Integrated analysis of plasma and single immune cells uncovers metabolic changes in individuals with COVID-19

Jihoon W. Lee [1,15], Yapeng Su [1,2,3,15 ✉], Priyanka Baloni [2], Daniel Chen [2], Ana Jimena Pavlovitch-Bedzyk[4], Dan Yuan[2], Venkata R. Duvvuri [2], Rachel H. Ng [2], Jongchan Choi [2], Jingyi Xie[2], Rongyu Zhang[2], Kim Murray[2], Sergey Kornilov[2], Brett Smith[2], Andrew T. Magis[2], Dave S. B. Hoon [5], Jennifer J. Hadlock [2], Jason D. Goldman[6,7,8], Nathan D. Price [2,9], Raphael Gottardo [3,10,11], Mark M. Davis [4,12,13], Leroy Hood[2,7], Philip D. Greenberg [1,14 ✉] and James R. Heath [2,9 ✉]

A better understanding of the metabolic alterations in immune cells during severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection may elucidate the wide diversity of clinical symptoms experienced by individuals with coronavirus disease 2019 (COVID-19). Here, we report the metabolic changes associated with the peripheral immune response of 198 individuals with COVID-19 through an integrated analysis of plasma metabolite and protein levels as well as single-cell multiomics analyses from serial blood draws collected during the first week after clinical diagnosis. We document the emergence of rare but metabolically dominant T cell subpopulations and find that increasing disease severity correlates with a bifurcation of monocytes into two metabolically distinct subsets. This integrated analysis reveals a robust interplay between plasma metabolites and cell-type-specific metabolic reprogramming networks that is associated with disease severity and could predict survival.

The COVID-19 pandemic has accounted for more than four million fatalities. Preventive and therapeutic measures of varying effectiveness have been deployed against COVID-19, but much remains to be learned about precisely how COVID-19 and other viral diseases induce and modulate immune responses in infected hosts. A detailed understanding of the pathogenesis of and immune response to COVID-19 and other serious viral infections is central for developing more effective treatments.

The metabolic needs of immune responses are intimately linked to their function[1]. Indeed, the metabolic status of individuals with COVID-19 significantly affects their prognosis[2]. To examine this, recent studies have gathered proteomic, transcriptomic and metabolomic measurements from individuals with COVID-19 (refs. [3,4]) and identified many associations between individual molecular changes and disease outcome. However, an integrated perspective of metabolic changes that may be critical to regulating immune function in COVID-19 and other acute viral diseases is lacking. Here, we analyze plasma metabolomics/proteomics and single-cell integrated transcriptomic/proteomic datasets collected from blood draws from 198 individuals with COVID-19 representing the full range of infection severities[4]. We find metabolic reprogramming highly specific to individual immune cell classes, and this reprogramming

is associated with changes in the plasma metabolome and with disease severity. We further identify circulating metabolite classifiers of disease severity and predictors of clinical outcomes. This integrated approach paves the way for understanding the metabolic mechanisms underlying the immune response against COVID-19.

Results

Multiomics analysis of the peripheral immune response. We evaluated plasma metabolomic profiles and transcriptional networks within circulating immune cells identified through single-cell RNA-seq (scRNA-seq) analysis (Methods). These profiles were collected from 374 blood samples from 198 individuals with COVID-19 with varying disease severities[4] (Fig. 1a and Supplementary Table 1; severity measured on the World Health Organization (WHO) ordinal scale for COVID-19, Supplementary Table 5). To define the early trajectory of acute illness, for each individual, data were collected near clinical diagnosis (T1) and several days later (T2) when individuals were still symptomatic in the acute phase.

Plasma metabolites distinguish varying disease severities. We first examined plasma metabolites to obtain a broad view of metabolic status. Of the 1,050 metabolites measured per individual,
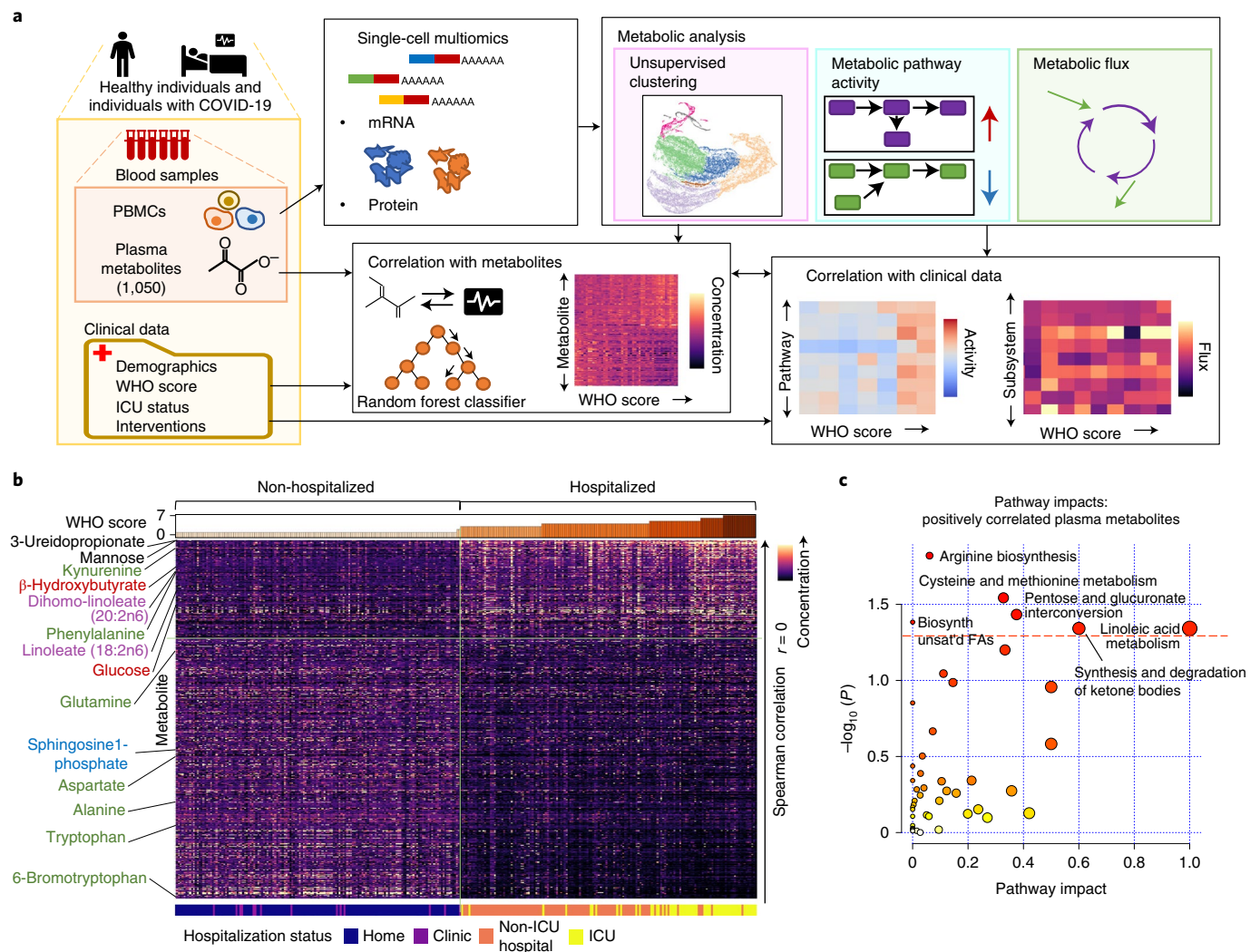
Fig. 1 | Overview of metabolic analysis of COVID-19 peripheral immune response. a, Overall workflow of metabolic analysis; PBMCs, peripheral blood mononuclear cells; ICU, intensive care unit; WHO, World Health Organization. **b**, Plasma metabolite levels per individual sample correlated with severity on the WHO ordinal scale per sample. Metabolites are listed in descending order of two-sided Spearman's rank correlation value. Metabolites of particular interest are labeled and colored by category; red, metabolites associated with canonical energy-producing metabolic pathways; green, amino acids and products of amino acid metabolism; purple, linoleic acid-containing lipids; blue, sphingolipids. The solid vertical line intersecting the heat map indicates separation between non-hospitalized versus hospitalized individuals, and the horizontal line indicates separation between metabolite levels positively and negatively correlated with WHO score. Only the most significantly correlated metabolites are shown (*P* < 0.001 with Benjamini–Hochberg false discovery rate (FDR) correction). Significant *P* values are summarized in Supplementary Table 2. **c**, Known metabolic pathways affected by plasma metabolites that are positively correlated with WHO score. The red dashed line indicates threshold of statistical significance on the *y* axis (*P* < 0.05) calculated by hypergeometric test. Exact *P* values are listed in Supplementary Table 7. Biosynth, biosynthesis; unsat'd, unsaturated; FAs, fatty acids.

227 and 510 were significantly (*P* < 0.05) positively and negatively correlated with disease severity, respectively (Fig. 1b and Supplementary Table 2). Phenylalanine, which is a mortality risk marker in severe infection[5], phenylalanine-related metabolites (for example, *N*-formylphenylalanine, also associated with rhinovirus[6]) and previously reported kynurenine[3] were each positively correlated with disease severity. Mannose also strongly correlated with severity and may be derived from oligomannose residues of SARS-CoV-2 spike protein[7], potentially reflecting elevated viral loads, or may indicate complement pathway activation[8]. Also positively correlated with disease severity were glucose, supporting the importance of glycolytic/gluconeogenic processes and anaerobic metabolism in COVID-19 (ref. [9]), and inflammation-associated lipids, such as linoleic acid[10]. Thus, increased COVID-19 severity is associated with metabolite alterations, suggesting increased immune-related activity.

We next performed metabolic pathway enrichment analysis of metabolites that correlated with severity. The positively correlated metabolites were enriched for metabolism of linoleic acid and ketone bodies (Fig. 1c). Ketone bodies increased with severity, potentially indicating a developing catabolic state. Metabolites that negatively correlated with severity were enriched for metabolism of sphingolipids and glycerophospholipids (Extended Data Fig. 1a), components of lipid membranes critical for cytokine synthesis[11] and stability against hypoxic stress[12]. Within the alanine, aspartate and glutamate metabolic pathway, glutamine is negatively correlated with disease severity, suggesting increased uptake that sustains proliferation of adaptive immune cells[13]. Thus, altered circulating metabolite levels may reflect altered metabolic processes generating components of a COVID-19 immune response.

We next tested whether plasma metabolite levels could distinguish disease severity between individuals by principal component analysis (PCA). A distinct separation of individuals with mild COVID-19 from individuals with more severe disease was evident along PC1 (Extended Data Fig. 1b), which comprises mannose, linoleic acid-containing glycerophospholipids and sphingomyelins as top contributors.

Taken together, COVID-19 severity is associated with many changes in plasma metabolite levels, which can be ascribed to distinct metabolic pathways that potentially impact immune responses. We next turn to integrating these disease-altered metabolite profiles with other omics measures.

Metabolism by immune cells at whole-PBMC resolution. We evaluated peripheral immune cell metabolism using single-cell gene expression-based inferences (Extended Data Fig. 1c) and metabolic flux. We considered all PBMCs together and identified metabolic pathway activities that correlated with disease severity (Fig. 2a). For example, we identified that glutathione metabolism within immune cells positively correlated with disease severity, suggesting that cellular metabolism may be compensating for increasing levels of oxidative stress to maintain immune cell self-renewal capacity[14]. Catabolism of arachidonic acid, a critical precursor to prostaglandins and leukotrienes[10], negatively correlated with disease severity, suggesting decreased inflammatory cytokine production and thus decreasingly effective immune responses. This analysis suggests that metabolic reprogramming of immune cells accompanies COVID-19.

Immune cell types are metabolically compartmentalized. To assess metabolic networks within individual PBMCs, we performed uniform manifold approximation and projection (UMAP) based on the 1,387 genes involved in known metabolic pathways (classical immune markers were excluded). Remarkably, this resolved major immune cell types, similar to an analysis using all genes from our scRNA-seq dataset (Fig. 2b). This indicates that each major immune cell type has a distinct metabolic signature.

We further explored immune cell-type-specific gene expression-based metabolic pathway activity[15] and metabolic flux by extracting pathway activity changes with respect to disease progression from T1 to T2. Individuals with increased disease severity at T2 showed a generally more activated metabolic profile, but the extent of activation was distinct per cell type (Fig. 2c). CD8+ T cells and B cells had the most increased metabolic activity, consistent with their major antiviral activities, while CD4+ T cells, natural killer (NK) cells and monocytes each exhibited relatively less elevation. Among clinically improved individuals, B cells had the greatest decrease in metabolic activity, potentially reflecting decreased activity with reduced antigen load, while NK cell activity slightly increased. Thus, examining metabolic activity at the resolution of individual cell types reveals insights not apparent from whole-PBMC analysis.

We further performed metabolic flux analysis[16] by integrating scRNA-seq data with genome-scale reconstruction of human metabolism (Methods). Reaction fluxes for each immune cell type per individual yielded three cell clusters (Extended Data Fig. 1d,e). Cluster 2 exclusively contained monocytes, suggesting distinct metabolic programs between the myeloid (monocytes) and lymphoid (T, B and NK cells) lineages. This cluster exhibited significantly lower nucleotide synthesis and increased eicosanoid metabolism (mediating inflammation[17]) and glycolysis/gluconeogenesis, suggesting reduced proliferation but sustained proinflammatory activity. Cluster 0 was composed of NK cells, B cells and CD4+ and CD8+ T cells (Extended Data Fig. 1e). Cluster 1, containing mostly CD4+ and CD8+ T cells, was metabolically similar to cluster 0 but had significantly higher extracellular transport (that is, *SLC7A6*-mediated amino acid transport) and nucleotide interconversion (Supplementary Fig. 1a), potentially indicating

increased intercellular signaling and cell proliferation, respectively. Thus, flux-defined clusters 0 and 1 may contain naive and activated lymphocytes, respectively. By contrast, when we calculated reaction fluxes at the whole-PBMC level rather than at the cell-type level, no discrete clusters formed (Supplementary Fig. 1b). This highlights major metabolic differences between immune cell types (and potential metabolic subtypes within each cell type) and underscores the importance of analyzing immune cell types individually. The relationship of these subtypes to immune functions is a key question.

Metabolically hyperactive CD8+ T cell subpopulation emerges. We further investigated the metabolic heterogeneity within individual cell types at single-cell resolution starting with CD8+ T cells. Clustering and UMAP projections with metabolic genes alone separated CD8+ T cells into seven metabolically defined clusters (Fig. 3a and Extended Data Fig. 2a). These metabolic clusters were consistent with clusters previously identified[4] using the expression levels of all measured genes (Fig. 3b) and so were phenotypically distinct as defined by classical phenotype markers (Fig. 3c and Extended Data Fig. 2b–e). As examples, metabolic clusters 0, 4 and 5 highly express effector genes *GZMB*, *GNLY* and *PRF1*; thus, we labeled them as effector-like cells. Cluster 4, while of low abundance, notably increases with disease severity (Fig. 3d).

We examined CD8+ T cell metabolic pathways with cluster-specific trends (Fig. 3e). For the effector-like metabolic clusters 0, 4 and 5, we resolved that almost all metabolic pathways positively correlated with WHO score. By contrast, relatively few metabolic pathways correlated when analyzing all CD8+ T cells pooled together. The energy-producing pathways of glycolysis/gluconeogenesis, oxidative phosphorylation and fatty acid degradation and biosynthesis were activated among these effector-like cells (especially cluster 4) but not among all CD8+ T cells. (Fig. 3f and Supplementary Table 6). Indeed, metabolic cluster 4 was by far the most active relative to all other clusters (Fig. 3g and Extended Data Fig. 2f) and so was defined as metabolically dominant. This was corroborated by flow cytometry, where metabolic cluster 4 resembled Ki-67+ proliferative-exhausted cells and exhibited significantly elevated hexokinase 2 (HK2) protein levels compared to other CD8+ T cells (Extended Data Fig. 3a–f). While all effector-like clusters 0, 4 and 5 expressed *PRF1* and *PRDM1*, indicating effector differentiation[18], only metabolic cluster 4 had high expression of *MKI67* and low expression of *B3GAT1* (encoding CD57 protein), suggesting minimal replicative senescence and antigen-induced T cell apoptosis[19] (Fig. 3h) despite also expressing high levels of PD-1. Comparing T cell antigen receptor (TCR) sequences from this hyperactive cluster 4 with previously published SARS-CoV-2-specific TCR sequences (Fig. 3i) suggested that this cluster is SARS-CoV-2 specific. Thus, metabolic effector-like clusters 0 and 5 were relatively terminally differentiated, while cluster 4 comprised proliferative-exhausted, effector-like cells that were likely SARS-CoV-2 specific. In summary, CD8+ T cells include a small, metabolically hyperactive subpopulation, presumably SARS-CoV-2 specific, that exhibits increasing metabolic activity as disease severity increases (Extended Data Fig. 3g).

