expressed *CD3E* and *CD4*; then, we sub-divided these cells into four clusters: naïve $CD4^{+}$ T cells (T1), which expressed high levels of *CCR7*, *LEF1*, and *TCF7*; central memory $CD4^{+}$ T cells (T2, $CD4$ Tcm), which expressed high levels of *CCR7*, but more *AQP3* and *CD69* compared to naïve $CD4^{+}$ T cells; effector memory $CD4^{+}$ T cells (T3, $CD4$ Tem), which expressed high levels of *CCR6*, *CXCR6*, *CCL5*, and *PRDM1*; and regulatory T cells (T4, Treg), which expressed *FOXP3*. $CD8^{+}$ T cells expressed *CD8A* and *CD8B* and were sub-divided into three clusters: naïve $CD8^{+}$ T cells (T5), which expressed high levels of *CCR7*, *LEF1*, and *TCF7*, similar to naïve $CD4^{+}$ T cells; effector memory $CD8^{+}$ T cells (T6, $CD8$ Tm), which expressed high levels of *GZMK*; and cytotoxic $CD8^{+}$ lymphocytes ($CD8^{+}$ CTL) (T7), which expressed high levels of *GZMB*, *GNLY*, and *PRF1*. Proliferating T cells (T8, T_{prol}) were *TYMS^{+}MKI67^{+} cells.*

The composition of the T and NK cell subsets differed significantly among the HCs and COVID-19 patients (Fig. 4c). The absolute number of $CD8^{+}$ T cells especially the effector memory $CD8^{+}$ T cell subgroup and NK cells decreased in COVID-19 patients, whereas the relative ratio of NK cells in ERS was higher than that in the HCs. The ratio of $CD4^{+}$ T cells was stable, but the composition of the $CD4^{+}$ T cell subset differed significantly between the HCs and COVID-19 patients. Among $CD4^{+}$ T cells, the ratio of central memory $CD4^{+}$ T cells was significantly higher, whereas the ratio of naïve $CD4^{+}$ T cells, Tregs and effector memory $CD4^{+}$ T cell was lower than that in the HCs especially in ERS group. Notably, genes associated with $CD4^{+}$ T cells had relatively high expression levels of inflammation-related genes and were significantly upregulated in the COVID-19 patients (Fig. 4d). $CD4^{+}$ T cells had high expression levels of inflammatory genes, including *FOS*, *JUN*, *KLF6*, and *S100A8* in patients in the ERS of COVID-19 (Fig. 4e). In contrast, anti-inflammatory genes associated with $CD4^{+}$ T cells were downregulated in COVID-19 patients relative to that in the HCs (Fig. 4d, e). This suggested that $CD4^{+}$ T cells were the main participants in the virus infection. Comparison of the DEGs in the $CD4^{+}$ T cells revealed the enrichment of genes participating in the cytokine pathway and inflammation activation, including *IFITM3* and *IFI6*

(see figure on previous page)

Fig. 2 An overview of NK and T, B, and myeloid cells in the blood of convalescent patients with COVID-19. **a** The t-SNE plot shows a comparison of the clustering distribution across HCs as well as early recovery stage (ERS) and late recovery stage (LRS) patients with COVID-19. **b** The bar plot shows the relative contributions of myeloid, NK and T, and B cells by individual samples, including five HCs, five ERS patients, and five LRS patients. **c** The pie chart shows the percentages of myeloid, NK and T, and B cells across HCs as well as ERS and LRS patients with COVID-19. **d** The heatmap shows the DEGs of myeloid cells among the HCs and the ERS and LRS COVID-19 patients. **e** The heatmap shows the DEGs of NK and T cells among the HCs and the ERS and LRS COVID-19 patients. **f** The heatmap shows the DEGs of B cells among the HCs and the ERS and LRS COVID-19 patients.

and *IL1B*, *JUN*, *FOS*, *JUNB*, and *KLF6* (Fig. 4f). Further studies are needed to elucidate the IFN pathways involved in COVID-19 pathogenesis.

TCR-seq analysis showed that T cell expansion was obviously decreased in the ERS group than in the HC group (Fig. 4g). Moreover, naïve or central memory T cells showed little clonal expansion, while effector memory T cells, terminal effector CD8⁺ T cells (CTLs), and proliferating T cells showed higher expansion levels (Fig. 4h). In addition, the most highly expanded (maximum) clone in the ERS group was TRAV8-6-TRAJ45:TRAV7-8-TRBJ2-1 (Supplementary Fig. S5d). The decreased ratio of CD8⁺ T cells in COVID-19 patients may implicate the role of CD8⁺ T cells in virus clearance (Fig. 4c). Moreover, the CD8⁺ CTL with expanded clones also exhibited overactivated inflammation and antiviral activity compared to those in HCs (Fig. 4i and Supplementary Fig. S4b). Together, these findings show that clonally expanded CD8⁺ T cells in the peripheral blood of COVID-19 patients help control the virus. We also performed DEG analysis via Seurat *FindAllMarkers* analysis and found similar results in T_{prol} cells (Supplementary Fig. S4c). Next, we took the average of inflammatory genes for each NK and T cell subset scRNA-seq subset in the COVID-19 patients versus normal RNA-seq data (Supplementary Fig. S4d).

Characterization of single-cell B cells in COVID-19 patients

By projecting the gene expression data of B cells using diffusion maps, we identified four B cell clusters using scRNA-seq: naïve B cells (B1) expressing *CD19*, *CD20* (*MS4A1*), *IGHD*, *IGHM*, *IL4R*, and *TCL1A*; memory B cells (B2) expressing *CD27*, *CD38*, and *IGHG*; immature B cells (B3) only expressing *CD19* and *CD20* (*MS4A1*); and plasma cells (B4) expressing high levels of *XBPI* and *MZBI* (Fig. 5a, b and Supplementary Fig. S5a).

In comparison with that in the HCs, the percentage of plasma cells increased significantly in COVID-19 patients, whereas naïve B cells decreased significantly in the COVID-19 patients (Fig. 5c). Memory B cells and plasma cells (MPB) might play an important role in the control of viral infection and the development of adoptive immunity as they synergistically work and induce specific antibodies. Moreover, compared to that in the HCs, B cell activation-

related genes, including *S100A8*, *IGLL5*, *SSR3*, *IGHA1*, *XBPI*, and *MZBI* were primarily expressed in the MPB of the ERS group (Fig. 5d). We also found similar results in the plasma cells, the antibody-secreting cells (ASC) (Supplementary Fig. S5b, c), suggesting a key role for ASC in viral control. Next, we took the average of the inflammatory genes for each B cell subset of the COVID-19 patients versus the normal RNA-seq data (Fig. 5e). The difference in the genes between the ERS and HCs indicated enhanced B cell reaction and antibody secretion in COVID-19 patients. GO analysis revealed that *IGHA1*, *XBPI*, *MZBI*, *JUN*, *POLR2L*, and *ZFP36* were over-presented in MPBs, which suggests enhanced B cell proliferation and viral transcription in COVID-19 patients (Fig. 5f). Single-cell BCR-seq analysis indicated that the IgA isotype was over-represented in COVID-19 patients compared to that in the HC (Fig. 5g). This corresponded with an increase in the levels of serum IgA, which was also pronounced in other coronavirus infections. Moreover, the ratio of (IgA+IgG+IgE) to (IgD+IgM) increased significantly in the ERS patients and showed a downward trend with recovery time (Fig. 5h).

