nature genetics

Article

https://doi.org/10.1038/s41588-023-01375-1

Single-cell analyses and host genetics highlight the role of innate immune cells in COVID-19 severity

Received: 31 January 2022

Accepted: 15 March 2023

Published online: 24 April 2023

Check for updates

Ryuya Edahiro[1,2,163], Yuya Shirai[1,2,3,163], Yusuke Takeshima[4], Shuhei Sakakibara[5], Yuta Yamaguchi[2,6], Teruaki Murakami[2,6], Takayoshi Morita[2,6], Yasuhiro Kato[2,6], Yu-Chen Liu[7], Daisuke Motooka[7,8,9], Yoko Naito[8], Ayako Takuwa[7], Fuminori Sugihara[10], Kentaro Tanaka[8], James B. Wing[11,12], Kyuto Sonehara[1,9,13,14], Yoshihiko Tomofuji[1,9,13], Japan COVID-19 Task Force*, Ho Namkoong[15], Hiromu Tanaka[16], Ho Lee[16], Koichi Fukunaga[16], Haruhiko Hirata[2], Yoshito Takeda[2], Daisuke Okuzaki[7,8,9,12,17], Atsushi Kumanogoh[2,6,9,12,17,164] ✉ & Yukinori Okada[1,3,9,12,13,14,164] ✉

Mechanisms underpinning the dysfunctional immune response in severe acute respiratory syndrome coronavirus 2 infection are elusive. We analyzed single-cell transcriptomes and T and B cell receptors (BCR) of >895,000 peripheral blood mononuclear cells from 73 coronavirus disease 2019 (COVID-19) patients and 75 healthy controls of Japanese ancestry with host genetic data. COVID-19 patients showed a low fraction of nonclassical monocytes (ncMono). We report downregulated cell transitions from classical monocytes to ncMono in COVID-19 with reduced *CXCL10* expression in ncMono in severe disease. Cell–cell communication analysis inferred decreased cellular interactions involving ncMono in severe COVID-19. Clonal expansions of BCR were evident in the plasmablasts of patients. Putative disease genes identified by COVID-19 genome-wide association study showed cell type-specific expressions in monocytes and dendritic cells. A COVID-19-associated risk variant at the *IFNAR2* locus (rs13050728) had context-specific and monocyte-specific expression quantitative trait loci effects. Our study highlights biological and host genetic involvement of innate immune cells in COVID-19 severity.

Coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) represents a serious global public health issue[1]. The clinical presentation of COVID-19 is highly variable, ranging from asymptomatic infection to fatal respiratory/multi-organ failure[2]. Although effective vaccines have successfully reduced both viral transmission and disease burden[3,4], there exists an urgent need to elucidate the mechanism of severe COVID-19 to predict its severity and develop new treatments.

Multiple studies have highlighted dysregulation of complex networks of peripheral blood immune responses in COVID-19, using single-cell RNA-sequencing (scRNA-seq) analysis[5–14]. Monocytes[5–9], antigen-presenting cells[10], natural killer (NK) cells[5,6,11], T cells[5–7,12] and B cells[5–7] are all reported to be related to the severity of COVID-19, while a dysregulated interferon (IFN) response[8,14,15], which has a key role on innate immune response[16], is closely associated with the pathogenesis of COVID-19 severity. Although these studies give

A full list of affiliations appears at the end of the paper. ✉e-mail: kumanogo@imed3.med.osaka-u.ac.jp; yokada@sg.med.osaka-u.ac.jp

us important aspects of the immunopathology of COVID-19, the immune response of the host to SARS-CoV-2 still remains unclear.

In addition, genome-wide association studies (GWASs) of COVID-19, highlighted as an achievement by the COVID-19 host genetics initiative (HGI), have demonstrated that the host genetic backgrounds influence susceptibility to and/or severity of COVID-19 (refs. 17–19). Multiple genetic variants associated with the COVID-19 risk are shared across different populations, while population-specific risk variants also have been reported[20,21]. Considering the different prognoses by ancestry[22], integrated analysis of transcriptome and genetic data at single-cell resolution by ancestry should provide new insights.

Here we performed a detailed scRNA-seq analysis of over 895,000 peripheral blood mononuclear cells (PBMCs) of 73 patients with COVID-19 as well as 75 healthy controls of Japanese ancestry, and then context-specific and cell type-specific expression quantitative trait loci (eQTL) analysis by integrating scRNA-seq and host genetics data.

The proportion of nonclassical monocytes (ncMono) decreased in COVID-19 patients and RNA velocity analysis revealed the downregulation of the cellular transitions from classical monocytes (cMono) to ncMono in COVID-19 patients. We found that *CXCL10* expression was downregulated in ncMono during severe COVID-19, and cell–cell communication analysis inferred that the cellular interactions involving ncMono and plasmacytoid dendritic cells (pDC) were reduced in severe COVID-19. The putative disease genes identified by the GWAS of severe COVID-19 were enriched in monocytes and dendritic cells (DC), and COVID-19-associated variants had context and cell type-specific eQTL effects, with the *IFNAR2* variant (rs13050728) in particular having COVID-19-specific and monocytes-specific eQTL effect. In summary, our data linked innate immune cell dysfunction, especially ncMono, with severe COVID-19 and demonstrated the enrichment of host genetic risk in innate immune cells, indicating biological and host genetic critical involvement of innate immune cells in COVID-19 severity.

Results

Single-cell transcriptional profiling of COVID-19 PBMC

To investigate the immunopathogenesis and host genetics mechanism of SARS-CoV-2 infections in COVID-19 patients, PBMCs were collected from 73 COVID-19 patients and 75 healthy controls of Japanese ancestry at Osaka University. The 73 patients with COVID-19 were classified into two conditions as follows: moderate ($n = 9$) and severe ($n = 64$) disease according to the WHO classification[23] (Fig. 1a and Supplementary Table 1). No significant difference was noted in age distribution and sex composition between moderate and severe disease groups. The clinical characteristics are summarized in Supplementary Table 1.

After the unified single-cell analysis pipeline (Methods), we obtained 895,460 high-quality cells from PBMCs of all the samples. Cells were manually annotated based on the RNA expression of known marker genes to discriminate subpopulations[6,7,24,25]. We defined 13 cell subsets (Fig. 1b and Extended Data Fig. 1a) and further identified 25 cell states by following subclustering (Figs. 2a, 3a and 4a, Extended Data Figs. 2a, 3a and 4a and Supplementary Table 2). Cell annotation was validated using Azimuth[26] (Extended Data Fig. 1b).

To reveal the compositional changes between COVID-19 and healthy controls, we applied Milo[27], identifying 39,170 neighborhoods, of which 21,279 showed evidence of differential abundance (FDR < 0.1, Fig. 1c). We found a prominent decrease in T cells, NK cells and DC for COVID-19 patients compared to healthy controls, consistent with the previous reports[13,28] (Fig. 1d and Extended Data Fig. 1c). Next, we evaluated differential abundance with disease severity, identifying 14,350 neighborhoods, of which 76 showed significant differential abundance (Extended Data Fig. 1d). While few neighborhoods showed significant differential abundance, there was a trend toward increased B cells and plasmablasts (PB) and decreased DC in severe COVID-19 (Extended Data Fig. 1e).

The dysregulated IFN response has been suggested in COVID-19 (refs. 8,14,15,29), leading us to evaluate IFN response across PBMCs. We first defined a score of response to Type I IFN and that of response to IFN-γ for each cell based on the expression of a published gene list. We observed high expressions of Type I IFN response genes in NK and $CD16^+$ monocytes and higher response for moderate COVID-19 compared to the other conditions, especially in monocytes and DC (Fig. 1e, Extended Data Fig. 1f and Supplementary Table 3). We also found high expressions of IFN-γ response genes in $CD16^+$ monocytes and conventional dendritic cells (cDC), and higher response for moderate COVID-19, especially in $CD16^+$ monocytes and pDC (Fig. 1f, Extended Data Fig. 1g and Supplementary Table 3). These data are consistent with the previous reports that systemic IFN response is higher for nonsevere disease[14,29], suggesting the potential importance of innate immune cells in immunopathology of COVID-19.

Monocytes and DC in COVID-19

To characterize transcriptome dynamics in 116,944 cells annotated as monocytes and DC across disease conditions, we subclustered monocytes into five subsets and DC into two subsets according to the expression of canonical gene markers as follows: cMono_S100A, cMono_IL1B, cMono_CCL3, intermediate monocytes (intMono), ncMono, cDC and pDC (Fig. 2a and Extended Data Fig. 2a).

To reveal the compositional changes between COVID-19 and healthy controls in this subset, we performed differential abundance analysis using Milo[27], identifying 6,721 neighborhoods, of which 1,265 showed evidence of differential abundance (Fig. 2b). The proportion of ncMono, cDC and pDC declined prominently in COVID-19 patients compared to healthy controls (Fig. 2c and Extended Data Fig. 2b). Next, we evaluated differential abundance with disease severity, identifying 3,003 neighborhoods, of which 19 showed significant differential abundance (Extended Data Fig. 2c). While few neighborhoods showed significant differential abundance, there was a trend toward a decrease in pDC and ncMono in severe compared to moderate COVID-19 (Extended Data Fig. 2d), implying that ncMono might contribute to immunopathology of COVID-19 severity because decreased cell proportion of ncMono is one of the COVID-19-specific features[13,30,31].

To gain insight into functions of different cell subsets according to COVID-19 severity, we performed differential expression analyses and a set of Gene Ontology (GO) analyses across the groups (that is, all COVID-19 patients versus healthy, moderate disease versus healthy and severe disease versus healthy) in the five subsets (cMono, intMono, ncMono, cDC and pDC). The top 20 enriched pathways upregulated in COVID-19 versus healthy group were related to innate immunity or antiviral response, and almost all were shared among the five subsets (Extended Data Fig. 2e). We next compared the enriched pathways upregulated in moderate severity versus healthy group and those in severe versus healthy group. The top ten enriched pathways were also almost the same in each subset, while there existed several pathways with different enrichment patterns between the moderate disease and severe group (Fig. 2d). Although 'response to IFN-γ' pathway was enriched in moderate COVID-19 of each subset, it was not evident in three monocyte subsets in severe COVID-19, suggesting that decreased IFN-γ response in monocyte subsets might contribute to the severity of COVID-19. Enrichment of 'response to Type I IFN' pathway was specifically depleted in ncMono of severe COVID-19.

To further elucidate the mechanism of COVID-19 severity, we conducted differential expression analysis between moderate disease versus severe disease group in ncMono. *PLD4*, which digests ssRNA and ssDNA[32], was the most downregulated in severe compared to moderate disease group (Fig. 2e). In addition, the expression of *CXCL10*, which belongs to IFN-γ-induced gene and critical in response to various infectious pathogens[33] and has been reported to be involved in COVID-19

https://doi.org/10.1038/s41588-023-01375-1

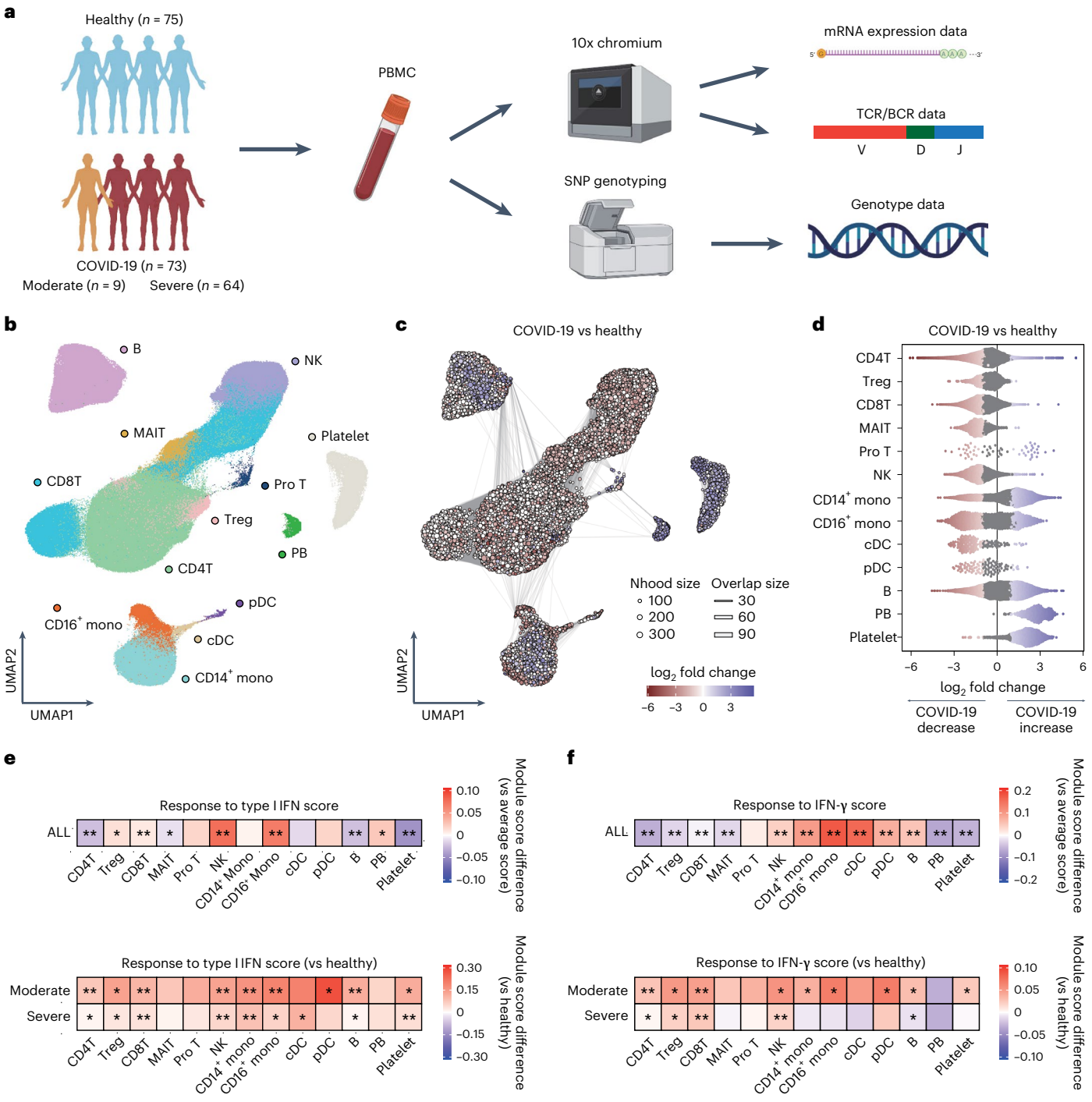


Fig. 1 | Study design and single-cell transcriptional analysis of PBMCs from COVID-19 patients and healthy controls. a, Overview of the study design. The image was created using BioRender.com. **b**, UMAP embedding of scRNA-seq data for all 895,460 cells. Thirteen cell types were defined by RNA expression of marker genes (Extended Data Fig. 1a). **c**, Graph representation of Nhoods identified by Milo. Nodes are Nhoods, colored by their $\log_2$ FC between COVID-19 patients ($n = 72$) and healthy controls ($n = 75$) adjusted by age and sex. Nondifferential abundance Nhoods (FDR ≥ 0.1) are colored white, and sizes correspond to the number of cells in a Nhood. Graph edges depict the number of cells shared between adjacent Nhoods. **d**, Beeswarm plot showing the distribution of adjusted $\log_2$ FC in abundance between COVID-19 patients and healthy controls in Nhoods according to 13 cell types. Colors are represented in the same way as in **c**. **e**,**f**, The module score of Type I IFN response and IFN-γ response in PBMCs. The score was calculated using a gene set termed 'GOBP_RESPONSE_TO_TYPE_I_INTERFERON' (GO:0034340) and 'GOBP_RESPONSE_TO_INTERFERON_GAMMA' (GO:0034341), respectively. Upper heatmaps depicting the difference between average scores of 13 cell types and that of all single cells. The module scores of cells in each cell type were compared with the average score of all PBMCs using a two-sided one-sample *t*-test. Lower heatmaps depicting the difference between average scores of moderate or severe disease group and those of the healthy group in each of 13 cell types ($n = 75$ healthy, $n = 9$ moderate, $n = 64$ severe). The module scores of cells of moderate or severe disease group were compared to those of healthy group in each cell type using a two-sided Welch's *t*-test. *$P_{uncorrected} < 1.0 \times 10^{-50}$, **$P_{uncorrected} < 1.0 \times 10^{-300}$. Mono, monocytes; Pro, proliferative; Nhood, neighborhood.

https://doi.org/10.1038/s41588-023-01375-1

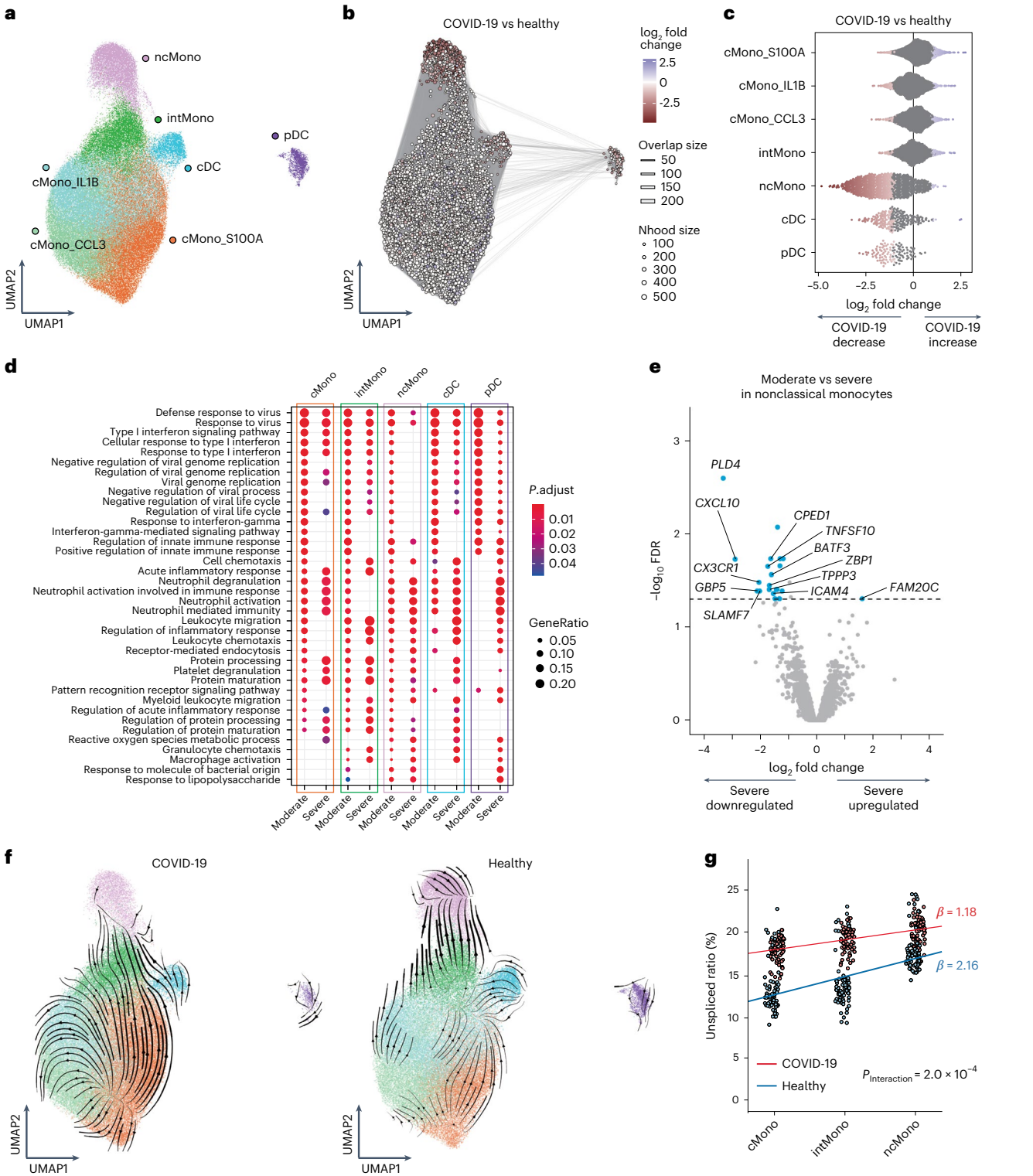


Fig. 2 | Defective IFN-γ response and reduced transition potential to ncMono in monocytes of severe COVID-19. a, UMAP embedding of 116,944 monocytes and DC. Seven cell types were defined by RNA expression of marker genes (Extended Data Fig. 2a). **b**, Graph representation of Nhoods identified by Milo. Nodes are Nhoods, colored by their $\log_2$ FC between COVID-19 ($n = 72$) and healthy controls ($n = 75$) adjusted by age and sex. Nondifferential abundance Nhoods (FDR ≥ 0.1) are colored white. **c**, Beeswarm plot showing the distribution of adjusted $\log_2$ FC in abundance between COVID-19 and healthy controls in Nhoods according to seven cell types. Colors are represented in the same way as in **b**. **d**, The top ten enriched biological processes by GO analysis of upregulated DEGs of moderate and severe disease compared to healthy group in five cell types. Dot color indicates the statistical significance of the enrichment (adjusted *P* values via the Benjamini–Hochberg method), and dot size represents gene ratio annotated to each term. **e**, The differential gene expression analysis between moderate ($n = 8$) and severe ($n = 64$) COVID-19 in ncMono. DEGs (FDR < 0.05 and FC > 2) are colored in light blue and labeled by gene symbols if $\log_2$ FC > 1.5. **f**, Velocities derived from the dynamical model for monocytes and DC cluster from COVID-19 and healthy group are projected into a UMAP-based embedding. Colors indicate the same clusters as in **a**. **g**, Average unspliced ratio of each sample stratified by three monocytes clusters, colored by COVID-19 ($n = 73$) and healthy ($n = 75$) groups. Condition-specific regression lines are shown. *P* value for the interaction effect between three monocyte clusters and two clinical conditions is uncorrected and reflects two-sided test.

