



ARTICLE OPEN

Single-cell immune profiling reveals distinct immune response in asymptomatic COVID-19 patients

Xiang-Na Zhao¹, Yue You^{2,3}, Xiao-Ming Cui⁴, Hui-Xia Gao⁵, Guo-Lin Wang⁴, Sheng-Bo Zhang^{2,3}, Lin Yao⁴, Li-Jun Duan⁴, Ka-Li Zhu⁴, Yu-Ling Wang⁵, Li Li⁵, Jian-Hua Lu⁵, Hai-Bin Wang⁵, Jing-Fang Fan⁵, Huan-Wei Zheng⁵, Er-Hei Dai⁵✉, Lu-Yi Tian^{2,3}✉ and Mai-Juan Ma⁴✉

While some individuals infected by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) present mild-to-severe disease, many SARS-CoV-2-infected individuals are asymptomatic. We sought to identify the distinction of immune response between asymptomatic and moderate patients. We performed single-cell transcriptome and T-cell/B-cell receptor (TCR/BCR) sequencing in 37 longitudinal collected peripheral blood mononuclear cell samples from asymptomatic, moderate, and severe patients with healthy controls. Asymptomatic patients displayed increased CD56^{bri}CD16⁻ natural killer (NK) cells and upregulation of interferon-gamma in effector CD4⁺ and CD8⁺ T cells and NK cells. They showed more robust TCR clonal expansion, especially in effector CD4⁺ T cells, but lack strong BCR clonal expansion compared to moderate patients. Moreover, asymptomatic patients have lower interferon-stimulated genes (ISGs) expression in general but large interpatient variability, whereas moderate patients showed various magnitude and temporal dynamics of the ISGs expression across multiple cell populations but lower than a patient with severe disease. Our data provide evidence of different immune signatures to SARS-CoV-2 in asymptomatic infections.

Signal Transduction and Targeted Therapy (2021)6:342

; <https://doi.org/10.1038/s41392-021-00753-7>

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of the coronavirus disease 2019 (COVID-19), has rapidly caused a worldwide pandemic with ever-increasing cases and COVID-19-related deaths.¹ COVID-19 patient exhibit a broad spectrum of clinical manifestations, ranging from mild or even asymptomatic infection to severe disease or death.² Therefore, understanding the host immune response involved in the disease course is of supreme importance to developing effective therapies.

In severe COVID-19 patients, hyper-inflammation responses referred to as cytokine storm^{3,4} and lymphopenia^{5,6} have been considered risk factors associated with the detrimental progression of COVID-19 patients. Elevated pro-inflammatory cytokines (e.g., IL-1 β , IL-6, and TNF- α) and inflammatory monocytes and neutrophils, and a sharp decrease in lymphocytes have also been reported in severe patients.^{3,5–12} Further, single-cell RNA sequencing (scRNA-seq) studies in peripheral blood mononuclear cells (PBMCs) or bronchoalveolar lavages of moderate and severe patients have revealed that moderate disease was associated with more protective T cell-dependent response, with exacerbated systemic inflammation and less effect T cells in severe disease.^{9,13–17} Longitudinal immune responses of moderate and severe COVID-19 patients have been analyzed by flow cytometry,¹⁰ while unbiased longitudinal single-cell transcriptome profiling is still missing. On the other hand, the contribution of asymptomatic individuals to the transmission of SARS-CoV-2 raises a significant public health concern.¹⁸ Despite the

clinical and immunological assessment of asymptomatic individuals,¹⁹ transcriptome profiles of asymptomatic individuals are lacking, which might help us understand the nature of the asymptomatic COVID-19 disease.

To explore characteristics that might lead to immunopathology in asymptomatic and moderate COVID-19, we performed scRNA-seq together with single-cell V(D)J sequencing using longitudinal PBMCs from 16 hospitalized COVID-19 patients and three healthy controls (HCs) to identify immunological profiles between distinct immune phenotype and disease severity.

RESULTS

Single-cell transcriptomes profiling of PBMCs from COVID-19 patients

A total of 16 laboratory-confirmed COVID-19 patients by real-time reverse transcription-polymerase chain reaction (rRT-PCR) and three HCs were enrolled. The demographics and clinical features of these subjects are shown in Table S1. Of the 16 patients, seven patients were asymptomatic (Pa), eight presented moderate (Pm) disease, and one exhibited severe (Ps) disease. Their ages ranged from 17 to 62 years old, and 12 of them were male. No significant differences in age were found between patient's groups. A total of 37 blood samples were collected from 16 patients, and eight of 16 patients provided more than or equal to three blood samples at different time points during hospitalization (Fig. 1a).

¹Center for Disease Control and Prevention of Chinese People's Liberation Army, Beijing, China; ²The Walter and Eliza Hall Institute of Medical Research, Parkville, Vic, Australia;

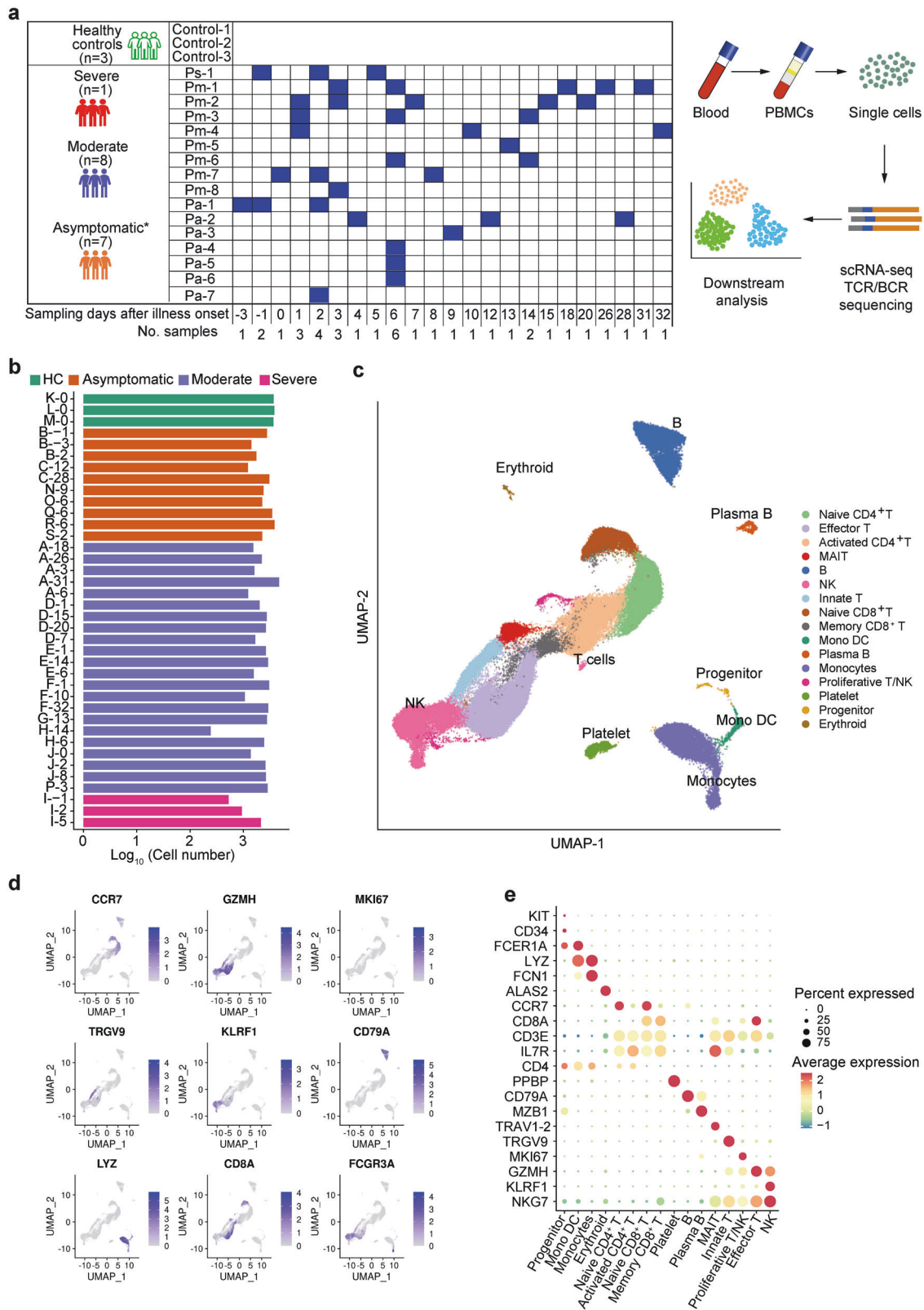
³Department of Medical Biology, The University of Melbourne, Parkville, Vic, Australia; ⁴State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing, China and ⁵The Fifth Hospital of Shijiazhuang, Hebei Medical University, Shijiazhuang, China

Correspondence: Er-Hei Dai (daieh2008@126.com) or Lu-Yi Tian (tian.l@wehi.edu.au) or Mai-Juan Ma (mjma@163.com)