Expanded BCR clones and biased usage of VDJ genes observed in COVID-19 patients

Using sc-BCR-seq to assess the status of clonal expansions in the blood of patients, we found that IL4R⁺ naïve B cells showed little clonal expansion, whereas CD27⁺CD38⁺ memory B cells showed the highest expansion levels among diverse B cell subsets (Fig. 6a). At the individual level, we found that COVID-19 patients had significantly expanded clones compared to that in the HCs, supporting the assumption that B cells had experienced unique clonal VDJ rearrangements under SARS-CoV-2-infection. We also found that a higher B cell clonality consistently remained in the ERS compared with that in the LRS patients (Fig. 6b). Moreover, quantification of the most highly expanded (maximum) clone for each subject showed that the ratios of the maximum clones were higher in the ERS group than in the HCs (Fig. 6c). To understand the functional status of expanded cloned B cells, we performed DEG analysis between the cloned memory B cells and the other B cells. Our results revealed increased expression of B cell genes, including

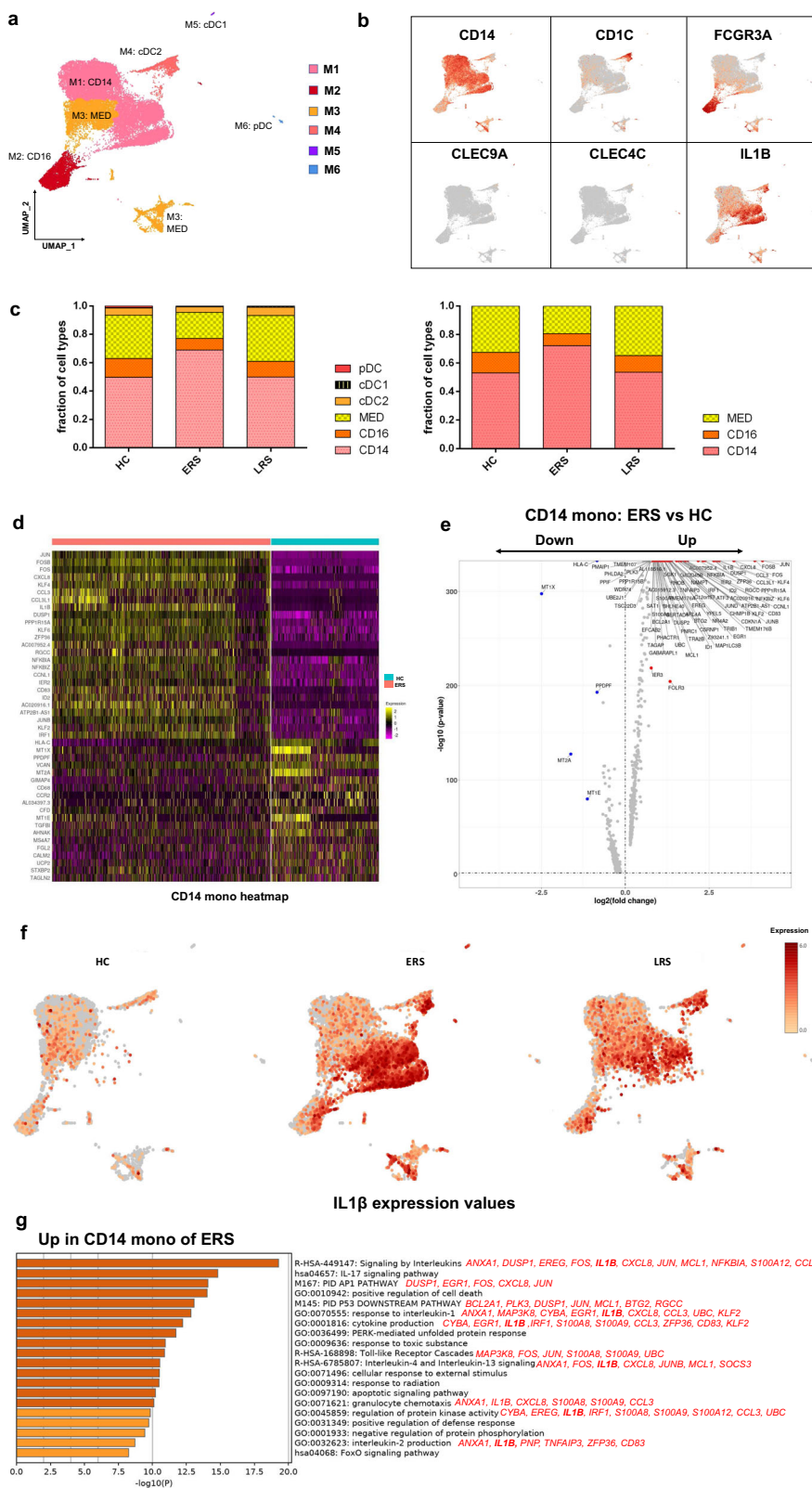


Fig. 3 (See legend on next page.)

(see figure on previous page)

Fig. 3 Myeloid cell subsets and their states in the blood of convalescent patients with COVID-19. **a** Six clusters of myeloid cells were displayed according to marker gene expression levels. Uniform manifold approximation and projection (UMAP) presentation of the heterogeneous clusters of peripheral myeloid cells. Classical CD14⁺⁺ monocytes (M1), non-classical CD16⁺⁺ (FCGR3A) CD14^{-/+} monocytes (M2), intermediate CD14⁺⁺ CD16⁺ monocytes (M3), CD11c⁺ cDC2 (M4), CLEC9A⁺ cDC1 (M5), and pDC (CLEC4E⁺CD123⁺) (M6). **b** The UMAP plot shows subtype-specific marker genes of myeloid cells, including *CD14*, *FCGR3A*, *CD11C*, *CLEC9A*, *CLEC4E*, and *IL-1β*. **c** Bar chart of the relative frequencies of the six sub-clusters of myeloid cells and three sub-clusters of monocytes across the HCs and the ERS and LRS patients. **d** The heatmap shows the top DEGs between COVID-19 patients and HCs in CD14⁺⁺ monocytes. **e** Volcano plot of fold change between COVID-19 patients and HCs in CD14⁺⁺ monocytes. *P* values were calculated using a paired, two-sided Wilcoxon test and FDR corrected using the Benjamini-Hochberg procedure. **f** The UMAP plot shows that IL-1β was highly expressed in the ERS patients vs. the LRS patients and HCs in myeloid cells. **g** GO BP enrichment analysis of the DEGs of CD14⁺⁺ monocytes upregulated in COVID-19 patients. *P* value was derived by a hypergeometric test.

CD27, *SSR4*, *IGHG1*, *MZB1*, and *XBPI*, which further supports the superior effector functions of the expanded cloned B cells (Fig. 6d). Moreover, the differential genes for expanded B cells significantly subsided over time and reduced in LRS patients (Fig. 6d).

To study the unique changes and preference genes of BCR in COVID-19 patients, we compared the usage of VDJ genes in COVID-19 patients with that in the HCs. We identified an over-representation of the IGHV3 family, especially the IGHV3-7, IGHV3-15, IGHV3-21, IGHV3-23, and IGHV3-30 in COVID-19 patients compared to that in the HCs (Fig. 6e). The preferred IGKVs were IGKV1-17, IGKV2-28, and IGKV3-15, whereas the preferred IGLVs were IGLV1-44, IGLV2-8, and IGLV3-27 (Fig. 6e). Moreover, the top two pairing frequencies in ERS patients were IGHV3-23-IGHJ4 and IGHV3-7-IGHJ6 (Fig. 6f). These cells showed IGH subunit pairing with the IGK/L subunit encoded by IGLV1-44-IGLJ3 and IGKV1-17-IGKJ1, respectively, which indicated expanded states associated with SARS-CoV-2 specificity. Individually, ERS-4 and ERS-5 had the maximum clones, referring to IGHV3-23-IGHJ4 (Supplementary Fig. S5e) and IGHV3-7-IGHJ6 (Supplementary Fig. S5f), respectively.

In summary, an increase in clonality in COVID-19, which was dominated by the IgA and IgM isotypes, together with a skewed use of the IGHV gene, suggested the contribution of SARS-CoV-2 to pathogenesis. Notably, the biased usage of dominated IGV genes, especially the IGHV3-23 and IGHV3-7 in COVID-19 patients, provides a framework for the rational design of SARS-CoV-2 vaccines.