https://doi.org/10.1038/s41588-023-01375-1

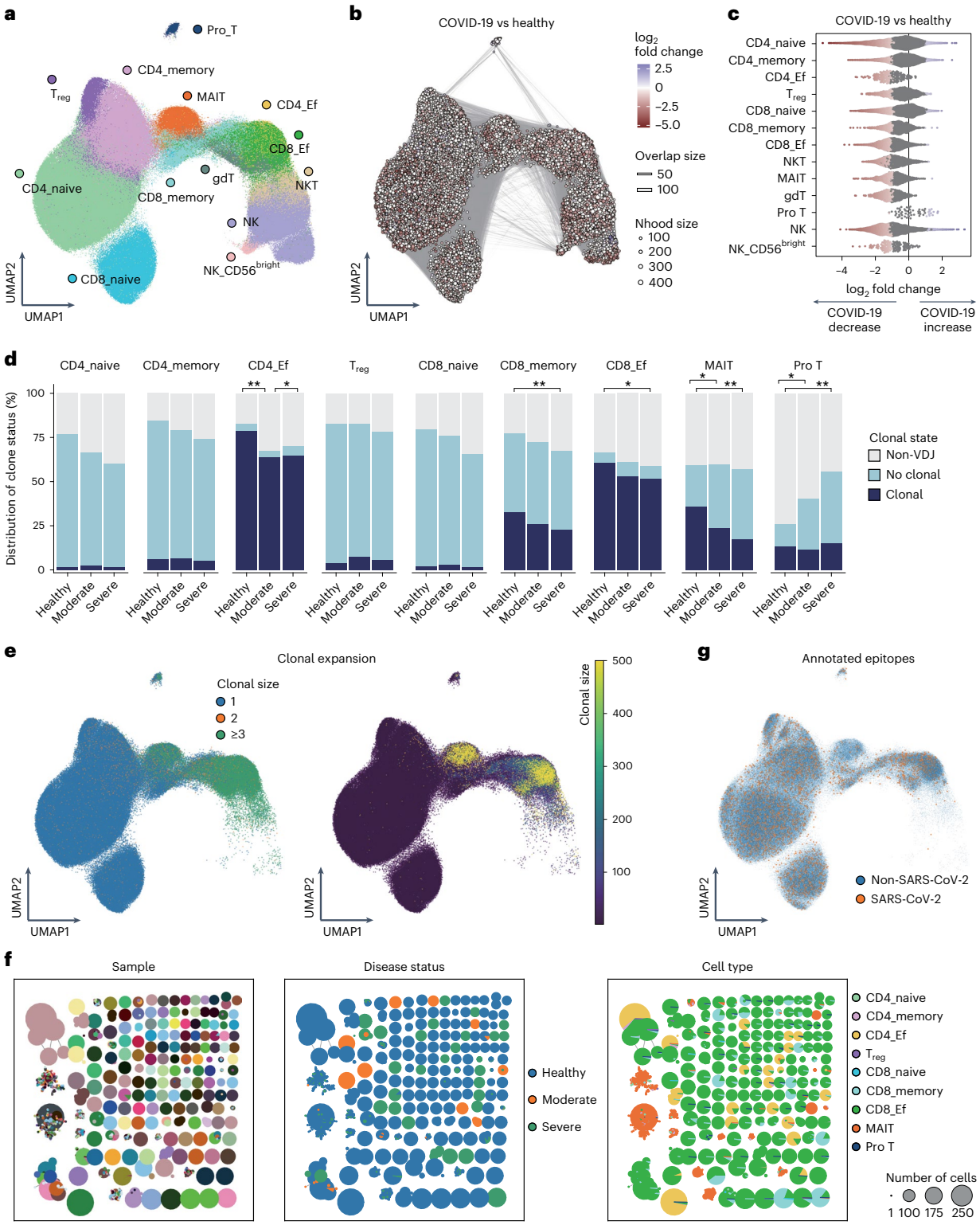


Fig. 3 | Differential abundance analysis of T cells and NK cells and TCR analysis. a, UMAP embedding of 628,715 T and NK cells. Thirteen cell types were defined by RNA expression of marker genes (Extended Data Fig. 3a). **b**, Graph representation of Nhoods identified by Milo. Nodes are Nhoods, colored by their $\log_2$ FC between COVID-19 ($n = 72$) and healthy controls ($n = 75$) adjusted by age and sex. Nondifferential abundance Nhoods (FDR ≥ 0.1) are colored white. **c**, Beeswarm plot showing the distribution of adjusted $\log_2$ FC in abundance between COVID-19 and healthy controls in Nhoods according to 13 cell types. Colors are represented in the same way as in **b**. **d**, The distribution of the clone state of T cells in each cluster according to disease status. Differences of average clonal expansion rate of each sample between clinical conditions were evaluated in each cluster using two-sided Welch's *t*-test (*$P_{uncorrected} < 0.05$, **$P_{uncorrected} < 0.005$), where only cells mapped with TCRs were included in the analysis ($n = 75$ healthy, $n = 9$ moderate, $n = 64$ severe). **e**, UMAP embedding of T cells (nine cell types) colored by clonal expansion size. Left panel shows clonal expansion divided into three categories, and right panel shows clonal expansion sizes ranging from 0 to 500. **f**, Network plots showing similarity of TCRα and TCRβ CDR3 amino acid sequence for each sample, disease status and cell types. Clonotype clusters with clonal size ≥50 are selected. **g**, T cells that were suspected to be specific to SARS-CoV-2 based on CDR3 amino acid sequence were projected on UMAP. Ef, effector.

https://doi.org/10.1038/s41588-023-01375-1

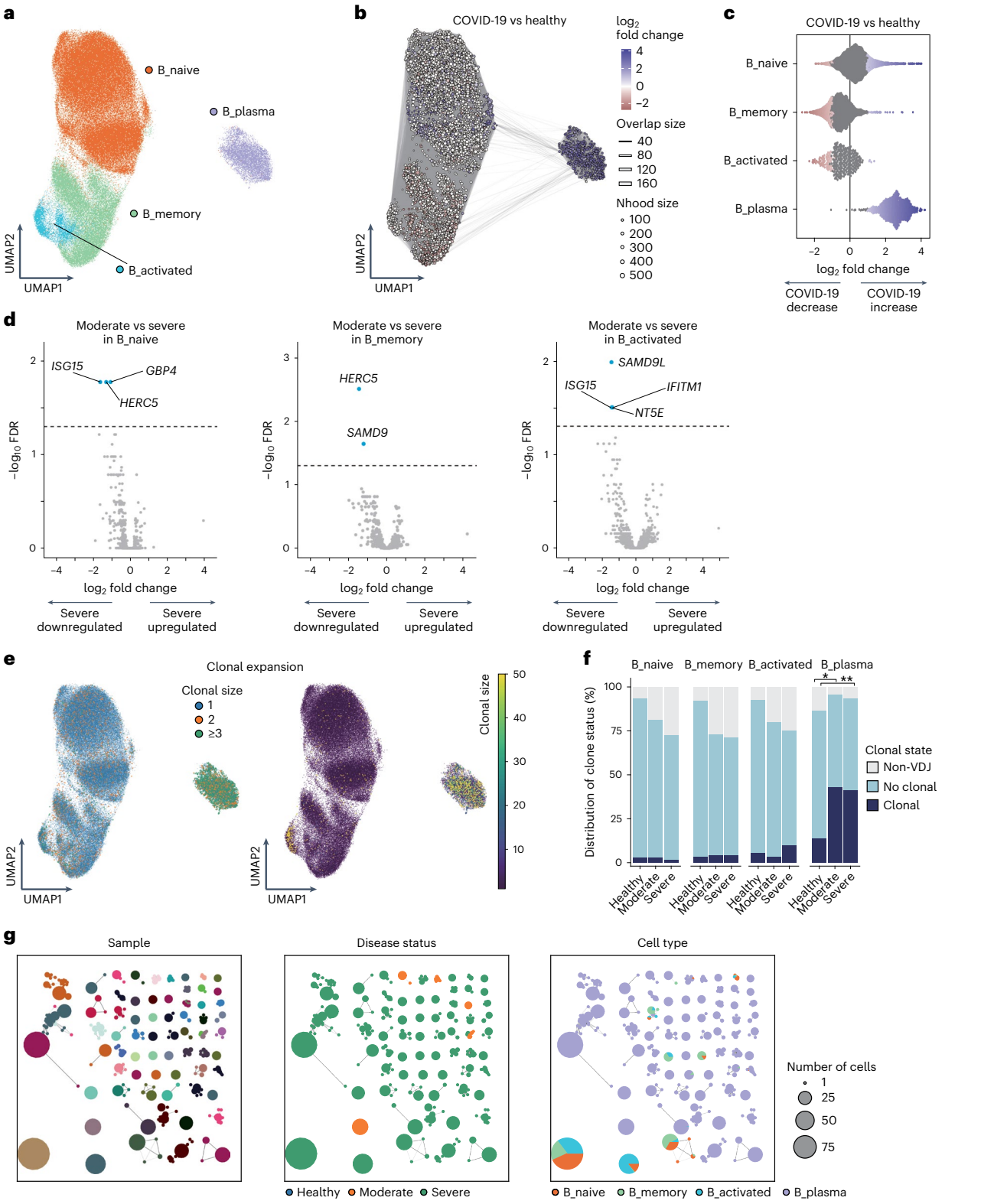


Fig. 4 | Differential abundance, gene expression and clonal analysis of B cells. **a**, UMAP embedding of four cell types of 123,728 B cells (Extended Data Fig. 4a). **b**, Graph representation of Nhoods identified by Milo. Nodes are Nhoods, colored by their $\log_2$ FC between COVID-19 ($n = 72$) and healthy controls ($n = 75$) adjusted by age and sex. Nondifferential abundance Nhoods (FDR ≥ 0.1) are colored white. **c**, Beeswarm plot showing the distribution of adjusted $\log_2$ FC in abundance between COVID-19 and healthy controls in Nhoods according to four cell types. **d**, The differential gene expression analysis between moderate ($n = 8$) and severe ($n = 64$) COVID-19 in B_naive, B_memory and B_activated. DEGs (FDR < 0.05 and FC > 2) are colored in light blue and labeled by gene symbols. **e**, UMAP embedding of B cells colored by clonal expansion size. Left panel shows clonal expansion divided into three categories, and right panel shows clonal expansion size from 0 to 50. **f**, The distribution of the clone state of B cells in each cluster according to disease status. Differences of average clonal expansion rate of each sample between disease status were evaluated in each cluster using two-sided Welch's t-test ($^*P_{\text{uncorrected}} = 0.012$, $^{**}P_{\text{uncorrected}} = 1.3 \times 10^{-7}$), where only cells mapped with BCRs were included in the analysis ($n = 75$ healthy, $n = 9$ moderate, $n = 64$ severe). **g**, Network plots showing similarity of CDR3 amino acid sequence in BCR heavy and light chain for each sample, disease status and cell types. Clonotype clusters with clonal size ≥10 are selected.

severity with proteomics analysis[5,13], was also prominently decreased in severe group.

To analyze how the dynamics of transcriptional activation and cell transition differ in disease status, we performed RNA velocity analysis[34]. The transition potential from intMono to ncMono was observed in healthy controls, while such transition was not observed in COVID-19 (Fig. 2f). To quantitatively compare the differences of estimated cell transition from cMono to ncMono through intMono[35] between COVID-19 and healthy controls, we analyzed the increase of unspliced fractions from cMono to ncMono. We found that condition-specific regression slope was lower in COVID-19 patients with significant interaction between three monocytes and two clinical conditions ($P_{\text{Interaction}} = 2.0 \times 10^{-4}$; Fig. 2g and Extended Data Fig. 2f). These data suggest that the decreased proportion of ncMono in COVID-19 patients is a consequence of the downregulation of the cellular transition from cMono to ncMono.

T cells and NK cells and T cell repertoires in COVID-19

We subclustered 628,715 cells manually annotated as T cells and NK cells from PBMCs and obtained 13 subsets according to RNA expression of canonical markers (Fig. 3a and Extended Data Fig. 3a).

To reveal the compositional changes between COVID-19 and healthy controls, we performed differential abundance analysis using Milo[27], identifying 27,182 neighborhoods, of which 10,035 showed evidence of differential abundance (Fig. 3b). The proportion of almost all clusters declined in COVID-19 patients (Fig. 3c and Extended Data Fig. 3b). We next evaluated differential abundance with disease severity, identifying 7,981 neighborhoods, of which nine showed significant differential abundance (Extended Data Fig. 3c). Although there were few neighborhoods showing significant differential abundance, the proportions of CD4T cells, regulatory T (T_{reg}) cells and $CD56^{bright}$ NK (NK_CD56^{bright}) cells tended to be lower in severe COVID-19 compared to moderate COVID-19, while those of natural killer T (NKT) cells and NK tended to be higher (Extended Data Fig. 3d).

To gain insight into the clonal relationship among individual T cells across three disease conditions, we performed T cell receptor (TCR) analysis for nine subsets except for γδ T (gdT) cells, NKT, NK and NK_CD56^{bright}. The detection percentage of TCRs was 73.1% (Fig. 3d and Extended Data Fig. 3e). Similar to the previous report[6,36], the large clonal expansions were observed particularly in $CD4^+$ effector T (CD4_Ef) cells, $CD8^+$ effector T (CD8_Ef) cells and mucosal-associated invariant T (MAIT) cells (Fig. 3e,f). The proportion of clonally expanded CD4_Ef increased in moderate disease compared to healthy and severe group (Fig. 3d and Supplementary Table 4), suggesting that efficient clonal expansion of CD4_Ef might contribute to the prevention of serious COVID-19. We next examined whether expanded clonotypes were shared among each sample, between COVID-19 and healthy controls, and among cell types. The large expanded clonotypes exhibited parallel expansion in several cell types, particularly in CD4_Ef, $CD8^+$ memory T (CD8_memory) cells and CD8_Ef, whereas a vast majority of these expanded clonotypes were unique for individual patients (Fig. 3f and Extended Data Fig. 3f,g). In contrast, the large expanded clonotypes in MAIT were shared among individuals.

We compared each clonotype from our data against the currently known CDR3 sequences of SARS-CoV-2-specific TCR in VDJdb[37]. A small number of our TCRs ($n = 4{,}143$; 0.8%) shared their CDR3 with reported SARS-CoV-2-specific TCRs (Fig. 3g). The distribution of TCR specific to SARS-CoV-2 was relatively uniform across cell types, and the clonal expansions of such TCRs were observed in memory or activated T cells (Fig. 3g and Extended Data Fig. 3h). Considering the low percentage of SARS-CoV-2-specific TCRs and the low sharing of expanded TCRs among individuals, further accumulation of data on SARS-CoV-2-specific TCRs would be warranted.

B cells and B cell repertoires in COVID-19

We subclustered 123,728 cells manually annotated as B cells and PB from PBMCs and obtained four subsets according to RNA expression of canonical markers (Fig. 4a and Extended Data Fig. 4a).

To reveal the compositional changes between COVID-19 and healthy controls, we performed differential abundance analysis using Milo[27], identifying 8,169 neighborhoods, of which 2,120 showed evidence of differential abundance (Fig. 4b). The proportion of B_naive and B_plasma increased in COVID-19 patients, whereas that of B_memory and B_activated decreased (Fig. 4c and Extended Data Fig. 4b). We also evaluated differential abundance with disease severity, identifying 4,006 neighborhoods, of which none showed significant differential abundance. However, there was a trend toward a higher proportion of B_plasma in severe compared to the moderate disease group, while that of B_activated tended to be lower (Extended Data Fig. 4c,d).

We next performed differential expression and pathway enrichment analysis in four subsets. Pathway enrichment analysis showed that pathways related to antiviral response and immune response were enriched in COVID-19 patients compared to healthy controls (Extended Data Fig. 4e), and that the biological pathways related to IFN response were enriched in three subsets (B_naive, B_memory and B_activated) in the moderate disease group, while those pathways were not in severe group (Extended Data Fig. 4f), consistent with a previous report[5]. Conversely, pathways related to Type I IFN were enriched in B_plasma in severe group, but not in the moderate severity group (Extended Data Fig. 4f). We also conducted DE analysis between moderate versus severe COVID-19. *HERC5*, which has direct antiviral function by catalyzing ISGylation[38], was significantly downregulated in B_naive and B_memory of severe compared to moderate COVID-19 (Fig. 4d). IFN-related genes, such as *ISG15*, *IFITM1* and *GBP4*, were also downregulated in B_naive and B_activated of severe compared to moderate COVID-19 (Fig. 4d).

Finally, we performed B cell receptor (BCR) analysis. The detection percentage of BCRs was more than 80% in B cell subset (Fig. 4e,f). Clonal expansions were most evident in B_plasma, which showed the larger expansion in COVID-19 (Fig. 4e,f and Supplementary Table 5). We next examined whether expanded clonotypes were shared among each sample, between COVID-19 and healthy controls, and among cell types (Fig. 4g and Extended Data Fig. 4g,h). In contrast to TCR analysis, very few clonotypes were shared between COVID-19 and healthy controls (Fig. 4g and Extended Data Fig. 4g). In addition, while the expanded clonotypes in B_plasma were not shared with the other B cell subsets, the expanded clonotypes exhibited parallel expansion among B_naive, B_memory and B_activated (Fig. 4g). Together, our results suggest that in COVID-19 disease, a robust antibody response characterized by clonally expanded circulating PB occurs against a background of augmented IFN responses.

Changes of intercellular interactions in PBMC across clinical status

To map the cellular interaction differences between COVID-19 and healthy controls, we inferred all possible intercellular communications by the expression of ligand–receptor pairs in both cell populations using CellPhoneDB[39] and NATMI[40]. CellPhoneDB and NATMI revealed strong interactions particularly among monocytes and DC in both COVID-19 and healthy groups (Fig. 5a and Extended Data Fig. 5a).