These authors contributed equally: Xiang-Na Zhao, Yue You, Xiao-Ming Cui, Hui-Xia Gao

Received: 12 February 2021 Revised: 26 August 2021 Accepted: 26 August 2021

Published online: 16 September 2021



To explore the immunological changes in patients with asymptomatic, moderate, and severe disease, we analyzed the immune profiles of PBMCs from 16 patients and three HCs by scRNA-seq using the 10× Chromium platform (Fig. 1a). A total of 88,374 cells were included for analysis, including 77,168 cells from

16 patients and 11,206 cells from three HCs. On average, there were 2300 cells for each PBMCs sample (Fig. 1b). We identified 16 major cell types (Fig. 1c–e; Supplementary Fig. 1a–c), including mucosal-associated invariant T (MAIT) cells (*IL17R*⁺), innate T cells (*TRGV9*⁺), effector T cells (*GZMK*⁺), naive CD8⁺ T cells (*CCR7*⁺

Fig. 1 Single-cell RNA sequencing of peripheral blood cells from COVID-19 patients. **a** Timeline of blood sample collection in the 19 subjects (seven asymptomatic, eight moderate, one severe patient, and three healthy controls) and schematic outline of the study design. The days were recorded based on the time after symptom onset for moderate and severe patients and the first positive RT-PCR test for SARS-CoV-2 for asymptomatic patients. **b** Bar plot shows the \log_{10} transformed cell number of each sample for each subject at different time points. Green, orange, blue, and red represent samples collected from healthy controls, asymptomatic, moderate, and severe patients, respectively. The same color palette was used throughout the study. **c** Cell type UMAP representation of all merged samples. Sixteen cell types were identified by cluster gene signatures and color-coded. Each dot represents an individual cell. **d** Canonical cell markers that are used to annotate clusters as represented in the UMAP plot. Colored according to expression levels and legend labeled in log scale. **e** Dot plots of average expression and percentage of expressed cells of marker genes in each labeled cell type

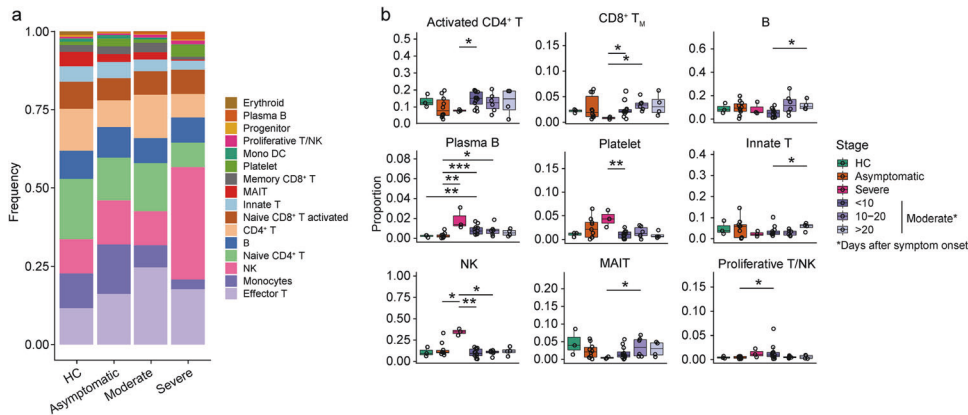


Fig. 2 Differences in major cell types compositions across disease conditions. **a** Proportion of cell types in PBMCs of healthy controls (HCs, $n = 3$), moderate ($n = 8$), severe ($n = 1$), and asymptomatic patients ($n = 7$). Colored according to cell type information. **b** Boxplots showing the percentages of each cell type to total cell number per PBMC sample in four disease conditions (HCs, asymptomatic, moderate, and severe) and stages. Boxes are colored according to disease conditions and stages of the moderate condition. The PBMC samples from moderate patients were classified into 3 stages (<10 days, 10–20 days, and >20 days) based on the days after symptom onset. Boxplots indicate the median and interquartile range (IQR); the whiskers represent 1.5 times the IQR. Each circle represents the proportion of each PBMC sample. Two-sided Kruskal–Wallis test was used for analysis, and a p value < 0.05 is considered significant. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. No asterisk indicates no statistical significance

SELL⁺), memory CD8⁺ T cells (*GPR183*⁺), naive CD4⁺ T cells (*CCR7*⁺ *SELL*⁺), activated CD4⁺ T cells (*IL7R*⁺ *CCR7*⁻), proliferative T/natural killer (NK) cells (*MKI67*⁺), NK cells (*NG2*⁺), progenitor cells (*CD34*⁺ *GATA2*⁺), B cells (*CD79A*⁺ *MS4A1*⁺), plasma B cells (*CD38*⁺ *MZB1*⁺), monocytes (*CD14*⁺ monocytes: *LYZ*⁺; *CD16*⁺ monocytes: *FCGR3A*⁺), platelet (*PPBP*⁺), monocyte-derived dendritic cells (Mono DC: *CD1C*⁺), and erythroid cells (*ALAS2*⁺). Further comparison of the proportions of 15 cell types among PBMCs that the proportions of T cell subsets were highly heterogeneous among different stages in moderate and asymptomatic patients, including activated CD4⁺ T cells and memory CD8⁺ T cells consistently lower abundance in severe and HCs (Fig. 2a, b; Supplementary Fig. 2a, b). We observed that one severe patient had an increased proportion of NK cells, plasma B cells, and platelets (Fig. 2b). There is no obvious difference in the abundance of major cell types between asymptomatic and moderate patients, except moderate patients had a higher proportion of plasma B cells.

Immune profiles of T cells and NK cells in COVID-19 patients
To further characterize the T and NK subsets, we extracted the data from T and NK cells. Fourteen cell subtypes were identified, including four CD4⁺ (naïve-*CCR7*⁺, central memory-*GPR183*⁺ *CCR7*⁺, effector memory-*CCR7*⁻ *SELL*⁻ *GZMA*⁺, and effector-*GZMA*⁺ *GZMB*⁺) and three CD8⁺ (naïve-*CCR7*⁺, central memory-*GPR183*⁺, and effector-*GZMA*⁺ *GZMB*⁺) T cell subsets, seven innate immune subsets (MAIT-*SLC4A10*⁺ *TRAV1*-2⁺, gamma-delta T ($\gamma\delta$ T) cells-*TRGV9*⁺ *TRDV2*⁺, immature NK cells, iNK-*KIT*⁺, *CD56*^{br}*CD16*⁻ NK cells, *CD56*^{dim}*CD16*⁺ NK cells, proliferative T/NK population-*MKI67*⁺), and a previously uncharacterized Th2-like lymphoid population (*CD4*⁻ *CD8A*⁻ *PTGDR2*⁺) (Fig. 3a, b, Supplementary Fig. 3).

We next compared the abundance of each cell type across disease conditions and stages (Fig. 3c; Supplementary Fig. S4b). We observed a decreased proportion of effector CD8⁺ T cells in asymptomatic patients compared to early-stage (<10 days post symptom onset) of moderate patients, and asymptomatic patients had a decreased proportion of central memory CD4⁺ T compared to early-stage of moderate patients and HCs (Fig. 3c). Of interest, the abundance of *CD56*^{br}*CD16*⁻ NK cells were significantly higher in asymptomatic patients than in severe and early stage of moderate patients, and there was an increasing trend in moderate patients over time (Fig. 3c). In contrast, the *CD56*^{dim}*CD16*⁺ NK, the most abundant NK subset, was substantially enriched in the severe patient. Like *CD56*^{br}*CD16*⁻ NK, its precursor iNK also increased in asymptomatic patients compared to moderate patients. These results indicated that asymptomatic patients had distinct T and NK cell responses during infection.

Next, we sought to identify the specific signature of the NK cells with distinct distribution in asymptomatic and severe conditions. We found that *CD56*^{br}*CD16*⁻ NK cells have high expression of *XCL1*, *XCL2*, and *IFNG* (Fig. 3d, f), consistent with our knowledge that these cells are efficient cytokine producers.²⁰ The Th2-like lymphoid cells were *TCR*⁻ *CD3*⁻ *CD4*⁻ *CD8*⁻ but expressed Th2 markers such as *PTGDR2* and *GATA3*; they were classified mainly as Th2-like cells (Supplementary Fig. 5a). We also found that *TNFRSF19* is uniquely expressed in the Th2-like cells in this dataset (Supplementary Fig. 5b). However, *TNFRSF19* is absent in most immune cells according to previous study²¹ and the human lung cell atlas database.²² It is highly expressed in epithelial cells such as ciliated cells, which express ACE2 and are considered as entry cells of SARS-CoV-2 (Supplementary Fig. 5c).