Cell-to-cell communication among immune cells in COVID-19 patients

An established computational approach¹⁷ was used to predict cell-to-cell interactions that may contribute to the distinct functional state of T cells, B cells, monocytes, and dendritic cells (DCs) in ERS and LRS (Fig. 7a, b). In ERS COVID-19 patients, we found adaptive signals involved in monocyte activation, proliferation, and inflammatory signaling (Fig. 7a, b). T cells expressed genes encoding

ligands of TNFSF8, LTA, IFNG, IL17A, CCR5, and LTB to TNFRSF8, TNFRSF1A/TNFRSF14, IFNGR1, IL-17RA, CCR1, and LTBR, which were expressed on monocytes and could contribute to the pro-inflammatory status. Other T cell-monocyte interactions involved the expression of CSF2 and CSF1. T cells might activate monocytes through the expression of CSF2 and CSF1, which bind to CSFRs (CSFR2/1) and contribute to inflammatory storm. A cluster of CD14⁺ monocytes exclusively expressed IL1β, which was predicted to bind to IL1RAP expressed by T cells. T cell-monocyte interaction may enhance immune response and be exclusive to COVID-19 patients (Fig. 7a, b). Furthermore, we found that monocytes highly expressed the poliovirus receptor, which serves as a cellular receptor for poliovirus in the first step of poliovirus replication and induction of the NF-kappa B signaling pathway. From the B cell-monocyte and B cell-T cell interactions, we found that B cells could secrete a large number of IL-6, LTA, and LTB, which are combined with IL-6R, LTAR, and LTBR expressed in monocytes, and a large amount of IL-6 was applied to T cells to promote the secretion of IFN-γ, IL-1β, and other inflammatory cytokines and chemokines. Thus, a cascade signature of inflammatory monocytes with high expression of IL-6 and their progeny were formed in the peak incidence of ERS COVID-19 patients (Fig. 7c). These activated immune cells may enter the circulation in the lung and other organs in large numbers and play an immune-damaging role. In LRS COVID-19 patients, DC ligands were predicted to interact with B and T cell receptors involved in cell proliferation and the production of antibodies. We discovered that the peripheral blood of LRS patients contains a diversity of antibodies; we found that IL18-IL18RAP, TNFSF13-TNFRSF13B, TNFSF13-TNFRSF17, TNFSF13B-TNFRSF17, TNFSF13B-TNFRSF13B, and TNFSF13B-TNFRSF13C were highly expressed in our analysis of DC-B cell interaction (Fig. 7d). Thus, we speculate that DCs produce IL-18, TNFSF13, and TNFSF13B to promote the proliferation of B cells and then secrete many antibodies into the blood in ERS. From the DC-T and T cell-B cell interactions, we discovered

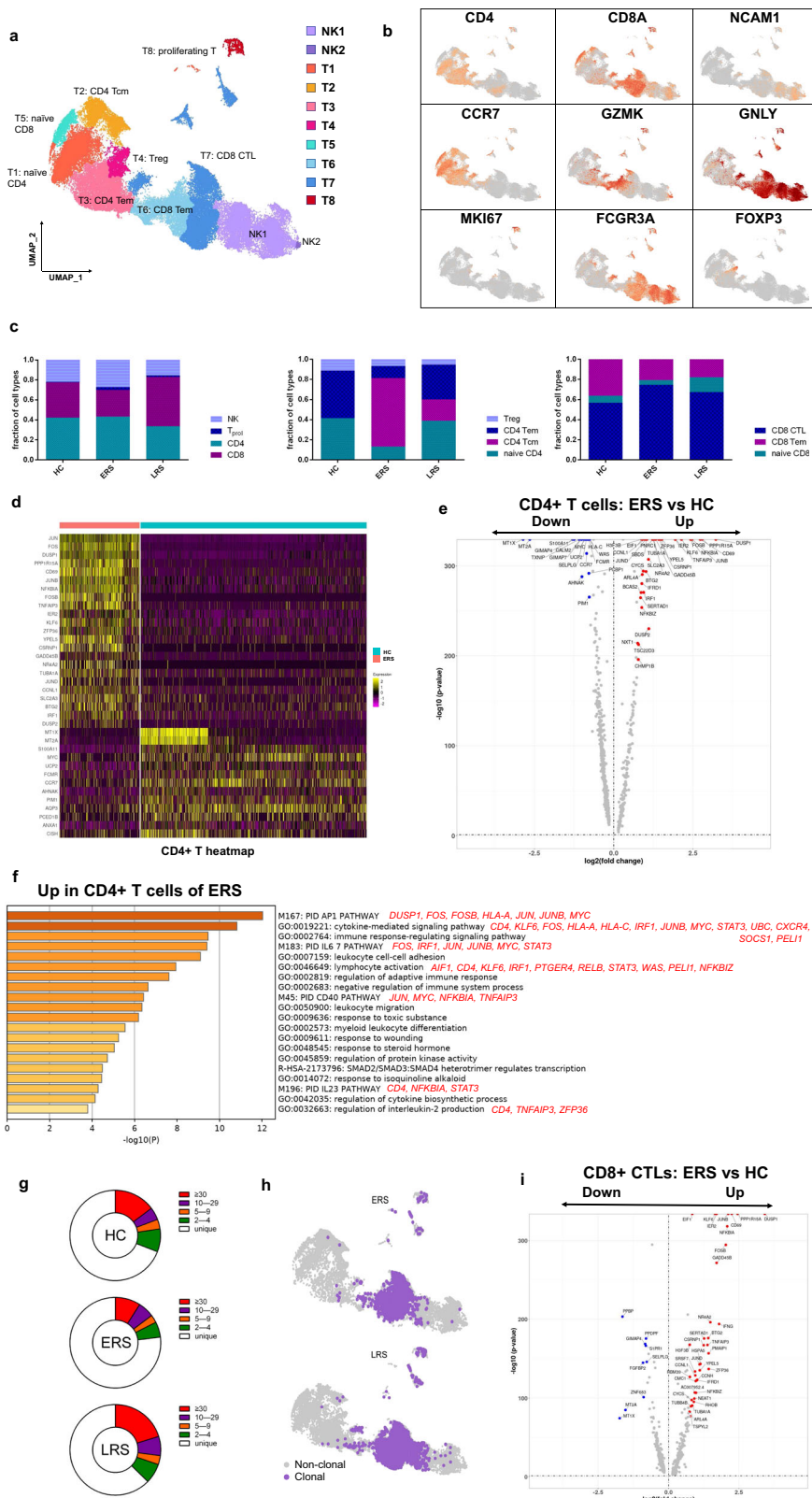


Fig. 4 (See legend on next page.)

(see figure on previous page)

Fig. 4 Characterization of T and NK cell responses in the blood of recovered COVID-19 patients. **a** Ten sub-clusters of NK and T lymphocytes were identified. The UMAP plot shows the clustering of T and NK cells. CD56⁺CD16⁻NK cells (NK1), C56⁺CD16⁺ NK cells (NK2), naïve CD4⁺ T cells (T1), central memory CD4⁺ T cells (T2), effector memory CD4⁺ T cells (T3), regulatory T cells (T4), naïve CD8⁺ T cells (T5), effector memory CD8⁺ T cells (T6), cytotoxic CD8⁺ T cells (T7), and proliferating T cells (T8). **b** UAMP plot showing subtype-specific marker genes of NK and T cells including *CD4*, *CD8A*, *NCAM1*, *CCR7*, *GZMK*, *GNLY*, *MKI67*, *FCGR3A*, and *IL-1β*. **c** The bar plot shows the percentages of four sub-clusters of NK and T cells, four sub-clusters of CD4⁺ T cells, and three sub-clusters of CD4⁺ T cells among the HCs and the ERS and LRS patients. **d** Heatmap of CD4⁺ T cells showing the DEGs between the COVID-19 patients and HCs. **e** The volcano plot shows the DEGs of CD4⁺ T cells between the COVID-19 patients and HCs. *P* values were calculated using a paired, two-sided Wilcoxon test and FDR corrected using the Benjamini–Hochberg procedure. **f** GO BP enrichment analysis of the DEGs of CD4⁺ T cells upregulated in the COVID-19 patients. *P* value was derived by a hypergeometric test. **g** The pie plot shows the TCR clone differences across the HCs and the ERS and LRS patients. **h** UAMP shows expanded TCR clones ($n \geq 2$) in the ERS and LRS patients. **i** The volcano plot shows the DEGs of CD8⁺ CTLs between the COVID-19 ERS group and HCs. *P* values were calculated using a paired, two-sided Wilcoxon test and FDR corrected using the Benjamini–Hochberg procedure.

that DCs produce not only IL-18 but also IL-7 to promote the proliferation of T cells; moreover, T cells produce IL-2 to promote the proliferation and antibodies production of B cells (Fig. 7d). Thus, cell-to-cell interactions help us to understand why COVID-19 patients manifested high rates of monocytes and low rates of lymphocytes and why the proportion of lymphocytes gradually increased in the peripheral blood of recovering patients.