In addition to simple edge count analysis, we examined the differences in the cell-connectivity-summary networks based on mean expression weight between COVID-19 and healthy controls using NATMI[40]. The cellular interactions involving pDC as the sender in COVID-19 patients were lower than in healthy controls, and those involving B_plasma as a receiver in COVID-19 patients were lower than in healthy controls (Fig. 5b).

https://doi.org/10.1038/s41588-023-01375-1

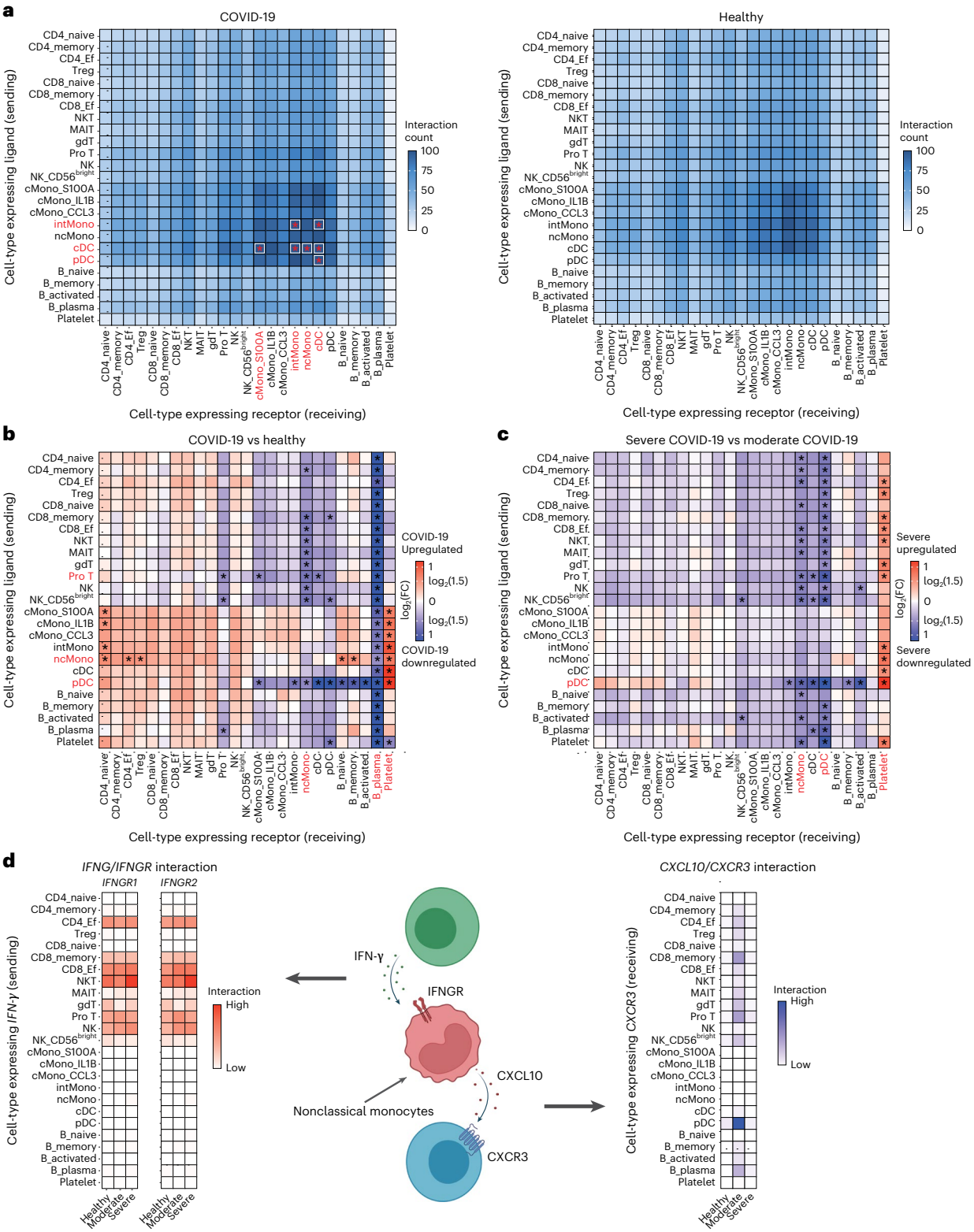


Fig. 5 | Differential cell–cell interactions between COVID-19 patients and healthy controls and within COVID-19 severity. a, Heatmaps depicting number of ligand–receptor pairs connecting cell–cell pairs for COVID-19 ($n = 73$) and healthy controls ($n = 75$), respectively. Rows indicate cells expressing the ligands and columns indicate cells expressing the receptors. An asterisk indicates cell–cell interactions with a number of ligand–receptor of more than 100, and cell types that share such interactions with at least one cell type are highlighted in red. **b**, Heatmap depicting the cell-connectivity-summary networks based on mean expression weight between COVID-19 ($n = 73$) and healthy controls ($n = 75$). An asterisk indicates the cell–cell interactions with FC of mean expression ≥1.5, and cell types that share cell–cell interactions with FC of mean expression ≥1.5 with five or more cell types are highlighted in red. **c**, Heatmap depicting the cell-connectivity-summary networks based on mean expression weight between moderate ($n = 9$) and severe ($n = 64$) COVID-19. An asterisk and red highlight mean the same as in **b**. **d**, Cell–cell interaction of *IFNG/IFNGR* and *CXCL10/CXCR3* around nonclassical monocytes. Heatmaps depicting the cell-connectivity-summary networks based on mean expression weight of *IFNG/IFNGR* (left) and *CXCL10/CXCR3* (right) according to three conditions, respectively ($n = 75$ healthy, $n = 9$ moderate, $n = 64$ severe). The image was created using BioRender.com. Ef, effector.

We next investigated the differences in the cell-connectivity-summary networks based on mean expression weight between moderate and severe COVID-19. Almost all intercellular interactions were reduced in severe compared to moderate disease group (Fig. 5c and Extended Data Fig. 5b). Notably, the intercellular interactions signaling from pDC and those signaling to ncMono and pDC were reduced in severe COVID-19, implying that the dysfunction of the intercellular interactions involving these two subsets might be related to the severity of COVID-19.

Finally, we explored the intercellular interactions centered on *CXCL10* in ncMono, which was significantly downregulated in severe COVID-19 (Fig. 2e). *CXCL10* is an IFN-γ-induced gene and exerts its biological effects by binding to *CXCR3* (ref. 33). Therefore, we investigated *IFNG*/*IFNGR* interactions of ncMono as receiver and *CXCL10*/*CXCR3* interaction of ncMono as sender using NATMI (Fig. 5d). Activated T cells and NK cells showed the strong interactions of *IFNG*/*IFNGR* with ncMono, which were enhanced as the severity of COVID-19 increased. On the other hand, pDC and to a lesser extent activated T cells showed the *CXCL10*/*CXCR3* interaction with ncMono in moderate COVID-19, which was not observed in healthy controls and severe COVID-19 (Fig. 5d). These intercellular interaction analyses computationally infer the possibility that dysfunction of ncMono and the consequently reduced interaction of *CXCL10*/*CXCR3* might be one of the factors responsible for driving COVID-19 severity.

Host genetics risks of the severity of COVID-19 in PBMC

Elucidating interaction between host genetics and transcriptional dynamics resolves causal biological mechanism of infection. To evaluate genome-wide host genetics risk of COVID-19 and identify subpopulations of disease-associated cells in PBMCs, we integrated information from our scRNA-seq data with polygenic signals from COVID-19 GWAS using scDRS[41].

First, we computed a disease score for each cell observed at our COVID-19 scRNA-seq datasets according to the COVID-19 case-control GWAS summary statistics of COVID-19 HGI (round 6; ref. 18), and projected the scores on Uniform manifold approximation and projection (UMAP; Extended Data Fig. 6a). The disease scores at the COVID-19 scRNA-seq dataset were similar across any cell types when using the GWAS summary statistics of self-reported COVID-19 (C2, n_{Case} = 112,612), while the cells annotated as monocytes and DC showed the higher disease scores than the other cell types when GWAS cases were restricted to severe ones (that is, hospitalized COVID-19 (B2, n_{Case} = 24,274) and very severe COVID-19 (A2, n_{Case} = 8,779); Fig. 6a). When comparing disease score for each phenotype, disease scores from severe GWASs were higher than those from self-reported GWAS, particularly at cells annotated as monocytes and DC, which was more prominent in very severe GWAS (Fig. 6b, Extended Data Fig. 6b and Supplementary Table 6). Next, we assessed associations between the six major cell types (Methods) and three COVID-19 GWAS phenotypes, and also within-cell type association heterogeneity using scDRS[41]. No cell type was enriched in the self-reported infection GWAS, whereas monocytes were associated with very severe GWAS, and DC was associated with hospitalization and very severe GWASs (Fig. 6c and Supplementary Table 7), demonstrating that polygenic risks involved in the severity of COVID-19 were enriched in the cells responsible for innate immunity. In addition, three cell types showed heterogeneity in association with hospitalization and critical illness (Fig. 6c). To investigate the host genetics association of monocytes and DC with severe COVID-19 in more detail, we examined cell type-disease association and its heterogeneity of the five innate immune subsets (cMono, intMono, ncMono, cDC and pDC). All of these subsets were associated with GWAS for severe disease, with stronger associations in very severe cases (Fig. 6c and Supplementary Table 7). Some of them showed significant heterogeneity within the subset (Fig. 6c and Supplementary Table 7). The subset analysis of manually annotated 25 clusters revealed the significant association of MAIT with GWAS for severe disease in addition to monocytes and DC subsets (Extended Data Fig. 6c and Supplementary Table 7). MAIT functions as innate sensors of viral infection[42], again implicating the involvement of host genetic risk of severe COVID-19 with innate immunity.

Context and cell-type-specific eQTLs of COVID-19-related variants

To gain a better understanding of transcriptional variability and dynamics regulated by the GWAS-identified COVID-19-associated variants, we examined eQTL effects of the replicated variants at the GWAS of COVID-19 in the Japanese population[20], separately for COVID-19 (n = 67) and healthy controls (n = 75; Extended Data Fig. 7a).

First, we performed eQTL analysis for the six major cell types (Methods). COVID-19-associated variants showed different cell type distributions with significant eQTL effects between the COVID-19 patients and healthy controls, demonstrating context-specific eQTL effects (Fig. 7a and Supplementary Table 8). Among them, monocytes of the COVID-19 patients had significant eQTL effects in multiple variants (FDR < 0.02). Given the previous analyses demonstrating the involvement of monocytes with COVID-19 severity in this dataset, we examined eQTL effects in the three subsets of monocytes. The two variants with eQTL effects in monocytes of the COVID-19 patients had significant eQTL effects specifically in cMono (FDR = 2.6×10^{-6} for *ABO* and FDR = 0.017 for *IFNAR2*), and no such eQTL effect of the *IFNAR2* variant was observed in the healthy controls (FDR = 0.66; Fig. 7b and Supplementary Table 9).

IFNAR2 has a key role in Type I IFN signaling pathway[16], leading us to investigate eQTL effect of the *IFNAR2* variant in more detail. We found COVID-19 context-specific increasing dosage effect of the risk allele (rs13050728-T) on *IFNAR2* expression levels in monocytes (β = 0.030, 95% CI = 0.011–0.049, $P = 2.7 \times 10^{-3}$ for COVID-19 and $\beta = 4.9 \times 10^{-4}$, 95% CI = −0.010 to 0.011, P = 0.93 for healthy controls), especially in cMono (β = 0.030, 95%CI = 0.012–0.049, $P = 2.3 \times 10^{-3}$ for COVID-19 and $\beta = 3.7 \times 10^{-3}$, 95% CI = -5.2×10^{-3} to 0.013, P = 0.42 for healthy controls; Fig. 7c, Extended Data Fig. 7b and Supplementary Tables 8 and 9). While many cell types showed a larger eQTL effect in COVID-19 than healthy controls, monocytes, in particular cMono, specifically had a significant interaction of eQTL effect between COVID-19 and healthy controls ($P_{Interaction}$ = 0.010 for monocytes and $P_{Interaction}$ = 0.012 for cMono; Fig. 7c and Extended Data Fig. 7b,c). DE analysis revealed increased *IFNAR2* expression in cMono of COVID-19 patients than healthy controls ($P = 1.7 \times 10^{-5}$; Extended Data Fig. 7d), consistent with previous findings[5]. Taken together, the risk allele of rs13050728 might contribute to severe COVID-19 by increasing expression of *IFNAR2* in cMono, highlighting the importance of context and cell type-specific eQTL analysis to elucidate host genetical effects in pathophysiology of COVID-19.

Discussion

Here we reported comprehensive analyses of single-cell transcriptome and TCR and BCR of PBMCs from the COVID-19 patients and healthy controls in Japanese, integrated with host genetics data. Our data presented the dysfunction of monocytes or DC, particularly ncMono, in severe COVID-19, the enrichment of host genetics COVID-19 risks in monocytes or DC, as well as COVID-19 context-specific eQTL effect of the *IFNAR2* variant (rs13050728) in monocytes. These highlighted the biological and host genetic critical involvement of innate immune cells in COVID-19 severity.

While previous studies on IFN and SARS-CoV-2 have focused on Type I IFN due to their robust capacity to interfere with viral replication[43], several studies have indicated that IFN-γ is also an essential component in the severity of COVID-19 (refs. 29,44,45). We found the depleted enrichment of IFN-γ response pathways in monocyte subsets

https://doi.org/10.1038/s41588-023-01375-1

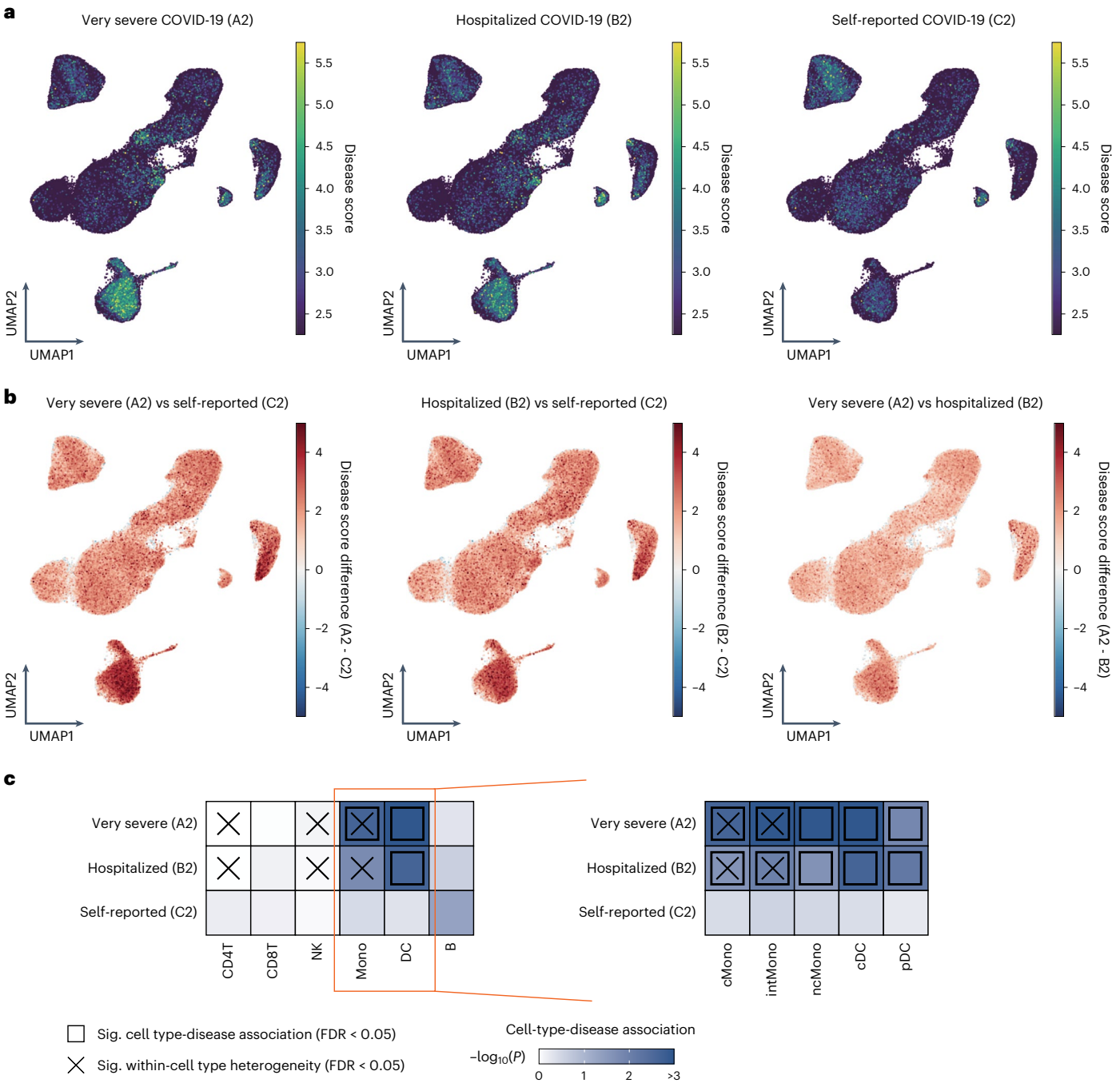


Fig. 6 | Associations of PBMC cell types with host genetic risk of COVID-19. **a**, UMAP embedding of COVID-19 PBMCs ($n = 72$) colored by scDRS disease score calculated from GWAS summary statistics of three phenotypes from COVID-19 HGI (round 6). **b**, Differences of disease score in individual cell-level among three phenotypes. **c**, Heatmaps depicting each cell type-disease association for three phenotypes, respectively. Heatmap colors denote uncorrected *P* value of cell-type-disease association evaluated using scDRS. Squares denote significant cell type-disease associations (FDR < 0.05), and cross symbols denote significant heterogeneity in association with disease across individual cells within a given cell type (FDR < 0.05). FDR was calculated via the Benjamini–Hochberg method across all pairs of cell types and three phenotypes.

from patients with severe COVID-19, again demonstrating the importance of IFN-γ response for COVID-19. One of the COVID-19-specific features is decreased cell fraction of ncMono in severe disease[13,31], whereas its fraction generally increases in patients with sepsis and inflammatory disease[46]. This phenomenon was also observed in this study, and we revealed the downregulation of the cellular transitions from cMono to ncMono in COVID-19 patients using RNA velocity analysis. In addition to decreased cell fraction, differential expression analysis revealed the severely downregulated expression of *CXCL10*, which is IFN-γ-induced gene and reported to be involved in COVID-19 severity with proteomics analysis[5,13], in ncMono for severe COVID-19, and cell–cell communication analysis inferred the possibility that *CXCL10/CXCR3* interaction between ncMono and pDC was depleted whereas ncMono firmly received IFN-γ signal from activated T cells in severe COVID-19. Thus, these findings indicate that ncMono might contribute to immunopathology of COVID-19 severity via decreased cell fraction as well as biological dysfunction. However, the mechanisms of the reduced differentiation to ncMono and the dysfunction

https://doi.org/10.1038/s41588-023-01375-1

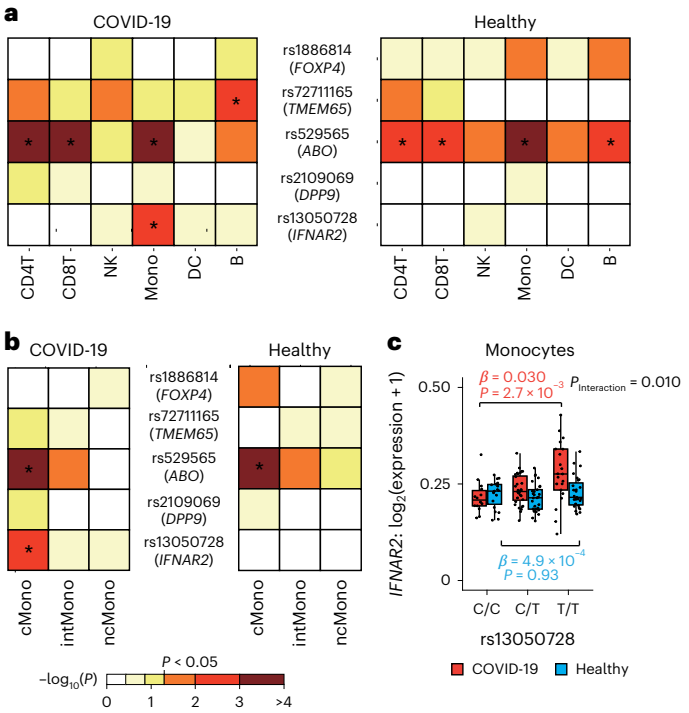


Fig. 7 | Context-specific and cell type-specific eQTL analysis of COVID-19-associated risk variants. a, Heatmaps depicting eQTL effect of COVID-19-associated risk variants on V2G, highest gene prioritized by the V2G score of Open Target Genetics, by six major cell types, separately for COVID-19 ($n = 67$) and healthy controls ($n = 75$). Heatmap colors denote uncorrected P value of eQTL effect. **b**, Heatmaps depicting eQTL effect of COVID-19-associated risk variants by three cell types of monocytes, separately for COVID-19 ($n = 67$) and healthy controls ($n = 75$). **c**, rs13050728 eQTL effect on *IFNAR2* expression in monocytes. The box plot shows the eQTL effect in COVID-19 (red) and healthy controls (blue). P values are uncorrected and reflect two-sided tests. Boxes denote the IQR, and the median is shown as horizontal bars. Whiskers extend to 1.5 times the IQR. All samples are shown as individual points ($n = 67$ COVID-19, $n = 75$ healthy). In **a** and **b**, an asterisk indicates FDR < 0.05, and FDR was calculated via the Benjamini–Hochberg method across all pairs of the cell types and five variants, separately for COVID-19 and healthy controls. IQR, interquartile range.

of ncMono are still elusive. Multimodal single-cell analysis and in vivo experiments should be warranted in the future.