We interrogated for similar, metabolically dominant CD8+ T cells in sepsis[20] and human immunodeficiency virus (HIV)[21] using public scRNA-seq data (Supplementary Table 1). The metabolically hyperactive, proliferative-exhausted subpopulation comprises a small fraction of CD8+ T cells in both COVID-19 and sepsis but constitutes a larger, albeit less, metabolically active fraction in HIV (Extended Data Fig. 4a,b). This is consistent with the dysfunctional state of expanded CD8+ T cells in HIV[22]. Differential expression analysis revealed 151 uniquely highly expressed metabolic genes in this subpopulation (Extended Data Fig. 4c and Supplementary Table 3), with elevated processes related to amino acid metabolism and protein synthesis and targeting (Extended Data Fig. 4d).
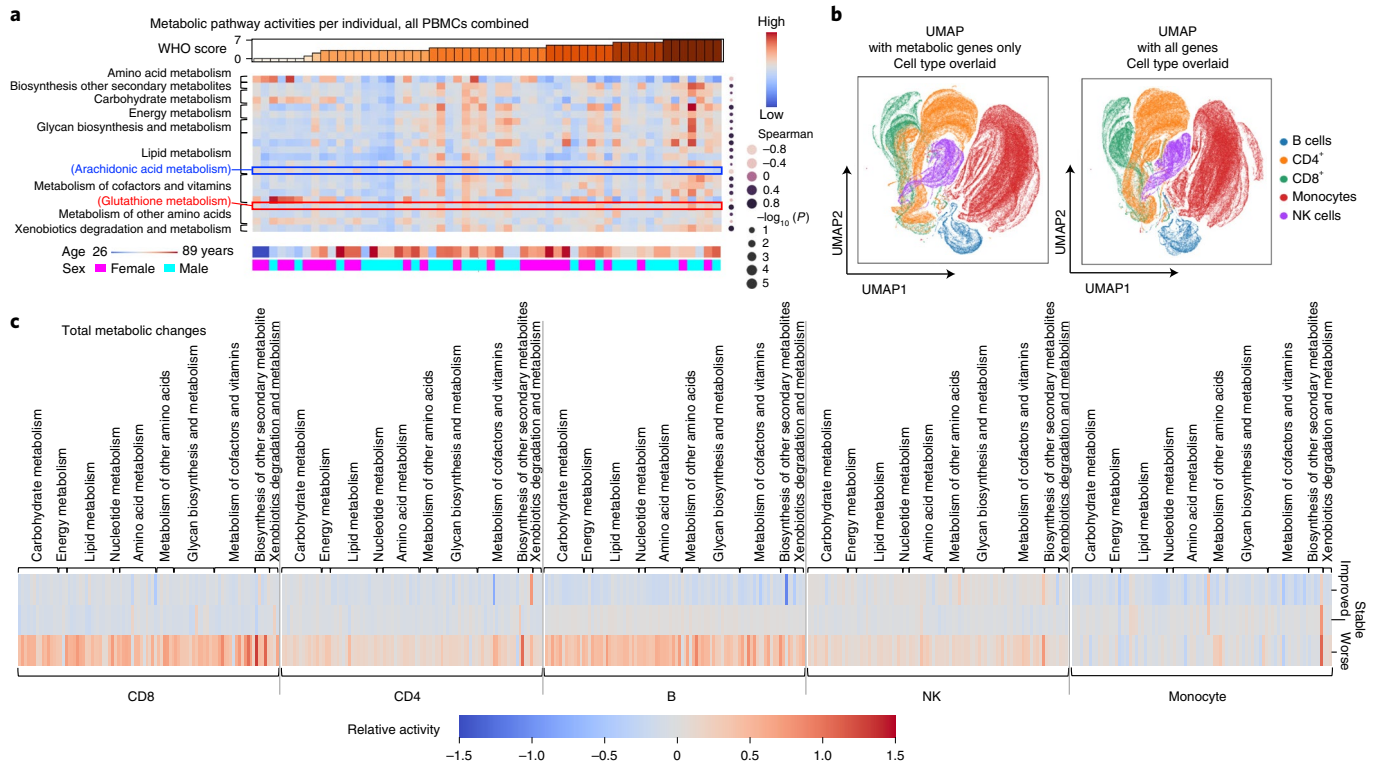
Fig. 2 | Total metabolic changes with changes in disease severity from T1 to T2. a, Metabolic pathway activities for all PBMCs pooled together that are significantly correlated with WHO score by two-sided Spearman's rank correlation ($P < 0.05$; labeled on right). Exact P values are listed in Supplementary Table 7. **b**, UMAP for dimensional reduction of scRNA-seq data of peripheral immune cells collected from each individual. UMAP was performed separately using only genes included in metabolic pathways (left) and using all genes (right). **c**, Changes in metabolic pathway activities over time (from T1 to T2) per peripheral immune cell type for individuals whose WHO scores worsened, improved or remained stable. A relative activity of 0 indicates no change in the activity of a metabolic pathway from T1 to T2.

Metabolic balance between CD4+ T cell subpopulations. CD4+ T cells also exhibited marked metabolic heterogeneity with eight metabolically defined clusters (Extended Data Fig. 5a) associated with distinct phenotypes. Similar to metabolic cluster 4 CD8+ T cells, metabolic cluster 6 CD4+ T cells were the most metabolically active in severe disease (Extended Data Fig. 5d–g) and exhibited intermediate expression of metabolic regulators *MYC*, *HIF1A*, *MAPK1* and *PRKAA1* as well as proliferative markers *MKI67* and *TYMS* (Extended Data Fig. 5a–d). A population of cytotoxic CD4+ T cells, metabolic cluster 4, exhibited a high degree of clonal expansion (Extended Data Fig. 5c), intermediately expressed *HIF1A*, *MAPK1* and *PRKAA1* and highly expressed cytolytic markers *GZMB*, *GNLY* and *PRF1* akin to effector CD8+ T cells.

We examined metabolic trends among CD4+ T cell subpopulations (Extended Data Fig. 5g). Proliferative cluster 6 exhibited increased activity compared to the pooled population of CD4+ T cells, including the core energy-producing pathways of glycolysis/gluconeogenesis, oxidative phosphorylation and fatty acid biosynthesis and degradation (Extended Data Fig. 5e–g and Supplementary Table 6). Meanwhile, the cytolytic cluster 4 remained metabolically stable with increased disease severity, despite increasing in percentages of the overall CD4+ population. Regulatory T cells (T$_{reg}$ cells), within metabolic cluster 1, exhibited mildly increased metabolism with increased severity, potentially countering increased proinflammatory activity. Overall, CD4+ T cells exhibit a similar pattern as CD8+ T cells in which a small subpopulation becomes metabolically dominant with increased activity with disease severity. Thus, the average increase in CD4+ T cell metabolic activity in more severe disease may be driven by this hyperactive subpopulation. On the contrary, the cytolytic CD4+ T cells, which are the most clonally

expanded, increase as a fraction of the CD4+ T cell population but are relatively stable in per cell metabolic activity. Thus, different subpopulations of CD4+ T cells exhibit divergent metabolic activities with increasing disease severity, an observation missed by bulk analyses.

Metabolic uniqueness and dominance of plasma B cells. Next, we examined B cells and found that metabolic clustering and UMAP projections isolate plasma B cells as a metabolically and phenotypically distinct cell cluster (Supplementary Fig. 2a–c,g), consistent with activated antibody-producing cells having a distinct metabolic program[23]. Compared to non-plasma B cells, this plasma cell cluster (metabolic cluster 3) exhibited globally increased metabolism (Supplementary Fig. 2d–e), including positive correlation of amino acid metabolic activity with disease severity (Supplementary Fig. 2f,h and Supplementary Table 6), perhaps reflecting the antibody-secreting functional demands in severe disease.

A metabolically dominant NK cell subpopulation emerges. NK cells were similarly resolved into six metabolic clusters (Supplementary Fig. 3a–c), among which cluster 4 was distinguished as a proliferative, metabolically active phenotype by high levels of *MKI67* and *HIF1A* and low levels of *B3GAT1* (encoding CD57 protein) (Supplementary Fig. 3d,g). Indeed, this proliferative cluster had increased metabolic activity in key pathways such as folate biosynthesis, allowing for proliferation[24], and sphingolipid metabolism, impacting immune regulation[25] (Supplementary Fig. 3e,h and Supplementary Table 6). The metabolic activities of these proliferative cells also positively correlated with disease severity (Supplementary Fig. 3f).
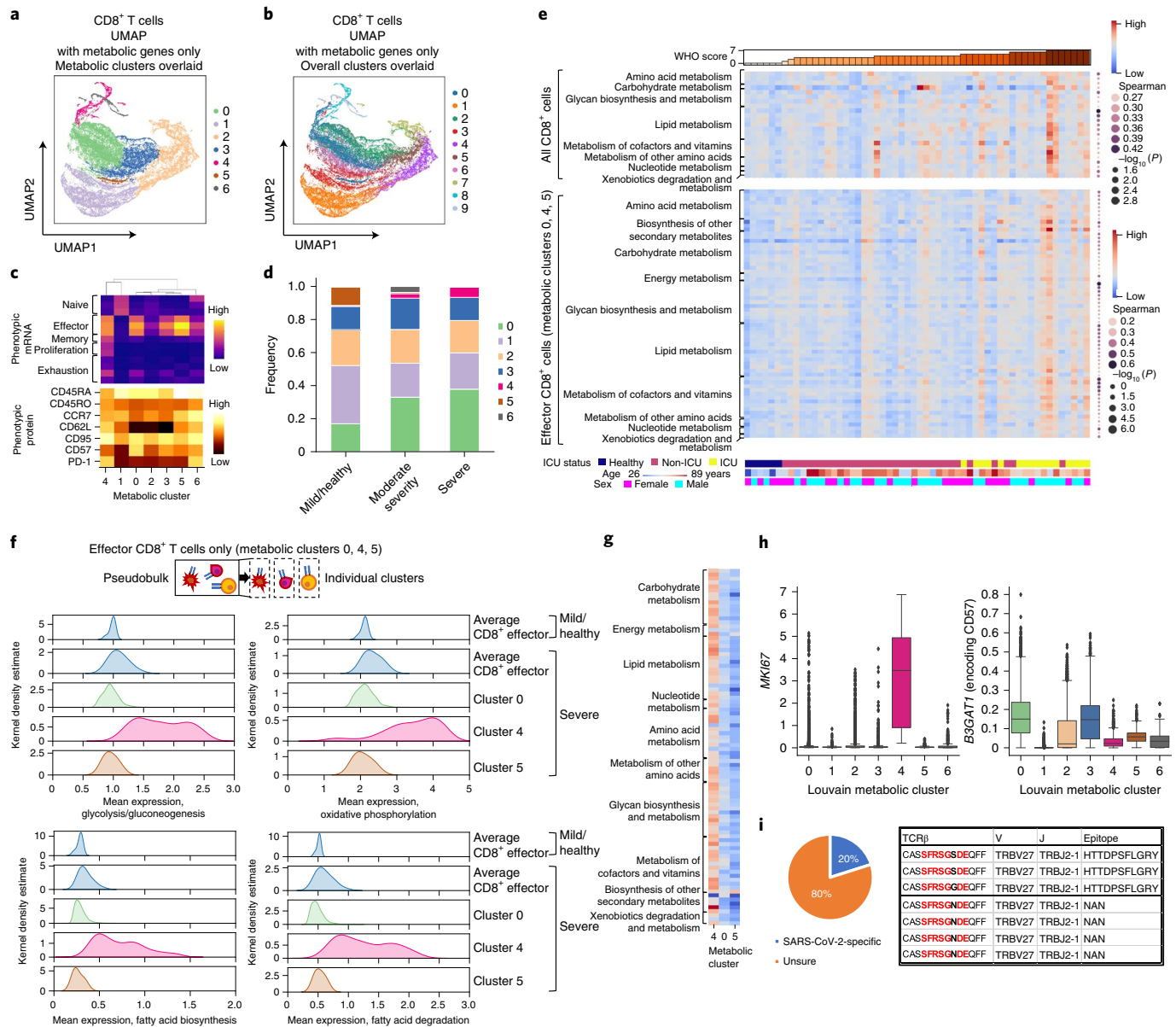
Fig. 3 | Heterogeneous metabolic activity of CD8[+] T cells is dominated by small, highly active subpopulations. a,b, Unsupervised clustering and UMAP of CD8[+] T cells using the expression of only genes involved in known metabolic pathways. Clusters defined from the expressions of only metabolic genes (**a**) and from the expressions of all genes (**b**) are overlaid. **c,** Heat map of mean expression of key phenotypic genes and proteins for each metabolic cluster in pseudobulk. **d,** Frequency of metabolically defined CD8[+] T cell clusters per individual grouped by WHO score-based categories of disease severity; mild/healthy (WHO 0–2), moderate disease (WHO 3–4) and severe disease (WHO 5–7). Statistical significances of population increases or decreases were calculated by two-sided Spearman's rank correlation. **e,** Metabolic pathway activities that are significantly correlated with WHO score by two-sided Spearman's rank correlation ($P < 0.05$; right) for all CD8[+] T cells per individual and for only effector CD8[+] T cells (metabolic clusters 0, 4 and 5), each in pseudobulk. Exact P values are in Supplementary Table 7. **f,** Kernel density estimates of expressions of genes included in key energy-producing metabolic pathways for effector CD8[+] T cells (metabolic clusters 0, 4 and 5) among individuals with mild disease and/or healthy individuals (WHO 0–2) and individuals with severe disease (WHO 5–7) in pseudobulk and separately for each of the metabolic clusters 0, 4 and 5. **g,** Comparison of metabolic pathway activities between CD8[+] T cell metabolic clusters 0, 4 and 5 for all samples combined. **h,** Expression of T cell proliferation marker *MKI67* and replicative senescence marker *B3GAT1* (encoding CD57) in each metabolically defined cluster of CD8[+] T cells (metabolic cluster 0, $n = 8,525$ cells; 1, $n = 5,983$; 2, $n = 5,270$; 3, $n = 4,230$; 4, $n = 1,008$; 5, $n = 475$; 6, $n = 475$). Boxes indicate 25th, 50th and 75th percentiles. Whiskers indicate 1.5 interquartile ranges below and above the 25th and 75th percentiles, respectively. **i,** Grouping of lymphocyte interactions by paratope hotspots (GLIPH) analysis of cluster 4 TCRs with published antigen-specific TCRs; left, pie chart illustrating fractions of TCRs from cluster 4 cells that are within the same GLIPH group with published SARS-CoV-2-specific TCRs; right, one representative GLIPH group that contains SARS-CoV-2-specific TCRs and cluster 4 cell TCRs.