Discussion

The clinical presentation of COVID-19 varies from asymptomatic to severe ARDS. This has been similarly observed in severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV), and influenza infections^{12,14}. In viral infection, it is generally accepted that host immune responses determine both protection against viral infections and the pathogenesis of respiratory injury^{18,19}. A coordinated response in innate and adaptive immune cells working in concert may lead to the rapid control of the virus, whereas a failed immune response might lead to viral spreading, cytokine storm, and a high mortality rate²⁰. Despite belonging to same group of viruses, recent studies have highlighted differences between COVID-19, SARS, and MERS, such as the speed of transmission, treatment scheme, and mortality rate. Moreover, this difference may also exist in the key immune players and the underlying molecular mechanisms related to these diseases. The lack of knowledge regarding the immune impact of COVID-19 has now become a critical issue in view of its rapid spread and the shortage of specific therapy²¹. Using single-cell sequencing, we profiled the complexity of immune populations in the blood and analyzed 70,858 cells from 10 patients. We identified a hyper-inflammatory response in ERS patients, which may explain why some patients fell sick after being discharged, and suggest that the current criteria for hospital discharge should be re-evaluated. In addition, we identified unique signatures of myeloid, NK and T, and B cells and pinpointed the changes in the epitopes of TCR and BCR. Our

findings helped elucidate the antiviral immune mechanisms and revealed promising opportunities for developing immunotherapies using vaccines and neutralizing antibodies.

Inflammation is a vital part of the immune system's response to COVID-19 invasion; previous and latest studies have reported significantly higher levels of inflammatory cytokines associated with disease severity in SARS, MERS, and COVID-19 patients^{22,23}. Among the various inflammatory cells, monocytes and their subsets (including classical, intermediate, and non-classical monocytes) may play a critical role because they are known to fuel inflammation^{24–27}. In our study, compared with the HCs, ERS patients demonstrated a significantly higher ratio of monocytes, and these cells expressed higher levels of inflammatory genes. Intriguingly, the ratio of classical CD14⁺ monocytes was high in ERS but remained normal in LRS. Furthermore, CD14⁺IL1β monocytes, which were absent in HCs, could be observed in ERS, and they declined in number in LRS. Notably, our cell-to-cell interaction analysis indicated that IL1β, CSF1, IL6, and CSF2 may be associated with cytokine storm. The CD14⁺IL1β subpopulation appeared to be part of the inflammatory landscape of COVID-19, since these cells increased in ERS stage. Virus-induced IL-1β production in monocytes is mediated via a caspase-1 pathway. Multiple microbial components, including viral RNA, are thought to trigger assembly of the inflammasome and consequent caspase-1 activation²⁸, which may give a reasonable explanation for the presence of IL1β⁺ monocytes in ERS patients. Previous studies with a pathogenic influenza A virus revealed impaired neutrophil and CD4⁺ T cell activation in IL-1R1^{-/-} mice, greatly diminished lung inflammatory infiltrates, reduced IgM levels in both serum and at mucosal sites and decreased activation of CD4⁺ T “helpers” in secondary lymphoid tissue²⁹, indicating IL1β is responsible for virus-induced lung immunopathology establishment. CD14⁺ IL1β is also expected to become an important detection marker for monitoring COVID-19 disease recovery. Collectively, our data provide important

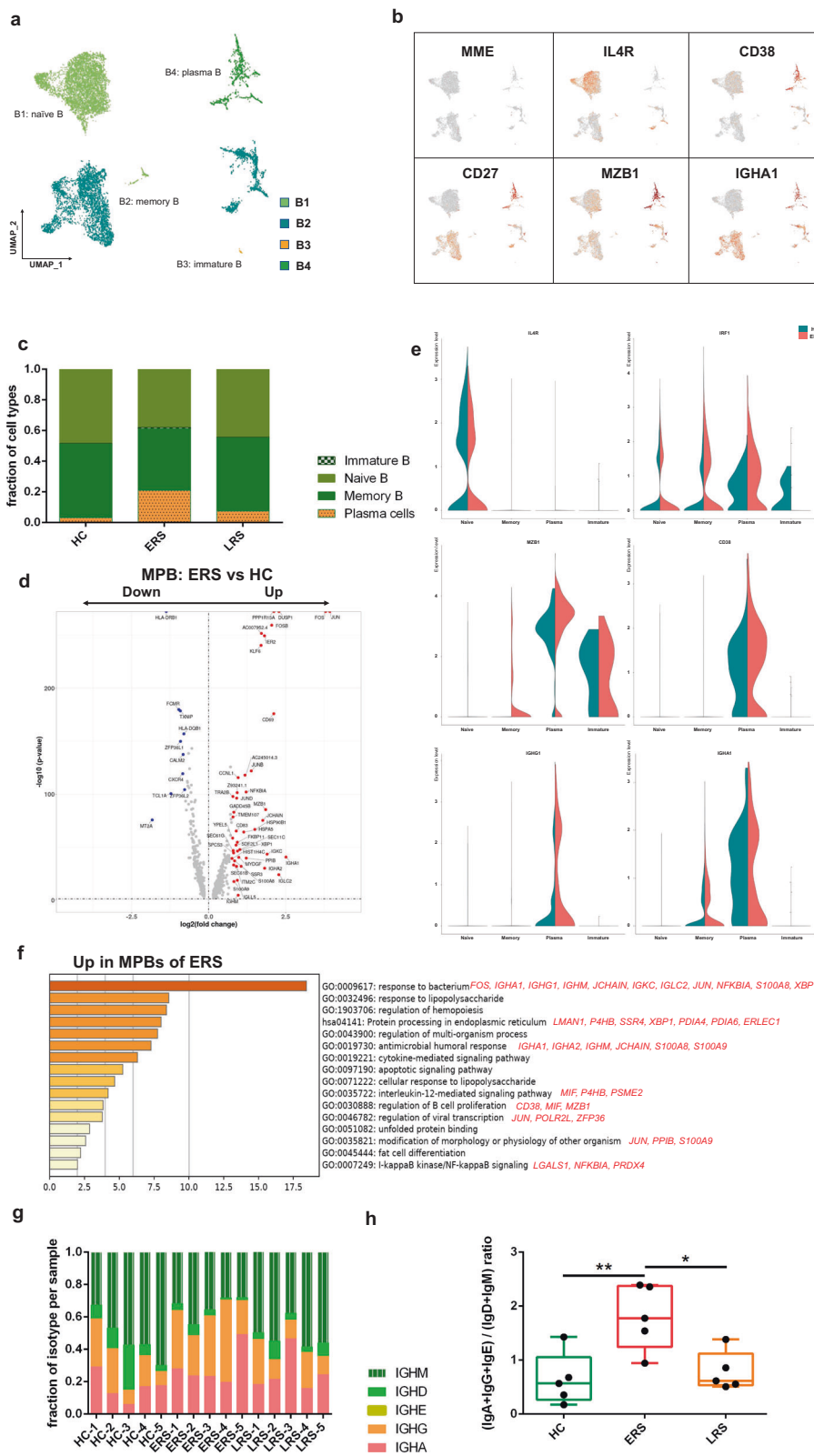


Fig. 5 (See legend on next page.)