Human genetic background has been demonstrated to influence the pathogenesis of COVID-19 (refs. 17–20). Functional analysis has mostly focused on the variant at *LZTFL1* on 3p21, which showed the strongest severity association in Europeans but conferred a rare frequency of the risk allele in East Asian[47,48], suggesting the importance of additional studies in non-Europeans. We found that the putative disease genes identified by the GWAS of severe phenotypes showed cell type-specific expressions in monocytes and DC. We also showed that eQTL effects of the COVID-19-associated variants replicated in the Japanese GWAS were context and cell type-specific, with the *IFNAR2* variant, in particular, having COVID-19-specific and monocytes-specific eQTL effect, indicating the host genetic involvement of innate immune cells in COVID-19 severity. Because the context-specific eQTL effects were confirmed, collecting more COVID-19 cases and comparing them with healthy controls will help to clarify the pathogenesis of COVID-19 from the perspective of the host genome.

Collectively, our results motivate us for a detailed examination of ncMono function in the context of COVID-19, as well as to increase sample size to perform integrated analysis with genetic data on a larger scale.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-023-01375-1.

References

1. Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
2. Huang, C. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506 (2020).
3. Polack, F. P. et al. Safety and efficacy of the BNT162b2 mRNA COVID-19 vaccine. *N. Engl. J. Med.* **383**, 2603–2615 (2020).
4. Bar-On, Y. M. et al. Protection of BNT162b2 vaccine booster against COVID-19 in Israel. *N. Engl. J. Med.* **385**, 1393–1400 (2021).
5. Stephenson, E. et al. Single-cell multi-omics analysis of the immune response in COVID-19. *Nat. Med.* **27**, 904–916 (2021).
6. Su, Y. et al. Multi-omics resolves a sharp disease-state shift between mild and moderate COVID-19. *Cell* **183**, 1479–1495 (2020).
7. Ren, X. et al. COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. *Cell* **184**, 1895–1913 (2021).
8. Hadjadj, J. et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science* **369**, 718–724 (2020).
9. Schulte-Schrepping, J. et al. Severe COVID-19 is marked by a dysregulated myeloid cell compartment. *Cell* **182**, 1419–1440 (2020).
10. Saichi, M. et al. Single-cell RNA sequencing of blood antigen-presenting cells in severe COVID-19 reveals multi-process defects in antiviral immunity. *Nat. Cell Biol.* **23**, 538–551 (2021).
11. Krämer, B. et al. Early IFN-α signatures and persistent dysfunction are distinguishing features of NK cells in severe COVID-19. *Immunity* **54**, 2650–2669 (2021).
12. Meckiff, B. J. et al. Imbalance of regulatory and cytotoxic SARS-CoV-2-reactive CD4$^+$ T cells in COVID-19. *Cell* **183**, 1340–1353 (2020).
13. Ahern, D. J. et al. A blood atlas of COVID-19 defines hallmarks of disease severity and specificity. *Cell* **185**, 916–938 (2022).
14. Yoshida, M. et al. Local and systemic responses to SARS-CoV-2 infection in children and adults. *Nature* **602**, 321–327 (2022).
15. Yao, C. et al. Cell-type-specific immune dysregulation in severely Ill COVID-19 patients. *Cell Rep.* **34**, 108590 (2021).
16. Ivashkiv, L. B. & Donlin, L. T. Regulation of type I interferon responses. *Nat. Rev. Immunol.* **14**, 36–49 (2014).
17. Niemi, M. E. K. et al. Mapping the human genetic architecture of COVID-19. *Nature* **600**, 472–477 (2021).
18. Pathak, G. A. et al. A first update on mapping the human genetic architecture of COVID-19. *Nature* **608**, E1–E10 (2022).
19. Pairo-Castineira, E. et al. Genetic mechanisms of critical illness in COVID-19. *Nature* **591**, 92–98 (2021).
20. Namkoong, H. et al. DOCK2 is involved in the host genetics and biology of severe COVID-19. *Nature* **609**, 754–760 (2022).
21. Wu, P. et al. Trans-ethnic genome-wide association study of severe COVID-19. *Commun. Biol.* **4**, 1034 (2021).
22. Mathur, R. et al. Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the OpenSAFELY platform. *Lancet* **397**, 1711–1724 (2021).
23. World Health Organization. COVID-19 therapeutic trial synopsis. https://www.who.int/publications/i/item/covid-19-therapeutic-trial-synopsis (2020).

24. Turner, J. S. et al. Human germinal centres engage memory and naive B cells after influenza vaccination. *Nature* **586**, 127–132 (2020).
25. Mulder, K. et al. Cross-tissue single-cell landscape of human monocytes and macrophages in health and disease. *Immunity* **54**, 1883–1900 (2021).
26. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
27. Dann, E., Henderson, N. C., Teichmann, S. A., Morgan, M. D. & Marioni, J. C. Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat. Biotechnol.* **40**, 245–253 (2022).
28. 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).
29. Chen, G. et al. Clinical and immunological features of severe and moderate coronavirus disease 2019. *J. Clin. Invest.* **130**, 2620–2629 (2020).
30. Gatti, A., Radrizzani, D., Viganò, P., Mazzone, A. & Brando, B. Decrease of non-classical and intermediate monocyte subsets in severe acute SARS-CoV-2 infection. *Cytometry A* **97**, 887–890 (2020).
31. Silvin, A. et al. Elevated calprotectin and abnormal myeloid cell subsets discriminate severe from mild COVID-19. *Cell* **182**, 1401–1418 (2020).
32. Gavin, A. L. et al. Cleavage of DNA and RNA by PLD3 and PLD4 limits autoinflammatory triggering by multiple sensors. *Nat. Commun.* **12**, 5874 (2021).
33. Liu, M. et al. CXCL10/IP-10 in infectious diseases pathogenesis and potential therapeutic implications. *Cytokine Growth Factor Rev.* **22**, 121–130 (2011).
34. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020).
35. Patel, A. A. et al. The fate and lifespan of human monocyte subsets in steady state and systemic inflammation. *J. Exp. Med.* **214**, 1913–1923 (2017).
36. Zhang, J. Y. et al. Single-cell landscape of immunological responses in patients with COVID-19. *Nat. Immunol.* **21**, 1107–1118 (2020).
37. Bagaev, D. V. et al. VDJdb in 2019: database extension, new analysis infrastructure and a T-cell receptor motif compendium. *Nucleic Acids Res.* **48**, D1057–D1062 (2020).
38. Tang, Y. et al. Herc5 attenuates influenza A virus by catalyzing ISGylation of viral NS1 protein. *J. Immunol.* **184**, 5777–5790 (2010).
39. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
40. Hou, R., Denisenko, E., Ong, H. T., Ramilowski, J. A. & Forrest, A. R. R. Predicting cell-to-cell communication networks using NATMI. *Nat. Commun.* **11**, 5011 (2020).
41. Zhang, M. J. et al. Polygenic enrichment distinguishes disease associations of individual cells in single-cell RNA-seq data. *Nat. Genet.* **54**, 1572–1580 (2022).
42. Provine, N. M. & Klenerman, P. MAIT cells in health and disease. *Annu. Rev. Immunol.* **38**, 203–228 (2020).
43. Wang, B. X. & Fish, E. N. Global virus outbreaks: interferons as 1st responders. *Semin. Immunol.* **43**, 101300 (2019).
44. Nile, S. H. et al. COVID-19: pathogenesis, cytokine storm and therapeutic potential of interferons. *Cytokine Growth Factor Rev.* **53**, 66–70 (2020).
45. Zhang, F. et al. IFN-γ and TNF-α drive a $CXCL10^{+}$ $CCL2^{+}$ macrophage phenotype expanded in severe COVID-19 lungs and inflammatory diseases with tissue inflammation. *Genome Med.* **13**, 64 (2021).
46. Kratofil, R. M., Kubes, P. & Deniset, J. F. Monocyte conversion during inflammation and injury. *Arterioscler. Thromb. Vasc. Biol.* **37**, 35–42 (2017).
47. Downes, D. J. et al. Identification of LZTFL1 as a candidate effector gene at a COVID-19 risk locus. *Nat. Genet.* **53**, 1606–1615 (2021).
48. Kasela, S. et al. Integrative approach identifies SLC6A20 and CXCR6 as putative causal genes for the COVID-19 GWAS signal in the 3p21.31 locus. *Genome Biol.* **22**, 1–10 (2021).

[1]Department of Statistical Genetics, Osaka University Graduate School of Medicine, Suita, Japan. [2]Department of Respiratory Medicine and Clinical Immunology, Osaka University Graduate School of Medicine, Suita, Japan. [3]Laboratory of Statistical Immunology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Suita, Japan. [4]Laboratory of Experimental Immunology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Suita, Japan. [5]Laboratory of Immune Regulation, Immunology Frontier Research Center, Osaka University, Suita, Japan. [6]Department of Immunopathology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Suita, Japan. [7]Laboratory of Human Immunology (Single Cell Genomics), WPI Immunology Frontier Research Center, Osaka University, Suita, Japan. [8]Genome Information Research Center, Research Institute for Microbial Diseases, Osaka University, Suita, Japan. [9]Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Suita, Japan. [10]Core Instrumentation Facility, Immunology Frontier Research Center and Research Institute for Microbial Diseases, Osaka University, Suita, Japan. [11]Laboratory of Human Immunology (Single Cell Immunology), Immunology Frontier Research Center, Osaka University, Suita, Japan. [12]Center for Infectious Disease Education and Research (CiDER), Osaka University, Suita, Japan. [13]Laboratory for Systems Genetics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan. [14]Department of Genome Informatics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. [15]Department of Infectious Diseases, Keio University School of Medicine, Tokyo, Japan. [16]Division of Pulmonary Medicine, Department of Medicine, Keio University School of Medicine, Tokyo, Japan. [17]Japan Agency for Medical Research and Development – Core Research for Evolutional Science and Technology (AMED–CREST), Osaka University, Osaka, Japan. [163]These authors contributed equally: Ryuya Edahiro, Yuya Shirai. [164]These authors jointly supervised this work: Atsushi Kumanogoh, Yukinori Okada. *A list of authors and their affiliations appear at the end of the paper. ✉e-mail: kumanogo@imed3.med.osaka-u.ac.jp; yokada@sg.med.osaka-u.ac.jp

Japan COVID-19 Task Force

Qingbo S. Wang[1,3], Ryuya Edahiro[1,2,163], Ho Namkoong[15], Takanori Hasegawa[18], Yuya Shirai[1,2,3,163], Kyuto Sonehara[1,9,13,14], Hiromu Tanaka[16], Ho Lee[16], Ryunosuke Saiki[19], Takayoshi Hyugaji[20], Eigo Shimizu[20], Kotoe Katayama[20], Masahiro Kanai[21], Tatsuhiko Naito[1], Noah Sasa[1,22], Kenichi Yamamoto[1,3,23], Yasuhiro Kato[2,6], Takayoshi Morita[2,6], Kazuhisa Takahashi[24], Norihiro Harada[24], Toshio Naito[25], Makoto Hiki[26,27], Yasushi Matsushita[28], Haruhi Takagi[24], Masako Ichikawa[24], Ai Nakamura[24], Sonoko Harada[24,29], Yuuki Sandhu[24], Hiroki Kabata[16], Katsunori Masaki[16], Hirofumi Kamata[16], Shinnosuke Ikemura[16], Shotaro Chubachi[16], Satoshi Okamori[16], Hideki Terai[16], Atsuho Morita[16], Takanori Asakura[16], Junichi Sasaki[30], Hiroshi Morisaki[31], Yoshifumi Uwamino[32], Kosaku Nanki[33], Sho Uchida[15], Shunsuke Uno[15], Tomoyasu Nishimura[15,34], Takashi Ishiguro[35], Taisuke Isono[35], Shun Shibata[35], Yuma Matsui[35], Chiaki Hosoda[35], Kenji Takano[35], Takashi Nishida[35], Yoichi Kobayashi[35], Yotaro Takaku[35], Noboru Takayanagi[35], Soichiro Ueda[36], Ai Tada[36], Masayoshi Miyawaki[36], Masaomi Yamamoto[36], Eriko Yoshida[36], Reina Hayashi[36], Tomoki Nagasaka[36], Sawako Arai[36], Yutaro Kaneko[36], Kana Sasaki[36], Etsuko Tagaya[37], Masatoshi Kawana[38], Ken Arimura[37], Kunihiko Takahashi[18], Tatsuhiko Anzai[18], Satoshi Ito[18], Akifumi Endo[39], Yuji Uchimura[40], Yasunari Miyazaki[41], Takayuki Honda[41], Tomoya Tateishi[41], Shuji Tohda[42], Naoya Ichimura[42], Kazunari Sonobe[42], Chihiro Tani Sassa[42], Jun Nakajima[42], Yasushi Nakano[43], Yukiko Nakajima[43], Ryusuke Anan[43], Ryosuke Arai[43], Yuko Kurihara[43], Yuko Harada[43], Kazumi Nishio[43], Tetsuya Ueda[44], Masanori Azuma[44], Ryuichi Saito[44], Toshikatsu Sado[44], Yoshimune Miyazaki[44], Ryuichi Sato[44], Yuki Haruta[44], Tadao Nagasaki[44], Yoshinori Yasui[45], Yoshinori Hasegawa[44], Yoshikazu Mutoh[46], Tomoki Kimura[47], Tomonori Sato[47], Reoto Takei[47], Satoshi Hagimoto[47], Yoichiro Noguchi[47], Yasuhiko Yamano[47], Hajime Sasano[47], Sho Ota[47], Yasushi Nakamori[48], Kazuhisa Yoshiya[48], Fukuki Saito[48], Tomoyuki Yoshihara[48], Daiki Wada[48], Hiromu Iwamura[48], Syuji Kanayama[48], Shuhei Maruyama[48], Takashi Yoshiyama[49], Ken Ohta[49], Hiroyuki Kokuto[49], Hideo Ogata[49], Yoshiaki Tanaka[49], Kenichi Arakawa[49], Masafumi Shimoda[49], Takeshi Osawa[49], Hiroki Tateno[50], Isano Hase[50], Shuichi Yoshida[50], Shoji Suzuki[50], Miki Kawada[51], Hirohisa Horinouchi[52], Fumitake Saito[53], Keiko Mitamura[54], Masao Hagihara[55], Junichi Ochi[53], Tomoyuki Uchida[55], Rie Baba[56], Daisuke Arai[56], Takayuki Ogura[56], Hidenori Takahashi[56], Shigehiro Hagiwara[56], Genta Nagao[56], Shunichiro Konishi[56], Ichiro Nakachi[56], Koji Murakami[57], Mitsuhiro Yamada[57], Hisatoshi Sugiura[57], Hirohito Sano[57], Shuichiro Matsumoto[57], Nozomu Kimura[57], Yoshinao Ono[57], Hiroaki Baba[58], Yusuke Suzuki[59], Sohei Nakayama[59], Keita Masuzawa[59], Shinichi Namba[1], Takayuki Shiroyama[2], Yoshimi Noda[2], Takayuki Niitsu[2], Yuichi Adachi[2], Takatoshi Enomoto[2], Saori Amiya[2], Reina Hara[2], Yuta Yamaguchi[2,6], Teruaki Murakami[2,6], Tomoki Kuge[2], Kinnosuke Matsumoto[2], Yuji Yamamoto[2], Makoto Yamamoto[2], Midori Yoneda[2], Kazunori Tomono[60], Kazuto Kato[61], Haruhiko Hirata[2], Yoshito Takeda[2], Hidefumi Koh[62], Tadashi Manabe[62], Yohei Funatsu[62], Fumimaro Ito[62], Takahiro Fukui[62], Keisuke Shinozuka[62], Sumiko Kohashi[62], Masatoshi Miyazaki[62], Tomohisa Shoko[63], Mitsuaki Kojima[63], Tomohiro Adachi[63], Motonao Ishikawa[64], Kenichiro Takahashi[65], Takashi Inoue[66], Toshiyuki Hirano[66], Keigo Kobayashi[66], Hatsuyo Takaoka[66], Kazuyoshi Watanabe[67], Naoki Miyazawa[68], Yasuhiro Kimura[68], Reiko Sado[68], Hideyasu Sugimoto[68], Akane Kamiya[69], Naota Kuwahara[70], Akiko Fujiwara[70], Tomohiro Matsunaga[70], Yoko Sato[70], Takenori Okada[70], Yoshihiro Hirai[71], Hidetoshi Kawashima[71], Atsuya Narita[71], Kazuki Niwa[72], Yoshiyuki Sekikawa[72], Koichi Nishi[73], Masaru Nishitsuji[73], Mayuko Tani[73], Junya Suzuki[73], Hiroki Nakatsumi[73], Takashi Ogura[74], Hideya Kitamura[74], Eri Hagiwara[74], Kota Murohashi[74], Hiroko Okabayashi[74], Takao Mochimaru[75,76], Shigenari Nukaga[75], Ryosuke Satomi[75], Yoshitaka Oyamada[75,76], Nobuaki Mori[77], Tomoya Baba[78], Yasutaka Fukui[78], Mitsuru Odate[78], Shuko Mashimo[78], Yasushi Makino[78], Kazuma Yagi[79], Mizuha Hashiguchi[79], Junko Kagyo[79], Tetsuya Shiomi[79], Satoshi Fuke[80], Hiroshi Saito[80], Tomoya Tsuchida[81], Shigeki Fujitani[82], Mumon Takita[82], Daiki Morikawa[82], Toru Yoshida[82], Takehiro Izumo[83], Minoru Inomata[83], Naoyuki Kuse[83], Nobuyasu Awano[83], Mari Tone[83], Akihiro Ito[84], Yoshihiko Nakamura[85], Kota Hoshino[85], Junichi Maruyama[85], Hiroyasu Ishikura[85], Tohru Takata[86], Toshio Odani[87], Masaru Amishima[88], Takeshi Hattori[88], Yasuo Shichinohe[89], Takashi Kagaya[90], Toshiyuki Kita[90], Kazuhide Ohta[90], Satoru Sakagami[90], Kiyoshi Koshida[90], Kentaro Hayashi[91], Tetsuo Shimizu[91], Yutaka Kozu[91], Hisato Hiranuma[91], Yasuhiro Gon[91], Namiki Izumi[92], Kaoru Nagata[92], Ken Ueda[92], Reiko Taki[92], Satoko Hanada[92], Kodai Kawamura[93], Kazuya Ichikado[93], Kenta Nishiyama[93], Hiroyuki Muranaka[93], Kazunori Nakamura[93], Naozumi Hashimoto[94], Keiko Wakahara[94], Sakamoto Koji[94], Norihito Omote[94], Akira Ando[94], Nobuhiro Kodama[95], Yasunari Kaneyama[95], Shunsuke Maeda[95], Takashige Kuraki[96], Takemasa Matsumoto[96], Koutaro Yokote[97], Taka-Aki Nakada[98], Ryuzo Abe[98], Taku Oshima[98], Tadanaga Shimada[98], Masahiro Harada[99], Takeshi Takahashi[99], Hiroshi Ono[99], Toshihiro Sakurai[99], Takayuki Shibusawa[99], Yoshifumi Kimizuka[100], Akihiko Kawana[100], Tomoya Sano[100], Chie Watanabe[100], Ryohei Suematsu[100], Hisako Sageshima[101], Ayumi Yoshifuji[102], Kazuto Ito[102], Saeko Takahashi[103], Kota Ishioka[103], Morio Nakamura[103], Makoto Masuda[104], Aya Wakabayashi[104], Hiroki Watanabe[104], Suguru Ueda[104], Masanori Nishikawa[104], Yusuke Chihara[105], Mayumi Takeuchi[105], Keisuke Onoi[105], Jun Shinozuka[105], Atsushi Sueyoshi[105], Yoji Nagasaki[106], Masaki Okamoto[107,108], Sayoko Ishihara[109], Masatoshi Shimo[109], Yoshihisa Tokunaga[107,108], Yu Kusaka[110], Takehiko Ohba[110], Susumu Isogai[110], Aki Ogawa[110], Takuya Inoue[110], Satoru Fukuyama[111], Yoshihiro Eriguchi[112], Akiko Yonekawa[112], Keiko Kan-o[111], Koichiro Matsumoto[111], Kensuke Kanaoka[113], Shoichi Ihara[113], Kiyoshi Komuta[113], Yoshiaki Inoue[114], Shigeru Chiba[115], Kunihiro Yamagata[116], Yuji Hiramatsu[117], Hirayasu Kai[116], Koichiro Asano[118], Tsuyoshi Oguma[118], Yoko Ito[118], Satoru Hashimoto[119], Masaki Yamasaki[119], Yu Kasamatsu[120], Yuko Komase[121], Naoya Hida[121], Takahiro Tsuburai[121], Baku Oyama[121], Minoru Takada[122], Hidenori Kanda[122], Yuichiro Kitagawa[123], Tetsuya Fukuta[123], Takahito Miyake[123], Shozo Yoshida[123], Shinji Ogura[123], Shinji Abe[124], Yuta Kono[124], Yuki Togashi[124], Hiroyuki Takoi[124], Ryota Kikuchi[124], Shinichi Ogawa[125], Tomouki Ogata[125], Shoichiro Ishihara[125], Arihiko Kanehiro[126,127], Shinji Ozaki[126], Yasuko Fuchimoto[126], Sae Wada[126], Nobukazu Fujimoto[126], Kei Nishiyama[128], Mariko Terashima[129], Satoru Beppu[129], Kosuke Yoshida[129], Osamu Narumoto[130], Hideaki Nagai[130], Nobuharu Ooshima[130], Mitsuru Motegi[131], Akira Umeda[132], Kazuya Miyagawa[133], Hisato Shimada[134], Mayu Endo[135], Yoshiyuki Ohira[132], Masafumi Watanabe[136], Sumito Inoue[136], Akira Igarashi[136], Masamichi Sato[136], Hironori Sagara[137], Akihiko Tanaka[137], Shin Ohta[137], Tomoyuki Kimura[137], Yoko Shibata[138], Yoshinori Tanino[138], Takefumi Nikaido[138], Hiroyuki Minemura[138], Yuki Sato[138], Yuichiro Yamada[139], Takuya Hashino[139], Masato Shinoki[139], Hajime Iwagoe[140], Hiroshi Takahashi[141], Kazuhiko Fujii[141], Hiroto Kishi[141], Masayuki Kanai[142], Tomonori Imamura[142], Tatsuya Yamashita[142], Masakiyo Yatomi[143], Toshitaka Maeno[143], Shinichi Hayashi[144], Mai Takahashi[144],