To further investigate the difference of transcriptomes for each cell type of T and NK cells across different conditions, we

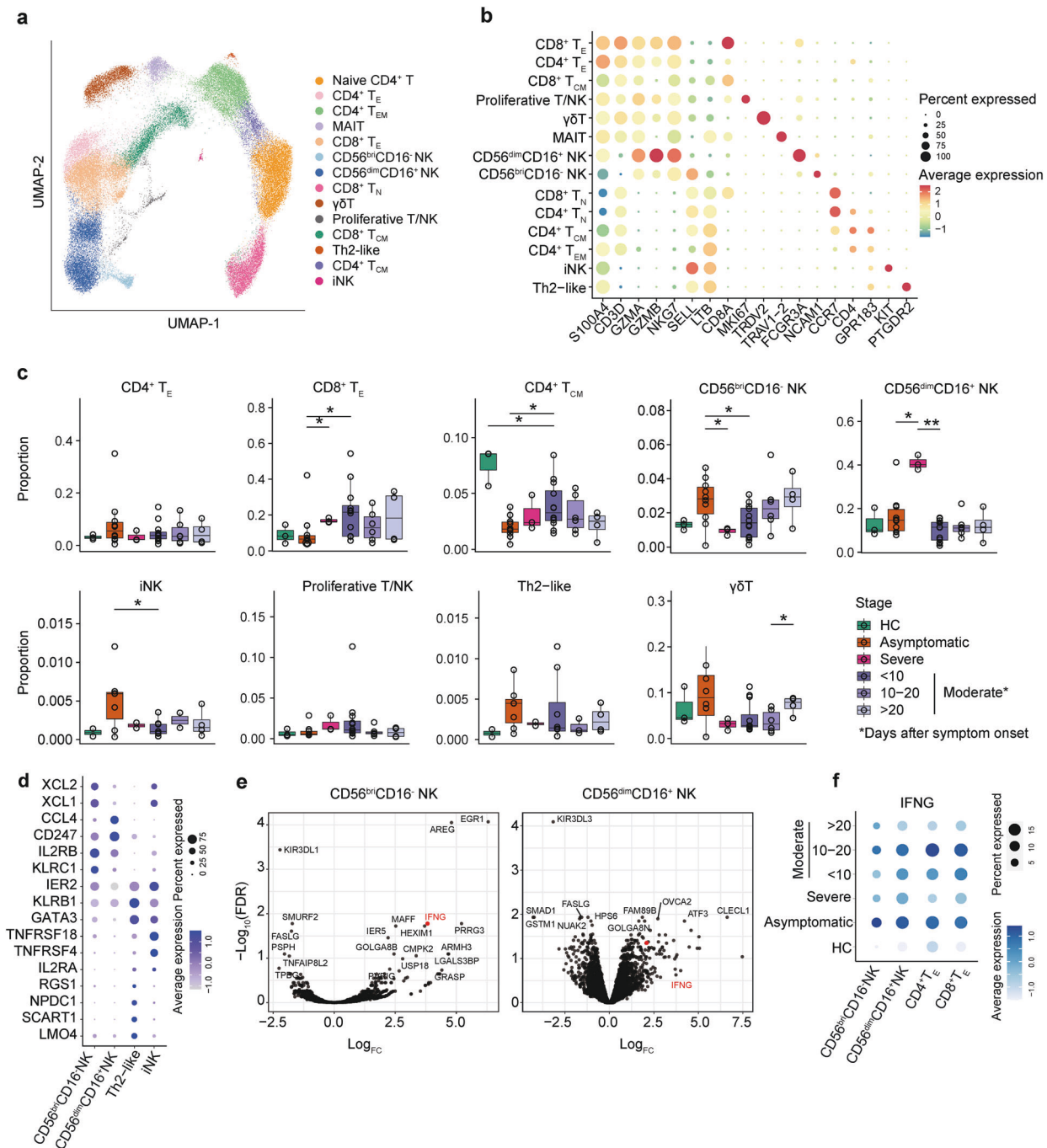


Fig. 3 Identification and characterization of the subpopulation of T and innate immune cells in COVID-19 patients. **a** UMAP of T and NK cells by Seurat. Cell types were identified by the marker genes. Each circle represents an individual cell. A total of seven T cell subtypes and seven innate immune cell subtypes were identified and color-coded. **b** Dot plot of average expression and percentage of expressed cells of selected canonical markers in each labeled cell subtype. **c** Boxplot showing the proportions of cell subtypes to the number T and NK cells in each PBMC sample in different disease conditions. Boxes are colored according to disease conditions and stages of the moderate condition. The PBMC samples from moderate patients were classified into 3 stages (<10 days, 10–20 days, and >20 days) based on the days after symptom onset. Boxplots indicate the median and interquartile range (IQR); the whiskers represent 1.5 times the IQR. Each circle represents the proportion of each sample. Two-sided Kruskal–Wallis test was used for analysis, and a p value < 0.05 is considered significant. **d** Dot plot showing the average expression and percentage of expressed cells of selected differentially expressed genes (DEGs) between three NK subtypes and Th2-like subset. **e** Volcano plots of top DEGs between asymptomatic and healthy samples for CD56^{br}CD16⁻ (left) and CD56^{dim}CD16⁺ (right) NK cell subtypes. Genes with a log fold change above 1 and false discovery rate (FDR, Benjamini–Hochberg) less than 0.05 were selected. **f** Dot plot showing the expression of *IFNG* on NK and effect T cells for different disease severity and stages. * p < 0.05, ** p < 0.01, *** p < 0.001. HC healthy control. No asterisk indicates no statistical significance

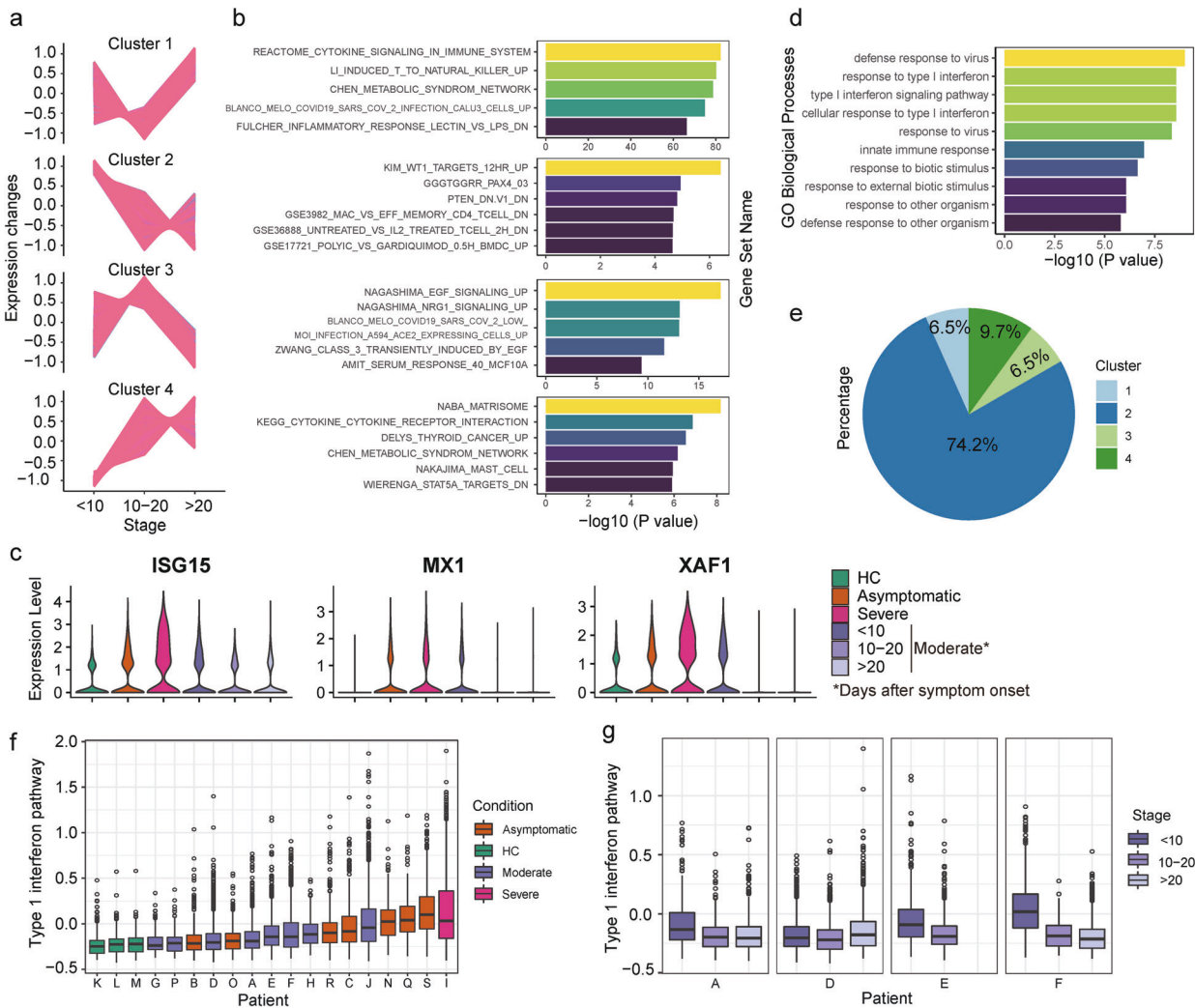


Fig. 4 Increased IFN-1 signaling pathway in severe and early stage moderate patients. **a** Gene clusters generated by Mfuzz R package based on genes' expression pattern along the progression of the disease in moderate patients using NK cells and effector T cells. **b** Top gene sets in gene set enrichment analysis for the gene in each cluster in panel (a). Gene sets are colored by \log_{10} transformed p -values representing genes in the overlap. **c** Violin plots of selected DEGs generated by comparing different conditions. HC, healthy control. **d** Bar plots of top ten enriched GO terms from genes upregulated in severe conditions in CD56^{dim}CD16⁺ NK cells. **e** The percentage of IFN-1 related genes found in four clusters in (a). **f** Box plots of the IFN-1 pathway activity score calculated across patients on NK and effector T cells. Colors denote disease conditions. **g** Boxplots of the IFN-1 pathway activity score for each moderate patient and summarised at each stage. Colors denote defined disease stages