(see figure on previous page)

Fig. 5 Characterization of single-cell B cells in COVID-19 patients. **a** Four clusters of B cells were identified. The UMAP plot shows the clustering of B cells. Naïve B cells (B1), memory B cells (B2), immature B cells (B3), and plasma cells (B4). **b** UAMP plot showing subtype-specific marker genes of B cells, including *MME*, *IL4R*, *CD38*, *CD27*, *MZB1*, and *IGHA1*. **c** The bar plot shows the percentages of B clusters across the HCs and the ERS and LRS patients. **d** The volcano plot shows the DEGs of MPB cells between the COVID-19 patients and HCs. *P* values were calculated using a paired, two-sided Wilcoxon test and FDR corrected using the Benjamini–Hochberg procedure. **e** The violin plot shows that *MZB1*, *IGHG1*, and *IGHA1* were highly expressed in COVID-19 patients vs. the HCs in the B cell sub-clusters. **f** GO BP enrichment analysis of the DEGs of MPB cells between the COVID-19 patients vs. the HCs. *P* value was derived by a hypergeometric test. **g** The bar plot shows the relative percentage of each isotype by individual sample. **h** The bar plot shows the ratio of (IgA+IgG+IgE) to (IgD+IgM) among the HCs and the ERS and LRS patients. Statistical analysis used One-Way ANOVA test. Values are mean \pm SD. **P* < 0.05, ***P* < 0.01.

insights into the role of monocytes in the immunopathogenesis of COVID-19.

The adaptive immune system harbors the ability to recognize and remember specific pathogens through antibody and T cell responses³⁰. Inducing adaptive immunity is the aim of vaccination³¹. Previous SARS studies have identified binding and neutralizing antibodies elicited by SARS-CoV infection. Their therapeutic effect is unclear³², although robust antibody responses could be induced³³. In COVID-19 infection, although several lines of evidence have consistently indicated a decline in lymphocyte counts, the distinct immune characteristics at single-cell resolution are unclear. Irani thevarajan et al.³⁴ reported that in the blood of mild-to-moderate COVID-19 hospitalized patients, the antibody-secreting cells, follicular T-helper cells, activated CD4⁺ and CD8⁺ T-cells and IgM/IgG SARS-CoV-2-binding antibodies were increased by using flow cytometry, and they found the changes persisted for at least 7 days following full resolution of symptoms. Our scRNA-seq analysis showed that, compared with the HCs, ERS patients who recovered less than 7 days have a lower ratio of T and NK cells, and these patients' T cells express higher levels of inflammatory genes, such as *JUN*, *FOS*, *JUNB*, and *KLF6*. Several studies have reported that lymphopenia is a prominent part of SARS-CoV2 infection and lymphocyte counts are useful in predicting the severity and clinical outcomes^{35–37}. We found the number of NK and T cell decreased but no significant change in B cell in COVID-19 patients especially in ERS patients. Possible reasons for it may be the direct infection of lymphocytes by SARS-CoV2, cytokine-mediated lymphocyte trafficking in the infected tissue or lymphocytes exhaustion in the peripheral blood and sequestration in the lung induced by cytokine storm³⁵. There may also be immune-mediated lymphocyte destruction, bone marrow or thymus suppression, or apoptosis³⁸, as is reported in other virus infection, which requires further study. In addition, high-throughput TCR sequencing identified expanded T cell clones in ERS patients. In LRS patients, the immunophenotype was different. In particular, LRS patients would have an increase in T and NK cells, with a lower expression of inflammatory genes. We also performed a detailed

analysis of B cells in patients and identified a higher population of plasma cells than that in the HCs. We found that BCR contained highly expanded clones, indicating their SARS-CoV-2 specificity. Importantly, we found several loci unique to COVID-19 infection. The strongest pairing frequencies, IGHV3-23-IGHJ4, indicated a monoclonal state associated with SARS-CoV-2 specificity, which has not been reported yet. Notably, numerous studies have reported biased usage of VDJ genes related to virus-specific antibodies. For example, IGHV3-30 and IGKV3-11 have been involved in encoding primary antibodies to neutralize human cytomegalovirus^{39,40}. In addition, IGHV3-30 and IGHV3-21 have been utilized to isolate influenza virus antibodies and used for the production of virus vaccines^{41,42}. Moreover, a recent study demonstrated that antibodies combining the IGHV3-15/IGLV1-40 segments had superior neutralizing activities against the Zaire Ebola virus⁴³. In addition, we observed lower expression of inflammatory genes in ERS patients than in the HCs. We envision that our results will provide direction for the development of vaccines and antibodies for COVID-19 patients.

Interaction between immune cells may help expedite or defer recovery from COVID-19 infection. Our cell-to-cell prediction analysis utilizing scRNA-seq data indicated that, in ERS patients, B cell-derived IL-6, T cell-derived CSF1 (M-CSF), and CSF2 (GM-CSF) may promote monocyte proliferation and activation. As a result, monocytes may produce a larger number of inflammatory mediators, including IL-1 β and IL-6, contributing to inflammatory storm. In LRS patients, both DCs-derived TNFSF13 and IL-18 and T cell-derived IL-2, IL-4 may promote B cell survival, proliferation, and differentiation. Consequently, B cells produce numerous SARS-COV-2-specific antibodies to clear viruses, which is in agreement with the research of Zhou Y, et al.⁴⁴ who reported CD4⁺ T cells are activated into T-helper (Th) 1 cells and generate GM-CSF etc. to induced inflammatory CD14⁺CD16⁺ monocytes with high expression of IL-6 and accelerates the inflammation after 2019-nCoV infection.

The immune system comprises a network of cells, tissues and organs that mediate host defense against