Mizuki Kuramochi[144], Isamu Kamimaki[144], Yoshiteru Tominaga[144], Tomoo Ishii[145], Mitsuyoshi Utsugi[146], Akihiro Ono[146], Toru Tanaka[147], Takeru Kashiwada[147], Kazue Fujita[147], Yoshinobu Saito[147], Masahiro Seike[147], Hiroko Watanabe[148], Hiroto Matsuse[149], Norio Kodaka[149], Chihiro Nakano[149], Takeshi Oshio[149], Takatomo Hirouchi[149], Shohei Makino[150], Moritoki Egi[150], Yosuke Omae[151], Yasuhito Nannya[19], Takafumi Ueno[152], Tomomi Takano[153], Kazuhiko Katayama[154], Masumi Ai[155], Atsushi Kumanogoh[2,6,9,12,17,164], Toshiro Sato[156], Naoki Hasegawa[15], Katsushi Tokunaga[151], Makoto Ishii[16], Ryuji Koike[157], Yuko Kitagawa[158], Akinori Kimura[159], Seiya Imoto[20], Satoru Miyano[18], Seishi Ogawa[19,160,161], Takanori Kanai[33,162], Koichi Fukunaga[16] & Yukinori Okada[1,3,9,12,13,14,164]

[18]M&D Data Science Center, Tokyo Medical and Dental University, Tokyo, Japan. [19]Department of Pathology and Tumor Biology, Kyoto University, Kyoto, Japan. [20]Division of Health Medical Intelligence, Human Genome Center, the Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [21]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA. [22]Department of Otorhinolaryngology—Head and Neck Surgery, Osaka University Graduate School of Medicine, Suita, Japan. [23]Department of Pediatrics, Osaka University Graduate School of Medicine, Suita, Japan. [24]Department of Respiratory Medicine, Juntendo University Faculty of Medicine and Graduate School of Medicine, Tokyo, Japan. [25]Department of General Medicine, Juntendo University Faculty of Medicine and Graduate School of Medicine, Tokyo, Japan. [26]Department of Emergency and Disaster Medicine, Juntendo University Faculty of Medicine and Graduate School of Medicine, Tokyo, Japan. [27]Department of Cardiovascular Biology and Medicine, Juntendo University Faculty of Medicine and Graduate School of Medicine, Tokyo, Japan. [28]Department of Internal Medicine and Rheumatology, Juntendo University Faculty of Medicine and Graduate School of Medicine, Tokyo, Japan. [29]Atopy (Allergy) Research Center, Juntendo University Graduate School of Medicine, Tokyo, Japan. [30]Department of Emergency and Critical Care Medicine, Keio University School of Medicine, Tokyo, Japan. [31]Department of Anesthesiology, Keio University School of Medicine, Tokyo, Japan. [32]Department of Laboratory Medicine, Keio University School of Medicine, Tokyo, Japan. [33]Division of Gastroenterology and Hepatology, Department of Medicine Keio University School of Medicine, Tokyo, Japan. [34]Keio University Health Center, Tokyo, Japan. [35]Department of Respiratory Medicine, Saitama Cardiovascular and Respiratory Center, Kumagaya, Japan. [36]Department of Internal Medicine, Japan Community Health Care Organization (JCHO) Saitama Medical Center, Saitama, Japan. [37]Department of Respiratory Medicine, Tokyo Women's Medical University, Tokyo, Japan. [38]Department of General Medicine, Tokyo Women's Medical University, Tokyo, Japan. [39]Clinical Research Center, Tokyo Medical and Dental University Hospital of Medicine, Tokyo, Japan. [40]Department of Medical Informatics, Tokyo Medical and Dental University Hospital of Medicine, Tokyo, Japan. [41]Respiratory Medicine, Tokyo Medical and Dental University, Tokyo, Japan. [42]Clinical Laboratory, Tokyo Medical and Dental University Hospital of Medicine, Tokyo, Japan. [43]Kawasaki Municipal Ida Hospital, Department of Internal Medicine, Kawasaki, Japan. [44]Department of Respiratory Medicine, Osaka Saiseikai Nakatsu Hospital, Osaka, Japan. [45]Department of Infection Control, Osaka Saiseikai Nakatsu Hospital, Osaka, Japan. [46]Department of Infectious Diseases, Tosei General Hospital, Seto, Japan. [47]Department of Respiratory Medicine and Allergy, Tosei General Hospital, Seto, Japan. [48]Department of Emergency and Critical Care Medicine, Kansai Medical University General Medical Center, Moriguchi, Japan. [49]Japan Anti-Tuberculosis Association (JATA) Fukujuji Hospital, Kiyose, Japan. [50]Department of Pulmonary Medicine, Saitama City Hospital, Saitama, Japan. [51]Department of Infectious Diseases, Saitama City Hospital, Saitama, Japan. [52]Department of General Thoracic Surgery, Saitama City Hospital, Saitama, Japan. [53]Department of Pulmonary Medicine, Eiju General Hospital, Tokyo, Japan. [54]Division of Infection Control, Eiju General Hospital, Tokyo, Japan. [55]Department of Hematology, Eiju General Hospital, Tokyo, Japan. [56]Saiseikai Utsunomiya Hospital, Utsunomiya, Japan. [57]Department of Respiratory Medicine, Tohoku University Graduate School of Medicine, Sendai, Japan. [58]Department of Infectious Diseases, Tohoku University Graduate School of Medicine, Sendai, Japan. [59]Department of Respiratory Medicine, Kitasato University Kitasato Institute Hospital, Tokyo, Japan. [60]Division of Infection Control and Prevention, Osaka University Hospital, Suita, Japan. [61]Department of Biomedical Ethics and Public Policy, Osaka University Graduate School of Medicine, Suita, Japan. [62]Tachikawa Hospital, Tachikawa, Japan. [63]Department of Emergency and Critical Care Medicine, Tokyo Women's Medical University Medical Center East, Tokyo, Japan. [64]Department of Medicine, Tokyo Women's Medical University Medical Center East, Tokyo, Japan. [65]Department of Pediatrics, Tokyo Women's Medical University Medical Center East, Tokyo, Japan. [66]Department of Internal Medicine, Sano Kosei General Hospital, Sano, Japan. [67]Japan Community Health care Organization Kanazawa Hospital, Kanazawa, Japan. [68]Department of Respiratory Medicine, Saiseikai Yokohamashi Nanbu Hospital, Yokohama, Japan. [69]Department of Clinical Laboratory, Saiseikai Yokohamashi Nanbu Hospital, Yokohama, Japan. [70]Department of Internal Medicine, Internal Medicine Center, Showa University Koto Toyosu Hospital, Tokyo, Japan. [71]Department of Respiratory Medicine, Japan Organization of Occupational Health and Safety, Kanto Rosai Hospital, Kawasaki, Japan. [72]Department of General Internal Medicine, Japan Organization of Occupational Health and Safety, Kanto Rosai Hospital, Kawasaki, Japan. [73]Ishikawa Prefectural Central Hospital, Kanazawa, Japan. [74]Kanagawa Cardiovascular and Respiratory Center, Yokohama, Japan. [75]Department of Respiratory Medicine, National Hospital Organization Tokyo Medical Center, Tokyo, Japan. [76]Department of Allergy, National Hospital Organization Tokyo Medical Center, Tokyo, Japan. [77]Department of General Internal Medicine and Infectious Diseases, National Hospital Organization Tokyo Medical Center, Tokyo, Japan. [78]Department of Respiratory Medicine, Toyohashi Municipal Hospital, Toyohashi, Japan. [79]Keiyu Hospital, Yokohama, Japan. [80]Department of Respiratory Medicine, KKR Sapporo Medical Center, Sapporo, Japan. [81]Division of General Internal Medicine, Department of Internal Medicine, St. Marianna University School of Medicine, Kawasaki, Japan. [82]Department of Emergency and Critical Care Medicine, St. Marianna University School of Medicine, Kawasaki, Japan. [83]Japanese Red Cross Medical Center, Tokyo, Japan. [84]Matsumoto City Hospital, Matsumoto, Japan. [85]Department of Emergency and Critical Care Medicine, Faculty of Medicine, Fukuoka University, Fukuoka, Japan. [86]Department of Infection Control, Fukuoka University Hospital, Fukuoka, Japan. [87]Department of Rheumatology, National Hospital Organization Hokkaido Medical Center, Sapporo, Japan. [88]Department of Respiratory Medicine, National Hospital Organization Hokkaido Medical Center, Sapporo, Japan. [89]Department of Emergency and Critical Care Medicine, National Hospital Organization Hokkaido Medical Center, Sapporo, Japan. [90]National Hospital Organization Kanazawa Medical Center, Kanazawa, Japan. [91]Nihon University School of Medicine, Department of Internal Medicine, Division of Respiratory Medicine, Tokyo, Japan. [92]Musashino Red Cross Hospital, Musashino, Japan. [93]Division of Respiratory Medicine, Social Welfare Organization Saiseikai Imperial Gift Foundation, Inc., Saiseikai Kumamoto Hospital, Kumamoto, Japan. [94]Department of Respiratory Medicine, Nagoya University Graduate School of Medicine, Nagoya, Japan. [95]Fukuoka Tokushukai Hospital, Department of Internal Medicine, Kasuga, Japan. [96]Fukuoka Tokushukai Hospital, Respiratory Medicine, Kasuga, Japan. [97]Department of Endocrinology, Hematology and Gerontology, Chiba University Graduate School of Medicine, Chiba, Japan. [98]Department of Emergency and Critical Care Medicine, Chiba University Graduate School of Medicine, Chiba, Japan. [99]National Hospital Organization Kumamoto Medical Center, Kumamoto, Japan. [100]Division of Infectious Diseases and Respiratory Medicine, Department of Internal Medicine, National Defense Medical College, Tokorozawa, Japan. [101]Sapporo City General Hospital, Sapporo, Japan. [102]Department of Internal Medicine, Tokyo Saiseikai Central Hospital, Tokyo, Japan. [103]Department of Pulmonary Medicine, Tokyo Saiseikai Central Hospital, Tokyo, Japan. [104]Department of Respiratory Medicine, Fujisawa City Hospital, Fujisawa, Japan. [105]Uji-Tokushukai Medical Center, Uji, Japan. [106]Department of Infectious Disease and Clinical Research Institute, National Hospital Organization Kyushu Medical Center, Fukuoka, Japan.

[107]Department of Respirology, National Hospital Organization Kyushu Medical Center, Fukuoka, Japan. [108]Division of Respirology, Rheumatology, and Neurology, Department of Internal Medicine, Kurume University School of Medicine, Kurume, Japan. [109]Department of Infectious Disease, National Hospital Organization Kyushu Medical Center, Fukuoka, Japan. [110]Ome Municipal General Hospital, Ome, Japan. [111]Research Institute for Diseases of the Chest, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan. [112]Department of Medicine and Biosystemic Science, Kyushu University Graduate School of Medical Sciences, Fukuoka, Japan. [113]Daini Osaka Police Hospital, Osaka, Japan. [114]Department of Emergency and Critical Care Medicine, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan. [115]Department of Hematology, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan. [116]Department of Nephrology, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan. [117]Department of Cardiovascular Surgery, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan. [118]Division of Pulmonary Medicine, Department of Medicine, Tokai University School of Medicine, Isehara, Japan. [119]Department of Anesthesiology and Intensive Care Medicine, Kyoto Prefectural University of Medicine, Kyoto, Japan. [120]Department of Infection Control and Laboratory Medicine, Kyoto Prefectural University of Medicine, Kyoto, Japan. [121]Department of Respiratory Internal Medicine, St. Marianna University School of Medicine, Yokohama-City Seibu Hospital, Yokohama, Japan. [122]KINSHUKAI Hanwa The Second Hospital, Osaka, Japan. [123]Emergency and Disaster Medicine, Gifu University School of Medicine Graduate School of Medicine, Gifu, Japan. [124]Department of Respiratory Medicine, Tokyo Medical University Hospital, Tokyo, Japan. [125]JA Toride Medical Hospital, Toride, Japan. [126]Okayama Rosai Hospital, Okayama, Japan. [127]Himeji St. Mary's Hospital, Himeji, Japan. [128]Emergency & Critical Care, Niigata University, Niigata, Japan. [129]Emergency & Critical Care Center, National Hospital Organization Kyoto Medical Center, Kyoto, Japan. [130]National Hospital Organization Tokyo Hospital Hospital, Kiyose, Japan. [131]Fujioka General Hospital, Fujioka, Japan. [132]Department of General Medicine, School of Medicine, International University of Health and Welfare Shioya Hospital, Ohtawara, Japan. [133]Department of Pharmacology, School of Pharmacy, International University of Health and Welfare Shioya Hospital, Ohtawara, Japan. [134]Department of Respiratory Medicine, International University of Health and Welfare Shioya Hospital, Ohtawara, Japan. [135]Department of Clinical Laboratory, International University of Health and Welfare Shioya Hospital, Ohtawara, Japan. [136]Department of Cardiology, Pulmonology, and Nephrology, Yamagata University Faculty of Medicine, Yamagata, Japan. [137]Division of Respiratory Medicine and Allergology, Department of Medicine, School of Medicine, Showa University, Tokyo, Japan. [138]Department of Pulmonary Medicine, Fukushima Medical University, Fukushima, Japan. [139]Kansai Electric Power Hospital, Osaka, Japan. [140]Division of Infectious Diseases, Kumamoto City Hospital, Kumamoto, Japan. [141]Department of Respiratory Medicine, Kumamoto City Hospital, Kumamoto, Japan. [142]Department of Emergency and Critical Care Medicine, Tokyo Metropolitan Police Hospital, Tokyo, Japan. [143]Department of Respiratory Medicine, Gunma University Graduate School of Medicine, Maebashi, Japan. [144]National Hospital Organization Saitama Hospital, Wako, Japan. [145]Tokyo Medical University Ibaraki Medical Center, Inashiki, Japan. [146]Department of Internal Medicine, Kiryu Kosei General Hospital, Kiryu, Japan. [147]Department of Pulmonary Medicine and Oncology, Graduate School of Medicine, Nippon Medical School, Tokyo, Japan. [148]Division of Respiratory Medicine, Tsukuba Kinen General Hospital, Tsukuba, Japan. [149]Division of Respiratory Medicine, Department of Internal Medicine, Toho University Ohashi Medical Center, Tokyo, Japan. [150]Division of Anesthesiology, Department of Surgery, Kobe University Graduate School of Medicine, Kobe, Japan. [151]Genome Medical Science Project (Toyama), National Center for Global Health and Medicine, Tokyo, Japan. [152]Department of Biomolecular Engineering, Graduate School of Tokyo Institute of Technology, Tokyo, Japan. [153]Laboratory of Veterinary Infectious Disease, School of Veterinary Medicine, Kitasato University, Aomori, Japan. [154]Laboratory of Viral Infection, Department of Infection Control and Immunology, Ōmura Satoshi Memorial Institute & Graduate School of Infection Control Sciences, Kitasato University, Tokyo, Japan. [155]Department of Insured Medical Care Management, Tokyo Medical and Dental University Hospital of Medicine, Tokyo, Japan. [156]Department of Organoid Medicine, Keio University School of Medicine, Tokyo, Japan. [157]Medical Innovation Promotion Center, Tokyo Medical and Dental University, Tokyo, Japan. [158]Department of Surgery, Keio University School of Medicine, Tokyo, Japan. [159]Institute of Research, Tokyo Medical and Dental University, Tokyo, Japan. [160]Institute for the Advanced Study of Human Biology (WPI-ASHBi), Kyoto University, Kyoto, Japan. [161]Department of Medicine, Center for Hematology and Regenerative Medicine, Karolinska Institute, Stockholm, Sweden. [162]AMED-CREST, Japan Agency for Medical Research and Development, Tokyo, Japan.

Methods

Ethics and specimen collection of PBMC for scRNA-seq

Peripheral blood samples were obtained from patients with COVID-19 ($n = 73$) and healthy controls ($n = 75$) at Osaka University Hospital. Patients with COVID-19 were further categorized into groups of moderate ($n = 9$) and severe ($n = 64$) according to disease severity based on the highest score on the World Health Organization (WHO) Ordinal Scale for Clinical Improvement ever-present (WHO, R&D Blueprint—new coronavirus—COVID-19 Therapeutic Trial Synopsis, 2020; ref. 23). Almost all cases were patients who were transferred from nearby general hospitals because of severe or potentially severe illness during treatment and already initiated with systemic corticosteroids therapy at others hospitals according to RECOVERY study[49]. The detailed clinical data are summarized in Supplementary Table 1. Part of the subjects ($n_{COVID\text{-}19} = 30$, $n_{Control} = 31$) are described elsewhere[20]. One patient with COVID-19 had the karyotype abnormality and was excluded from the analyses including sex as a covariate. This study strictly follows the principles according to the Declaration of Helsinki, with written informed consent obtained from all participants before sample collection according to regular principles. Ethical approvals were gained from the Ethics Committees of Osaka University. There was no compensation for participants.

Preparation of single-cell suspensions

For both patients with COVID-19 and healthy controls, blood was collected into heparin tubes and PBMCs were isolated using Leucosep (Greiner Bio-One) density gradient centrifugation according to the manufacturer's instructions. Blood was processed within 3 h of collection for all samples, and stored at −80 °C until use.

Droplet-based single-cell sequencing

Single-cell suspensions were processed through the 10X Genomics Chromium Controller following the protocol outlined in the Chromium Single Cell V(D)J Reagent Kits (v1.1 Chemistry) User Guide. Chromium Next GEM Single Cell 5′ Library & Gel Bead Kit v1.1 (PN-1000167), Chromium Next GEM Chip G Single Cell Kit (PN-1000127) and Single Index Kit T Set A (PN-1000213) were applied during the process. Oil droplets of encapsulated single cells and barcoded beads (GEMs) were subsequently reverse-transcribed in a Veriti Thermal Cycler (Thermo Fisher Scientific), resulting in cDNA tagged with a cell barcode and unique molecular index (UMI). Next, cDNA was then amplified to generate single-cell libraries according to the manufacturer's protocol. Quantification was made with an Agilent Bioanalyzer High-Sensitivity DNA assay (Agilent, High-Sensitivity DNA Kit, 5067-4626). Subsequently amplified cDNA was enzymatically fragmented, end-repaired and polyA tagged. Cleanup/size selection was performed on amplified cDNA using SPRIselect magnetic beads (Beckman-Coulter; SPRIselect, B23317). Next, Illumina sequencing adapters were ligated to the size-selected fragments and cleaned up using SPRIselect magnetic beads. Finally, sample indices were selected and amplified, followed by a double-sided size selection using SPRIselect magnetic beads. Final library quality was assessed using an Agilent Bioanalyzer High-Sensitivity DNA assay. Samples were then sequenced on an Illumina NovaSeq 6000 as paired-end mode.