performed systematic differential gene expression (DGE) analysis (Data S1). We found distinct signatures expressed in severe and asymptomatic patient samples in NK cells and effector T cells (Supplementary Fig. 6a–d). *IFNG* was upregulated in CD56^{dim}CD16⁺ NK and CD56^{brn}CD16⁻ NK cells in asymptomatic conditions compared to HCs (Fig. 3d–f; Supplementary Fig. 6c, d). We also found *IFNG* expression showed stage-specific expression in moderate samples, with the highest expression at 10–20 days after symptom onset.

To further explore the IFN-I pathway activity at the single-cell level, we analyzed the stage-dependent expression patterns and explored related gene sets based on both NK cells and effector T cells of moderate patients. We identified four stage-dependent expression patterns (Fig. 4a) and their related gene sets (Fig. 4b, Data S2). We found upregulation of genes such as *EGR1* and *NR4A1* (Data S2) in later stages as shown in cluster 3 and is known to be induced by TCR stimulation and enhance T cell functions.^{23,24} Genes expression in cluster 2 decreases with time, and one enriched gene set contains genes downregulated with *PTEN* knockdown, and *PTEN* was demonstrated with function

promoting type 1 interferon responses and antiviral innate immunity.²⁵ Meanwhile, we found that type 1 interferon (IFN-I)-related genes, *ISG15*, *MX1*, and *XAF1*, expressed more severe and moderate patients at the early stage (Fig. 4c), and a group of widely upregulated signatures in disease states are genes in IFN-I signaling pathway, especially in severe condition (Fig. 4d, e, Data S3). We then summarized IFN-I pathway activities per cell and found heterogeneous results among different patients (Fig. 4f, Supplementary Fig. 7a). Apart from HCs who show the stable low activity of IFN-I pathway and cells from severe patients with the highest activity on average, the moderate and asymptomatic patients have a wide spectral of IFN-I activity ranging from HCs to severe conditions. In addition to intra-patient variability, the IFN-I pathway activity also varies on time, decreases in later stages in moderate patients, suggesting a recovery of disease (Fig. 4g, Supplementary Fig. 7b). The disease progression stage is hard to define for asymptomatic patients. However, the highly heterogeneous IFN-I activity suggests the asymptomatic patients that we profiled in this study were in different disease stages.

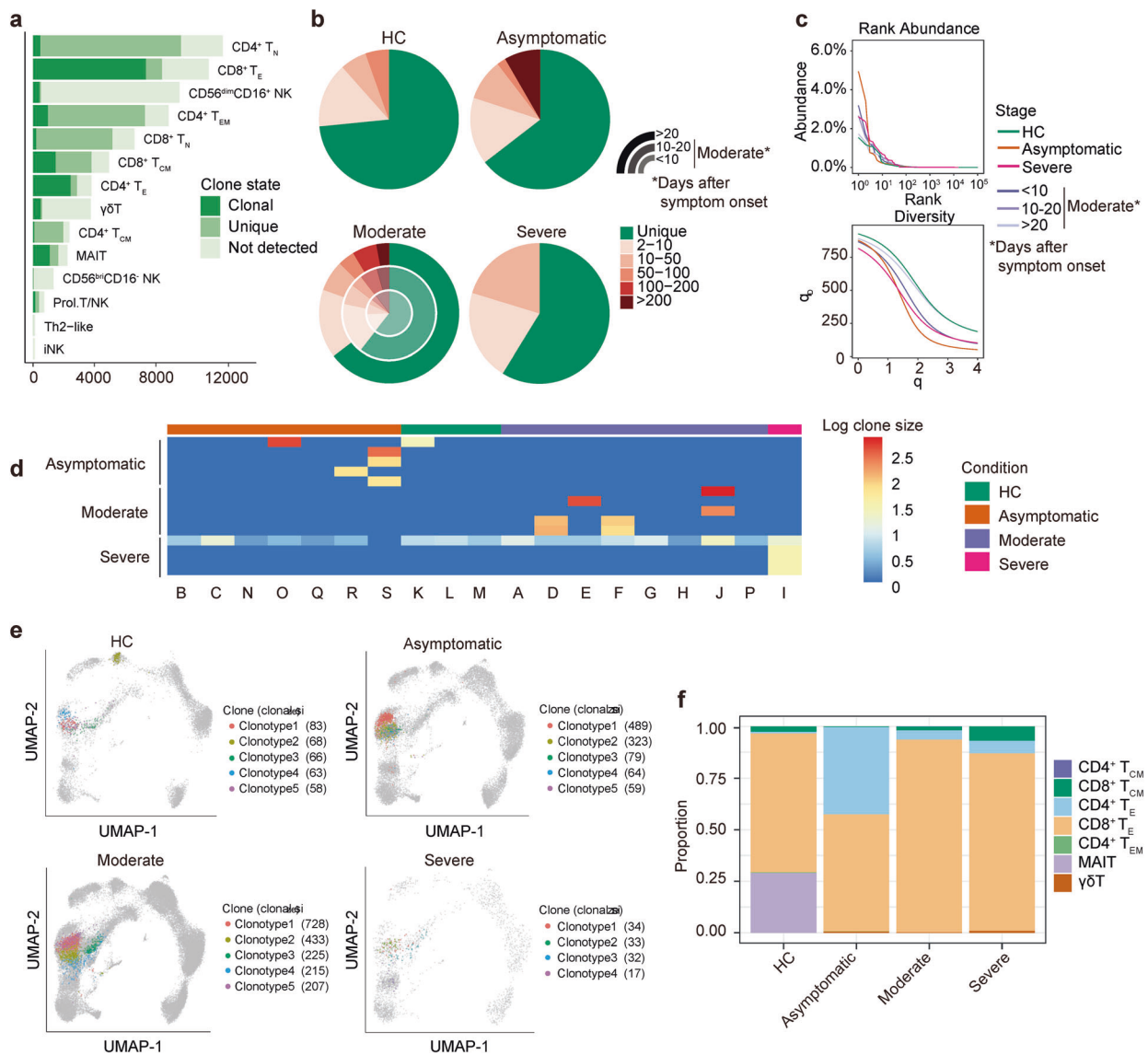


Fig. 5 Expanded TCR clones and selective usage of V(D)J genes. a Bar plots showing the cells from T and NK cell subtypes whether they have TCR and their TCR clonal status. **b** The percentage of the clonal status of T cells with TCR. The clonal status was defined by clone size in four disease conditions. **c** TCR abundance and diversity on CD4⁺ and CD8⁺ T cells across disease conditions and stages generated by *alakazam* package. The 95% confidence interval is estimated via bootstrapping ($B = 200$). **d** Heatmap of \log_{10} -transformed clone sizes of top three abundant clonotypes within each condition (not including healthy controls [HC]) across patients. **e** UMAP of T cells derived from PBMCs for different conditions. Clusters are denoted by colors labeled with TCR clones with the top 5 largest clone sizes in each condition. **f** Proportions of cell types that produce the top five abundant clonotypes within each condition. Color denotes cell types