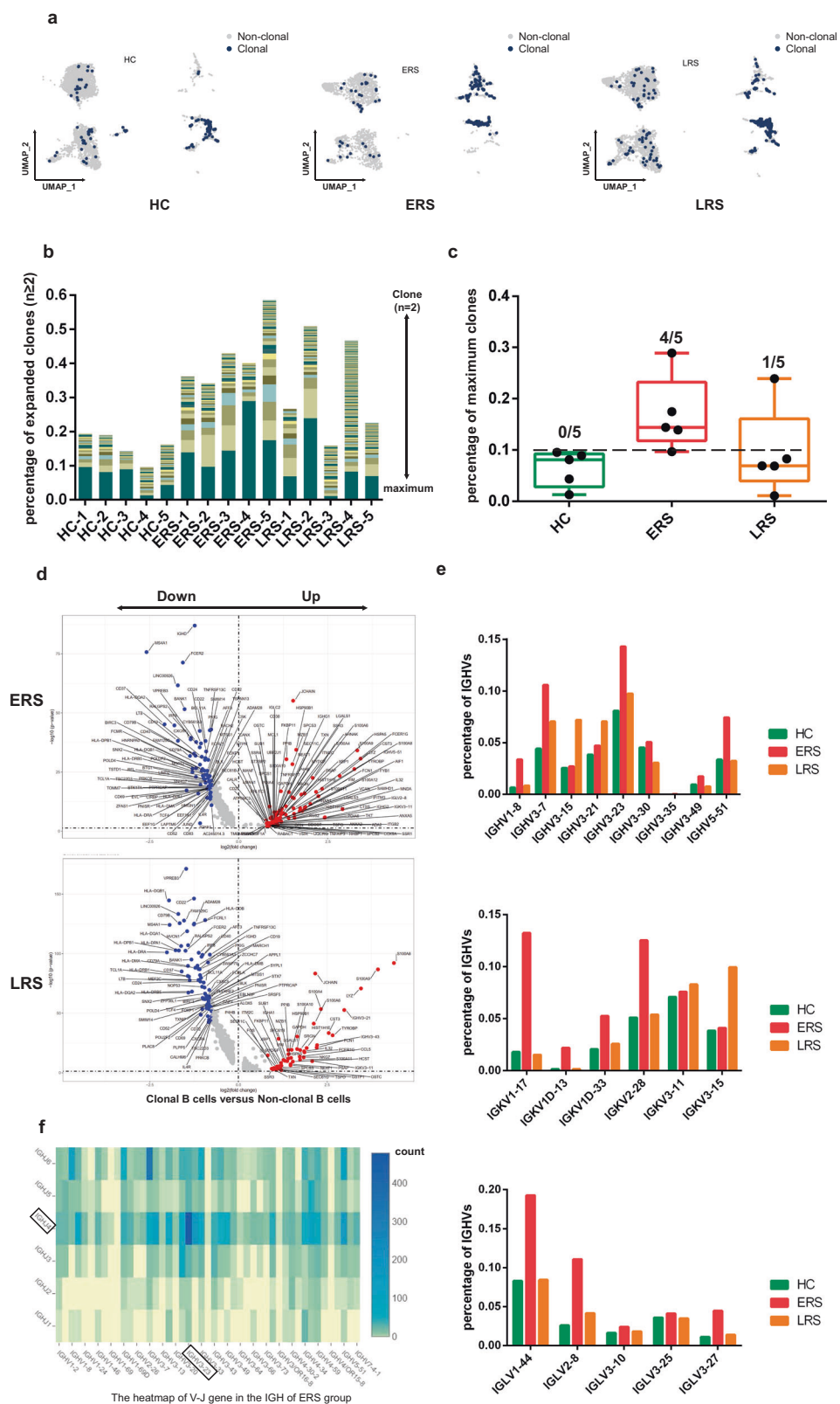


Fig. 6 (See legend on next page.)

(see figure on previous page)

Fig. 6 Expanded BCR clones and biased usage of VDJ genes observed in the COVID-19 patients. **a** The UMAP plot shows the B cell expansion status in the HCs and the ERS and LRS COVID-19 patients. **b** The bar plots show the clonal expansion status of B cells in peripheral blood from each individual sample. The number of color blocks represents the complexity of the clonal states. **c** Separate analysis of HC, ERS and LRS group by percentages of maximum clones revealed an enrichment of highly expanded clones (defined as comprising 10% or more of all BCR sequences; indicated by dotted line) in each group. None of healthy subjects had a highly expanded clone, versus four out of five patients in ERS, one out of five patients in LRS. Values are mean \pm SD. **d** The volcano plot shows the DEGs of expanded vs. non-expanded B cells in ERS and LRS patients. P values were calculated using a paired, two-sided Wilcoxon test and FDR corrected using the Benjamini–Hochberg procedure. **e** The bar plots show specific IGHV, IGKV, IGLV usage in the HCs and the ERS and LRS COVID-19 patients. **f** Heatmap showing IGH rearrangements in peripheral blood samples from ERS group.

pathogens. Immune cells can be classified into distinct types based on specific surface markers with the aid of flow cytometry and microscopy. However, not all immune cell types can be completely addressed by a separate analysis of phenotypic markers, as many markers are expressed by multiple cell lineages or are regulated differently during inflammation. In recent years, sequencing technology has been widely used in biological research. On this base, scRNA-seq is used in immunological research to seek to address previously unrecognized cellular heterogeneity, and to reveal key pathways in gene regulatory networks that predict immune function⁴⁵. In the present study, we applied single-cell technology to comprehensively characterize transcriptional changes in peripheral blood mononuclear cells during the recovery stage of COVID-19. scRNA-seq is a powerful tool to identify novel cell subsets during disease progression. In our study, CD14⁺IL1 β subpopulation was mapped using this method, which uncovered the originator cells in ERS patients. In conclusion, our study provided the first immune atlas of patients who have recovered from COVID-19 and identified adaptive immune dysregulation after discharge. The clonal expansion of both T and B cells indicated that the immune system has gradually recovered; however, the sustained hyper-inflammatory response for more than 7 days after discharge suggested the need for medical observation after patients are discharged from hospital. Longitudinal studies of recovered patients in a larger cohort might help to understand the consequences of the disease. The novel BCRs identified in our study may advance our understanding of B cell mechanisms and have potential clinical utility in COVID-19 immunotherapies.

Materials and methods

Patients

Ten COVID-19 patients diagnosed with by real-time fluorescent RT-PCR were collected in the Wuhan Hankou Hospital China. Patients were divided into early recovery stage (ERS) group and late recovery stage (LRS) group according to the days from first negative nucleic acid transfer date to blood sampling date. We defined the ERS group of five cases as the date of nucleic acid

turning negative to blood sampling is less than seven days and LRS group of five cases as is more than fourteen days. The 10 patients consisted of five males and five females and ranged from ages 30–80 years old, with a median of 58 years old in ERS, a median of 49 years old in LRS and a median of 55 years old in healthy controls (HCs). No significant differences were detected between HCs, ERS group and LRS group. The demographic characteristics of these patients and HCs are provided in Supplementary Fig. S1a-c. A written informed consent was regularly obtained from all patients. The study was approved by the Ethics Committee of Wuhan Hankou Hospital, China.

Quantitative reverse transcription polymerase chain reaction

The throat swab, sputum from the upper respiratory tract and blood were collected from patients at various time-points after hospitalization. Sample collection, processing, and laboratory testing complied with WHO guidance. Viral RNA was extracted from samples using the QIAamp RNA Viral Kit (Qiagen, Heiden, Germany) according to the manufacturer's instructions. SARS-CoV-2-infected patients were confirmed by use of qRT-PCR kit (TaKaRa, Dalian, China) as recommended by China CDC.

Single-cell collection and scRNA-seq

The peripheral blood mononuclear cell (PBMCs) were isolated from heparinized venous blood of patients or healthy donors using a Ficoll-Hypaque density solution according to standard density gradient centrifugation methods. For each sample, the cell viability exceeded 80%.

The single-cell suspensions of scRNA-seq samples were converted to barcoded scRNA-seq libraries using the Chromium Single Cell 5' Library, Gel Bead and Multiplex Kit, and Chip Kit (10x Genomics). The Chromium Single Cell 5' v2 Reagent (10x Genomics, 120237) kit was used to prepare single-cell RNA libraries according to the manufacturer's instructions. The FastQC software was used for quality check. The Cell Ranger software (version 3.1.0) was used for initial processing of the sequencing data.