Inflammatory and non-classical monocytes split metabolically. Monocyte dysregulation is prominently featured in COVID-19 (ref. [26]). In fact, we found several metabolically defined monocyte clusters (Fig. 4a and Extended Data Fig. 6a–d) that also separate

phenotypes (Fig. 4b). For instance, metabolic clusters 0 and 1, which are CD14[+]CD16[–]HLA-DR[mid] monocytes highly expressing M1 macrophage-associated metabolic regulator *STAT1* (ref. [27]) and *HIF1A* (ref. [28]), represent an inflammatory subpopulation
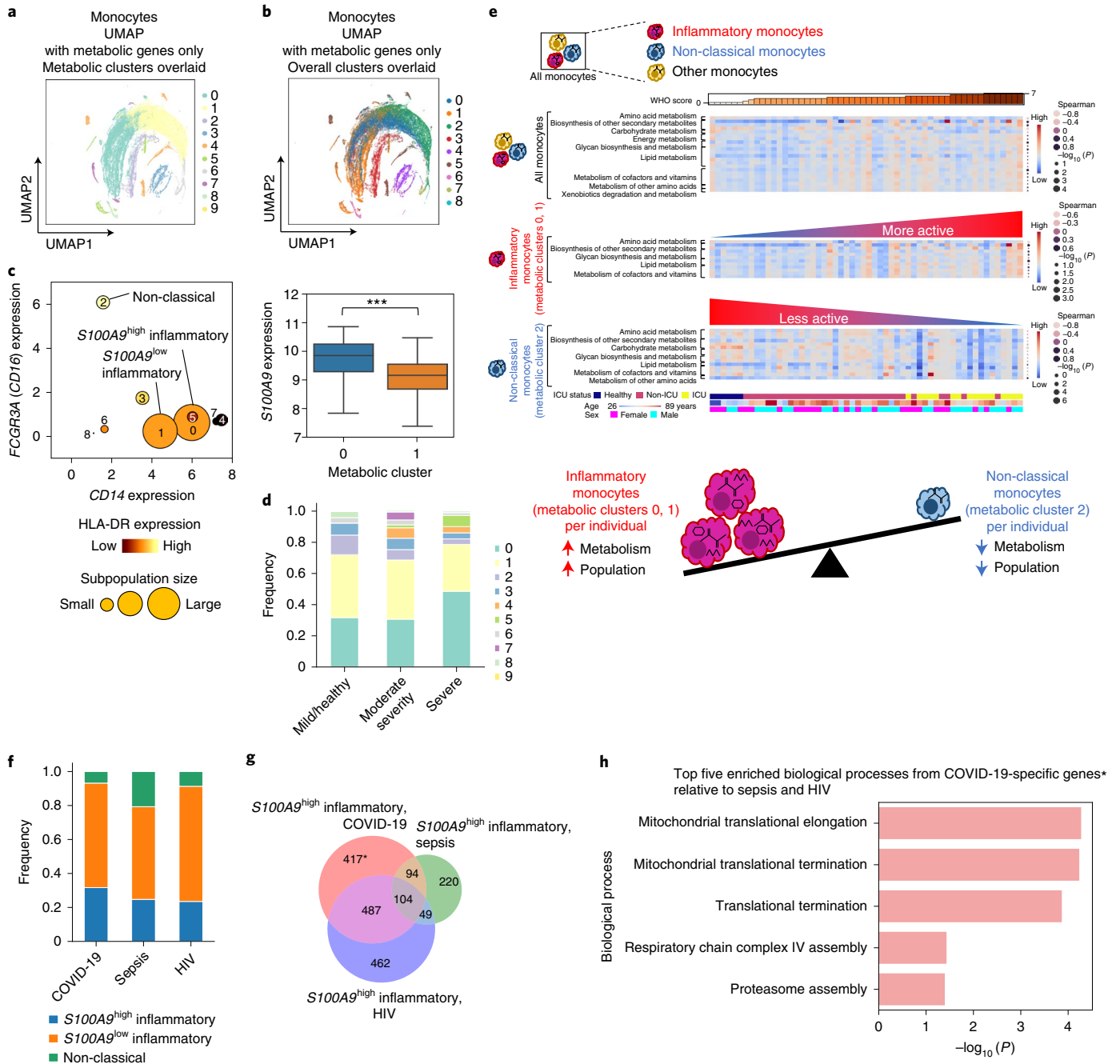
Fig. 4 | Inflammatory monocytes become metabolically dominant with increased COVID-19 severity. a,b, Unsupervised clustering and UMAP of monocytes using the expression of only genes involved in known metabolic pathways. Clusters defined from expressions of only metabolic genes (**a**) and from expressions of all genes (**b**) are depicted. **c,** Summary of markers for defining inflammatory and non-classical monocyte subpopulations. Left, *CD14*, *FCGR3A* and HLA-DR expression per metabolically defined monocyte cluster. Clusters are labeled by number. *S100A9*high and *S100A9*low inflammatory monocytes and non-classical monocytes are specifically labeled. Cluster 9 is not shown due to its lack of valid gene expression data. Right, comparison of *S100A9* expression between the two inflammatory monocyte subpopulations (metabolic cluster 0, $n = 36{,}651$ cells; 1, $n = 33{,}242$). Boxes indicate 25th, 50th and 75th percentiles. Whiskers indicate 1.5 interquartile ranges below and above the 25th and 75th percentiles, respectively. Statistical significance was calculated by two-sided Mann–Whitney U test; ***$P < 0.001$. **d,** Frequency of metabolically defined monocyte clusters per individual, grouped by WHO score-based disease severity categories: mild/healthy (WHO 0–2), moderate disease (WHO 3–4) and severe disease (WHO 5–7). **e,** Metabolic pathway activities that are significantly correlated with WHO score by two-sided Spearman's rank correlation ($P < 0.05$; right) for all monocytes per individual, for only inflammatory monocytes (metabolic clusters 0 and 1) and for only non-classical monocytes (metabolic cluster 2), each in pseudobulk. Exact P values are in Supplementary Table 7. **f,** Fraction of monocytes per disease (peripheral monocytes in COVID-19, sepsis and HIV) that are *S100A9*high inflammatory, *S100A9*low inflammatory or non-classical. **g,** Venn diagram of metabolic pathway-related genes that are uniquely elevated in *S100A9*high inflammatory monocytes of each disease. **h,** Top five most highly enriched biological processes by gene ontology for metabolic genes uniquely elevated in *S100A9*high inflammatory monocytes of COVID-19 relative to sepsis and HIV. The x axis indicates $-\log_{10}$ (P value) for enrichment of each process. None of the displayed processes were significantly enriched in the sepsis or HIV conditions. Asterisks in **g** and **h** indicate the relation of the 417 genes in **g** to the biological processes in **h**.
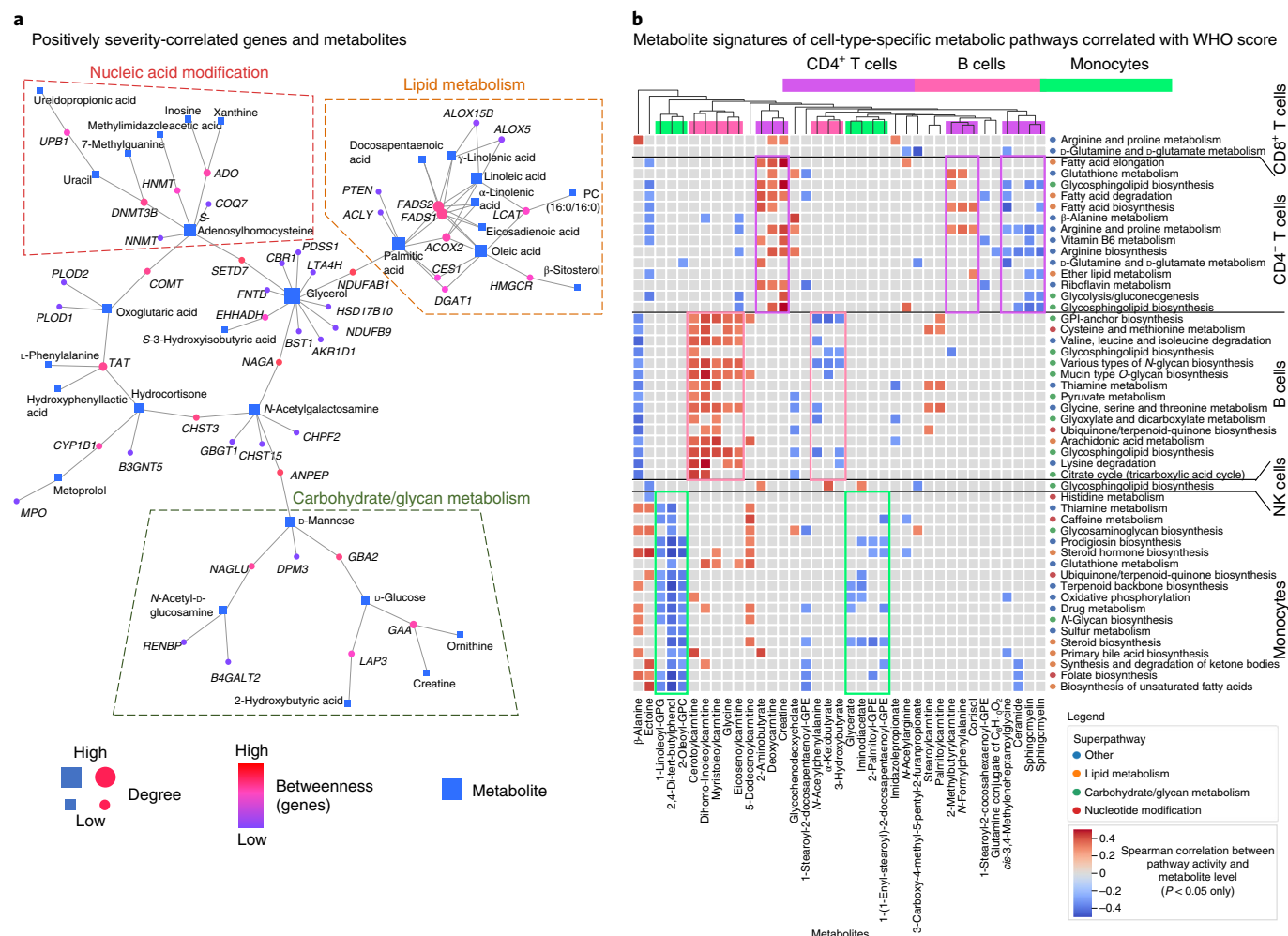
Fig. 5 | Integrated analysis of metabolites and cell-type-specific pathway activities reveals cell-type-specific metabolite signatures in COVID-19.
a, Metabolic pathway-related gene expression was calculated at the whole-individual level, and transcripts significantly positively correlated with WHO score were integrated with plasma metabolites that also positively correlated with WHO score to create a gene–metabolite network in which genes are connected to metabolites in known metabolic pathways; PC, phosphatidylcholine. **b**, Selected correlations of plasma metabolites with peripheral immune cell-type-specific metabolic pathway activities. Correlation was calculated by two-sided Spearman's rank correlation. Hierarchical clustering was performed for metabolites to resolve metabolite signatures; TCA, tricarboxylic acid.

consistent with previous clustering[4] (Fig. 4c and Extended Data Fig. 6a–d). Cluster 2, consisting of CD14^lowCD16^+HLA-DR^high cells, was composed of non-classical monocytes that were potentially immunomodulatory[29]. With more severe disease, the inflammatory monocytes increased not only in population percentage (Fig. 4d) (as reported in[26]) but also in metabolic activity per cell (Fig. 4e and Extended Data Fig. 6e). The non-classical monocytes decreased in both aspects (Fig. 4d,e).

Within inflammatory monocytes, cluster 0 is distinct from cluster 1 by higher expression of *S100A9* transcript and *HK2* (both confirmed at the protein level by flow cytometry) (Extended Data Fig. 7a–f). This suggests that the *S100A9*^high subset may be responsible for the increased monocyte metabolic activity in severe COVID-19.

We compared the three major subpopulations (*S100A9*^high inflammatory, *S100A9*^low inflammatory and non-classical) in COVID-19 to other acute and chronic diseases (sepsis[20] and HIV[21]) (Fig. 4f–h, Extended Data Fig. 8a,b and Supplementary Table 3). Intriguingly, the *S100A9*^high inflammatory monocytes in COVID-19 uniquely highly expressed 417 metabolic genes enriched in mitochondrial protein synthesis-related processes, particularly respiratory chain assembly (Fig. 4h, Extended Data Fig. 8b and Supplementary

Table 3), suggesting that there is increased oxidative phosphorylation for these monocytes in COVID-19.

Taken together, monocytes in individuals with COVID-19 diverge into two metabolically and functionally heterogeneous groups: an inflammatory, metabolically activated subpopulation and a potentially immunomodulatory, metabolically repressed non-classical subpopulation. These monocyte subpopulations further bifurcate by relative cell count and by metabolic reprogramming.

Integrated multiomics reveal cell-type-specific signatures. We next integrated the plasma metabolic signatures with metabolic pathways of each immune cell class and phenotype. To obtain a general view of such interactions, we generated networks of known gene–metabolite interactions that included transcripts in our metabolic pathways and plasma metabolites and proteins that were significantly correlated with disease severity (Fig. 5a and Extended Data Fig. 9a–c). Positively correlated metabolic transcripts and metabolites were integrated into gene–metabolite networks broadly classified into categories of lipid metabolism, nucleic acid modification and carbohydrate/glycan metabolism. Based on the resulting associations, we next asked whether metabolite levels correlated with

cell-type-specific metabolic programs. Indeed, cell-type-specific metabolic pathways, especially among CD4[+] T cells, B cells and monocytes, associated with unique signatures of metabolites (Fig. 5b). All three metabolic categories shown in Fig. 5a were found within the immune cell-type-specific metabolic changes shown in Fig. 5b, indicating the relevance of these metabolic categories' networks to the COVID-19 immune response.

The above associations suggested that integrating information from both plasma metabolites and immune cell metabolic pathway activities would yield a higher-resolution classification of individuals with COVID-19 compared to plasma metabolites alone. We thus performed PCA using both plasma metabolites and immune metabolic activities on a set of samples from individuals with COVID-19 whose single-cell transcriptomes were used to calculate metabolic pathway activities (Extended Data Fig. 9d) and then projected 242 samples that had both immune cell gene expression and plasma metabolite levels available onto the PC space. In contrast to our classification based on plasma metabolites alone (Extended Data Fig. 1b), even moderate and severe disease were distinguished. In fact, k-means clustering yielded three clusters separating individuals with mild (WHO score 1–2; cluster 0), moderate (WHO score 3–4; cluster 1) and severe (WHO score 5–7; cluster 2) disease (Fig. 6a,b and Extended Data Fig. 9e). Other clinical metrics, such as hospitalization status and level of respiratory support, were also well separated (Extended Data Fig. 9e). Furthermore, the top two PCs both correlated with clinical biomarkers from these individuals, including immune activation (cell counts) and inflammation (for example, C-reactive protein and fibrinogen) (Fig. 6c). PC2 correlated more strongly with markers of immune response to severe infections, such as peripheral immature granulocyte count. Both plasma metabolites and cell-type-specific metabolic pathway activities, especially the pathway activities that correlated with multiple metabolite patterns (Fig. 5b), contributed to PC1 and PC2 (Fig. 6d and Extended Data Fig. 9f), suggesting their potential roles in driving the overall immune response. Altogether, each immune cell class has unique metabolic signatures. Integrating cell-type-specific metabolic activities with plasma metabolite levels classified individuals with COVID-19 at higher resolution with respect to severity than what is achieved using metabolites alone.

Identifying plasma metabolites to predict clinical outcomes. We next sought to identify a small number of key plasma metabolites measured near the time of clinical diagnosis (T1) to predict COVID-19 survival among individuals who were hospitalized. We first used our integrated plasma and immune cell class-specific analysis in Fig. 6a–d to select top metabolite contributors to each

PC (Fig. 6d) as inputs for a machine learning model. A random forest classifier was used to refine a small panel of T1 metabolites for predicting survival within hospitalized INCOV individuals (Fig. 6e and Methods). We split our INCOV dataset into separate training and test sets with repeated randomized trials then used these datasets to train and test our model to predict death using levels of these selected T1 metabolites (Fig. 6f, left). Only five metabolite levels from T1 were sufficient to predict survival: palmitoyl-linoleoyl-glycerol (16:0/18:2), α-ketobutyrate, 1-palmitoyl-2-oleoyl-GPE (16:0/18:1), acetoacetate and maltose (Fig. 6g,h). To further validate this signature, we applied the model to an independent cohort of samples from individuals with COVID-19 collected and processed from individuals that were hospitalized at a different site by different investigators (Supplementary Table 4). Indeed, the predictive model predicted survival in the independent cohort with comparable accuracy (Fig. 6f, right).

Discussion

Understanding metabolic programs associated with immune regulation can provide insights into the mechanisms of the COVID-19 immune response and potentially yield effective tools to evaluate and treat this disease. Using multiomics data from a large cohort of healthy donors and individuals with COVID-19, we resolved diverse metabolic reprogramming in the COVID-19 peripheral immune response (Extended Data Fig. 10). We integrated cell-type-specific metabolic shifts with the plasma metabolome, which yielded intriguing, cell class-specific associations between these metabolically focused multiomics layers and precise disease classification.

This single-cell, multiomics approach revealed insights into COVID-19-associated metabolic reprogramming of the immune response at single-cell resolution. Highly active subpopulations, albeit small, dominated metabolic trends. For example, among CD8[+] and CD4[+] T cells, metabolic activation corresponding with COVID-19 severity was driven by small, metabolically hyperactive subpopulations not identifiable through analyses of averaged behavior[23,30]. Similar trends were also observed for B cells, NK cells and monocytes. For monocytes, we documented metabolic changes between two diverging subpopulations. Inflammatory monocytes increased in number and metabolic activity per cell, while nonclassical monocytes behaved oppositely in both respects (Fig. 4e). These changes occurred not only during the initial polarization of monocytes to these phenotypes as previously reported[27] but also continuously along the spectrum of disease severity.