Clonal expansion in T cells and usage of V(D)J genes COVID-19 patients

To evaluate the relationship of clonal expansion among individual T cells and usage of V(D)J genes across different conditions, we analyzed single-cell TCR sequencing data and reconstructed high-quality TCR sequence in 70.5% of the T cells with various degrees of clonal expansion. We observed that both effector CD8⁺ and CD4⁺ T cell subsets displayed more TCR clonal expansion (Fig. 5a; Supplementary Fig. 8a). Both asymptomatic and moderate patients had high clonal expansion compared to the severe patient and HCs (Fig. 5b). We performed a quantitative analysis of TCR clonal abundance and diversity to mitigate the difference in sample size between different conditions. The TCR expansion decreases in moderate patients from early (<10 days) to late-stage (>20 days), indicating recovery of the disease. We found more TCR expansion in asymptomatic patients than moderate patients at the early stage (Fig. 5c). We summarized the distribution of top

clone types per condition and found that most of them are COVID-19 specific (Fig. 5d). The clustering of the CDR3 sequence showed a similar sequence enriched in multiple patients, suggest they are reactive to the SARS-CoV-2 virus (Supplementary Fig. 8b). The UMAP visualization of these TCR clones showed that the most abundant clones in asymptomatic and moderate patients have larger clone sizes, consistent with the TCR diversity analysis. It also highlighted the different cell populations in different conditions that enriched for most abundant TCR clones (Fig. 5e). We found that asymptomatic patients have significantly more CD4⁺ effector T cells with the most abundant TCR clones, while most of the abundant TCR clone types were in CD8⁺ effector T cells for moderate and severe patients (Fig. 5f). We then compared the usage of V(D)J genes across disease conditions and disease stages of moderate patients (Supplementary Fig. 8c–e). The preferred *TRBJ* gene in asymptomatic patients was *TRBJ2-2*, *TRBJ2-1* for moderate patients, and *TRBJ2-7* for one severe patient

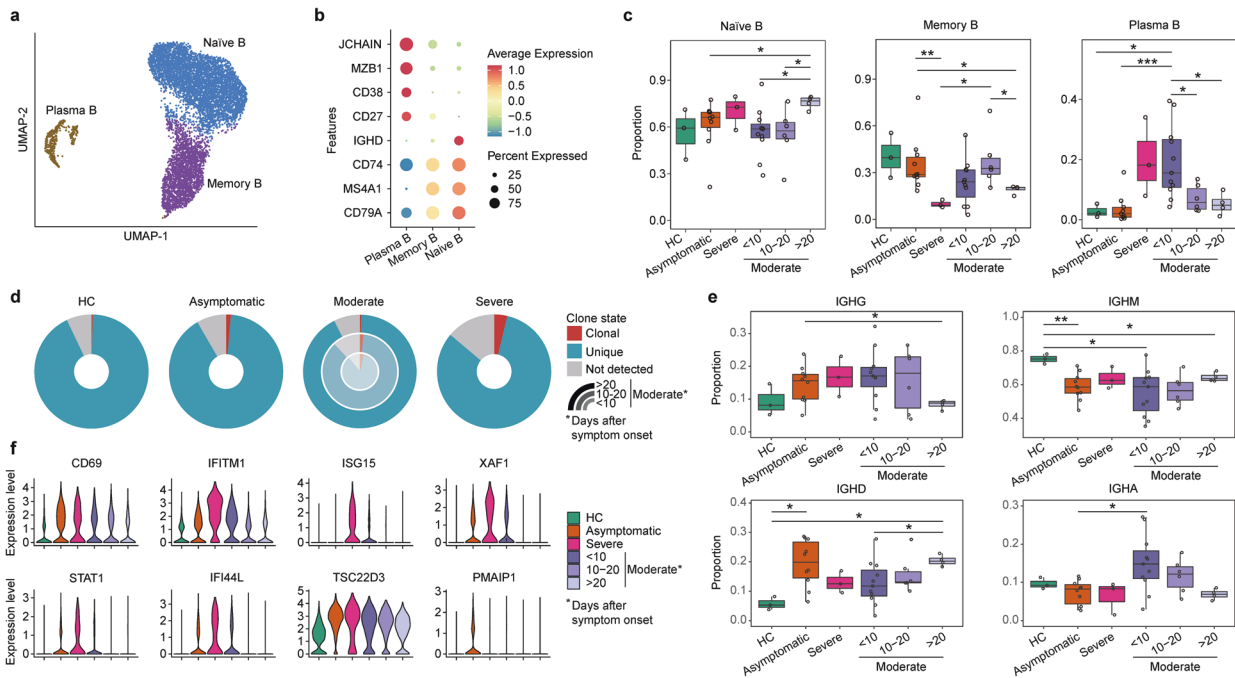


Fig. 6 Subpopulation analysis of B cells. **a** UMAP projection of B cells. Each dot represents a single cell, colored according to cell type. **b** Dot plot of canonical cell markers used to annotate clusters in the UMAP plot. **c** Boxplots showing the differences in percentages of each cell type to the number of B cells comparing different disease conditions to healthy controls (HCs). Boxes are colored according to disease conditions and stages of the moderate condition. The PBMC samples from moderate patients were classified into 3 stages (<10 days, 10–20 days, and >20 days) based on the days after symptom onset. Boxplots indicate the median and interquartile range (IQR); the whiskers represent 1.5 times the IQR. Each circle represents the proportion of each sample. Two-sided unpaired Mann–Whitney *U* test was used for analysis, and a *p* value < 0.05 is considered significant. **d** The percentage of clonal status in B cells that have BCR, across disease conditions and stages. **e** Boxplots show the proportion of IgG, IgM, IgD, and IgA, comparing different disease conditions to HCs. Boxes are colored according to disease conditions and stages of the moderate condition. The PBMC samples from moderate patients were classified into 3 stages (<10 days, 10–20 days, and >20 days) based on the days after symptom onset. Boxplots indicate the median and interquartile range (IQR); the whiskers represent 1.5 times the IQR. Each circle represents the proportion of each sample. Two-sided Kruskal–Wallis test was used for analysis, and a *p* value < 0.05 is considered significant. **f** Violin plots showing the gene expression levels of selected DEGs generated genes in B cells by comparing different conditions. **p* < 0.05, ***p* < 0.01, ****p* < 0.001. No asterisk indicates no statistical significance

(Supplementary Fig. 8c). Collectively, the different patterns of TCR clonal expansion and diversity and the selective usage of V(D)J genes indicated that different immunodominant epitopes may drive the molecular composition of T cell responses and may be associated with SARS-CoV-2-specific infection.

Features of B cells and expansion and specific rearrangements of V (D)J genes

We extracted single-cell B cells sequencing data and identified three B cell subsets according to the expression of canonical B cell markers, including naïve B cells (*MS4A1*⁺ *IGHD*⁺), memory B cells (*CD27*⁺), and plasma B cells (*MZB1*⁺ *CD38*⁺) (Fig. 6a, b). Moderate patients at an early stage (<10 days after symptom onset) had a higher proportion of plasma B than HCs and asymptomatic patients, and a declining trend of the plasma B proportion was observed over time for the samples of moderate patients (Fig. 6c). We further reconstructed high-quality BCR sequences in more than 80% of the B cells using single-cell BCR sequencing data and found asymptomatic patients displayed less BCR clonal expansion. In contrast, an obvious BCR clonal expansion was observed in moderate patients at the early stage but decreased over time, suggesting humoral immune responses declined at the convalescent stage (Fig. 6d). The severe patient has even more clonal expansion, suggesting a strong humoral immune response (Fig. 6d). Further analysis of the distribution of IgG, IgM, IgD, and IgA at different disease conditions and stages revealed that asymptomatic and moderate patients had lower IgM compared to HCs (Fig. 6e). The IgG was highly variable at the early moderate patient

sample and returned to similar levels of HCs at the late stage (>20 days post symptom onset).

We then assessed V(D)J rearrangements of the BCR and analyzed the usage of V(D)J genes across different disease conditions and stages (Supplementary Fig. 9b). We found different skewing of BCR usage in different conditions. The top two paired V–J frequencies in asymptomatic patients were *IGHV3–23/IGHJ4* and *IGLV1–44/IGLJ3*, whereas *IGLV1–44/IGLJ3* and *IGLV2–14/IGLJ2* for moderate patients and *IGLV1–51/IGKJ3* and *IGHV3–33/IGHJ4* for the severe patient (Supplementary Fig. 9b). Collectively, increased B cell clonality from asymptomatic to severe condition and skewed usage of the IGHV and IGKJ genes in different disease conditions suggest that SARS-CoV-2 infection perturbs V(D)J rearrangements of B cells.

To further investigate transcriptomic changes of B cells after SARS-CoV-2 infection, we performed a DGE analysis comparing different conditions in each B cell subsets (Fig. 6f). We found an increased expression of activation marker gene *CD69* in patients, regardless of disease severity, compared to HCs. Like T and NK cells, we observed an increased expression of *IFITM1*, *ISG15*, *XAF1*, *STAT1*, and *IFI44L* involving the IFN-I signaling pathway in the severe and moderate patients, but moderate patients displayed relative low-level expression (Fig. 6f and Supplementary Fig. 9b). By contrast, asymptomatic patients displayed a lack or less expression of these genes.