Alignment, quantification and quality control of scRNA-seq

Droplet libraries were processed using Cell Ranger 5.0.0 (10X Genomics). Sequencing reads were aligned with STAR (v2.7.2a)[50] using the GRCh38 human reference genome. Filtered expression matrices generated using Cell Ranger count were used to perform the analysis. We excluded cells that had fewer than the first percentile of UMIs or greater than 99th percentile of UMIs in each sample. We also excluded cells with <200 genes expressed or >10% of reads from mitochondrial genes or hemoglobin genes. Additionally, putative doublets were removed using Scrublet (v0.2.1)[51] and scds (v1.10.0)[52] for each sample.

scRNA-seq computational pipelines and analysis

The R package Seurat (v4.1.0) was used for data scaling, transformation, clustering, dimensionality reduction and most visualization[26]. Data were scaled and transformed using the SCTransform() function, and linear regression was performed to remove unwanted variation due to cell quality (% mitochondrial reads). For integration, we identified 3,000 shared highly variable genes (HVGs) using SelectIntegrationFeatures() function. Principal component analysis (PCA) was run on gene expression, followed by batch correction using harmony (v0.1)[53]. UMAP dimension reduction was generated based on the first 30 harmony-adjusted principal components[54]. A nearest-neighbor graph using the first 30 harmony-adjusted principal components was calculated using FindNeighbors() function, followed by clustering using FindClusters() function.

Cellular identity was determined by finding differentially expressed genes (DEGs) for each cluster using FindMarkers() function with parameter 'test.use=wilcox', and comparing those markers to known cell type-specific genes (Extended Data Fig. 1a). At the first round of clustering, we identified 13 cell subsets (Fig. 1b). To identify clusters within each major cell type, we performed a second round of clustering on monocytes/DC (CD14$^+$ monocytes, CD16$^+$ monocytes, cDC and pDC), T/NK cells (CD4T, T_{reg}, CD8T, MAIT, proliferative T cells (Pro_T) and NK) and B cells (B and PB), separately, and then we obtained 25 cell subsets (Extended Data Figs. 1a, 2a, 3a and 4a). The final annotation of PBMCs was compared to the PBMC annotation of Azimuth using Seurat[26]. For each of the 25 clusters in our data, the percentage of cells mapped to each Azimuth annotation was calculated (Extended Data Fig. 1b).

To perform polygenic GWAS signal analysis (Fig. 6c) and single-cell eQTL analysis (Fig. 7a), six major cell types were defined from 25 clusters as follows: CD4_naive, CD4_memory, CD4_Ef and T_{reg} were annotated as CD4$^+$ T cells (CD4T); CD8_naive, CD8_memory and CD8_Ef were annotated as CD8$^+$ T cells (CD8T); NK and NK_CD56^{bright} were annotated as NK; cMono_S100A, cMono_IL1B, cMono_CCL3, intMono and ncMono were annotated as monocytes (Mono); cDC and pDC were annotated as DC; B_naive, B_memory and B_activated and B_plasma were annotated as B cells (B).

Differential abundance analysis

We used Milo[27] (v1.2.0) to test for the differential abundance of cells within defined neighborhoods, between two conditions (that is, COVID-19 versus healthy controls or moderate COVID-19 versus severe COVID-19). We first used the buildGraph function to construct a KNN graph with $k = 30$, using 30 principal components ($d = 30$). Next, we used the make neighborhoods function to assign cells to neighborhoods based on their connectivity over the KNN graph. For computational efficiency, we subsampled 5% for all PBMCs and T cells and 10% for mono/DC and B cells. To test for differential abundance, Milo fit an NB GLM to the counts for each neighborhood, accounting for different numbers of cells across samples using TMM normalization, and use the QL F-test with a specified contrast to compute a P value for each neighborhood. We included age, sex, days since symptom onset and duration of systemic steroids treatment at the time of specimen collection (the last two variables included only in moderate COVID-19 versus severe COVID-19) as covariates in testNhoods function. To control for multiple testing, we adapted the spatial FDR implemented in Milo and used 10% spatial FDR as a threshold for significance. The spatial FDR and $\log_2$ foldchange (FC) of number of cells between two conditions in each neighborhood were used for visualization.

IFN response scoring

To evaluate IFN response across PBMCs, we downloaded a gene set termed 'GOBP_RESPONSE_TO_TYPE_I_INTERFERON (GO:0034340)' and 'GOBP_RESPONSE_TO_INTERFERON_GAMMA (GO:0034341)' from MSigDB. IFN response scores were evaluated using AddModuleScore() function implemented in Seurat with default parameter. To identify cell

types with high IFN response, we calculated the average module scores across each of 13 cell types and compared them with that of all PMBCs. The module scores of cells in each cell type were also compared with the average score of all PBMCs using two-sided one-sample *t*-test. To evaluate IFN response of COVID-19 patients in each cell type, we calculated the average module scores across each of 13 cell types by three clinical statuses and compared the scores of moderate or severe group with those of healthy group. The module scores of cells of moderate or severe group were also compared to those of healthy group in each cell type using two-sided Welch's *t*-test.

Differential gene expression analysis and GO enrichment

Differential gene expression analysis was performed among (1) all COVID-19 patients versus healthy controls, (2) moderate patients versus healthy controls, (3) severe patients versus healthy controls and (4) moderate versus severe patients. Donor pseudobulk matrices were first created by aggregating gene counts for each cell type, within each sample. Genes were considered for the analysis if they were expressed (UMI count > 0) in more than 10% of cells per cell type. Samples with more than five cells in a cell type were considered in the analysis of the corresponding cell type. Differential gene expression testing was performed using an NB GLM implemented in the Bioconductor package edgeR (v3.32.0)[55]. We included age, sex, days since symptom onset and duration of systemic steroids treatment at the time of specimen collection (the last two variables included only in moderate COVID-19 versus severe COVID-19) in the model as covariates. Statistically significant DEGs were defined with FDR < 0.05 and FC > 2. To find the function of upregulated DEGs, we used the function compareCluster (fun = "enrichGO," pvalueCutoff = 0.05, pAdjustMethod = "BH," OrgDb = "org.Hs.eg.db", ont = "BP") of Clusterprofiler (v3.14.3)[56].

Estimation of RNA velocity

Spliced and unspliced transcripts were quantified using dropEst (v0.8.6)[57]. Monocytes and DC clusters were evaluated by RNA velocity analysis using scVelo (v0.2.3)[34], separately for each of the two conditions. dropEst-derived counts were processed, filtered and normalized before velocity estimation on the basis of the top 2,000 HVGs with at least 20 UMI for both spliced and unspliced transcripts across all cells. The moments facilitated the RNA velocity estimation implemented in function scv.tl.velocity with mode set to 'dynamical'. The estimated velocities were used to construct a velocity graph representing the transition probabilities among cells by function scv.tl.velocity_graph. Finally, the velocity graph was used to embed the RNA velocities into the UMAP by the function scv.pl.velocity_embedding_stream. The fractions of unspliced counts were adopted to quantitatively compare the differences in estimated cell transition in monocytes between COVID-19 and healthy groups. The average unspliced ratio of each sample for each of the three monocyte clusters was calculated. The difference of increase in unspliced ratio across three monocyte clusters (ordered from cMono < intMono < ncMono) was evaluated using a linear regression model in each of the two conditions. The interaction between three monocyte clusters and two conditions (ordered from healthy < COVID-19) was also evaluated using a linear regression model with cluster, condition and cluster × condition as covariates.

TCR and BCR analysis

Droplet-based sequencing data for TCR sequences and BCR sequences were aligned and quantified using 5.0.0 (10X Genomics) against the GRCh38 human VDJ reference genome. Filtered annotated contigs for TCR sequences and BCR sequences were analyzed using Scirpy (v0.10.0)[58]. For TCR analysis, we selected T cells that were annotated as following nine cell types via single-cell RNA-seq analysis: CD4_naive, CD4_memory, CD4_Ef, T_{reg}, CD8_naive, CD8_memory, CD8_Ef, MAIT and Pro_T (Extended Data Fig. 3a). Only cells with both TCR α-chain (TRA) and TCR β-chain (TRB) remained for the downstream analysis. Each unique TRA(s)–TRB(s) pair was defined as a clonotype. Similarly, for BCR analysis, we selected B cells which were annotated as following four cell types via scRNA-seq analysis: B_naive, B_memory, B_activated and B_plasma (Extended Data Fig. 4a). Only cells with both heavy chain (IGH) and light chain (IGK or IGL) were kept for further analysis. Each unique IGH(s)–IGK/IGL(s) pair was defined as a clonotype.

For TCR and BCR data, clonotypes were defined based on CDR3 amino acid sequences with receptor_arms = 'any', metric = 'alignment' and default cutoff of ten. If one clonotype was present in at least two cells, cells harboring this clonotype were considered to be clonal and the number of cells with such pairs indicated the degree of clonality of the clonotype. Using barcode information, T cells with prevalent TCR clonotypes and B cells with prevalent BCR clonotypes were projected on UMAP embedding. To evaluate differences of clonal state between disease status in each cluster, the average clonal expansion rate of each sample was evaluated using two-sided Welch's *t*-test, where only cells mapped with TCRs/BCRs were included in the analysis.

We downloaded VDJdb[37], a curated database of TCR sequences with known antigen specificities, and then investigated the TCR that was specific to SARS-CoV-2, based on CDR3 amino acid sequences with the same parameter as above.

Cell–cell interaction analysis in PBMC

At first, to reduce the influence of individual samples contributing a larger number of cells and to speed up computation, we capped the number of cells per sample at 2,500 randomly sampled cells. This was done using the SubsetData() function in Seurat. Genes were adopted if they were expressed in more than 1% of all PBMCs. Putative cell–cell interactions of COVID-19 and healthy were quantified using CellPhoneDB (v2.0.0) and NATMI with default settings[39,40]. To investigate the differences in cell–cell interactions between COVID-19 versus healthy and between moderate versus severe COVID-19, we evaluated FC of mean expression weight using DiffEdges.py implemented in NATMI with default settings. As *CXCL10*/*CXCR3* interaction is not listed on connectomeDB2020, we manually added *CXCL10* as ligand and *CXCR3* as receptor. Then, cell–cell interactions of *CXCL10*/*CXCR3* and *IFNG*/*IFNGR* were evaluated for each of the three conditions by mean expression weight using NATMI.

Polygenic GWAS signals on PBMC

We used MAGMA (v1.07)[59] to compute gene-level association *P* values and *z*-score from GWAS summary statistics of COVID-HGI (round 6)[18]. We used a reference panel based on individuals of European ancestry in the 1000 Genomes Project and used a 10-kb window around the gene body to map SNPs to genes. We selected the top 100 genes based on MAGMA *P* values as putative disease genes.

We used scDRS (v1.0.1)[41] to quantify the aggregate expression of putative disease genes derived from GWAS summary statistics using MAGMA (each putative disease gene is weighted by its GWAS MAGMA *z*-score and inversely weighted by its gene-specific technical noise level in the single-cell data) in each cell of COVID-19 scRNA-seq data to generate cell-specific raw disease scores. A 1,000 sets of cell-specific raw control scores were calculated from matched control gene sets (matching the gene set size, mean expression and expression variance of the putative disease genes). Then, we normalized the raw disease score and raw control scores for each cell, producing the normalized disease score and normalized control scores. To compute the scores described above, we used the function scdrs compute-score (--n_ctrl = 1000, --cov-file = age, sex, number of genes per cell and disease severity).

For downstream analysis, we performed cell type-level analyses to associate predefined cell types to disease and assess heterogeneity in association to disease across cells within a predefined cell type using the function scdrs perform-downstream with default settings. To correct multiple testing, FDR was calculated via the Benjamini–Hochberg method across all pairs of cell types and three GWAS phenotypes. To compare scDRS disease scores between COVID-19 phenotypes in six

major cell types, the differences in average disease scores of each sample were evaluated using two-sided paired *t*-test (adjusted for multiple comparisons using Bonferroni's correction).

Genotype data, quality control and genotype imputation

We performed GWAS genotyping of COVID-19 cases and healthy controls using Infinium Asian Screening Array (Illumina) through collaboration with Japan COVID-19 Task Force (https://www.covid19-taskforce.jp/en/home/). We applied stringent quality control filters to the samples (sample call rate < 0.98, related samples with PI_HAT > 0.175 or outlier samples from East Asian clusters in PCA with HapMap project samples), and variants (variant call rate < 0.99, deviation from Hardy–Weinberg equilibrium with $P < 1.0 \times 10^{-6}$, or minor allele count < 5). We used SHAPEIT4 software (version 4.2.1)[60] for haplotype phasing of autosomal genotype data. After phasing, we used Minimac4 software (version 1.0.1)[61] for genome-wide genotype imputation. We used the population-specific imputation reference panel of Japanese ($n = 1{,}037$) combined with 1,000 Genomes Project Phase3v5 samples ($n = 2{,}504$)[62].

Single-cell eQTL analysis

We applied pseudobulk approach for single-cell eQTL analysis. First, we performed single-cell-level normalization using scran (v1.18.5)[63]. Gene expression per cell type per sample was then calculated as the mean of log2-transformed counts-per-cell-normalized expression across cells. Samples with more than five cells in a cell type were considered in the analysis of the corresponding cell type. We examined eQTL effects of the replicated variants in Japanese COVID-19 GWAS on V2G, the highest gene prioritized by the V2G score of Open Target Genetic, separately for COVID-19 and healthy controls[20,64]. rs35081325 and rs77534576 were excluded from eQTL analysis due to low allele frequency and low expression of V2G (*FLJ45513*), respectively. For PCA, genes were adopted if they were expressed in more than 1% of all PBMCs.

In the eQTL analysis of COVID-19-associated variants, dosage effects of the variants on the gene expression mean were evaluated using linear regression models with the top two PCs of the genotype data, the top two PCs of the gene expression, age, sex, days since symptom onset, duration of systemic steroids treatment at the time of specimen collection and disease severity (the last three variables included only in COVID-19 analysis) as covariates. In the interaction eQTL analysis of the *IFNAR2* variant (rs13050728), the top two PCs of the genotype data, the top two PCs of the gene expression, age and sex were included as covariates. *R* statistical software (version 4.0.2) was used for the analysis. To correct multiple testing, FDR was calculated via the Benjamini–Hochberg method across all pairs of cell types and five variants, separately for COVID-19 and healthy controls.

Statistics and reproducibility

No statistical method was used to predetermine sample size. No data were excluded from the analyses. We did not use any study design that required randomization or blinding.

Reporting summary

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

Data availability

Raw sequencing data of scRNA-seq are available at the Japanese Genotype-phenotype Archive (JGA) with accession codes JGAS000593 (https://ddbj.nig.ac.jp/resource/jga-study/JGAS000593)/JGAD000722 (https://ddbj.nig.ac.jp/resource/jga-dataset/JGAD000722). A part of the raw scRNA-seq data ($n_{\text{COVID-19}} = 30$, $n_{\text{Control}} = 31$) has already been deposited[20] and is available under controlled access at JGA with accession codes JGAS000543 (https://ddbj.nig.ac.jp/resource/jga-study/JGAS000543)/JGAD000662 (https://ddbj.nig.ac.jp/resource/jga-dataset/JGAD000662). All the raw sequencing data of scRNA-seq can also be accessed through application at the NBDC with the accession code hum0197 (https://humandbs.biosciencedbc.jp/en/hum0197-latest). Genotype data of the subjects are available at European Genome-Phenome Archive (EGA) with the accession code EGAS00001006950 (https://ega-archive.org/studies/EGAS00001006950). Raw sequencing data of scRNA-seq and genotype data are potentially identifiable and therefore under controlled access at JGA and EGA. The GWAS summary statistics of COVID-19 HGI (release 6) were obtained from https://www.covid19hg.org/results/r6/. The reference for cell type annotation of PBMC in scRNA-seq (pbmc_multimodal.h5seurat) was obtained from https://satijalab.org/seurat/articles/multimodal_reference_mapping.html.

Code availability

The code used in the paper is available at https://github.com/REdahiro/JPN_COVID-19_scRNAseq.

References

49. The RECOVERY Collaborative Group. Dexamethasone in hospitalized patients with Covid-19. *N. Engl. J. Med.* **384**, 693–704 (2021).
50. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
51. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **8**, 281–291 (2019).
52. Bais, A. S. & Kostka, D. Scds: computational annotation of doublets in single-cell RNA sequencing data. *Bioinformatics* **36**, 1150–1158 (2020).
53. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
54. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: uniform manifold approximation and projection. *J. Open Source Softw.* **3**, 861 (2018).
55. 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).
56. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. ClusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
57. Petukhov, V. et al. dropEst: pipeline for accurate estimation of molecular counts in droplet-based single-cell RNA-seq experiments. *Genome Biol.* **19**, 78 (2018).
58. Sturm, G. et al. Scirpy: a Scanpy extension for analyzing single-cell T-cell receptor-sequencing data. *Bioinformatics* **36**, 4817–4818 (2020).
59. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
60. Delaneau, O., Zagury, J.-F., Robinson, M. R., Marchini, J. L. & Dermitzakis, E. T. Accurate, scalable and integrative haplotype estimation. *Nat. Commun.* **10**, 5436 (2019).
61. Das, S. et al. Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
62. Okada, Y. et al. Deep whole-genome sequencing reveals recent selection signatures linked to evolution and disease risk of Japanese. *Nat. Commun.* **9**, 1631 (2018).
63. Lun, A. T. L., Bach, K. & Marioni, J. C. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biol.* **17**, 75 (2016).
64. Ghoussaini, M. et al. Open targets genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. *Nucleic Acids Res.* **49**, D1311–D1320 (2021).

Acknowledgements

We would like to sincerely thank all the participants involved in this study. We thank and acknowledge T. Shiroyama, K. Miyake, Y. Suga and Y. Naito for contribution to clinical practice. We are grateful to Y. Noda, T. Niitsu, Y. Adachi, T. Enomoto, S. Amiya, R. Hara, M. Yamamoto, T. Kuge, K. Matsumoto, M. Yoneda, Y. Yamamoto, Y. Yoshimine, S. Minoda, T. Hirayama, K. Funakoshi, Y. Okita, S. Kawada, D. Nakatsubo, T. Tada, M. Okamoto and H. Shimagami for contribution to clinical practice and sample collection.

We thank all the members of the Japan COVID-19 Task Force for their support and COVID-19 HGI for publicly sharing the GWAS summary statistics. This study was supported by AMED (JP20fk0108415 and JP20fk0108452 to K. Fukunaga, JP21km0405211, JP21km0405217, JP21ek0410075, JP223fa627002, JP223fa627010, JP233fa627011 and JP23zf0127008 to Y. Okada), JST CREST (JPMJCR20H2 to Y. Okada), JST PRESTO (JPMJPR21R7 to H. Namkoong), JST Moonshot R&D (JPMJMS2021 and JPMJMS2024 to Y. Okada), MHLW (20CA2054 to Y. Okada), JSPS KAKENHI (22H00476 to Y. Okada and JP18H05282 to A. Kumanogoh), AMED–CREST (22gm1810003h0001 to A. Kumanogoh), Takeda Science Foundation, the Mitsubishi Foundation, the Team Osaka University Research Project in The Nippon Foundation—Osaka University Project for Infectious Disease Prevention and Bioinformatics Initiative of Osaka University Graduate School of Medicine, Institute for Open and Transdisciplinary Research Initiatives, Center for Infectious Disease Education and Research (CiDER), and Center for Advanced Modality and DDS (CAMaD), Osaka University. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the paper.

Author contributions

R.E., Y. Shirai, D. Okuzaki, A. Kumanogoh and Y. Okada conceived and designed the study; R.E., Y. Shirai, Y. Yamaguchi, T. Murakami, Y.L., D. Motooka, Y. Naito, A. Takuwa, F. Sugihara and K. Tanaka performed the experiments; R.E., Y. Shirai, K. Sonehara, Y. Tomofuji, D. Okuzaki and Y. Okada conducted data analysis; R.E., Y. Shirai, D. Okuzaki, A. Kumanogoh and Y. Okada wrote the paper; H. Hirata and Y. Takeda took care of patients and provided the clinical information; Y. Takeshima, S. Sakakibara, T. Morita, Y. Kato, Y.L., J.W., H. Namkoong, H. Tanaka, H.L. and K. Fukunaga provided intellectual input into throughout the study, provided comments and helped edit the paper. All authors read and approved the final paper.

Competing interests

All authors declare no conflict of interest.