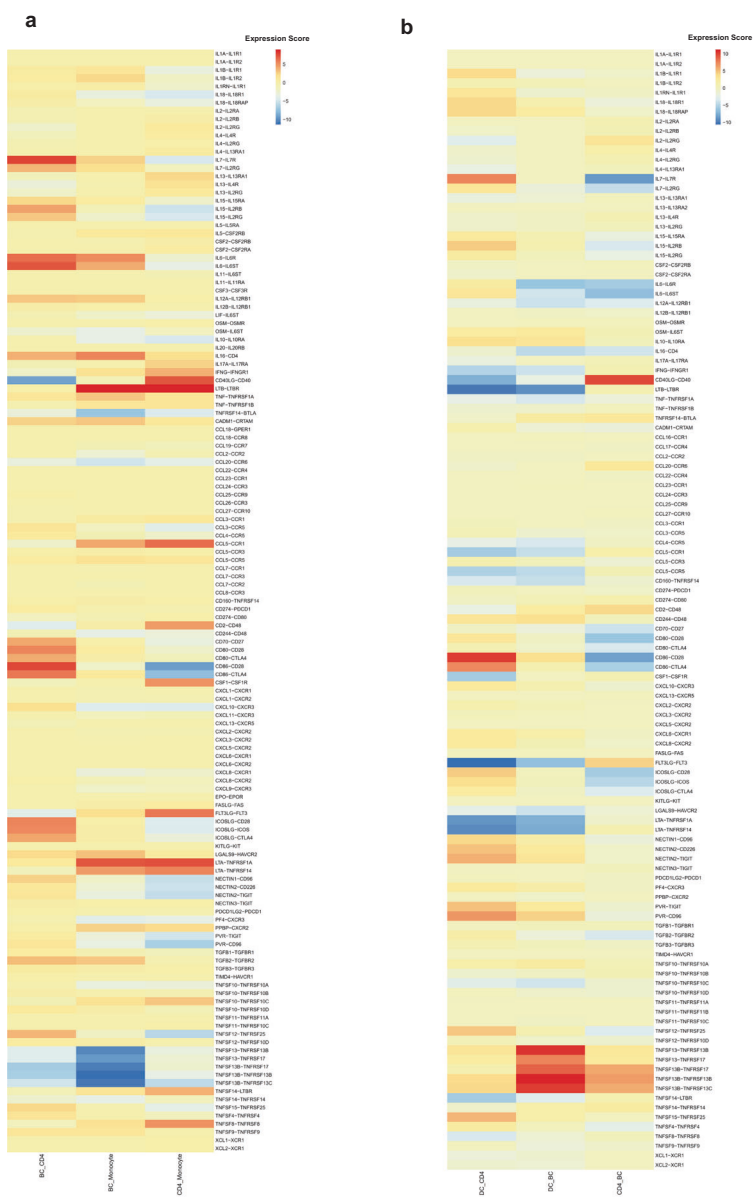


Fig. 7 Cell-to-cell communication among immune cells in the COVID-19 patients. **a** T cell-monocyte interactions, B cell-monocyte interactions, B cell-T cell interactions, and monocyte-T cell interactions in the ERS COVID-19 patients. **b** DC-T cell interactions, DC-B cell interactions, and T cell-B cell interactions in the LRS COVID-19 patients. **c, d** Schematics illustrating the key innate and adaptive immune cell functional alterations and main differences in cell-cell communications in the ERS (**c**) and LRS (**d**) COVID-19 patients.

ScRNA-seq data alignment and sample aggregating

We de-multiple and barcode the sample by using The Cell Ranger Software Suite (Version 3.1.0) (<https://support.10xgenomics.com>) and with command cell ranger count. After getting each sample gene counts, and aggregate them together. Finally, gene-barcode matrix of all ten patients and five HCs was integrated with Seurat v3⁴⁶ (<https://satijalab.org/>) and monocle3⁴⁷ (<https://cole-trapnell-lab.github.io/monocle3>). Following criteria were then applied to each cell, i.e., gene number between 200 and 7000. After filtering, a total of 128096 cells (13092/10035/13624/8329/12158 cells for HCs; 5163/7685/7171/10058/6581 cells for ERS; 3242/7895/7487/7164/8412 cells for LRS) were left for following analysis. The unique molecular identifier (UMI) count matrix was converted to Seurat objects using the R package Seurat v3.

Dimensionality reduction and clustering analysis

We handle the data with Log normalize before cluster and reduction, scale data with top 5000 most variable genes by using *FindVariableFeatures* function in R package Seurat v3. Clustering and dimensionality method mainly used in monocle3 package. For quality control, the genes used in PCA analysis have eliminated mitochondria (MT), and ribosomes (RPL and RPS) genes including MT-ND3, MT-ATP8, RPS15A, RPS28, RPS21, RPS27, RPS29, RPL36, RPL34, RPL37, RPL38, RPL39, RPL26 and et al. with 50 principal components, and then aligned together, followed by UMAP and t-SNE are both used after the results of the aligned. We used the default parameters with a shared nearest neighbor parameter optimized for each combined dataset inside Monocle3.

Differential analysis for clusters

Seurat package *FindAllMarkers* in Seurat v3 was used to perform differential analysis between the control and disease groups of the same cell type, the function parameters we used in Seurat v3 are default. For each cluster, differentially expressed genes (DEGs) were generated relative to all of the other cells.

Gene functional annotation

Gene ontology, gene-set enrichment analysis and KEGG pathway analyses from DEGs were performed using Metascape webtool⁴⁸ (www.metascape.org), which supports statistical analysis and visualization of functional profiles for genes and gene clusters.

TCR and BCR V(D)J sequencing and analysis

Full-length TCR/BCR V(D)J segments were enriched from amplified cDNA from 5' libraries via PCR amplification using a Chromium Single-Cell V(D)J Enrichment kit according to the manufacturer's protocol (10x Genomics). The TCR/BCR sequences for each single T/B cell

were assembled by Cell Ranger vDJ pipeline (v3.1.0), leading to the identification of CDR3 sequence and the rearranged TCR/BCR gene. Analysis was performed using Loupe V(D)J Browser v.2.0.1 (10x Genomics) (<https://support.10xgenomics.com>). In brief, a TCR/BCR diversity metric, containing clonotype frequency and barcode information, was obtained. Using barcode information, T/B cells with prevalent TCR/BCR clonotypes were projected on a t-SNE plot.

Cell-cell interaction analysis

The cell-cell interaction analysis was based on the expression of immune-related receptors and ligands.

The potential ligand-receptor interaction between one set of ligand-expressing cells and another set of receptor-expressing cells was calculated as the average of the product of ligand and receptor expression

(respectively, from set one and two) across all single-cell pairs:

$$I = \sum_i^n l_i \times \sum_j^m r_j \left(\frac{1}{m \times n} \right)$$

where I is the interaction score between ligand-expressing cells in set one and receptor-expressing cells in set two, l_i is the ligand expression of cell i in cell set one, r_j is the receptor expression of cell j in cell set two, n is the number of cells in set one and m is the number of cells in set two⁴⁹.

The gene list contained 168 pairs of well-annotated receptors and ligands, including cytokines, chemokines and co-stimulators. We estimated the potential interaction between two cell types mediated by a specific ligand-receptor pair by the product of the average expression levels of the ligand in one cell type and the corresponding receptor in the other cell type.

Acknowledgements

This study was supported by National Natural Science Foundation of China (81722034, 81670015, 81830054, 81530028, 91953122, 31871326, 81721003, 81988101, 91859205), the Ministry of Science and Technology Key Program (2018ZX09101002, 2017ZX100203205). Guangdong Natural Science Funds for Distinguished Young Scholar (2016A030306006), Guangdong Basic and Applied Basic Research Foundation (2020B1515020057), and Shanghai Pujiang Program (2019PJD059).

Author details

¹National Center for Liver Cancer, Second Military Medical University, 200438 Shanghai, China. ²State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, 510060 Guangzhou, China. ³Department of Respiratory and Critical Care Medicine, Changzheng Hospital, Second Military Medical University, 200003 Shanghai, China. ⁴Department of Critical Care, Wuhan Huoshenshan Hospital, 430113 Wuhan, Hubei, China. ⁵Department of Critical Care, Wuhan Hankou Hospital, 430000 Wuhan, Hubei, China. ⁶Laboratory of Vaccine and Antibody Engineering, Beijing Institute of Biotechnology, 100071 Beijing, China. ⁷GrandOmics Diagnosis Co. Ltd., 430014 Wuhan, Hubei, China. ⁸Berry Genomics Co. Ltd., 102206 Beijing, China. ⁹Eastern Hepatobiliary Surgery Hospital, Second Military Medical University, 200438 Shanghai, China. ¹⁰Ministry of Education (MOE) Key Laboratory of Signaling

Regulation and Targeting Therapy of Liver Cancer, Second Military Medical University, 200433 Shanghai, China

Author contributions

H.Y.W., W.C., C.L.X., and W. W. supervised the overall project design and execution. S.W.R. participated in data analysis and wrote the paper. Y.F.Z., L.W.D., and X.L.C. revised the paper and useful ideas in discussion. H. T., W.Q.L., and Y.S.M. collected the samples and clinical information. X.P.Z. performed the virus detection. J.M.L. measured the antibodies. X.X.L. and L.H.X. wrote the manuscript and created diagrams. J.G.Y., D.P.W., and J.T.D. participated in sequencing and bioinformatics analysis.

Data availability

The accession numbers for the sequencing raw data and processed data in this paper are GSA (Genome Sequence Archive in BIG Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences): HRA000069 and EGA: EGAS00001003449, respectively.

Conflict of interest

The authors declare that they have no conflict of interest.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information accompanies the paper at (<https://doi.org/10.1038/s41421-020-0168-9>).