DISCUSSION

Although the immune responses to SARS-CoV-2 infection have been studied in patients with moderate and severe disease,^{13–17}

the mechanisms under asymptomatic infection of COVID-19 are less studied. In this study, we performed scRNA-seq and V(D)J-sequencing on longitudinal PBMCs from asymptomatic individuals of COVID-19 and systematically compared their immunological characteristics with moderate and severe COVID-19 patients. Our study confirms previously published data on increased type-I interferon (IFN-I) response^{26,27} and the increased plasma B cells and decreased T subsets in severe patient,^{10,13–17} but not for monocytes, which we did not observe might due to the sampling time that only early stage severe samples.²⁸ In addition to that, we found that the asymptomatic condition is not just an intermediate state between healthy and moderate but has unique immunological features. Overall, PBMCs in asymptomatic patients had lower IFN-I related gene expression than in severe and early-stage moderate conditions. In contrast, moderate patients at the early stage of infection had an increased IFN-I response, but less than the severe condition and declined over time. However, the heterogeneous IFN-I activity was observed among different patients that 11 patients (4 asymptomatic and 7 moderate patients) had IFN-I activity blow or close to the average level of 0, and the rest of the patients showed more IFN-I activity, reflecting the heterogeneous response of IFN-I induction in patients and even their heterogeneous pathogenesis.²⁹ Importantly, a rapid decrease of IFN-I activity is associated with a more rapid recovery of the disease. IFN-I has direct antiviral activity, and their immunopathological role was also previously reported.³⁰ The IFN-I response induced the accumulation of pathogenic inflammatory monocytes–macrophages and vascular leakage, leading to death in a BALB/c mice model of acute respiratory distress syndrome (ARDS).³¹ Moreover, a delayed but considerable IFN-I response was proposed that is critical for the development of ARDS and increased lethality during pathogenic human coronavirus infection.^{32,33} Therefore, our data suggest that early moderate IFN-I response might effectively control viral replication and prevent severe COVID-19.

Effector CD4⁺ T cells play a key role in regulating the antiviral inflammatory response and mediating viral clearance through direct cytotoxic effects on virus-infected cells.³⁴ Although we observed similar frequencies of effector CD4⁺ T cells among different conditions, the effector CD4⁺ T cells in asymptomatic patients had upregulated *IFNG* compared to other disease conditions (Fig. 3f). Of note, a recent study has shown that the magnitude of T cell responses to SARS-CoV-2 was similar between asymptomatic infections and symptomatic COVID-19 patients,³⁵ but SARS-CoV-2-specific T cells produced a higher level of IFN- γ and IL-2 in asymptomatic patient,³⁵ which is in line with our study. A previous study also showed that SARS-CoV-2-specific CD4⁺ T cells in severe but not mild COVID-19 displayed low avidity.³⁶ In addition, while we observed fewer effector CD8⁺ T cells in asymptomatic patients than moderate and severe patients at the early stage of the infection, the effector CD8⁺ T cells in asymptomatic patients expressed a higher level of *IFNG* but not in severe patient and moderate patients until ten days post-symptoms onset. Because most PBMC samples from asymptomatic patients were collected <10 days after the first positive RT-PCR testing of SARS-CoV-2, we speculated that higher expression of *IFNG* in T cells plays an important role in antiviral infection, especially during the early stage of infection. These data suggest that early activation of effector CD4⁺ and CD8⁺ T cells expressing a higher level of *IFNG* may play an important role in protecting SARS-CoV-2 infection in asymptomatic patients.

Despite the difference in adaptive immune response, we also identified profound differences in innate immune responses in asymptomatic patients compared to moderate and severe patients. Asymptomatic patients had significantly increased CD56^{bri}CD16⁻ NK cell fractions than moderate and severe patients at the early stage. We found CD56^{bri}CD16⁻ NK subsets had upregulated cytokine-related genes such as *XCL1*, *XCL2*, and *IFNG*,

consistent with our knowledge that these regulatory cells act as potent cytokine and chemokine producers. In addition to the difference in abundance, there were significant differences in gene expressions between asymptomatic and moderate conditions, with cytokine-related genes such as *IFNG* upregulated in asymptomatic conditions, but not severe patient and moderate patients until ten days post-symptoms onset. CD56^{bri}CD16⁻ NK cells have been linked to virus infection. Infection of influenza A Virus-induced NK cell hyperresponsiveness and cytokine production, particularly in the CD56^{bri}CD16⁻ NK subset.³⁷ An asymptomatic hemophiliac patient co-infected with HIV/HCV also had increased CD56^{bri}CD16⁻ NK cells.³⁸ These results suggest that the CD56^{bri}CD16⁻ regulatory NK cells may play a critical role in protecting SARS-CoV-2 infection in asymptomatic patients.

TCR and BCR repertoire profiling are important to reflect the disease's adaptive immune status and develop new therapeutics for infectious disease. Therefore, another central part of our study is to assess the repertoire diversity of TCR and BCR and especially their clustering, enabling us to deduce COVID-19-relevant TCR or BCR signatures. Of note, we found more TCR clonal expansion in asymptomatic patients compared to moderate patients (also highly diverse than severe patient) at the early stage and severe patient, which is in line with previous studies that in patients with a moderate^{14,39} or mild^{36,40} clinical course and but not patients with a severe clinical course, T cell repertoires displayed high clonality. Moreover, the most abundant TCR clones were observed in CD4⁺ effector T cells of asymptomatic patients and CD8⁺ effector T cells of moderate and severe patients. These results indicate that the effective TCR diversification on CD4 or CD8 T cells may contribute the outcome and immune control in COVID-19.

While asymptomatic COVID-19 patients displayed more abundant TCR diversity, on the contrary, they showed less BCR clonal expansion compared to moderate and severe patients, which is in line with the previous studies^{28,40,41} and are likely to be associated with antibody-secreting B cells that produce antibodies to neutralize the infecting pathogen.^{41,42} We found that patients with different conditions shared a common IGLV1/2 and IGLJ2/3 usage pattern and IGHV3 and IGHJ4 usage. Moreover, we found a higher proportion of IGHV3 and IGHJ4 usage in asymptomatic patients. In agreement with our study, IGHV3 and IGHJ4 usage were identified in clusters specific for antibody-positive individuals with COVID-19⁴⁰ and moderate to severe patients.¹⁴ Notably, a high proportion of IGHV3 and IGHJ4 usage was also observed in HCs.¹⁴ Previous studies have shown extensive class switching to IgG and IgA subclasses with limited somatic hypermutation.^{40,43} Because our study produced few B cells (~300 cells per sample) and a low percentage of clonal BCR, we could not perform clustering antibody class switching, somatic hypermutation other analyses. Despite our study and other studies that have provided a relationship between BCR repertoire and disease infection, detailed work on BCR repertoire is needed to reveal its role in disease progression in asymptomatic patients.

There are several limitations to this study. One of the major limitations is only one patient with severe COVID-19 and paucity of later samples from the severe patient. These blood samples from patients with severe COVID-19 are challenging to obtain since very few newly diagnosed patients were reported in China when our study was initiated. Although the results from one severe patient are consistent with previous studies, it will be essential to gather such samples further to identify the immune characteristic between asymptomatic and severe COVID-19. Second, because of frequent blood draws of patients during their hospitalization for clinical lab testing, they only provide small volume blood samples, resulting in insufficient PBMCs to detect virus-specific T cells and B cells by flow cytometry to assess their relationship with results of scRNA-seq.

Our data clearly show the immune profiles in asymptomatic COVID-19 patients and highlight the difference of immune

response toward disease progression. The specific signatures in asymptomatic patients will increase understanding of COVID-19 disease severity and guide early prediction and therapeutics.

MATERIALS AND METHODS

Ethics statement

The study was conducted following the Declaration of Helsinki, and the Institutional Review Board of the Academy of Military Medical Sciences approved the study protocol (IRB number: AF/SC-08/02.46). All patients or their surrogates provided written informed consent.

Patients

Sixteen patients diagnosed with SARS-CoV-2 infection were enrolled from the Fifth Hospital of Shijiazhuang from March to April 2020. SARS-CoV-2 RNA was detected in the patient's nasopharyngeal swab or sputum specimens by real-time reverse-transcriptase PCR (RT-PCR) using the SARS-CoV-2 nucleic acid detection kit (Cat No. DA0930-DA0932, DAAN GENE Ltd., Guangzhou, China). Peripheral blood was collected from all patients during hospitalization, and blood draws from patients occurred in concert with usual care and the patient's willingness to avoid frequent blood sampling and unnecessary personal protective equipment usage. The patients' demographic, clinical features, laboratory findings, and chest radiographs were collected from their electronic medical records.

According to the diagnostic and treatment guidelines for SARS-CoV-2 issued by the Chinese National Health Committee (Trail Version 7), the disease severity was defined as asymptomatic, moderate, and severe. Asymptomatic infection was defined as an individual who had a positive SARS-CoV-2 by RT-PCR without any associated clinical symptoms in the preceding 14 days and during hospitalization. Moderate was defined according to the following criteria: (i) fever and respiratory symptoms; (ii) radiological signs of pneumonia. Severe was defined if satisfying at least one of the following items: (i) breathing rate $\geq 30/\text{min}$; (ii) pulse oximeter oxygen saturation (SpO_2) $\leq 93\%$ at rest; (iii) ratio of the partial pressure of arterial oxygen (PaO_2) to a fraction of inspired oxygen (FiO_2) ≤ 300 mm Hg (1 mm Hg = -0.133 kPa).