These collective observations revealed two independent modes of metabolic reprogramming that are likely fundamental to an immune system response. The first corresponds to changes in the

Fig. 6 | Integrated metabolomic analysis reveals high-resolution classification of individuals with COVID-19 and inputs for a machine learning model to predict clinical outcomes. a, PCA of integrated plasma metabolites and metabolic pathways per individual with COVID-19. PCs were applied from the PCs that were calculated for the samples that had scRNA-seq data (Extended Data Fig. 9d). Resulting k-means clusters are colored, accompanied by contours and kernel density estimates of each cluster. **b**, Violin plot distribution of WHO scores per k-means cluster shown in **a** (k-means cluster 0: n = 111 samples; 1: n = 101; 2: n = 30). Within each violin, thick black lines are bounded by 25th and 75th percentiles. White dots indicate 50th percentiles. Thin black lines indicate 1.5 interquartile ranges below and above the 25th and 75th percentiles, respectively. Differences between the WHO scores of each k-means cluster were tested for statistical significance by two-sided Mann–Whitney U tests. **c**, Correlation of resulting PC1 and PC2 with clinical measurements per individual sample; BUN, blood urea nitrogen; GFR, glomerular filtration rate; RBC, red blood cell; WBC, white blood cell; APTT, activated partial thromboplastin time; INR, international normalized ratio; ALT, alanine aminotransferase; AST, aspartate aminotransferase. **d**, Loading scores of cell-type-specific metabolic pathway activities and plasma metabolites to PC1 and PC2. Pathway activities for only the cell types that exhibited strong metabolite signatures are shown (CD4[+] T cells, B cells and monocytes); TCA, tricarboxylic acid. **e**, Workflow of machine learning model training and testing to predict survival versus death at a later timepoint using metabolite levels measured at an initial timepoint T1. **f**, Left, receiver operating characteristic (ROC) for iterations of the machine learning model to predict survival versus death, initially tested within random subsets of the INCOV dataset. Right, ROCs after applying the machine learning model to an independent cohort of samples from individuals with COVID-19; AUC, area under curve. For both plots, shaded bands indicate standard error of the mean ROC. **g**, Relative importance of each metabolite level or clinical information from a trained model, after feature selection, for predicting survival. **h**, Correlation of each selected metabolite with change in WHO score, T2 versus T1. Statistical significance was calculated by two-sided Spearman's rank correlation with Benjamini–Hochberg FDR correction; *P < 0.05 (palmitoyl-linoleoyl-glycerol, P = 0.41; α-ketobutyrate, P = 0.019; 1-palmitoyl-2-oleoyl-GPE, P = 0.41; acetoacetate, P = 0.019; maltose, P = 0.41).

quantity of the metabolically active immune cell subpopulations, while the second involves shifts in the metabolism within individual cells within a subpopulation ('quality'). (Fig. 4e, bottom). Understanding these two modes of metabolic reprogramming may yield more precise strategies for promoting more effective protective responses, because specific immune cell subpopulations are identified to be manipulated. Importantly, these two modes were only resolved through analyzing metabolism at single-cell

resolution. Kinetic studies may provide more insight on the trajectories leading to such reprogramming[31,32].

The integration of cell-type-specific COVID-19-associated metabolic alterations with changes in plasma metabolite levels proved useful for classifying disease severity and predicting clinical outcomes. Broader categories of metabolic activity in the peripheral immune response were revealed (Fig. 5a), and we found cell-type-specific interplay between cellular metabolic programs
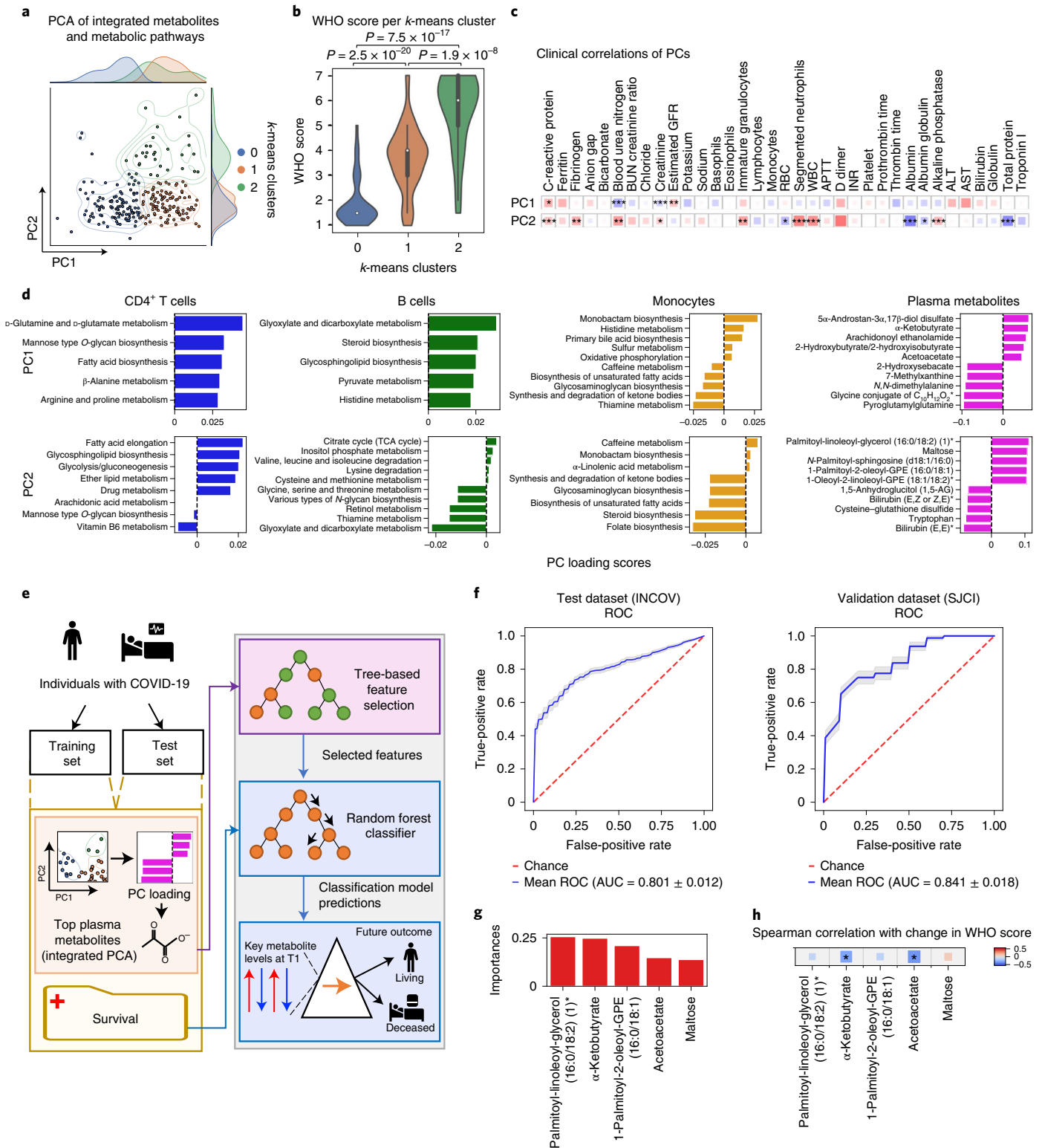
and distinct sets of metabolic products and byproducts (Fig. 5b). This analysis guided the classification of individuals with COVID-19 by disease severity or other clinical measurements (for example, level of respiratory support) with greater precision than by using individual omics layers alone, suggesting the importance of examining this interplay between plasma and cells. We identified subsets of metabolic pathways and plasma metabolites that correlated with clinical biomarkers associated with inflammation and immune activation.

Previous studies distilled individual features to predict concurrent disease status[3,33]. Here, we resolved a few plasma metabolites for predicting future outcome of newly diagnosed individuals with COVID-19. The clinical significances of these metabolites, such as acetoacetate and α-ketobutyrate, are known in other diseases, perhaps suggesting a pathophysiology shared with COVID-19. For instance, acetoacetate is produced in response to impaired cellular glucose uptake, while α-ketobutyrate is involved in the production of α-hydroxybutyrate[34], an early insulin resistance biomarker. Given the potential of these metabolites for predicting COVID-19 survival, we speculate that diabetes may affect COVID-19 (ref. [35]) by a related mechanism. For example, insulin resistance may impair glucose uptake needed by immune cells for upregulated metabolism and function. However, the metabolite signatures reported here may differ depending on comorbidities, interventions and other factors reflecting the complexity of human disease and thus require additional validations with multiple independent datasets.

An important consequence of this work is that it updates metabolic changes studied in bulk[14,36,37] with single-cell resolution to unveil what subpopulations are specifically responsible for driving these changes. Despite recent advances in single-cell methods, single-cell metabolomic analyses have been limited. For example, previous single-cell metabolic studies used custom-designed assays[38–40] or cytometry by time of flight[41], limiting the number of simultaneously quantifiable metabolic markers. Our approach combining large-scale, single-cell transcriptomic data with an extensive panel of plasma metabolites/proteins widens the view to permit the study of metabolic changes within the context of detailed, known metabolic pathways in the setting of acute viral infection and likely other diseases. Label-free methods of metabolite measurement, permitting general rather than selective detection of metabolites in live single cells[42], may provide another insightful omics layer.

Our analysis resolved significant alterations in metabolic pathways in the COVID-19 immune response and resolved their contrasts with PBMCs in sepsis and HIV (Fig. 4f–h and Extended Data Figs. 4 and 8). The function and molecular interactions of many of the involved individual metabolites and metabolic pathways, and how they relate to disease, are often poorly understood, but the data and methodology provided here should guide hypothesis development to further elucidate immune metabolic mechanisms for identifying therapeutic targets against COVID-19 and other diseases.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41587-021-01020-4.

References

1. Ganeshan, K. & Chawla, A. Metabolic regulation of immune responses. *Annu. Rev. Immunol.* **32**, 609–634 (2014).
2. Ayres, J. S. A metabolic handbook for the COVID-19 pandemic. *Nat. Metab.* **2**, 572–585 (2020).
3. Shen, B. et al. Proteomic and metabolomic characterization of COVID-19 patient sera. *Cell* **182**, 59–72 (2020).
4. Su, Y. et al. Multi-omics resolves a sharp disease-state shift between mild and moderate COVID-19. *Cell* **183**, 1479–1495 (2020).
5. Huang, S. S. et al. Phenylalanine- and leucine-defined metabolic types identify high mortality risk in patients with severe infection. *Int. J. Infect. Dis.* **85**, 143–149 (2019).
6. Stewart, C. J. et al. Respiratory syncytial virus and rhinovirus bronchiolitis are associated with distinct metabolic pathways. *J. Infect. Dis.* **217**, 1160–1169 (2018).
7. Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S. & Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **369**, 330–333 (2020).
8. Gao, T. et al. Highly pathogenic coronavirus N protein aggravates lung injury by MASP-2-mediated complement over-activation. Preprint at *medRxiv* https://doi.org/10.1101/2020.03.29.20041962 (2020).
9. Codo, A. C. et al. Elevated glucose levels favor SARS-CoV-2 infection and monocyte response through a HIF-1α/glycolysis-dependent axis. *Cell Metab.* **32**, 498–499 (2020).
10. Hanna, V. S. & Hafez, E. A. A. Synopsis of arachidonic acid metabolism: a review. *J. Adv. Res.* **11**, 23–32 (2018).
11. Maceyka, M. & Spiegel, S. Sphingolipid metabolites in inflammatory disease. *Nature* **510**, 58–67 (2014).
12. Xia, Z. et al. Multiple-omics techniques reveal the role of glycerophospholipid metabolic pathway in the response of *Saccharomyces cerevisiae* against hypoxic stress. *Front. Microbiol.* **10**, 1398 (2019).
13. Cruzat, V., Macedo Rogero, M., Noel Keane, K., Curi, R. & Newsholme, P. Glutamine: metabolism and immune function, supplementation and clinical translation. *Nutrients* **10**, 1564 (2018).
14. Vardhana, S. A. et al. Impaired mitochondrial oxidative phosphorylation limits the self-renewal of T cells exposed to persistent antigen. *Nat. Immunol.* **21**, 1022–1033 (2020).
15. Xiao, Z., Dai, Z. & Locasale, J. W. Metabolic landscape of the tumor microenvironment at single cell resolution. *Nat. Commun.* **10**, 3763 (2019).
16. Orth, J. D., Thiele, I. & Palsson, B. O. What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248 (2010).
17. Buczynski, M. W., Dumlao, D. S. & Dennis, E. A. Thematic Review sSeries: proteomics. An integrated omics analysis of eicosanoid biology. *J. Lipid Res.* **50**, 1015–1038 (2009).
18. Nutt, S. L., Fairfax, K. A. & Kallies, A. BLIMP1 guides the fate of effector B and T cells. *Nat. Rev. Immunol.* **7**, 923–927 (2007).
19. Brenchley, J. M. et al. Expression of CD57 defines replicative senescence and antigen-induced apoptotic death of CD8+ T cells. *Blood* **101**, 2711–2720 (2003).
20. Reyes, M. et al. An immune-cell signature of bacterial sepsis. *Nat. Med.* **26**, 333–340 (2020).
21. Kazer, S. W. et al. Integrated single-cell analysis of multicellular immune dynamics during hyperacute HIV-1 infection. *Nat. Med.* **26**, 511–518 (2020).
22. Ndhlovu, Z. M. et al. Magnitude and kinetics of CD8+ T cell activation during hyperacute HIV infection impact viral set point. *Immunity* **43**, 591–604 (2015).
23. Jellusova, J. Metabolic control of B cell immune responses. *Curr. Opin. Immunol.* **63**, 21–28 (2020).
24. Zheng, Y. et al. Mitochondrial one-carbon pathway supports cytosolic folate integrity in cancer cells. *Cell* **175**, 1546–1560 (2018).
25. Olivera, A. & Rivera, J. Sphingolipids and the balancing of immune cell function: lessons from the mast cell. *J. Immunol.* **174**, 1153–1158 (2005).
26. Yonggang Zhou, B. F. et al. Pathogenic T cells and inflammatory monocytes incite inflammatory storm in severe COVID-19 patients. *Natl Sci. Rev.* **7**, 998–1002 (2020).
27. Viola, A., Munari, F., Sanchez-Rodriguez, R., Scolaro, T. & Castegna, A. The metabolic signature of macrophage responses. *Front. Immunol.* **10**, 1462 (2019).
28. Cramer, T. et al. HIF-1α is essential for myeloid cell-mediated inflammation. *Cell* **112**, 645–657 (2003).
29. Narasimhan, P. B., Marcovecchio, P., Hamers, A. A. J. & Hedrick, C. C. Nonclassical monocytes in health and disease. *Annu. Rev. Immunol.* **37**, 439–456 (2019).
30. Bantug, G. R., Galluzzi, L., Kroemer, G. & Hess, C. The spectrum of T cell metabolism in health and disease. *Nat. Rev. Immunol.* **18**, 19–34 (2018).
31. Su, Y. et al. Single-cell analysis resolves the cell state transition and signaling dynamics associated with melanoma drug-induced resistance. *Proc. Natl Acad. Sci. USA* **114**, 13679–13684 (2017).
32. Su, Y. et al. Kinetic inference resolves epigenetic mechanism of drug resistance in melanoma. Preprint at *bioRxiv* https://doi.org/10.1101/724740 (2019).
33. Shu, T. et al. Plasma proteomics identify biomarkers and pathogenesis of COVID-19. *Immunity* **53**, 1108–1122 (2020).

34. Gall, W. E. et al. α-Hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. *PLoS ONE* **5**, e10883 (2010).

35. Rubino, F. et al. New-onset diabetes in Covid-19. *N. Engl. J. Med.* **383**, 789–790 (2020).

36. O'Sullivan, D. et al. Memory CD8+ T cells use cell-intrinsic lipolysis to support the metabolic programming necessary for development. *Immunity* **41**, 75–88 (2014).

37. Lee, M. K. S. et al. Glycolysis is required for LPS-induced activation and adhesion of human CD14+CD16- monocytes. *Front. Immunol.* **10**, 2054 (2019).

38. Xue, M. et al. Chemical methods for the simultaneous quantitation of metabolites and proteins from single cells. *J. Am. Chem. Soc.* **137**, 4066–4069 (2015).

39. Xue, M., Wei, W., Su, Y., Johnson, D. & Heath, J. R. Supramolecular probes for assessing glutamine uptake enable semi-quantitative metabolic models in single cells. *J. Am. Chem. Soc.* **138**, 3085–3093 (2016).

40. Su, Y. et al. Multi-omic single-cell snapshots reveal multiple independent trajectories to drug tolerance in a melanoma cell line. *Nat. Commun.* **11**, 2345 (2020).

41. Hartmann, F. J. et al. Single-cell metabolic profiling of human cytotoxic T cells. *Nat. Biotechnol.* **39**, 186–197 (2020).

42. Du, J. et al. Raman-guided subcellular pharmaco-metabolomics for metastatic melanoma cells. *Nat. Commun.* **11**, 4830 (2020).

Methods

Human participation in this study was obtained previously (INvestigating COVID-19 (INCOV))[4] and by PH&S (Providence Health and Services; St. John's Cancer Institute (SJCI)) and complies with all relevant ethical regulations. Procedures for the INCOV study, including informed consent for participants, were approved by the Institutional Review Board (IRB) at Providence St. Joseph Health (IRB study number STUDY2020000175) and the Western IRB (IRB study number 20170658). Plasma samples for SJCI were obtained from individuals consented under the PH&S IRB-approved protocol JWCI-18-0401, PH&S IRB STUDY2018000254. INCOV participants were compensated with a complimentary T-shirt. Individuals were not compensated for participation in SJCI. Among the INCOV cohort, there were 165 individuals with COVID-19 and 204 healthy volunteers. Individuals with COVID-19 were aged between 18 and 89 years (median 56 years), and 53% were female. Healthy volunteers were aged 28–70 years (median 56 years), and 50% were female. COVID-19 severity spanned the full range of the WHO ordinal scale from 0 to 7, including healthy volunteers. Among individuals with COVID-19, the median WHO score was 4 at T1. Comorbidities commonly included hypertension (40% of individuals with COVID-19) and diabetes mellitus (19%). The SJCI cohort contained 33 individuals. Participants were aged between 28 and 96 years with a median age of 66 years; 36% were female. Large fractions of the cohort had hypertension (67%) and/or diabetes mellitus (36%) among other comorbidities. COVID-19 severity spanned from 3 to 7 on the WHO ordinal scale (median 5).