Received: 23 March 2020 Accepted: 10 April 2020

Published online: 04 May 2020

References

- Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
- Li, Q. et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* <https://doi.org/10.1056/NEJMoa2001316> (2020).
- Livingston, E., Bucher, K. Coronavirus disease 2019 (COVID-19) in Italy. *JAMA* <https://doi.org/10.1001/jama.2020.4344> (2020).
- World Health Organization. Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV). WHO. [https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)) (2020).
- Huang, C. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506 (2020).
- Chan, J. F. et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* **395**, 514–523 (2020).
- Young, B. E. et al. Epidemiologic features and clinical course of patients infected with SARS-CoV-2 in Singapore. *JAMA* <https://doi.org/10.1001/jama.2020.3204> (2020).
- Cyranoski, D. et al. This scientist hopes to test coronavirus drugs on animals in locked-down Wuhan. *Nature* **577**, 607 (2020).
- Yao, X. et al. In vitro antiviral activity and projection of optimized dosing design of hydroxychloroquine for the treatment of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). *Clin. Infect.* <https://doi.org/10.1093/cid/ciaa237> (2020).
- Arabi, Y. M. et al. Critical care management of adults with community-acquired severe respiratory viral infection. *Intensive Care Med.* **46**, 315–328 (2020).
- Cao, B. et al. A trial of lopinavir-ritonavir in adults hospitalized with severe Covid-19. *N. Engl. J. Med.* <https://doi.org/10.1056/NEJMoa2001282> (2020).
- Assiri, A. et al. Epidemiological, demographic, and clinical characteristics of 47 cases of Middle East respiratory syndrome coronavirus disease from Saudi Arabia: a descriptive study. *Lancet Infect. Dis.* **13**, 752–761 (2013).
- Yin, Y. et al. MERS, SARS and other coronaviruses as causes of pneumonia. *Respirology* **23**, 130–137 (2018).
- Wang, D. et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* <https://doi.org/10.1001/jama.2020.1585> (2020).
- Zhao, J. et al. Airway memory CD4(+) T cells mediate protective immunity against emerging respiratory coronaviruses. *Immunity* **44**, 1379–1391 (2016).
- Fox, A. et al. Severe pandemic H1N1 2009 infection is associated with transient NK and T deficiency and aberrant CD8 responses. *PLoS ONE* **7**, e31535 (2012).
- Kumar, M. P. et al. Analysis of single-cell RNA-seq identifies cell-cell communication associated with tumor characteristics. *Cell Rep.* **25**, 1458–1468 (2018).
- Perlman, S. & Dandekar, A. A. Immunopathogenesis of coronavirus infections: implications for SARS. *Nat. Rev. Immunol.* **5**, 917–927 (2005).
- Rothenburg, S. & Brennan, G. Species-specific host-virus interactions: implications for viral host range and virulence. *Trends Microbiol.* **28**, 46–56 (2020).
- Wang, Z. et al. Recovery from severe H7N9 disease is associated with diverse response mechanisms dominated by CD8⁺ T cells. *Nat. Commun.* **6**, 6833 (2015).
- Walls, A. C. et al. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* <https://doi.org/10.1016/j.cell.2020.02.058> (2020).
- Channappanavar, R. & Perlman, S. Pathogenic human coronavirus infections: causes and consequences of cytokine storm and immunopathology. *Semin. Immunopathol.* **39**, 529–539 (2017).
- Liu, J. et al. Overlapping and discrete aspects of the pathology and pathogenesis of the emerging human pathogenic coronaviruses SARS-CoV, MERS-CoV, and 2019-nCoV. *J. Med. Virol.* **92**, 491–494 (2020).
- Channappanavar, R. et al. Dysregulated type I interferon and inflammatory monocyte-macrophage responses cause lethal Pneumonia in SARS-CoV-infected mice. *Cell Host Microbe* **19**, 181–193 (2016).
- Gu, J. et al. Multiple organ infection and the pathogenesis of SARS. *J. Exp. Med.* **202**, 415–424 (2005).
- Nicholls, J. M. et al. Lung pathology of fatal severe acute respiratory syndrome. *Lancet* **361**, 1773–1778 (2003).
- Villani, A. C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017).
- Thomas, P. G. et al. The intracellular sensor NLRP3 mediates key innate and healing responses to influenza A virus via the regulation of caspase-1. *Immunity* **30**, 566–575 (2009).
- Schmitz, N., Kurrer, M., Bachmann, M. F. & Kopf, M. Interleukin-1 is responsible for acute lung immunopathology but increases survival of respiratory influenza virus infection. *J. Virol.* **79**, 6441–6448 (2005).
- Bonilla, F. A. & Oettgen, H. C. Adaptive immunity. *J. Allergy Clin. Immunol.* **125**, S33–S40 (2010).
- Aoshi, T., Koyama, S., Kobiyama, K., Akira, S. & Ishii, K. J. Innate and adaptive immune responses to viral infection and vaccination. *Curr. Opin. Virol.* **1**, 226–232 (2011).
- Zhang, L. et al. Antibody responses against SARS coronavirus are correlated with disease outcome of infected individuals. *J. Med. Virol.* **78**, 1–8 (2006).
- Zhao, J. et al. Recovery from the Middle East respiratory syndrome is associated with antibody and T-cell responses. *Sci. Immunol.* **2**, eaan5393 (2017).
- Thevarajan, I. et al. Breadth of concomitant immune responses underpinning viral clearance and patient recovery in a non-severe case of COVID-19. *MedRxiv* <https://doi.org/10.1101/2020.02.20.20025841> (2020).
- Qin, C. et al. Dysregulation of immune response in patients with COVID-19 in Wuhan, China. *Clin. Infect. Dis.* <https://doi.org/10.1093/cid/ciaa248> (2020).
- Qiu, H. et al. Clinical and epidemiological features of 36 children with coronavirus disease 2019 (COVID-19) in Zhejiang, China: an observational cohort study. *Lancet Infect. Dis.* [https://doi.org/10.1016/S1473-3099\(20\)30198-5](https://doi.org/10.1016/S1473-3099(20)30198-5) (2020).
- Chen, G. et al. Clinical and immunologic features in severe and moderate Coronavirus Disease 2019. *J. Clin. Invest.* <https://doi.org/10.1172/JCI137244> (2020).
- Ryon, J. J., Moss, W. J., Monze, M. & Griffin, D. E. Functional and phenotypic changes in circulating lymphocytes from hospitalized zambian children with measles. *Clin. Diagn. Lab Immunol.* **9**, 994–1003 (2002).
- Thomson, C. A. et al. Germline V-genes sculpt the binding site of a family of antibodies neutralizing human cytomegalovirus. *EMBO J.* **27**, 2592–2602 (2008).
- Thomson, C. A., Little, K. Q., Reason, D. C. & Schrader, J. W. Somatic diversity in CDR3 loops allows single V-genes to encode innate immunological memories for multiple pathogens. *J. Immunol.* **186**, 2291–2298 (2011).

41. Fu, Y. et al. A broadly neutralizing anti-influenza antibody reveals ongoing capacity of haemagglutinin-specific memory B cells to evolve. *Nat. Commun.* **7**, 12780 (2016).
42. Goodwin, E. et al. Infants infected with respiratory syncytial virus generate potent neutralizing antibodies that lack somatic hypermutation. *Immunity* **48**, 339–349 (2018). e5.
43. Ehrhardt, S. A. et al. Polyclonal and convergent antibody response to Ebola virus vaccine rVSV-ZEBOV. *Nat. Med.* **25**, 1589–1600 (2019).
44. Zhou, Y. et al. Aberrant pathogenic GM-CSF+T cells and inflammatory CD14+CD16+monocytes in severe pulmonary syndrome patients of a new coronavirus. *bioRxiv* <https://doi.org/10.1101/2020.02.12.945576> (2020).
45. See, P., Lum, J., Chen, J. & Ginhoux, F. A single-cell sequencing guide for immunologists. *Front. Immunol.* **9**, 2425 (2018).
46. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).
47. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
48. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
49. Fernandez, D. M. et al. Single-cell immune landscape of human atherosclerotic plaques. *Nat. Med.* **25**, 1576–1588 (2019).