Isolation of PBMCs

PBMCs were isolated from whole blood using density gradient centrifugation with Lymphoprep in SepMate tubes (Stemcell Technologies) in a biosafety level 2 plus facility according to the manufactory's instruction. Briefly, the blood was centrifuged at $1200 \times g$ for 10 min. PBMCs were harvested and washed twice with PBS at $400 \times g$ for 10 min. Isolated PBMCs were frozen in cell recovery Media containing 10% DMSO (GIBCO), supplemented with 90% heat-inactivated fetal bovine serum, and stored liquid nitrogen before assays analyses.

The droplet-based single-cell RNA sequencing

Single-cell suspensions at a density of 1000 cells/ μl in PBS plus 0.04% bovine serum albumin were prepared for scRNA-seq using the Chromium Single Cell 5' Reagent version 2 kit and Chromium Controller (10 \times Genomics, Pleasanton, CA), aiming for an estimated 4000 cells per library following the manufacturer's instructions. Briefly, 9000 cells per reaction were loaded for gel bead-in-emulsion (GEM) generation and barcoding. GEM-RT, post-GEM-RT cleanup, and cDNA amplification were performed to isolate and amplify cDNA for library construction. Libraries were constructed using the Chromium Single Cell 5' Reagent kit (10 \times Genomics) and Gel Bead Kit, Single Cell V(D)J Enrichment Kit, Human T Cell (1000005) and a Single Cell V(D)J Enrichment Kit, Human B Cell (1000016) according to the manufacturer's protocol. Library quality and concentration were assessed according to the

manufacturer's instructions. Libraries were sequenced on an Illumina PE150.

Single-cell RNA-seq data processing

Reads from each sample were processed with Cell Ranger (3.0.1) separately. Human reference genome GRCh38 and genome of SARS-CoV-2 were merged with corresponding GTF files used to annotate genes. The filtered matrices were then delivered into R (3.6.2) for downstream analysis. In order to demultiplex samples pooled into one sequencing run, we applied SoupCell (2.0)⁴⁴ to separate them by individuals. Next, we used the *shared_samples.py* script in SoupCell to identify individuals. The script uses *vcf* files to compare shared variations when overlapping patients between the two runs and identifies the shared patient.

Quality control was performed using R's scatter (1.14.6)⁴⁵ to remove cells with: (1) more than three median absolute deviations (MADs) of the \log_{10} read counts below the median; (2) more than three MADs of the \log_{10} genes detected below the median; and (3) more than three MADs of the genes coming from mitochondria above the median. Size factors were then considered for calculating average counts per feature, and features with average counts above 0 were kept. COVID-19 genes were removed in this step as they are not detected in the data. Afterward, we used Seurat (3.2.0)⁴⁶ for data normalization and to identify highly variable genes.

Data integration and clustering

*RunFastMNN*⁴⁷ wrapped in Seurat was performed using the top 2000 highly variable genes to integrate data sets from each sample. The first 30 MNN dimension reductions were applied to construct an SNN graph and *FindClusters* with Louvain algorithms using a standard Seurat pipeline. UMAP was also generated with the first 30 MNN dimension reductions to embed the data sets into two dimensions for visualization. Doublets labeled by SoupCell and clusters enriched for doublets ($>15\%$) were removed from further analysis. We removed ribosomal and mitochondrial genes to explore the subtypes of T/NK cells and B cells, then performed the integration and clustering again using the same strategies.

Cell-type annotation

To annotate each cluster, we used *FindAllMarkers* in Seurat to find marker genes for each cluster and selected immune cells marker genes. SingleR (1.0.5)⁴⁸ was also applied to help interpretation with Monaco Immune Data⁴⁹ (Monaco Immune Cell Data (GSE107011)) used as reference data to annotate the Th2-like cell type. We applied *FindMarkers* in Seurat to compare the innate immune subsets that have distinct distribution in asymptomatic conditions. We searched the expression of ACE2 and TNFRSF19 using the human lung atlas visualization tool (https://cellxgene.cziscience.com/d/krasnow_lab_human_lung_cell_atlas_10x-1.cxg/). The heatmaps that present the proportions of each cell type were generated using the proportions calculated as detailed above and are scaled by row to be plotted on the same color scale.

DEG analysis

For subtypes of T cells, NK cells, and B cells with >200 cells, we aggregated the counts for cells in each sample and generated pseudobulk samples, following the data analysis workflow specified by a previous study.⁵⁰ Non-protein-coding genes and genes related to sex were removed in the counts before being analyzed by edgeR (3.28.1).⁵¹ *glmQLFit* and *glmQLFTest* from edgeR were used to find marker genes between each condition. Genes with a log fold change above 1 and FDR (Benjamini-Hochberg) less than 0.05 were selected. Then, genes with logCPM above five were shown in heatmaps in Fig. S6. GO analysis was conducted on upregulated genes using topGO (3.28.1),⁵² and the top 10 enriched GO terms were plotted. Mfuzz package⁵³ was

applied to cluster gene expression patterns on stages in moderate patients. Cells were aggregated by stages from on CD56^{bri}CD16⁻ NK cells, CD56^{dim}CD16⁺ NK cells, and CD4/CD8 effector T cells. Genes with more than 25% are missing values are excluded, normalized summed counts are standardized. We clustered the standardized genes into 4 clusters. Genes in each cluster are then used as identifiers to compute overlaps with C2 curated gene sets in MsigDB.⁵⁴ To calculate the percentage of IFN-1 related genes in the clusters, genes from GO: 0034340 were used. The IFN-1 score was calculated on CD56^{bri}CD16⁻ NK cells, CD56^{dim}CD16⁺ NK cells, and effector T cells using genes involved in the pathway as input, including *IFIT1*, *IFIT2*, *IFIT3*, *IRF9*, *OAS3*, *RSAD2*, *USP18*, *IFI27*, *ISG15*, *MX1*, and *XAF1*. The AddModuleScore function in Seurat was used to calculate the score.

TCR and BCR analysis

Raw fastq files were processed with the CellRanger (3.0.1) pipeline with default settings with the reference mentioned above. For TCR analysis, only cells with at least one TCR alpha chain (TRA) and one TCR beta chain (TRB) were considered as detection TCR. Moreover, each unique TRA–TRB pair was defined as a clone type. The following analyses were based on cells with detected TCR. To analyses the clonal abundance and diversity of cells from each stage, we use the alakazam (1.0.2)⁵⁵ package. For the clonal abundance curve, the 95% confidence interval was estimated via bootstrapping. For the diversity curve, special cases of the generalized diversity index correspond to the most popular diversity measures in ecology. At $q=0$, different clones weigh equally, regardless of sample size. As the parameter q increase from 0 to $+\infty$, the diversity index (D) depends less on rare clones and more on common (abundant) ones. For BCR analysis, only cells with at least one heavy chain (IGH) and one light chain (IGK or IGL) were considered high-quality BCR and kept for further analysis. Furthermore, each unique IGH-IGK/IGL pair was defined as a clone type. A clone type was considered clonal if it is detected in more than one cell. iSMART⁵⁶ was used to perform local alignment on CDR3 sequences for T cells and B cells separately, and then those CDR3 sequences were clustered into antigen-specific groups. We selected the top2 abundant CDR3 clusters across conditions, and we calculated the proportions of patients for each CDR3 sequence. Clustal Omega⁵⁷ was used to align CDR3 sequences and generated a guide tree. All plots were generated using ggplot2 (3.3.1),⁵⁸ and heatmaps were generated using pheatmap (1.0.12) unless otherwise specified.

Statistical analysis

The two-tailed Kruskal–Wallis test, followed by Dunn's post-test was used for multiple group comparisons. Because the data in this has a small sample size with large inpatient variances contributed from both patient-specific effect and sampling time, which is hard to align for asymptomatic patients, therefore the samples were treated as independent. The specific statistical tests and their resultant significance levels are also noted in each figure legend. The R packages Seurat, ggplot2 (version 3.1.0) (Wickham, 2016), GraphPad Prism, and Adobe Illustrator were used to generate figures. P-values were added to the plot by *stat_compare_means* function in ggpubr (0.3.0) package.

DATA AVAILABILITY

All Raw and processed data are available on CNGB Nucleotide Sequence Archive (CNSA) with accession number CNP0001250.

CODE AVAILABILITY

The codes supporting the current study are available from Github (https://github.com/YOU-k/covid_analysis).

ACKNOWLEDGEMENTS

We thank all patients for their participation in this study and for providing blood samples. We also thank Matt Ritchie, Peter Hickey, Gordon Smyth, and Yifan Zhan for their helpful discussions. This work was supported by grants from the Beijing Natural Science Foundation (L202038), the Natural Science Foundation of China (81773494 and 81621005), and the Emergency Science and Technology Project for Prevention and Control of COVID-19 (20277734D).