Metabolic data collection. PBMC scRNA-seq data and gas chromatography–mass spectrometry (GC–MS)-based plasma metabolite levels from healthy volunteers and individuals with COVID-19 were taken as raw read counts and concentrations. Metabolon measured metabolite levels from all participant plasma samples via Metabolon's ultra high-performance liquid chromatography/tandem mass spectrometry (UHPLC/MS/MS) Global Platform. Metabolon performed sample handling and quality control in their Clinical Laboratory Improvement Amendments (CLIA)-certified laboratory. Metabolon performed data extraction, biochemical identification, data curation, quantification and normalization. Samples were processed in batches including pooled quality control samples in each batch, which were consistent across batches. Metabolite levels were corrected for potential batch effects by dividing by the corresponding average values found in the pooled quality control samples within each batch.

We correlated plasma metabolite levels with WHO score by Spearman's rank correlation. Each individual with COVID-19 was assigned a clinical score on the WHO ordinal scale by clinicians, and we used the score to also classify individuals into broader categories of mild disease/healthy status (WHO 0–2), moderate disease (WHO 3–4) and severe disease (WHO 5–7) as described in Supplementary Table 5. We also generated data for flow cytometric analysis of CD8[+] T cells and monocyte metabolism. For comparisons of metabolic characteristics of CD8[+] T cells and monocytes with other diseases, we obtained published data for sepsis[20] and HIV[21]. Finally, for our independent validation cohort (SJCI) in our machine learning analysis, we used newly generated, hitherto unpublished clinical and metabolite data as summarized in Supplementary Table 4.

We have summarized all data sources, including published and new data, in Supplementary Table 1.

10X data processing. Droplet-based sequencing data were aligned and quantified using the Cell Ranger Single-Cell Software Suite (version 3.0.0, 10X Genomics) against the GRCh38 human reference genome. Cells from each demultiplexed sample were first filtered for cells that expressed a minimum of 200 genes and then filtered based on three metrics: (1) the total number of unique molecular identifier (UMI) counts per cell (library size), (2) the number of detected genes per cell and (3) the proportion of mitochondrial gene counts (UMIs from mitochondrial genes/ total UMIs). Exact cutoffs varied from sample to sample and were manually set based on pairwise plots of the aforementioned metrics but were typically set at a max of 10,000 UMIs per cell, 2,500 genes per cell and 15% mitochondrial reads per cell. Doublets were simultaneously identified in sample demultiplexing and were removed before filtering. After quality control metric filtering, a total of 221,748 cells were retained for downstream analysis. Scanpy was used to normalize cells via counts per million reads mapped (CPM) normalization (UMI total count of each cell was set to 10[6]) and log1p transformation (natural log of CPM plus one).

PCA was performed on the normalized, $\ln(\text{CPM}+1)$, gene expression matrix of all cells passing the quality control metrics, and 'arpack' was set as the singular value decomposition (SVD) solver. A neighborhood graph was built with n-neighbors set to 15 and all 50 calculated PCs as inputs. This neighborhood graph was used to calculate a UMAP, and clusters were determined via Louvain. Clusters identified in this first round of clustering were annotated based on the expression of canonical marker genes. Clusters that were not uniform in their expression of well-known marker genes were extracted, and a second round of dimension reduction and clustering was performed on these subsets (recalculation of UMAP and Leiden, which was mainly used to separate T cells and NK cells and rare cell type populations). Identified T cells from the first and second round of clustering were extracted for CD4[+] and CD8[+] T cell identification. CD4 and CD8 surface proteins were used for a two-dimensional scatterplot, and cells with a non-zero expression level of at least one of the two surface markers were classified as CD4[+]

or CD8[+] T cells using a linear cutoff. Other T cells were then defined as T cells that did not confidently categorize as CD4[+] or CD8[+] (that is, did not express either CD4 or CD8 surface protein). Scanpy was used for all dimensional reduction and clustering steps. After cell type identification, normalized, $\ln(\text{CPM}+1)$, gene expression matrices for each cell type were subject to MAGIC imputation[43] using the 'scanpy.external.magic' function from Scanpy with the 'random_state' parameter set to 10. For computational simplicity, we used a representative subset of 56 samples for metabolic pathway activity analyses. These preprocessed scRNA-seq datasets were further used for metabolic gene clustering. We used Scanpy to apply Louvain or Leiden clustering separately for each PBMC cell type (CD8[+] T cells, CD4[+] T cells, B cells, NK cells and monocytes/macrophages).

For comparison of metabolic genes between diseases, we defined subpopulations from each scRNA-seq dataset and then calculated differentially expressed metabolic genes between subpopulations using Scanpy. We performed gene ontology enrichment analysis on these differentially expressed gene sets using PANTHER[44].

Pathway activity analysis. For all PBMCs and for each PBMC cell type, we calculated the level of activity of specific metabolic pathways as described previously[15]. To summarize, we calculated the mean expression level (log1p from preprocessed 10X data as described above) of a gene among all cells within a given group of cells (for example, a cell cluster, cells belonging to a specific category of WHO scores and so on within a given group of PBMCs). We then calculated the ratio of this mean expression level compared to the average of mean expression levels of this gene across all cell groups being analyzed. Statistical significances of pathway activity scores were calculated by a random permutation test in which cell groups were randomly assigned to single cells 1,000 times, after which we calculated pathway activity scores for each set of randomized cell groups. Metabolic pathways, including gene sets involved in a specific metabolic process, were as defined by the Kyoto Encyclopedia of Genes and Genomes (KEGG; downloaded from https://www.kegg.jp/kegg-bin/get_htext?ko00001.keg; 14 July 2020). We excluded pathways that were entirely composed of genes not found in our final scRNA-seq data.

Pathway and network analysis of plasma metabolites/proteins. Plasma metabolite concentrations were correlated to each individual's severity score on the ordinal WHO scale ('WHO score'; Supplementary Table 5) via Spearman's rank correlation. We performed metabolic pathway enrichment analysis ('Pathway Analysis') with MetaboAnalyst 4.0 web-based software (https://www.metaboanalyst.ca; accessed 17 September 2020) using metabolites that were significantly correlated ($P < 0.05$) with individuals' WHO scores by Spearman's rank correlation. We combined significantly correlated metabolites and proteins with genes that had pseudobulk expression significantly correlated with individuals' WHO scores, and then we performed metabolic network analysis via the Network Explorer (set to create a 'gene–metabolite network') in MetaboAnalyst.

Flow cytometry analysis. Cryopreserved PBMCs were thawed and washed once with RPMI (Gibco). Cells were treated with 1× RBC Lysis Buffer to eliminate red blood cells. Cells were treated with Fc blocker and monocyte blocker before being stained in Cell Staining Buffer with the following antibodies: in the T cell panel, BV421-CD3, PerCP/Cy5.5-CD4 and Alexa Fluor 488-CD8; in the monocyte panel, BV421-HLA-DR, PE-CD14 and APC/Cy7-CD16. After cell surface staining, cells were fixed in Fixation Buffer and washed with 1× Intracellular Staining Perm Wash Buffer twice before resuspension in 1× Intracellular Staining Perm Wash Buffer with the following antibodies: in the T cell panel, PE-Ki-67 and Alexa Fluor 647-HK2 (Abcam); in the monocyte panel, FITC-S100A8/9 (Abcam) and Alexa Fluor 647-HK2 (Abcam). Cells were washed twice and analyzed on an Attune NxT. Data were analyzed using FlowJo software. All antibodies and reagents were purchased from BioLegend, unless otherwise specified. All antibodies were used at a 1:50 dilution except for FITC-S100A8/9 (1:100 dilution). Samples that had detectable measurements of HK2 for both T cells and monocytes on flow cytometry were used for comparison with gene expression data.

Generation of personalized immune cell metabolic networks. We integrated the immune cell transcriptomes of our samples with genome-scale metabolic reconstruction of human gene expression using Recon3D (ref. [45]). Using the iMAT algorithm[46], we generated context-specific personalized metabolic networks for each immune cell in the dataset. Human cells in general do not proliferate rapidly and tend to maintain their metabolic functions. We therefore chose the biomass maintenance reaction as the objective function for the immune cells. In total, 7,964 metabolic networks were generated for five different immune cell types in 56 samples. We performed flux variability analysis[47] to evaluate minimum and maximum flux for each reaction in the metabolic networks. COBRA toolbox v3.0 (ref. [48]) was used for metabolic analysis that was implemented in MATLAB R2018a, and academic licenses of Gurobi optimizer v7.5 and IBM CPLEX v12.7.1 were used to solve Linear Programming (LP) and Mixed Integer Linear Programming (MILP) problems. Maximum flux value of reactions was used for the analyses in this study.

Machine learning model analysis. To identify metabolites for predicting survival, we used the scikit-learn package (version 0.23.2) in Python. We identified

metabolites among the top five highest or lowest loading scores per PC1 and PC2 in our integrated PCA (thus, 20 starting metabolites; Fig. 6a–d). In addition to the dataset that we used for the integrated analysis (INCOV dataset), we newly collected an independent cohort of plasma metabolite data from a set of individuals that were hospitalized (SJCI dataset); this new dataset was collected and processed at a different site by different investigators. All metabolite data were winsorized, missing values were imputed via iterative imputation and then the data were normalized. To make an even comparison between the two datasets, we selected only individuals who were hospitalized from our initial dataset, then randomly resampled individuals within each dataset to achieve the same fractions of deceased individuals in each dataset. Using normalized metabolite levels measured at the first timepoint (T1; shortly after COVID-19 diagnosis) for individuals that had survival data and metabolite levels in the INCOV dataset, we performed feature selection via ExtraTreesClassifier. The selected metabolites were used as inputs in RandomForestClassifier using bootstrapping with out-of-box scoring and were trained and tested within our INCOV dataset via repeated, randomized splits of these data (repeated stratified k-fold cross-validation) to predict survival (living versus deceased) at a final timepoint. We then tested the INCOV-trained predictive model directly on the new independent SJCI dataset.

GLIPH analysis. We performed GLIPH analysis, which clusters TCRs by predicted epitope specificity[49], on the complementarity-determining region 3 (CDR3) sequences of individual TCRs and the database of epitope–TCR-matched sequences from Adaptive Biotechnologies (https://doi.org/10.21417/ADPT2020COVID; release 002)[50]. No reference was used to cluster these sequences. We then filtered through the GLIPH clusters to create new modified groups that could be assigned to only one group of epitopes from the Adaptive Biotechnologies database. These groups were used to determine which GLIPH clusters had sequences associated with each epitope.

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

Data availability
All of the COVID-19 single-cell data used in this study are either included in the manuscript or were downloaded from ArrayExpress E-MTAB-9357. Plasma metabolite and protein data are available from Mendeley Data at https://doi.org/10.17632/tzydswhhb5.5 and in Supplementary Tables 2 and 4 of this manuscript. Metabolic pathways, including gene sets involved in a specific metabolic process, were used as defined by KEGG (downloaded from https://www.kegg.jp/kegg-bin/get_htext?ko00001.keg; 14 July 2020). Epitope–TCR-matched sequences for GLIPH analysis were obtained from the Adaptive Biotechnologies database (https://doi.org/10.21417/ADPT2020COVID; release 002)[50]. Data for HIV[21] were downloaded from the Broad Institute Single Cell Project at https://singlecell.broadinstitute.org/single_cell/study/SCP256. Data for sepsis[20] were also downloaded from the Single Cell Project at https://singlecell.broadinstitute.org/single_cell/study/SCP548.

Code availability
Data analysis was performed using custom code available on https://github.com/jihoonlee0/SARS-CoV-2_metab.

References
43. van Dijk, D. et al. Recovering gene interactions from single-cell data using data diffusion. *Cell* **174**, 716–729 (2018).
44. Mi, H., Muruganujan, A., Ebert, D., Huang, X. & Thomas, P. D. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* **47**, D419–D426 (2019).
45. Brunk, E. et al. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nat. Biotechnol.* **36**, 272–281 (2018).
46. Zur, H., Ruppin, E. & Shlomi, T. iMAT: an integrative metabolic analysis tool. *Bioinformatics* **26**, 3140–3142 (2010).
47. Gudmundsson, S. & Thiele, I. Computationally efficient flux variability analysis. *BMC Bioinformatics* **11**, 489 (2010).
48. Heirendt, L. et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat. Protoc.* **14**, 639–702 (2019).
49. Glanville, J. et al. Identifying specificity groups in the T cell receptor repertoire. *Nature* **547**, 94–98 (2017).
50. Nolan, S. et al. A large-scale database of T-cell receptor β (TCRβ) sequences and binding associations from natural and synthetic exposure to SARS-CoV-2. Preprint at *Research Square* https://doi.org/10.21203/rs.3.rs-51964/v1 (2020).

Author contributions
J.W.L., Y.S., P.D.G., J.D.G. and J.R.H. conceptualized the study. J.R.H., J.D.G. and D.S.B.H. provided resources. J.W.L., Y.S., P.D.G. and J.R.H. developed the methodology. J.W.L., Y.S., P.B., D.C., A.J.P.B., D.Y., V.R.D., R.H.N., J.C., J.X., R.Z., K.M., S.K., B.S., A.T.M., D.S.B.H., J.J.H., J.D.G., N.D.P., R.G., M.M.D., L.H., P.D.G. and J.R.H. performed the experimental investigations. J.W.L., Y.S., P.B., D.C., A.J.P.B., V.R.D., R.H.N., S.K., B.S., A.T.M. and J.J.H. analyzed the data. J.W.L., Y.S., P.D.G. and J.R.H. wrote the original draft of the manuscript. J.W.L., Y.S., P.B., D.C., A.J.P.B., V.R.D., R.H.N., D.S.B.H., J.J.H., J.D.G., N.D.P., R.G., M.M.D., L.H., P.D.G. and J.R.H. reviewed and edited the manuscript.