Additional information

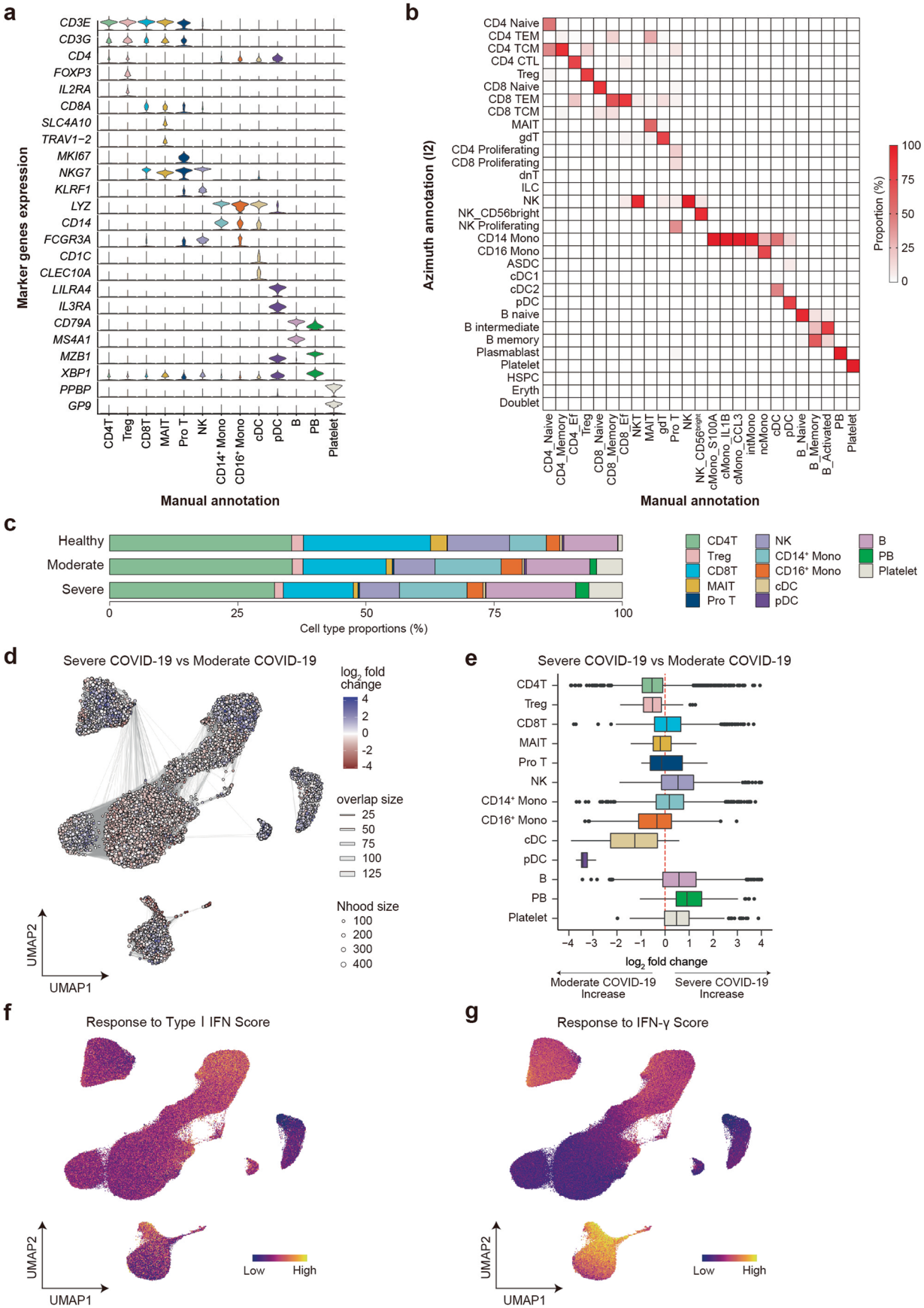
Extended data is available for this paper at https://doi.org/10.1038/s41588-023-01375-1.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41588-023-01375-1.

Correspondence and requests for materials should be addressed to Atsushi Kumanogoh or Yukinori Okada.

Peer review information *Nature Genetics* thanks Xianwen Renand the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

https://doi.org/10.1038/s41588-023-01375-1

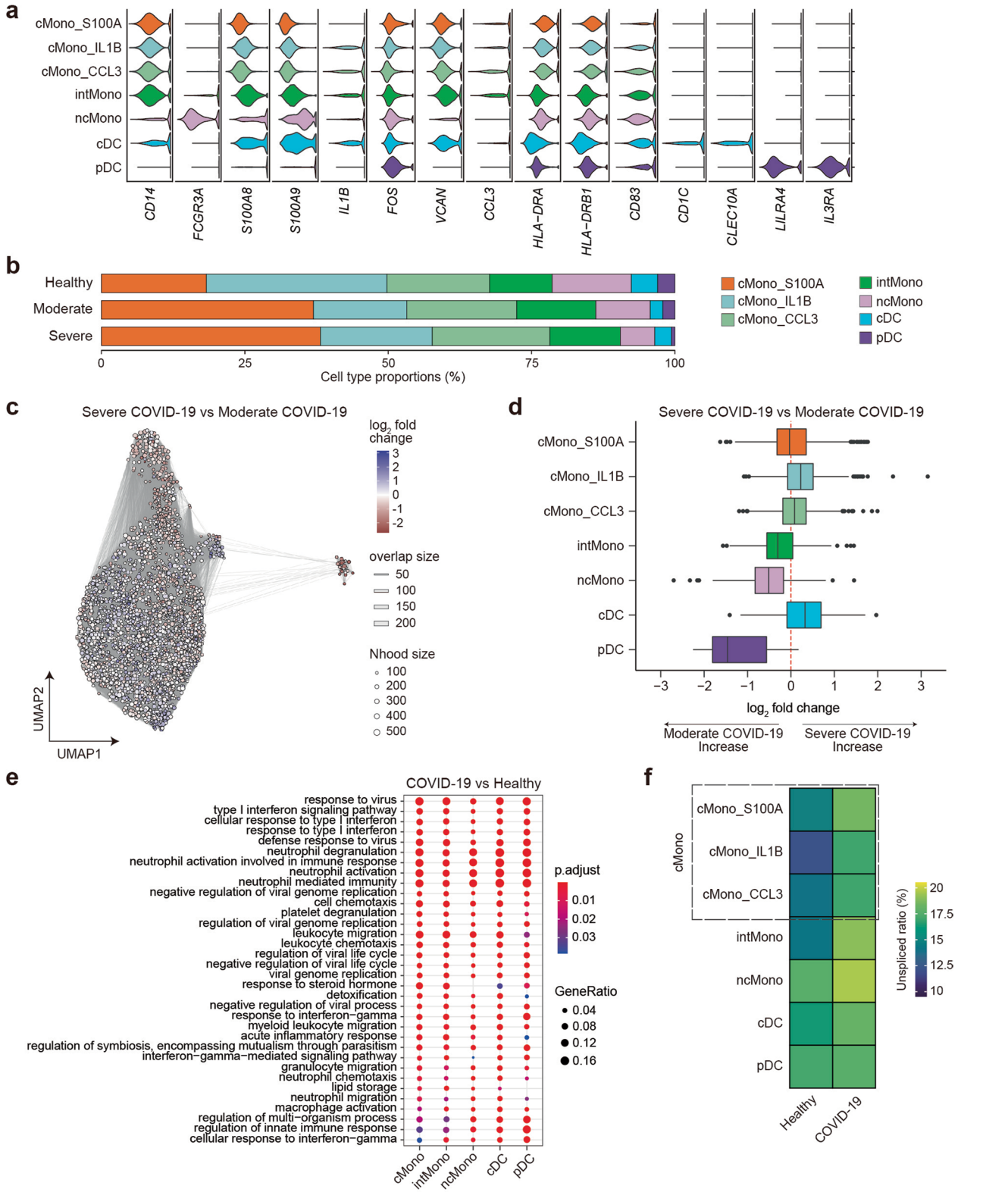


Extended Data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s41588-023-01375-1

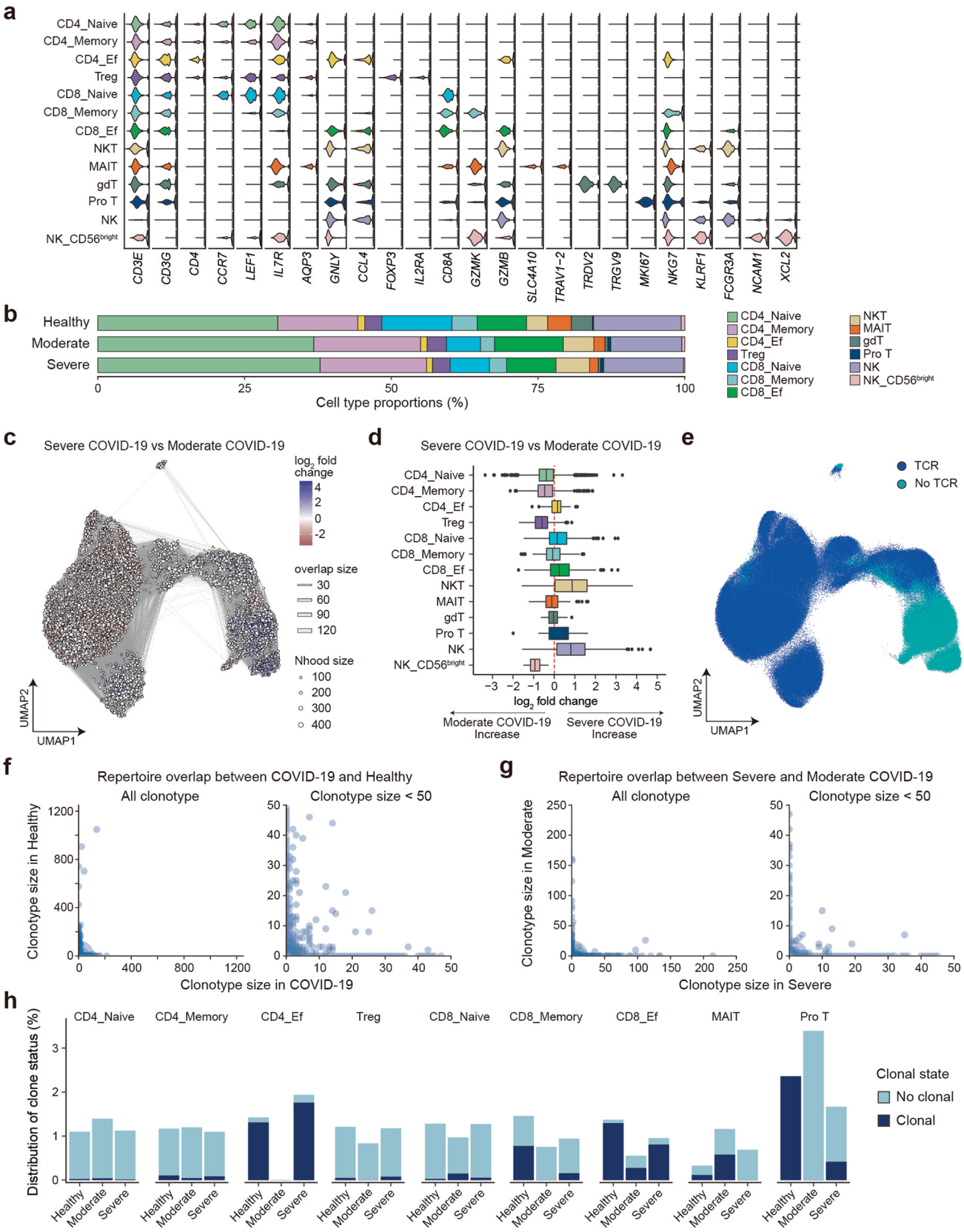
Extended Data Fig. 1 | Cell type annotation of PBMC, differential abundance analysis and IFN responses. (**a**) Violin plots showing the expression distribution of selected canonical cell markers in the 13 clusters. The rows represent selected marker genes and the columns represent clusters. (**b**) Tile plot showing percentage concordance between the manually annotated 25 clusters and Azimuth annotation. (**c**) A bar plot of the proportion of cell types shown in Fig. 1b, separated by three conditions ($n = 75$ healthy, $n = 9$ moderate, $n = 64$ severe). (**d**) Graph representation of neighborhoods identified by Milo in COVID-19 patients. Nodes are neighborhoods, colored by their $\log_2$ fold change between moderate ($n = 8$) and severe ($n = 64$) COVID-19 patients adjusted by age, sex, time since symptom onset and duration of systemic steroids treatment. Nhood, neighborhood. (**e**) Box plot showing the distribution of adjusted $\log_2$ fold change in abundance between moderate and severe COVID-19 in neighborhoods according to 13 cell types. Boxes denote the interquartile range (IQR), and the median is shown as horizontal bars. Whiskers extend to 1.5 times the IQR, and outliers are shown as individual points. (**f**,**g**) UMAP embedding of PBMCs colored by Type I IFN response score and IFN-γ response score. The score was calculated using a gene set termed 'GOBP_RESPONSE_TO_TYPE_I_INTERFERON' (GO:0034340) and 'GOBP_RESPONSE_TO_INTERFERON_GAMMA' (GO:0034341), respectively.

https://doi.org/10.1038/s41588-023-01375-1



Extended Data Fig. 2 | Immunological features of monocytes and dendritic cells. (**a**) Violin plots showing the expression distribution of selected canonical cell markers in the seven clusters. The rows represent clusters and the columns represent selected marker genes. (**b**) A bar plot of the proportion of cell types, separated by three conditions ($n = 75$ healthy, $n = 9$ moderate, $n = 64$ severe). (**c**) Graph representation of neighborhoods identified by Milo in COVID-19 patients. Nodes are neighborhoods, colored by their $\log_2$ fold change between moderate ($n = 8$) and severe ($n = 64$) COVID-19 patients adjusted by age, sex, time since symptom onset and duration of systemic steroids treatment. Nhood, neighborhood. (**d**) Box plot showing the distribution of adjusted $\log_2$ fold change in abundance between moderate and severe COVID-19 in neighborhoods according to seven cell types. Boxes denote the interquartile range (IQR), and the median is shown as horizontal bars. Whiskers extend to 1.5 times the IQR, and outliers are shown as individual points. (**e**) The top 20 enriched biological processes by GO analysis of upregulated DEGs of COVID-19 ($n = 72$) compared to healthy controls ($n = 75$) in five cell types. Dot color indicates the statistical significance of the enrichment (adjusted P-values via the Benjamini-Hochberg method), and dot size represents gene ratio annotated to each term. (**f**) Heatmaps depicting the average unspliced ratio of seven cell types stratified by COVID-19 ($n = 73$) and healthy controls ($n = 75$).

https://doi.org/10.1038/s41588-023-01375-1

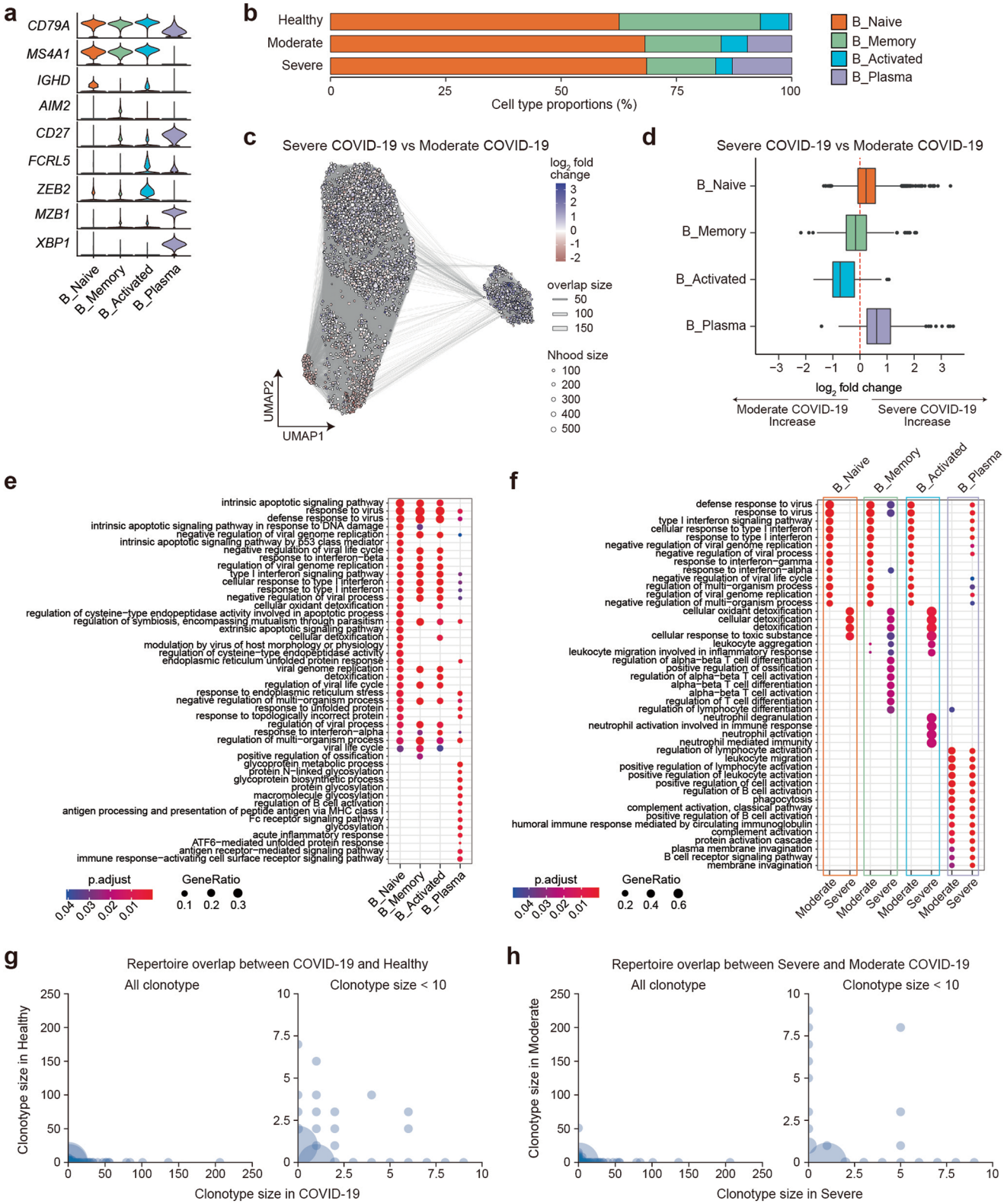


Extended Data Fig. 3 | See next page for caption.

https://doi.org/10.1038/s41588-023-01375-1

Extended Data Fig. 3 | Immunological features of T cells and NK cells and TCR analysis. (**a**) Violin plots showing the expression distribution of selected canonical cell markers in 13 clusters. The rows represent clusters and the columns represent selected marker genes. (**b**) A bar plot of the proportion of cell types, separated by three conditions ($n = 75$ healthy, $n = 9$ moderate, $n = 64$ severe). (**c**) Graph representation of neighborhoods identified by Milo in COVID-19 patients. Nodes are neighborhoods, colored by their $\log_2$ fold change between moderate ($n = 8$) and severe ($n = 64$) COVID-19 patients adjusted by age, sex, time since symptom onset and duration of systemic steroids treatment. Nhood, neighborhood. (**d**) Box plot showing the distribution of adjusted $\log_2$ fold change in abundance between moderate and severe COVID-19 in neighborhoods according to 13 cell types. Boxes denote the interquartile range (IQR), and the median is shown as horizontal bars. Whiskers extend to 1.5 times the IQR, and outliers are shown as individual points. (**e**) UMAP embedding of TCR detection. (**f**) Repertoire overlap according to clonotype size between COVID-19 ($n = 73$) and healthy controls ($n = 75$) in all clonotypes (left) and in clonotypes with size < 50 (right). (**g**) Repertoire overlap according to clonotype size between moderate ($n = 9$) and severe ($n = 64$) group in all clonotypes (left) and in clonotypes with size < 50 (right). (**h**) The distribution of the clonal state of T cells which were suspected to be specific to SARS-CoV-2 based on CDR3 amino acid sequence in each cluster according to disease status ($n = 75$ healthy, $n = 9$ moderate, $n = 64$ severe).