AUTHOR CONTRIBUTIONS

M.J.M., L.Y.T. and E.H.D. conceived and designed the study. H.X.G., Y.L.W., L.L., J.H.L., H.B. W., J.F.F. and H.W.Z. collected the clinical samples; M.J.M., X.N.Z., G.L.W., H.X.G., L.J.D. and L.Y. performed the experiments; M.J.M., X.N.Z., G.L.W., H.X.G., L.J.D., X.M.C. and L.Y. collected epidemiological and clinical data. L.Y.T., Y.Y. and S.B.Z. performed bioinformatic analyses, and M.J.M. and L.Y.T. drafted the paper. All authors reviewed and approved the final paper.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41392-021-00753-7>.

Competing interests: The authors declare no competing interests.

REFERENCES

1. World Health Organization. *Coronavirus disease (COVID-19) Situation Report—177*, https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200715-covid-19-sitrep-177.pdf?sfvrsn=b1a193f3_2 (2020).
2. Raouf, D., Zumla, A., Locatelli, F., Ippolito, G. & Kroemer, G. Coronavirus infections: epidemiological, clinical and immunological features and hypotheses. *Cell Stress* **4**, 66–75 (2020).
3. Jose, R. J. & Manuel, A. COVID-19 cytokine storm: the interplay between inflammation and coagulation. *Lancet Respir. Med.* **8**, e46–e47 (2020).
4. Mehta, P. et al. COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet* **395**, 1033–1034 (2020).
5. Huang, C. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506 (2020).
6. Chen, G. et al. Clinical and immunological features of severe and moderate coronavirus disease 2019. *J. Clin. Investig.* **130**, 2620–2629 (2020).
7. Giamarellos-Bourboulis, E. J. et al. Complex immune dysregulation in COVID-19 patients with severe respiratory failure. *Cell Host Microbe* **27**, 992–1000 e1003 (2020).
8. Zhou, Z. et al. Heightened innate immune responses in the respiratory tract of COVID-19 patients. *Cell Host Microbe* **27**, 883–890 e882 (2020).
9. Mathew, D. et al. Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science* **369**, eabc8511 (2020).
10. Lucas, C. et al. Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* **584**, 463–469 (2020).
11. Chen, N. et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* **395**, 507–513 (2020).
12. Skarica, M., Eckstein, C., Whartenby, K. A. & Calabresi, P. A. Novel mechanisms of immune modulation of natalizumab in multiple sclerosis patients. *J. Neuroimmunol.* **235**, 70–76 (2011).
13. 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.e3 (2020).
14. Zhang, J. Y. et al. Single-cell landscape of immunological responses in patients with COVID-19. *Nat. Immunol.* **21**, 1107–1118 (2020).
15. Liao, M. et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020).
16. 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).
17. Lee, J. S. et al. Immunophenotyping of COVID-19 and influenza highlights the role of type I interferons in development of severe COVID-19. *Sci. Immunol.* **5**, eabc8511 (2020).
18. Gandhi, M., Yokoe, D. S. & Havlir, D. V. Asymptomatic transmission, the Achilles' heel of current strategies to control covid-19. *N. Engl. J. Med.* **382**, 2158–2160 (2020).
19. Long, Q. X. et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat. Med.* **26**, 1200–1204 (2020).
20. Michel, T. et al. Human CD56bright NK cells: an update. *J. Immunol.* **196**, 2923–2931 (2016).

21. Kojima, T. et al. TROY, a newly identified member of the tumor necrosis factor receptor superfamily, exhibits a homology with Edar and is expressed in embryonic skin and hair follicles. *J. Biol. Chem.* **275**, 20742–20747 (2000).
22. Travaglini, K. J. et al. A molecular cell atlas of the human lung from single cell RNA sequencing. *Nature* **587**, 619–625 (2020).
23. Odagiu, L., May, J., Boulet, S., Baldwin, T. A. & Labrecque, N. Role of the orphan nuclear receptor NR4A family in T-cell biology. *Front. Endocrinol.* **11**, 624122. (2021).
24. Collins, S. et al. Opposing regulation of T cell function by Egr-1/NAB2 and Egr-2/Egr-3. *Eur. J. Immunol.* **38**, 528–536 (2008).
25. Cao, Y. et al. PTEN-L promotes type I interferon responses and antiviral immunity. *Cell. Mol. Immunol.* **15**, 48–57 (2018).
26. Pedersen, S. F. & Ho, Y. C. SARS-CoV-2: a storm is raging. *J. Clin. Investig.* **130**, 2202–2205 (2020).
27. Merad, M. & Martin, J. C. Pathological inflammation in patients with COVID-19: a key role for monocytes and macrophages. *Nat. Rev. Immunol.* **20**, 355–362 (2020).
28. Schulte-Schrepping, J. et al. Severe COVID-19 is marked by a dysregulated myeloid cell compartment. *Cell* **182**, 1419–1440.e23 (2020).
29. Zhang, F. et al. Adaptive immune responses to SARS-CoV-2 infection in severe versus mild individuals. *Signal Transduct. Target. Ther.* **5**, 156 (2020).
30. Davidson, S., Maini, M. K. & Wack, A. Disease-promoting effects of type I interferons in viral, bacterial, and coinfections. *J. Interferon Cytokine Res.* **35**, 252–264 (2015).
31. 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).
32. Channappanavar, R. & Perlman, S. Pathogenic human coronavirus infections: causes and consequences of cytokine storm and immunopathology. *Semin. Immunopathol.* **39**, 529–539 (2017).
33. Kindler, E. & Thiel, V. SARS-CoV and IFN: Too little, too late. *Cell Host Microbe* **19**, 139–141 (2016).
34. Swain, S. L., McKinstry, K. K. & Strutt, T. M. Expanding roles for CD4(+) T cells in immunity to viruses. *Nat. Rev. Immunol.* **12**, 136–148 (2012).
35. Le Bert, N. et al. Highly functional virus-specific cellular immune response in asymptomatic SARS-CoV-2 infection. *J. Exp. Med.* **218**, e20202617 (2021).
36. Bacher, P. et al. Low-avidity CD4(+) T cell responses to Sars-Cov-2 in unexposed individuals and humans with severe COVID-19. *Immunity* **53**, 1258–1271.e1255 (2020).
37. Scharenberg, M. et al. Influenza A virus infection induces hyperresponsiveness in human lung tissue-resident and peripheral blood NK cells. *Front. Immunol.* **10**, 1116 (2019).
38. Fregni, G. et al. High number of CD56(bright) NK-cells and persistently low CD4+ T-cells in a hemophiliac HIV/HCV co-infected patient without opportunistic infections. *Virology* **10**, 33 (2013).
39. Niu, X. et al. Longitudinal analysis of T and B cell receptor repertoire transcripts reveal dynamic immune response in COVID-19 patients. *Front. Immunol.* **11**, 582010 (2020).
40. Schultheiß, C. et al. Next-generation sequencing of T and B cell receptor repertoires from COVID-19 patients showed signatures associated with severity of disease. *Immunity* **53**, 442–455.e444 (2020).
41. Montague, Z. et al. Dynamics of B cell repertoires and emergence of cross-reactive responses in patients with different severities of COVID-19. *Cell Rep.* **35**, 109173 (2021).
42. Wrammert, J. et al. Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. *Nature* **453**, 667–671 (2008).
43. Nielsen, S. C. A. et al. Human B cell clonal expansion and convergent antibody responses to SARS-CoV-2. *Cell Host Microbe* **28**, 516–525.e515 (2020).
44. Heaton, H. et al. Souporecell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. *Nat. Methods* **17**, 615–620 (2020).
45. McCarthy, D. J., Campbell, K. R., Lun, A. T. & Wills, Q. F. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179–1186 (2017).
46. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
47. Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* **36**, 421–427 (2018).
48. Aran, D. et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
49. Monaco, G. et al. RNA-Seq signatures normalized by mRNA abundance allow absolute deconvolution of human immune cell types. *Cell Rep.* **26**, 1627–1640.e1627 (2019).
50. Crowell, H. L. et al. muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. *Nat Commun.* **11**, 6077 (2020).
51. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
52. Alexa, A., Rahnenfuhrer, J. & Lengauer, T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* **22**, 1600–1607 (2006).
53. Kumar, L. & M, E. F. Mfuzz: a software package for soft clustering of microarray data. *Bioinformatics* **2**, 5–7 (2007).
54. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
55. Gupta, N. T. et al. Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics* **31**, 3356–3358 (2015).
56. Zhang, H. et al. Investigation of antigen-specific T-cell receptor clusters in human cancers. *Clin. Cancer Res.* **26**, 1359–1371 (2020).
57. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
58. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. XVI, 260 (Springer International Publishing, 2016).



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/>.

© The Author(s) 2021