Extended Data Fig. 1 | Overview of metabolic analysis of COVID-19 peripheral immune response. (a) Known metabolic pathways affected by plasma metabolites that are negatively correlated with WHO score. Red dashed line indicates threshold of statistical significance on y-axis ($p < 0.05$), calculated by hypergeometric test. Exact p-values are in Supplementary Table 7. (**b**) Principal component analysis (PCA) of patient samples by plasma metabolite levels. Each point is a patient sample. Left: colored according to WHO score of the patient at the time of observation. Right: colored according to K-means clusters. (**c**) Workflow of calculating metabolic pathway activities. (**d**) UMAP of metabolic fluxes per immune cell type per patient. (**e**) Frequency of immune cell types within each metabolic flux-defined cluster in (d).
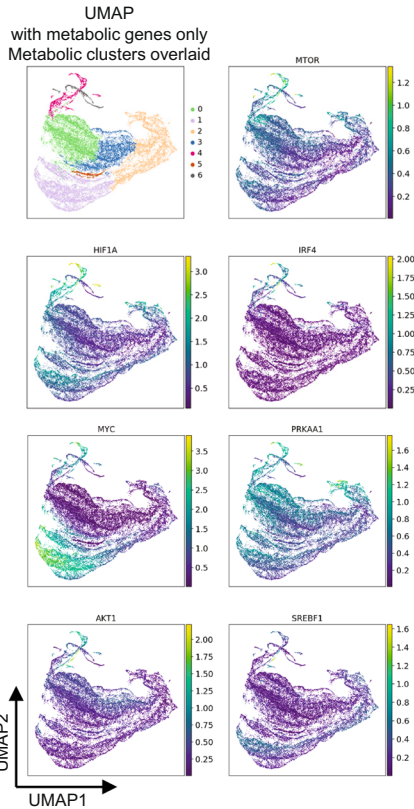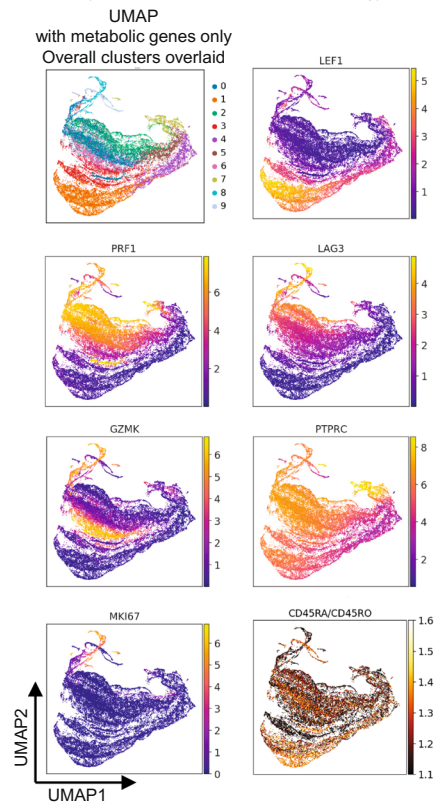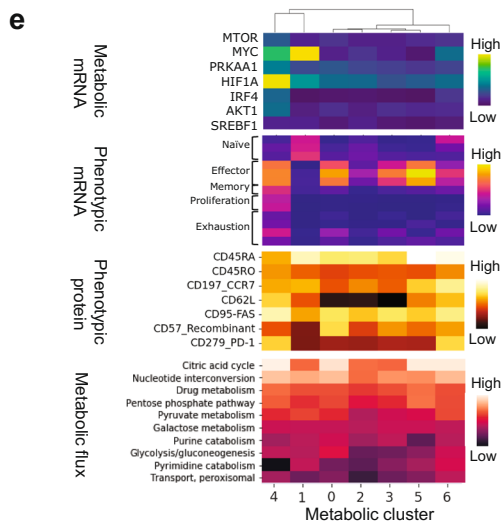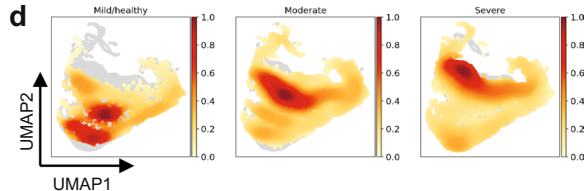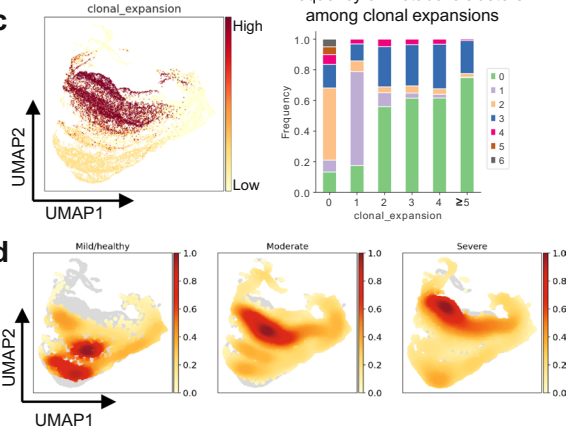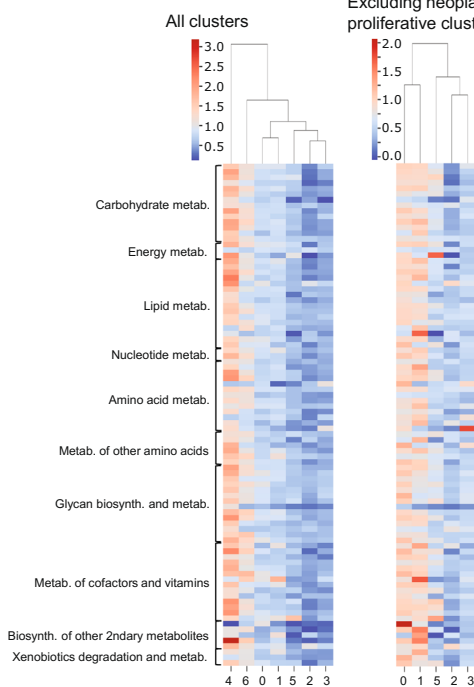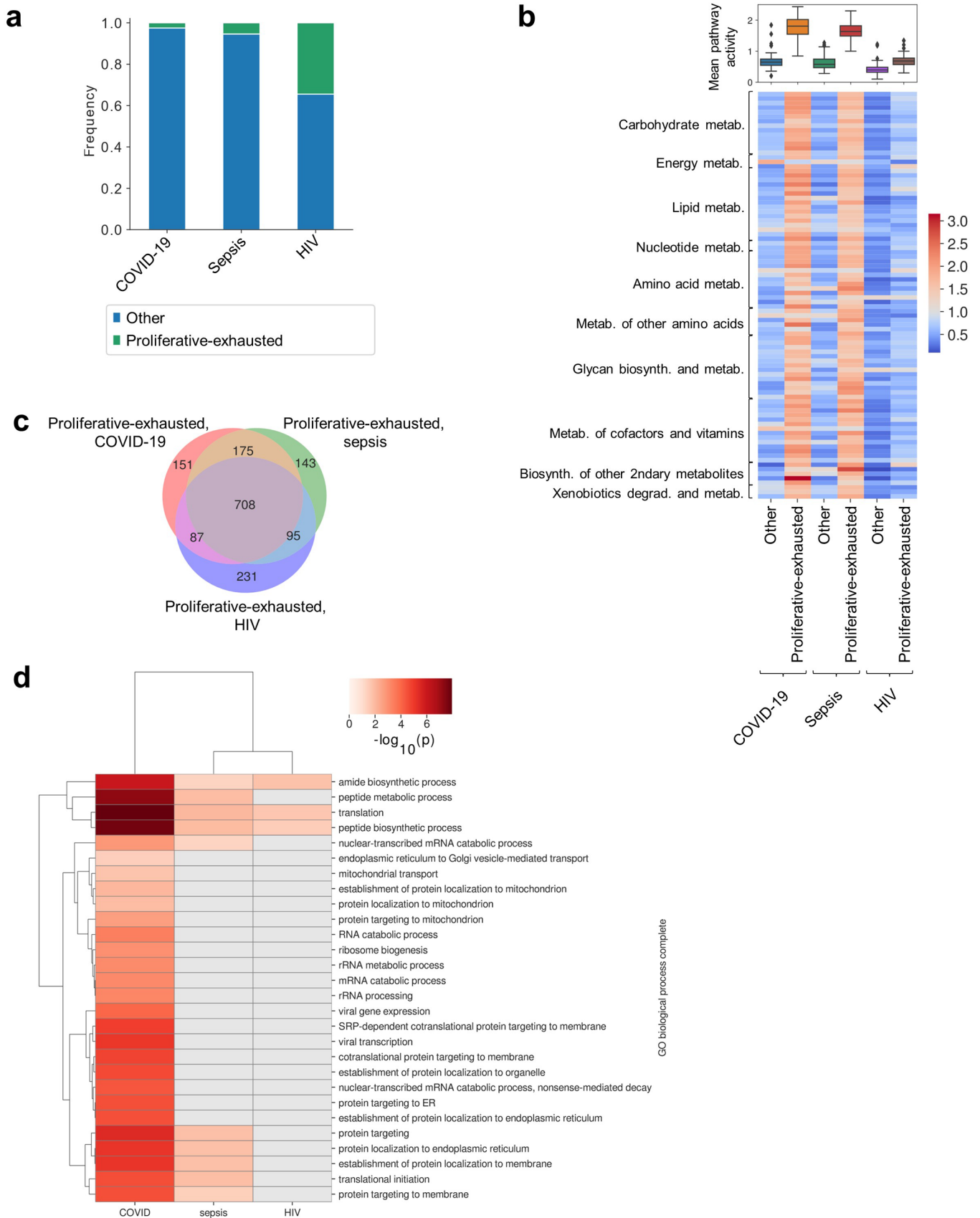
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Metabolically defined CD8+ T cell clusters have distinct metabolic activities and regulatory programs. (a) Unsupervised clustering and UMAP of CD8+ T cells using expression of only genes involved in metabolic pathways, colored by metabolic cluster and by expression of key metabolic regulator genes. (**b**) Unsupervised clustering and UMAP of CD8+ T cells, colored by all genes-defined cluster and by expression of phenotypic markers. (**c**) Left: T cell receptor (TCR) clonal expansion per cell overlaid onto UMAP plot. Right: frequency of TCR clonal expansions per metabolic cluster. (**d**) Density map of patient disease severity category per cell overlaid onto UMAP plot: mild/healthy (WHO 0–2), moderate disease (WHO 3–4), and severe disease (WHO 5–7). (**e**) Heatmap of mean expression of key metabolic regulator genes, phenotypic genes and proteins, and top 10 most active metabolic fluxes for each metabolic cluster in pseudobulk. (**f**) Metabolic pathway activities of CD8+ T cell metabolic clusters in pseudobulk among all patients combined, with and without inclusion of proliferative and neoplastic cell clusters (metabolic clusters 4 and 6, respectively).
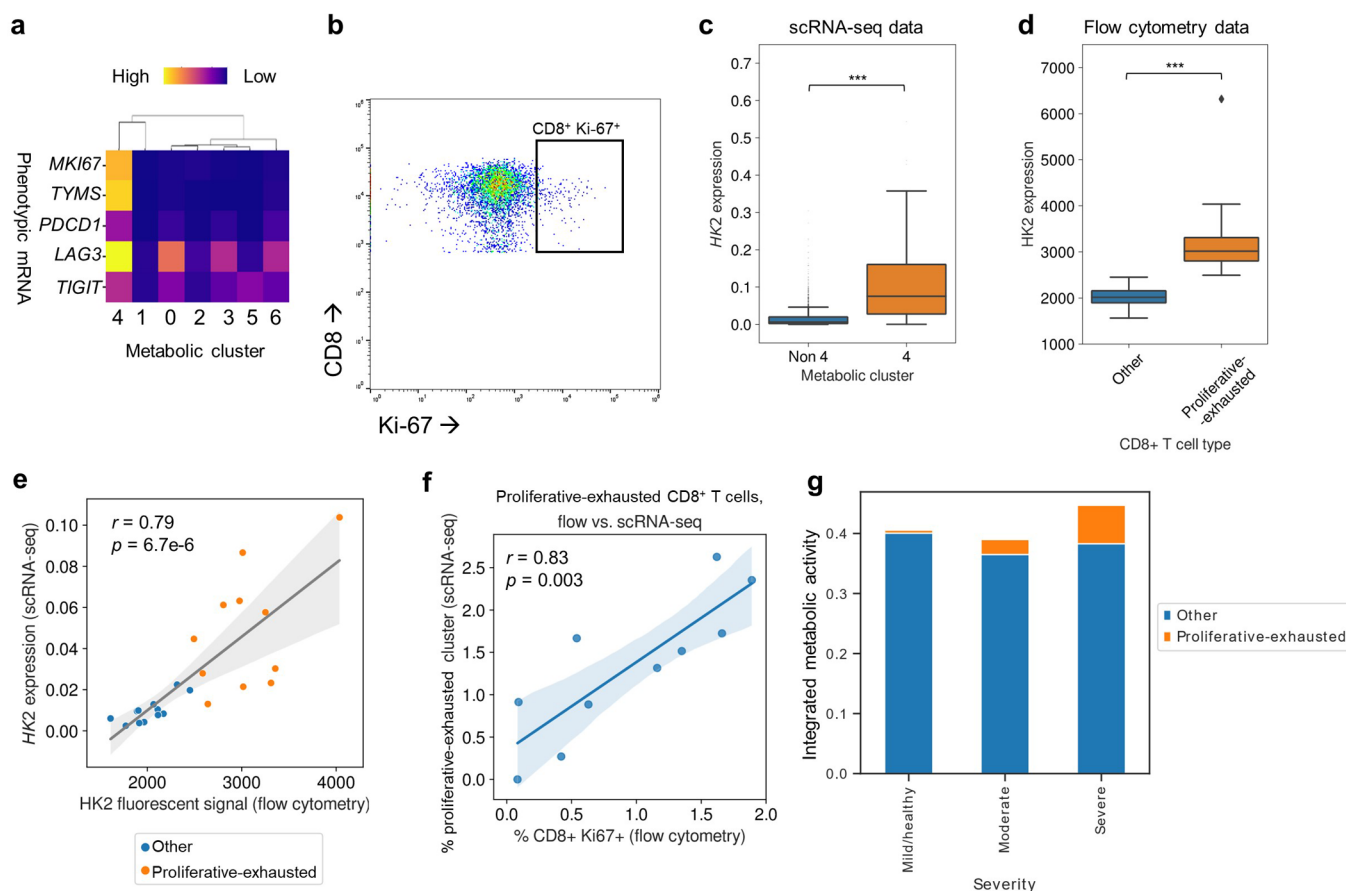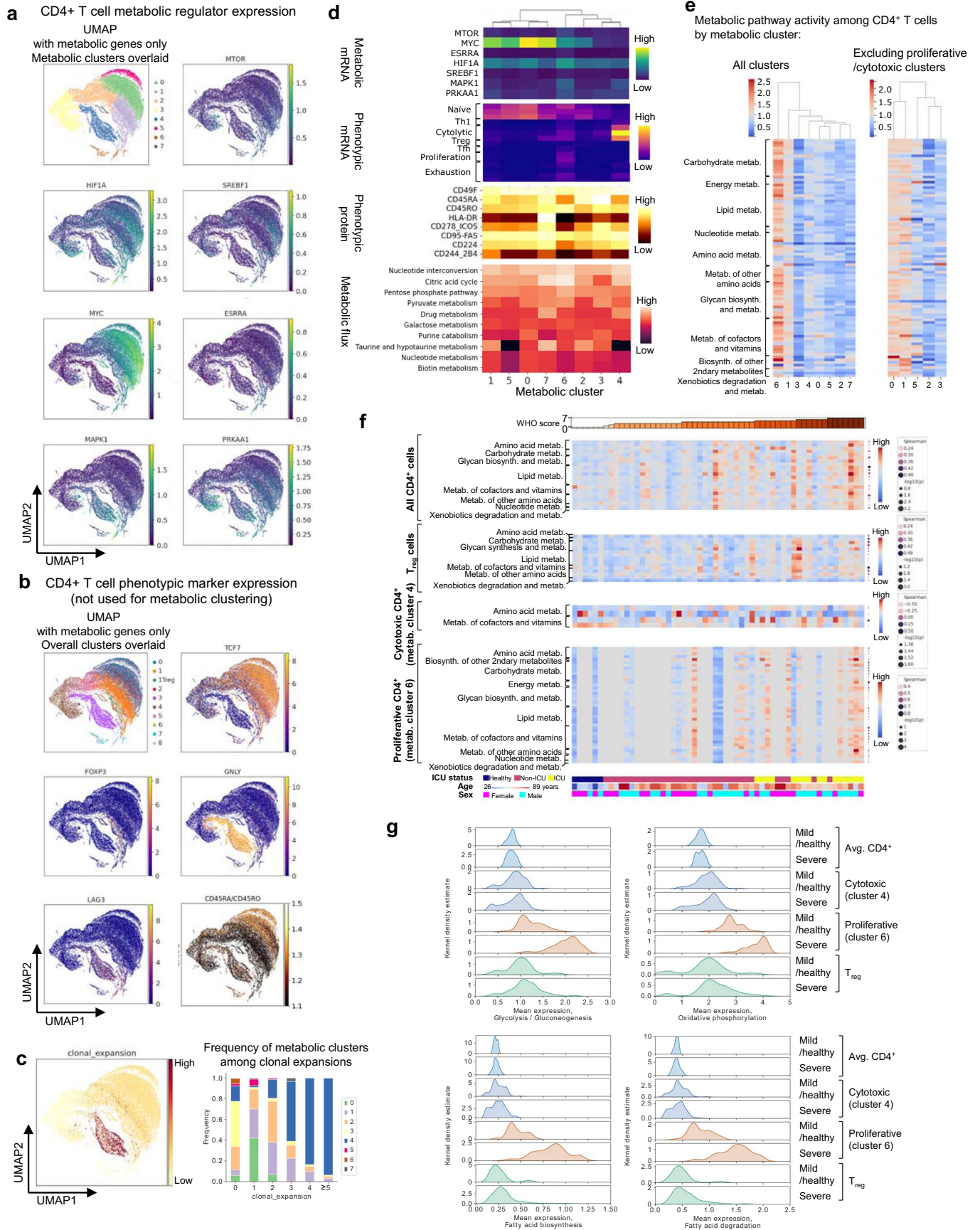
Extended Data Fig. 3 | Flow cytometric analysis of metabolism in the proliferative-exhausted subpopulation (cluster 4) of CD8$^+$ T cells. (**a**) Heatmap summary of proliferation and exhaustion markers across metabolically defined CD8$^+$ T cell clusters. Metabolic cluster 4 is the proliferative-exhausted subpopulation. (**b**) Representative flow cytometry plot of the gating strategy for proliferative-exhausted cells. Gating was performed on all CD8$^+$ T cells. Since all CD8$^+$ Ki-67$^+$ cells were also PD-1$^+$ in scRNA-seq analysis and because of flow cytometer color limitations, we gated on this population with Ki-67. (**c**) Expression of *HK2* mRNA from COVID-19 patient scRNA-seq data in the proliferative-exhausted CD8$^+$ T cell subpopulation (that is, metabolic cluster 4) vs. other CD8$^+$ T cells (non-metabolic cluster 4: n = 24,958 cells; metabolic cluster 4: n = 1,008). (**d**) Protein expression level of HK2 from flow cytometric measurements of COVID-19 patient samples in the gated, proliferative-exhausted CD8$^+$ T cell subpopulation vs. other CD8$^+$ T cells (other: n = 13 patient samples; proliferative-exhausted: n = 14). For (c) and (d), boxes indicate 25th, 50th, and 75th percentiles. Whiskers indicate 1.5 interquartile ranges below and above the 25th and 75th percentiles, respectively. Statistical significances were calculated by two-sided Mann-Whitney U test. ***: p < 0.001. (**e**) Two-sided Pearson correlation of *HK2* mRNA expression from scRNA-seq data and HK2 protein level from flow cytometry. (**f**) Two-sided Pearson correlation of % proliferative-exhausted cells among CD8$^+$ T cells as determined from scRNA-seq data vs. flow cytometry. Shaded bands in (e) and (f) each indicates 95% confidence interval of linear regression. (**g**) Integrated metabolic activity of CD8$^+$ T cells per disease severity category: mild/healthy (WHO 0–2), moderate (WHO 3–4), and severe (WHO 5–7). The contributions from CD8$^+$ T cell subpopulations are shown with cluster-specific color. Metabolic activity is defined as expression levels of metabolic pathway-related genes from each metabolically defined CD8$^+$ T cell subpopulation, multiplied by the fraction of CD8$^+$ T cells that are in each subpopulation.
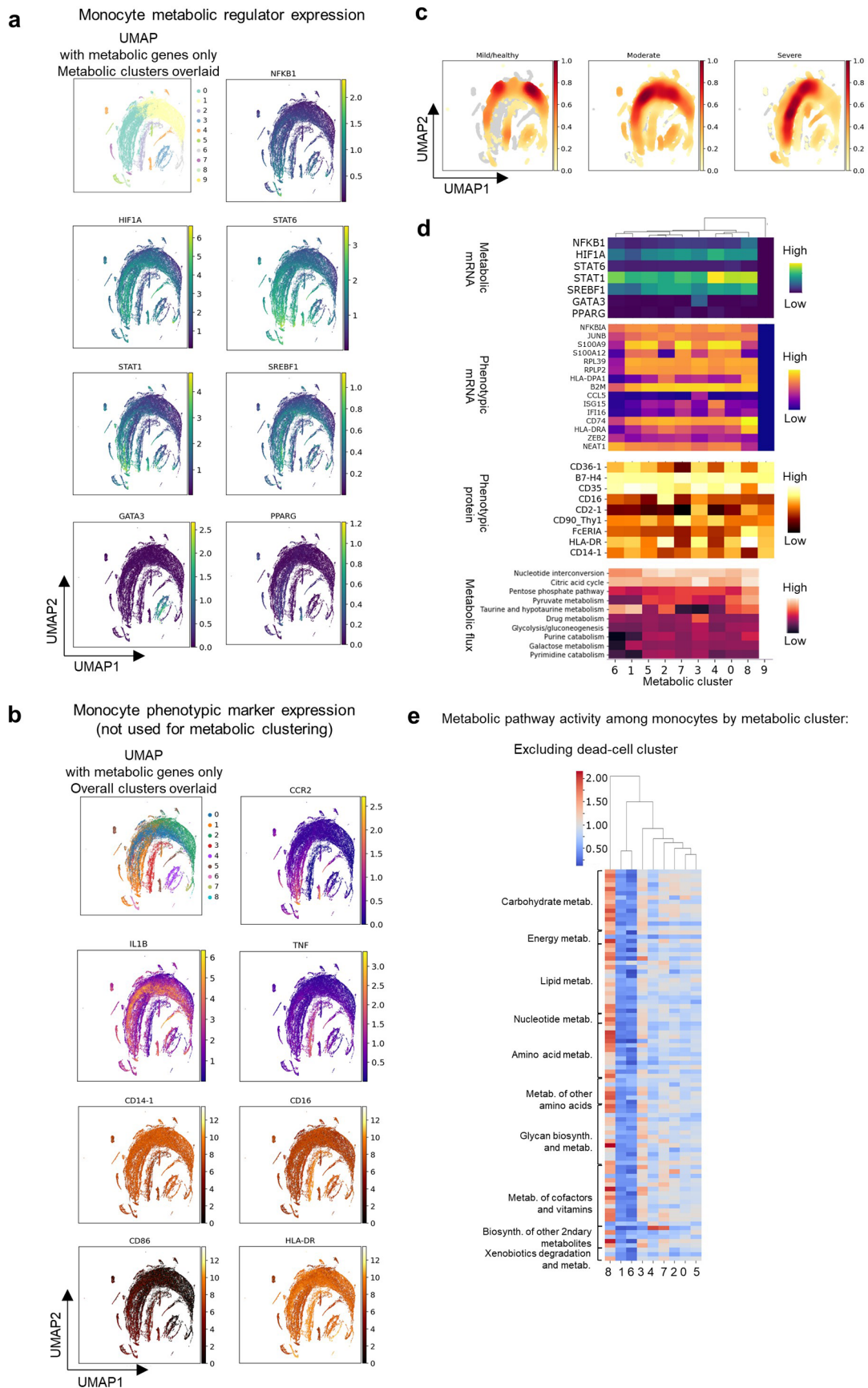
a