https://doi.org/10.1038/s41588-023-01375-1

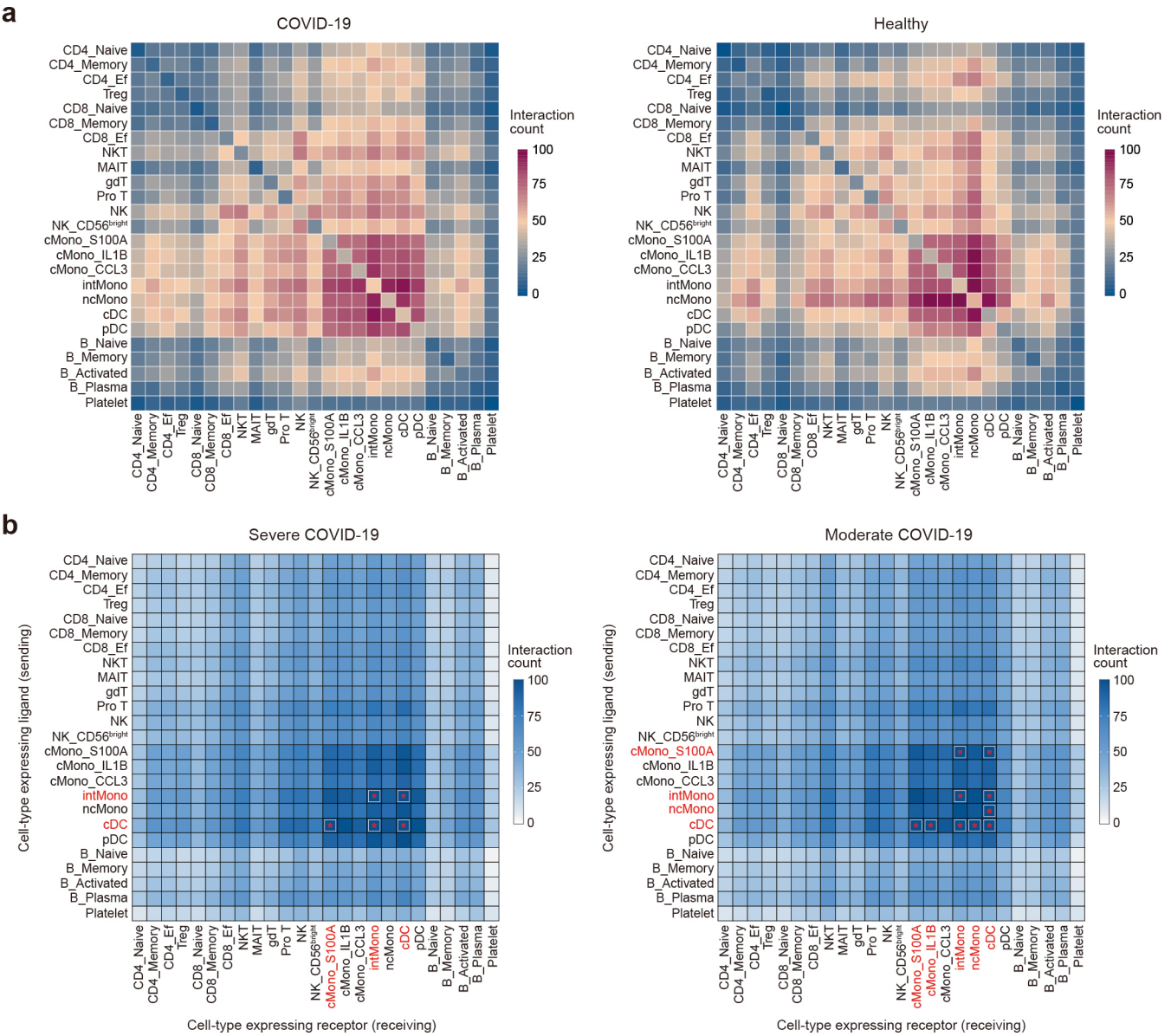


Extended Data Fig. 4 | See next page for caption.

https://doi.org/10.1038/s41588-023-01375-1

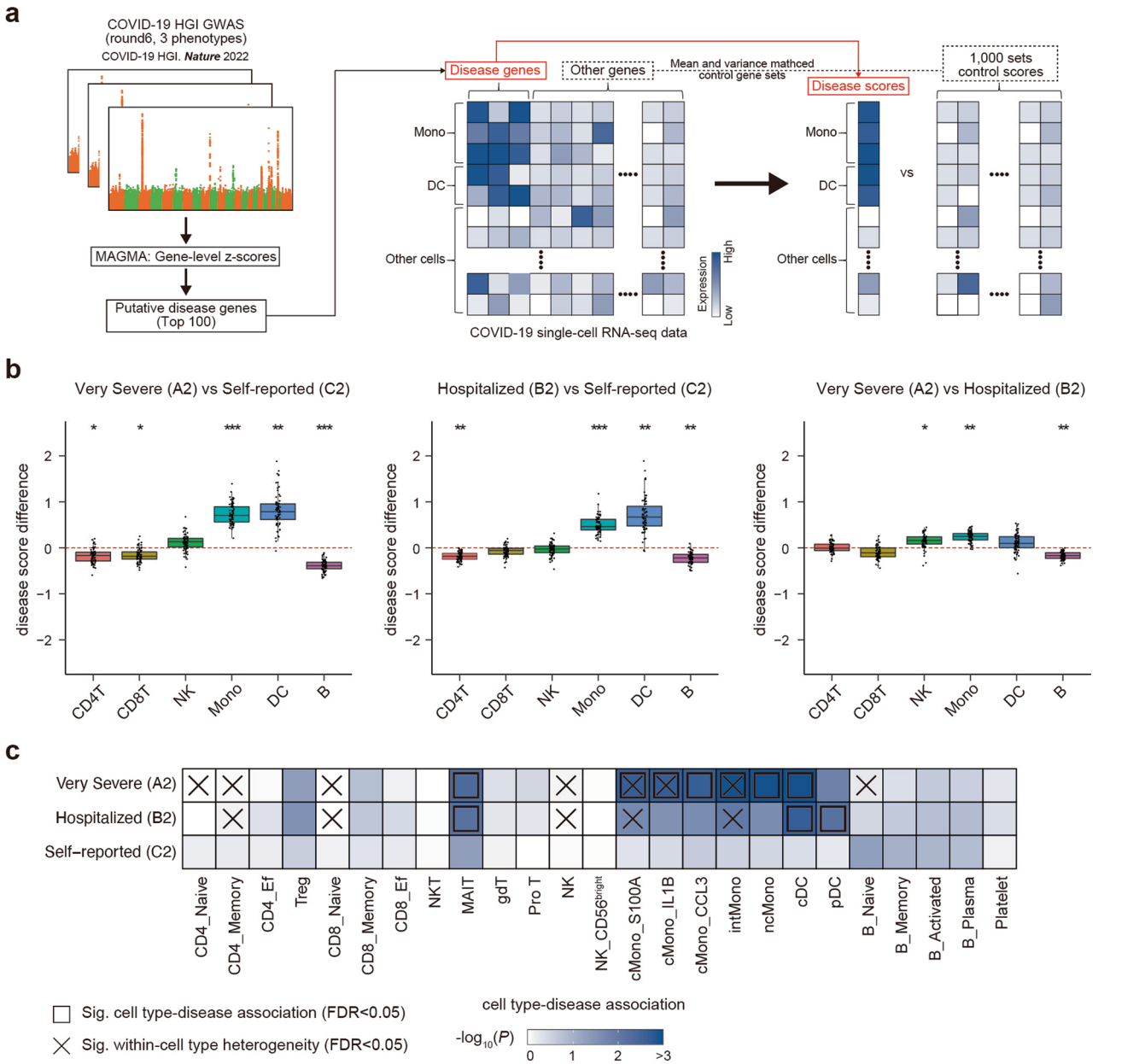
Extended Data Fig. 4 | Immunological features of B cells and BCR analysis. (**a**) Violin plots showing the expression distribution of selected canonical cell markers in four clusters. (**b**) A bar plot of the proportion of cell types, separated by three conditions ($n = 75$ healthy, $n = 9$ moderate, $n = 64$ severe). (**c**) Graph representation of neighborhoods identified by Milo in COVID-19 patients. Nodes are neighborhoods, colored by their $\log_2$ fold change between moderate ($n = 8$) and severe ($n = 64$) COVID-19 patients adjusted by age, sex, time since symptom onset and duration of systemic steroids treatment. Nhood, neighborhood. (**d**) Box plot showing the distribution of adjusted $\log_2$ fold change in abundance between moderate and severe COVID-19 in neighborhoods according to four cell types. Boxes denote the interquartile range (IQR), and the median is shown as horizontal bars. Whiskers extend to 1.5 times the IQR, and outliers are shown as individual points. (**e**) The top 20 enriched biological processes by GO analysis of upregulated DEGs of COVID-19 ($n = 72$) compared to healthy controls ($n = 75$) in four cell types. Dot color indicates the statistical significance of the enrichment (adjusted *P*-values via the Benjamini-Hochberg method), and dot size represents gene ratio annotated to each term. (**f**) The top 10 enriched biological processes by GO analysis of upregulated DEGs of moderate ($n = 8$) and severe ($n = 64$) compared to healthy group ($n = 75$). Dot color and dot size mean the same as in (**e**). (**g**) Repertoire overlap according to clonotype size between COVID-19 ($n = 73$) and healthy controls ($n = 75$) in all clonotypes (left) and in clonotypes with size < 10 (right). (**h**) Repertoire overlap according to clonotype size between moderate ($n = 9$) and severe ($n = 64$) group in all clonotypes (left) and in clonotypes with size < 10 (right).

https://doi.org/10.1038/s41588-023-01375-1



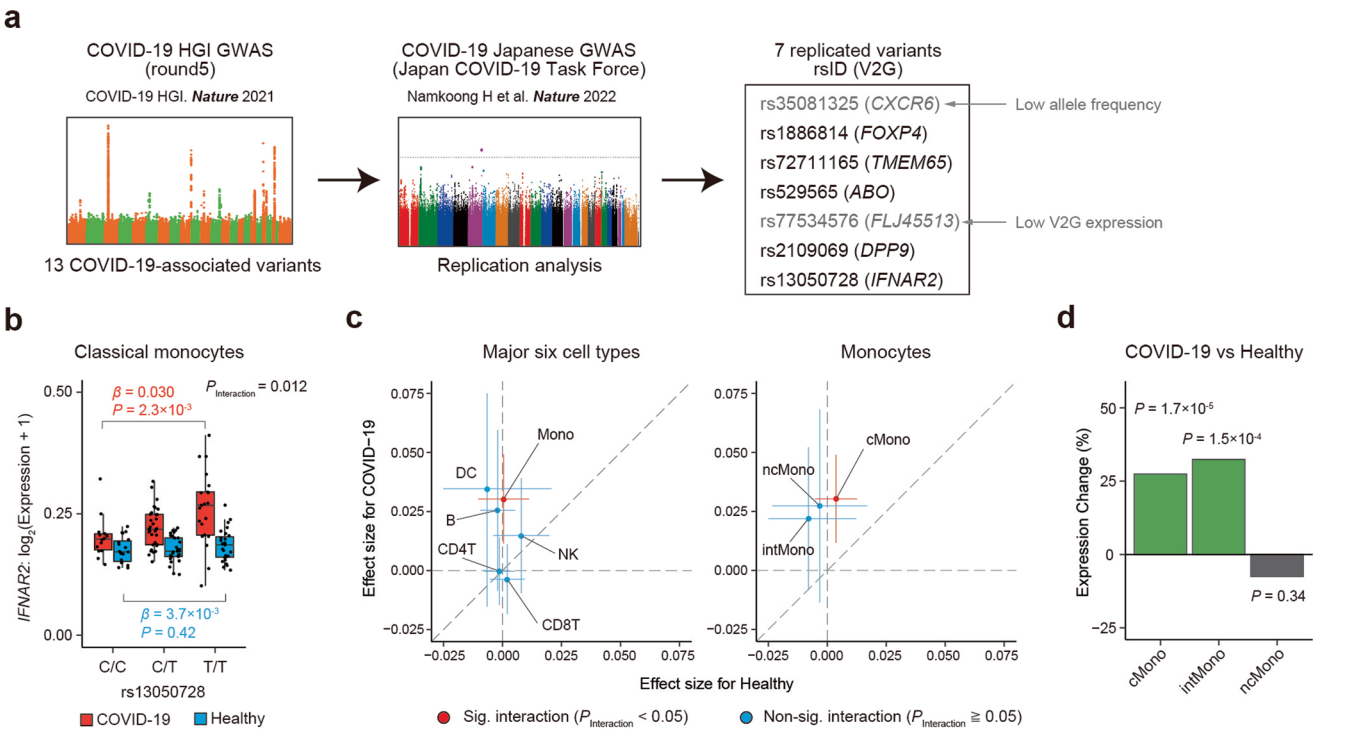
Extended Data Fig. 5 | Cell-cell interaction analysis using CellPhoneDB and NATMI. (**a**) Heatmaps depicting cell-cell communications quantified by CellPhoneDB in COVID-19 ($n = 73$) and healthy controls ($n = 75$), respectively. (**b**) Heatmaps depicting cell-cell communications quantified by NATMI in moderate ($n = 9$) and severe ($n = 64$) group, respectively. Rows indicate cells expressing the ligands and columns indicate cells expressing the receptors. Asterisk indicates cell-cell interactions with a number of ligand-receptor more than 100, and cell types which share more than 100 cell-cell interactions with at least one cell type are highlighted in red.

https://doi.org/10.1038/s41588-023-01375-1



Extended Data Fig. 6 | Assessing the association of PBMC cell types with host genetic risk. (**a**) Flowchart of integrated analysis of scRNA-seq data and COVID-19 HGI (round6) GWAS summary statistics using scDRS. The figure design is based on the original paper of scDRS[41]. (**b**) Differences of scDRS disease score between COVID-19 phenotypes in six major cell types. Differences of average disease scores of each sample were evaluated using two-sided paired *t*-test (adjusted for multiple comparisons using Bonferroni's correction across all pairs of six cell-types and three GWAS comparisons, $*P < 1\times10^{-10}$, $**P < 1\times10^{-20}$, $***P < 1\times10^{-30}$). Boxes denote the interquartile range (IQR), and the median is shown as horizontal bars. Whiskers extend to 1.5 times the IQR, and all COVID-19 patients ($n = 72$) are shown as individual points. (**c**) Heatmaps depicting cell type-disease association for three phenotypes in the manually annotated 25 clusters of COVID-19 scRNA-seq dataset. Heatmap colors denote uncorrected *P*-value of cell type-disease association evaluated using scDRS. Squares denote significant cell type-disease associations (FDR < 0.05), and cross symbols denote significant heterogeneity in association with disease across individual cells within a given cell type (FDR < 0.05). FDR was calculated via the Benjamini-Hochberg method across all pairs of 25 cell types and three GWAS phenotypes.

https://doi.org/10.1038/s41588-023-01375-1



Extended Data Fig. 7 | Single-cell eQTL analysis of rs13050728 and differential expression analysis of *IFNAR2*. (**a**) Flowchart of how to select COVID-19-associated variants in single-cell eQTL analysis. (**b**) rs13050728 eQTL for *IFNAR2* in classical monocytes. The box plot shows the eQTL in COVID-19 (red) and healthy controls (blue). Boxes denote the interquartile range (IQR), and the median is shown as horizontal bars. Whiskers extend to 1.5 times the IQR. All samples are shown as individual points ($n = 67$ COVID-19, $n = 75$ healthy). (**c**) Co-plots of the eQTL effect sizes of rs13050728 and 95% confidence intervals between COVID-19 and healthy controls. Dots represent the effect sizes and bars represent the 95 % confidence intervals. Cell types with significant interaction ($P_{Interaction} < 0.05$) in eQTL effect between COVID-19 and healthy controls are colored in red, and the rest are colored in blue. (**d**) Expression changes of *IFNAR2* in differential expression (DE) analysis between COVID-19 ($n = 72$) vs healthy controls ($n = 75$) are indicated for each cell cluster. Cell types with $P < 0.05$ in DE analysis are colored in green, and the rest is colored in gray. *P*-values are uncorrected and reflect two-sided tests in (**b**-**d**).

nature portfolio

Corresponding author(s): Yukinori Okada, Atsushi Kumanogoh

Last updated by author(s): Mar 9, 2023

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	No software was used in data collection.
Data analysis	Droplet libraries were processed using Cell Ranger 5.0.0 (10x Genomics). Sequencing reads were aligned with STAR (v2.7.2a) using the GRCh38 human reference genome. Filtered expression matrices generated using 'cellranger_count' were used to perform the analysis. Putative doublets were removed using Scrublet (v0.2.1) and scds (v1.10.0). Seurat (v4.1.0) was used for data scaling, transformation, clustering, dimensionality reduction and most visualization. Integration and batch correction was performed using harmony (v0.1). The final manual annotation of PBMCs was compared to the PBMC annotation of Azimuth. Differential abundance analysis was performed using Milo (v1.2.0).Differential gene expression testing was performed using edgeR (v3.32.0), and pathway enrichment analysis was performed using Clusterprofiler (v3.14.3). Spliced and unspliced transcripts were quantified using dropEst (v0.8.6). RNA velocity analysis was performed using scVelo (v0.2.3). Droplet-based sequencing data for TCR sequences and BCR sequences were analyzed using Scirpy (v0.10.0). Cell-cell interaction analysis among PBMC was performed using CellPhoneDB (v2.0.0) and NATMI (https://github.com/forrest-lab/NATMI). Gene-level association p-values from GWAS summary statistics were computed using MAGMA (v1.07), and integrating the information from scRNA-seq data with polygenic signals from COVID-19 GWAS was performed using scDRS (v1.0.1). Single-cell-level normalization was performed using scran (v1.18.5) before pseudo-bulk single-cell eQTL analysis. R statistical software (version 4.0.2) was used for the eQTL analysis of COVID-19-associated variants. SHAPEIT4 (v4.2.1) was used for haplotype phasing of genotype data. After phasing, Minimac4 (v1.0.1) was used for genome-wide genotype imputation. The code is available at https://github.com/REdahiro/JPN_COVID-19_scRNAseq.

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

Data

Policy information about availability of data

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

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

Raw sequencing data of scRNA-seq are available at the Japanese Genotype-phenotype Archive (JGA) with accession codes JGAS000593 (https://ddbj.nig.ac.jp/resource/jga-study/JGAS000593) /JGAD000722 (https://ddbj.nig.ac.jp/resource/jga-dataset/JGAD000722). Part of the raw scRNA-seq data (nCOVID-19 = 30, nControl = 31) has already been deposited and are available under controlled access at JGA with accession codes JGAS000543 (https://ddbj.nig.ac.jp/resource/jga-study/JGAS000543) /JGAD000662 (https://ddbj.nig.ac.jp/resource/jga-dataset/JGAD000662). All the raw sequencing data of scRNA-seq can also be accessed through application at the NBDC with the accession code hum0197 (https://humandbs.biosciencedbc.jp/en/hum0197-latest). Genotype data of the subjects are available at European Genome-Phenome Archive (EGA) with the accession code EGAS00001006950 (https://ega-archive.org/studies/EGAS00001006950). Raw sequencing data of scRNA-seq and genotype data are potentially identifiable and therefore under controlled access at JGA and EGA. The GWAS summary statistics of COVID-19 HGI (release 6) were obtained from https://www.covid19hg.org/results/r6/. The reference for cell type annotation of PBMC in scRNA-seq (pbmc_multimodal.h5seurat) was obtained from https://satijalab.org/seurat/articles/multimodal_reference_mapping.html.

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size	The sample size was not predetermined and all samples that were available were processed. We recruited 73 COVID-19 cases and 75 healthy controls.
Data exclusions	Low quality cells: Cells that had fewer than 1st percentile of UMIs or greater than 99th percentile of UMIs in each sample, as well as cells that contained greater than 10% of reads from mitochondrial genes or Hemoglobin genes, were considered low quality and removed from further analysis. Doublets: Cells labeled as doublets by Scrublet or scds were excluded from the analysis.
Replication	All findings were based on statistical analysis of a large patient cohort. There was no replication cohort.
Randomization	We investigated two major groups; "COVID-19 patients" and "Healthy controls". COVID-19 patients were further categorized into groups of moderate and severe according to disease severity based on WHO guidelines. Randomization was not applicable to this study.
Blinding	We did not apply blinding of the samples because no intervention was conducted in our study.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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

Human research participants

Policy information about studies involving human research participants

Population characteristics	COVID-19 patients (n = 73) are of Ease Asian ancestry, the median age was 68, 71.2% were male, and all of them were tested positive for PCR test. COVID-19 patients were further categorized into groups of moderate (n = 9) and severe (n = 64) according to disease severity based on WHO guidelines. Healthy controls (n = 75) are of Ease Asian ancestry, the median age was 34, 64.0% were male. Detailed patients characteristics are summarized in Supplementary Table 1.
Recruitment	We recruited the hospitalized cases diagnosed as COVID-19 by physicians using the clinical manifestation and PCR test results. We included all COVID-19 cases treated at Osaka University from July 2020 to September 2021 who gave consent for the study. Thus, there is little selection bias. All control participants were recruited at Osaka University.
Ethics oversight	This study was approved by the ethical committee of Osaka University Graduate School of Medicine.

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