b

c

d

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

Extended Data Fig. 4 | Comparison of proliferative-exhausted CD8$^+$ T cells between COVID-19, sepsis, and HIV. (**a**) Fraction of CD8$^+$ T cells per disease (COVID-19, sepsis, and HIV) that are proliferative-exhausted. (**b**) Summary of metabolic pathway activities for proliferative-exhausted vs. other CD8$^+$ T cells in each disease. For the box plots, $n = 90$ pathways for each condition. Boxes indicate 25th, 50th, and 75th percentiles. Whiskers indicate 1.5 interquartile ranges below and above the 25th and 75th percentiles, respectively. (**c**) Venn diagram of metabolic pathway-related genes that are uniquely elevated in proliferative-exhausted CD8$^+$ T cells of each disease. (**d**) Gene ontology analysis of metabolic genes uniquely elevated in proliferative-exhausted CD8$^+$ T cells of COVID-19, showing only the processes that have greater than fivefold enrichment in the COVID-19 condition.

a CD4+ T cell metabolic regulator expression

b CD4+ T cell phenotypic marker expression (not used for metabolic clustering)

c

d

e Metabolic pathway activity among CD4+ T cells by metabolic cluster:

f

g

Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Metabolically defined CD4$^+$ T cell clusters have distinct metabolic activities and regulatory programs. (a) Unsupervised clustering and UMAP of CD4$^+$ T cells using expression of only genes involved in metabolic pathways, colored by metabolic cluster and by expression of key metabolic regulator genes, and (**b**) colored by all genes-defined cluster and by expression of phenotypic markers. (**c**) Left: T cell receptor (TCR) clonal expansion per cell overlaid onto UMAP plot. Right: frequency of TCR clonal expansions per metabolic cluster. (**d**) Heatmap of mean expression of key metabolic regulator genes, phenotypic genes and proteins, and top 10 most active metabolic fluxes for each metabolic cluster in pseudobulk. (**e**) Metabolic pathway activities of CD4$^+$ T cell metabolic clusters in pseudobulk among all patients combined, with and without the proliferative and cytotoxic cell clusters (metabolic clusters 6 and 4, respectively). (**f**) Metabolic pathway activities that are significantly correlated with WHO score by two-sided Spearman's rank correlation ($p < 0.05$; labeled on right) for all CD4$^+$ T cells per patient, for only T$_{reg}$ cells, only cytotoxic cells (metabolic cluster 4), and only proliferative cells (metabolic cluster 6). Exact p-values are in Supplementary Table 7. (**g**) Kernel density estimates of expressions of genes included in key energy-producing metabolic pathways for CD4$^+$ T cells among patients with mild disease/healthy subjects (WHO 0–2) and patients with severe disease (WHO 5–7) in pseudobulk and separately for metabolic clusters 4 and 6 and T$_{reg}$ cells.

a Monocyte metabolic regulator expression

b Monocyte phenotypic marker expression (not used for metabolic clustering)

c

d

e Metabolic pathway activity among monocytes by metabolic cluster:

Excluding dead-cell cluster

Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Metabolically defined monocyte clusters have distinct metabolic activities and regulatory programs. (**a**) Unsupervised clustering and UMAP of monocytes using expression of only genes involved in metabolic pathways, colored by metabolic cluster and by expression of key metabolic regulator genes. (**b**) Unsupervised clustering and UMAP of monocytes, colored by all genes-defined cluster and by expression of phenotypic markers. (**c**) Density map of patient disease severity category per cell overlaid onto UMAP plot: mild/healthy (WHO 0–2), moderate disease (WHO 3–4), and severe disease (WHO 5–7). (**d**) Heatmap of mean expression of key metabolic regulator genes, phenotypic genes and proteins, and top 10 most active metabolic fluxes for each metabolic cluster in pseudobulk. (**e**) Metabolic pathway activities of monocyte metabolic clusters in pseudobulk among all patients combined, without inclusion of the dead-cell cluster (metabolic cluster 9).
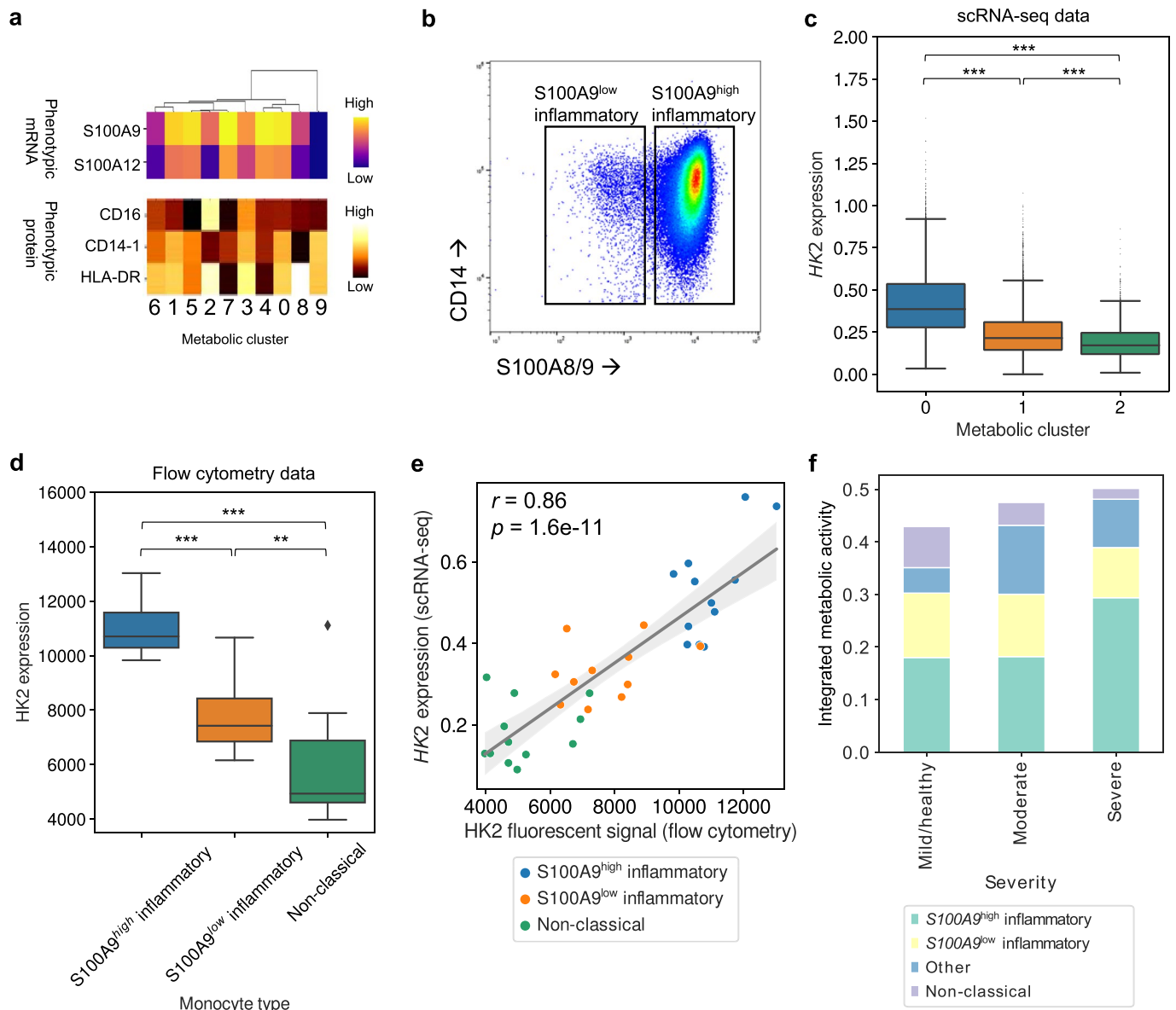
Extended Data Fig. 7 | Flow cytometric analysis of metabolism in inflammatory and non-classical monocyte subpopulations. (**a**) Heatmap summary of markers for defining inflammatory and non-classical monocyte subpopulations. Clusters 0 and 1 are inflammatory monocytes; cluster 2 is non-classical monocytes. (**b**) Representative flow cytometry plot for the gating strategy for S100A9high and S100A9low inflammatory monocytes. Gating was performed on classical monocytes. (**c-d**) Expression of HK2 in S100A9high and S100A9low inflammatory/classical monocytes and non-classical monocytes from (c) scRNA-seq analysis (metabolic cluster 0: $n = 36,651$ cells; 1: $n = 33,242$; 2: $5,601$) and (d) flow cytometry of COVID-19 patient samples ($n = 14$ patient samples for each condition). For (c) and (d), boxes indicate 25th, 50th, and 75th percentiles. Whiskers indicate 1.5 interquartile ranges below and above the 25th and 75th percentiles, respectively. Statistical significances were calculated by two-sided Mann-Whitney U test. **: $p < 0.01$. ***: $p < 0.001$. (**e**) Two-sided Pearson correlation of HK2 expression from scRNA-seq data and from flow cytometry. Shaded band indicates 95% confidence interval of linear regression. (**f**) Integrated metabolic activity of monocytes per disease severity category: mild/healthy (WHO 0–2), moderate (WHO 3–4), and severe (WHO 5–7). The contributions from monocyte subpopulations are shown with cluster-specific color. Metabolic activity is defined as expression levels of metabolic pathway-related genes from each metabolically defined monocyte subpopulation, multiplied by the fraction of monocytes that are in each subpopulation.
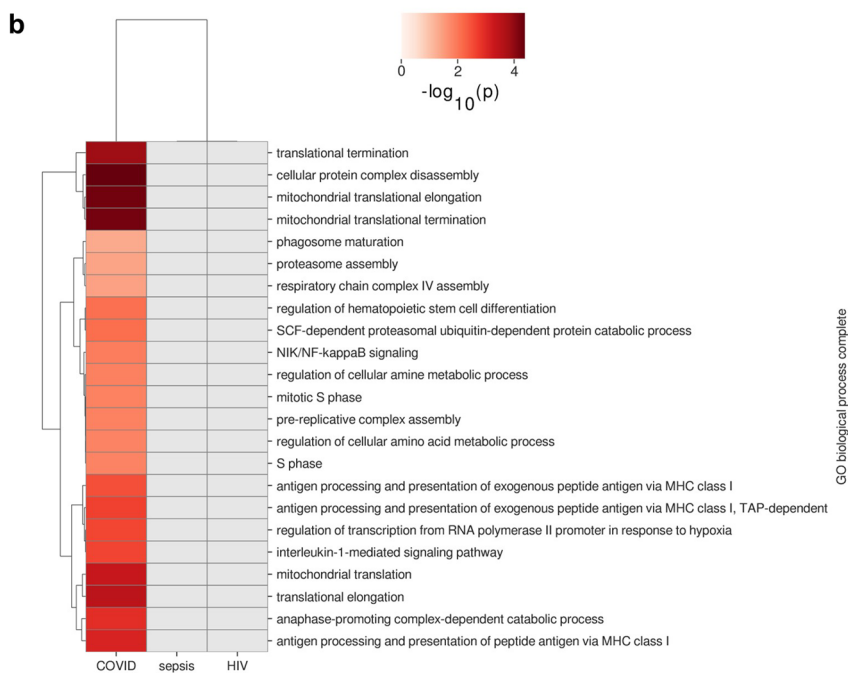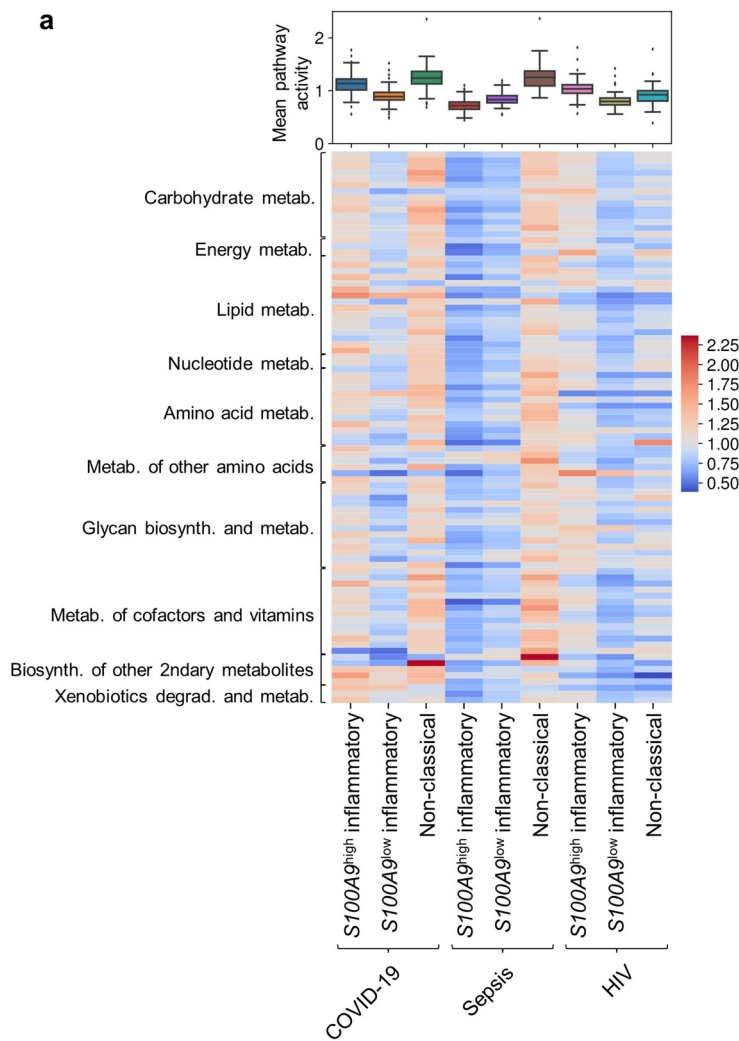
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Comparison of monocyte subtypes between COVID-19, sepsis, and HIV. (**a**) Summary of metabolic pathway activities for metabolically defined monocyte subtypes in each disease. For the box plots, $n = 90$ pathways for each condition. Boxes indicate 25th, 50th, and 75th percentiles. Whiskers indicate 1.5 interquartile ranges below and above the 25th and 75th percentiles, respectively. (**b**) Gene ontology analysis of metabolic genes uniquely elevated in *S100A9*high inflammatory monocytes of COVID-19, showing only the processes that have greater than fivefold enrichment in the COVID-19 condition.
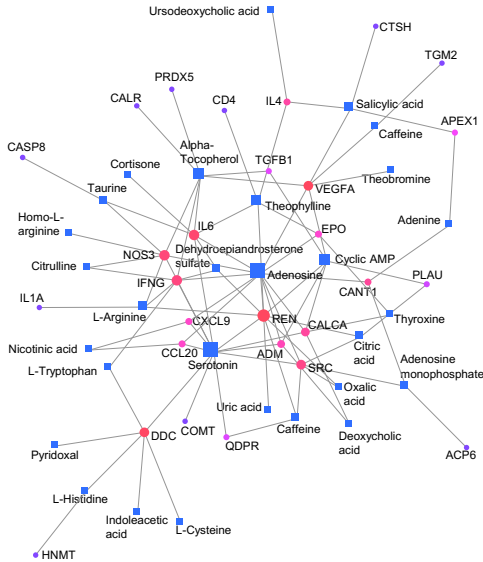
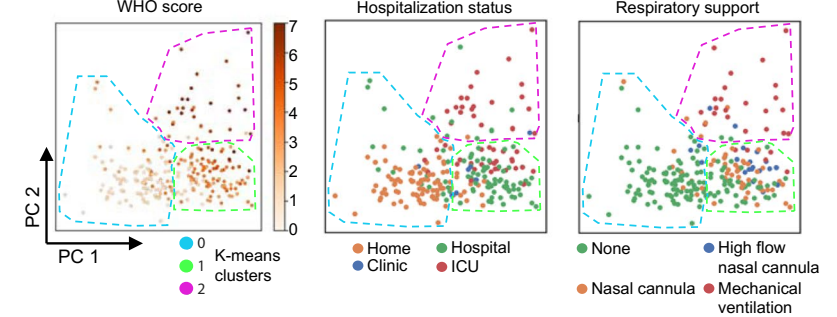**a** Positively severity-correlated proteins and metabolites
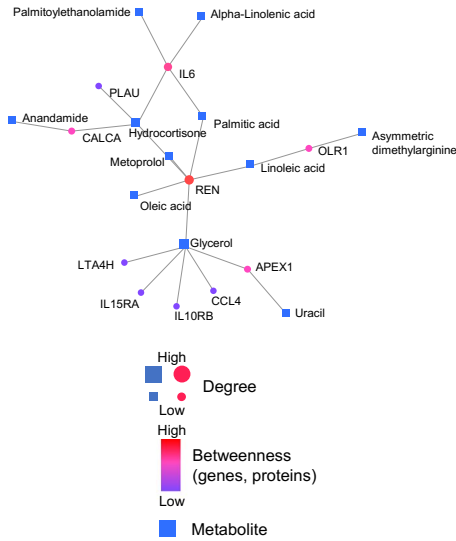
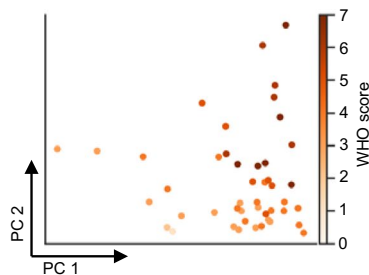**b** Positively severity-correlated transcripts and negatively correlated metabolites
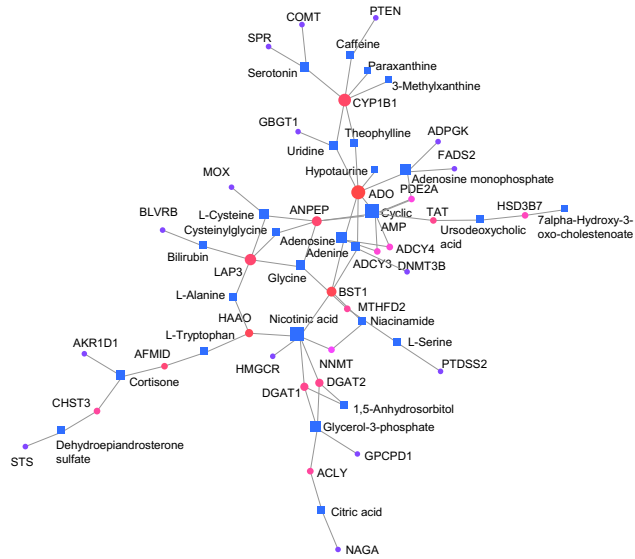
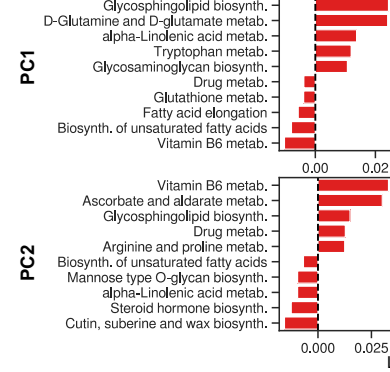**c** Negatively severity-correlated proteins and metabolites

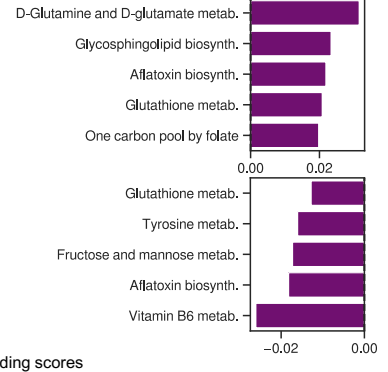**e** PCA of integrated metabolites and metabolic pathways

WHO score   Hospitalization status   Respiratory support

f CD8⁺ T cells   NK cells

Loading scores

d PCA of integrated metabolites and metabolic pathways
Initial cohort (50 samples)
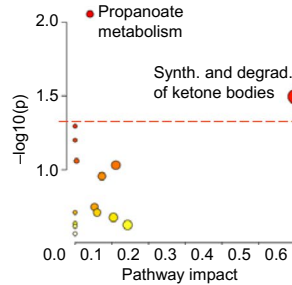

g Pathway enrichment, PCA top metabolites

Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Integrated network of plasma metabolites, plasma proteins, and genes involved in cell metabolic pathways. (**a**) Plasma proteins and metabolites that were positively correlated with disease severity to generate a protein-metabolite network. (**b**) Levels of metabolic pathway-related transcripts were ranked at the patient pseudobulk level, and transcripts significantly positively correlated with WHO score were integrated with plasma metabolites that negatively correlated with WHO score to generate a gene-metabolite network. (**c**) Plasma proteins and metabolites that were negatively correlated with disease severity to generate a protein-metabolite network. (**d**) PCA of per-patient integrated plasma metabolites and metabolic pathways, performed on the cohort of 50 patient samples that had scRNA-seq data available. (**e**) PCA of per-patient integrated plasma metabolites and metabolic pathways, performed on the full cohort of patient samples. Dashed lines indicate well-separated clusters of patient samples by K-means clustering. Samples are colored by WHO score (left), hospitalization status (middle), and level of respiratory support (right). (**f**) Loading scores of cell type-specific metabolic pathway activities for PC1 and PC2 from the integrated PCA for cell types that did not exhibit strong metabolite signatures: CD8[+] T cells and NK cells. (**g**) Pathway enrichment analysis of metabolites among the top five highest and lowest loading scores for each PC (Fig. 6d).
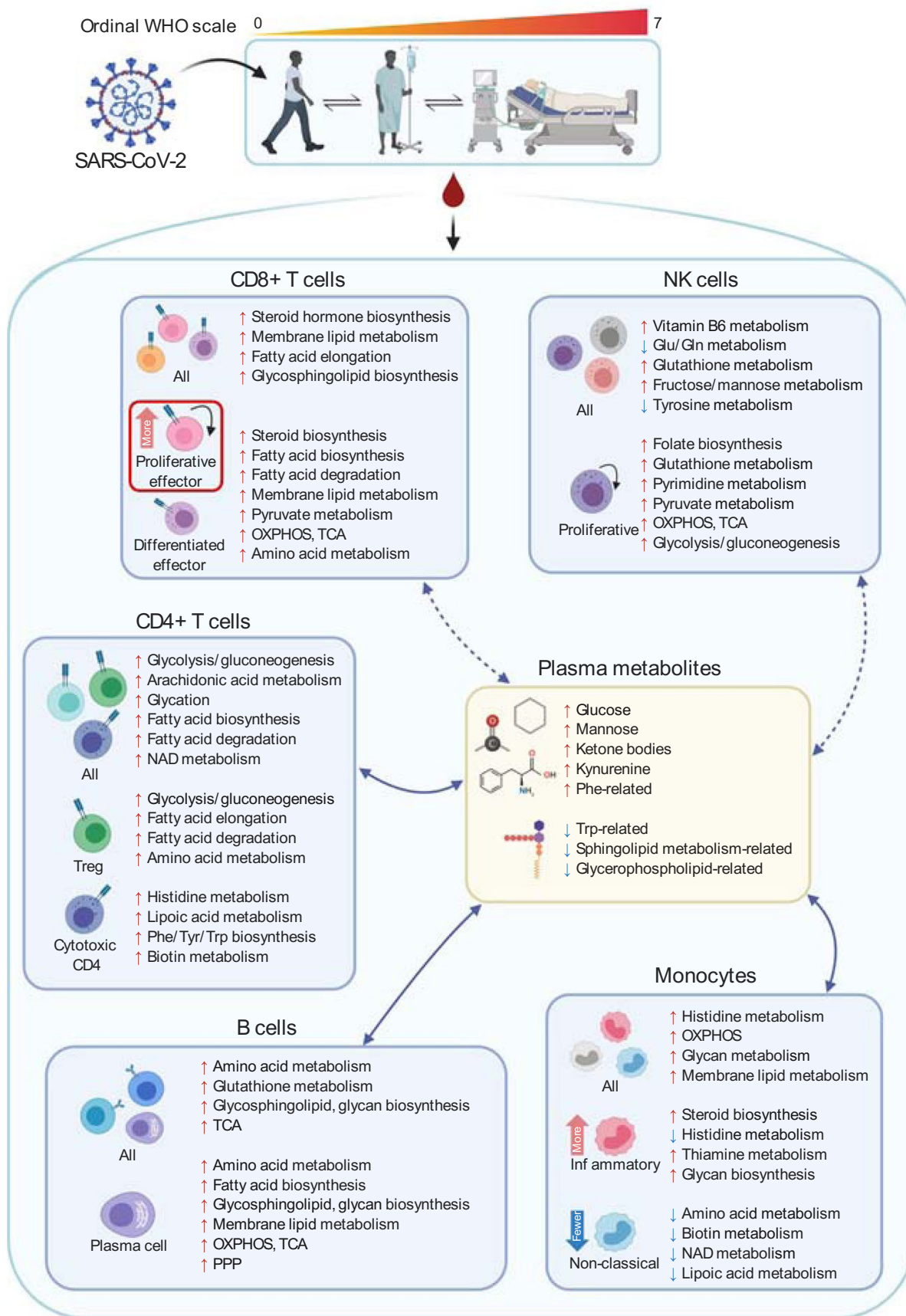
Extended Data Fig. 10 | Summary of metabolic reprogramming in the peripheral immune cell response to COVID-19. As COVID-19 develops in patients, dramatic metabolic changes manifest in the peripheral immune response and can be dissected per cell type. Furthermore, these diverse metabolic alterations can be connected to changes in plasma metabolite levels.

nature portfolio

Corresponding author(s):	James R. Heath, Philip D. Greenberg, Yapeng Su
Last updated by author(s):	Jun 30, 2021

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	All of the COVID-19 data used in this study are either included in the manuscript or were downloaded from ArrayExpress E-MTAB-9357. Data for HIV were downloaded from the Broad Institute Single Cell Project at https://singlecell.broadinstitute.org/single_cell/study/SCP256. Data for sepsis were also downloaded from the Single Cell Project at https://singlecell.broadinstitute.org/single_cell/study/SCP548.
Data analysis	Data analysis was performed using custom code. The custom code used in this study is available on: https://github.com/jihoonlee0/SARS-CoV-2_metab. Software/tools used for this analysis are Python 3, FlowJo 10.7.1, MetaboAnalyst 4.0, MATLAB R2018a, Gurobi optimizer 7.5, and IBM CPLEX 12.7.1.

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

All of the COVID-19 data used in this study are either included in the manuscript or were downloaded from ArrayExpress E-MTAB-9357. Plasma metabolite and protein data are available from Mendeley Data at http://dx.doi.org/10.17632/tzydswhhb5.5 and in Supplementary Tables 2 and 4 of this manuscript. Metabolic pathways, including gene sets involved in a specific metabolic process, were used as defined by the Kyoto Encyclopedia of Genes and Genomes (KEGG; downloaded

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	Not applicable: The manuscript analyzed existing data from previous publications.
Data exclusions	No data excluded
Replication	Not applicable: The manuscript analyzed existing data from previous publications.
Randomization	Not applicable: The manuscript analyzed existing data from previous publications.
Blinding	Not applicable: The manuscript analyzed existing data from previous publications.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used	BV421-CD3, Biolegend, Cat# 300434, clone UCHT1, 1:50 dilution PerCP/Cy5.5-CD4, Biolegend, Cat# 317428, clone OKT4, 1:50 dilution Alexa Fluor 488-CD8, Biolegend, Cat# 344716, clone SK1, 1:50 dilution BV421-HLA-DR, Biolegend, Cat# 307636, clone L243, 1:50 dilution PE-CD14, Biolegend, Cat# 301806, clone M5E2, 1:50 dilution APC/Cy7-CD16, Biolegend, Cat# 302018, clone 3G8, 1:50 dilution PE-Ki67, Biolegend, Cat# 350504, clone Ki-67, 1:50 dilution Alexa Fluor 647-HK2, Abcam, Cat# ab237314, clone EPR20839, 1:50 dilution FITC-S100A8/9, Abcam, Cat# ab7429, clone MAC387, 1:100 dilution
Validation	All antibodies have been tested for specificity for human by their manufacturers. All antibodies except the HK2 antibody have been tested by their manufacturers for flow cytometry; HK2 antibody was validated by manufacturer for ICC/IF. Validation data for all antibodies are available on each manufacturer's website.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Among the INCOV cohort, there were 165 COVID-19 patients and 204 healthy volunteers. COVID-19 patients were aged between 18-89 years (median 56 years) and 53% female. Healthy volunteers were aged 28-70 years (median 56 years) and

50% female. COVID-19 disease severity spanned the full range of the WHO ordinal scale from 0 to 7, including healthy volunteers. Among the COVID-19 patients, median WHO score was 4 at T1. Comorbidities commonly included hypertension (40% of COVID-19 patients) and diabetes mellitus (19%).

The JWCI/SJCI cohort contained 33 subjects. Participants were aged between 28 and 96 years with a median of 66 years. 36% were female. Large fractions of the cohort had hypertension (67%) and/or diabetes mellitus (36%) among other comorbidities. COVID-19 severity spanned from 3 to 7 on the WHO ordinal scale (median 5).

Recruitment	Adult patients were contacted about enrolling into the study after they had received a positive diagnosis of COVID-19 based upon qPCR analysis of nasal swabs. There was no bias except for a positive COVID-19 diagnosis. INCOV participants were compensated with a complimentary T shirt. Patients were not compensated for participation in SJCI.
Ethics oversight	Procedures for the INCOV study were approved by the Institutional Review Board (IRB) at Providence St. Joseph Health with IRB Study Number [STUDY2020000175] and the Western Institutional Review Board (WIRB) with IRB Study Number 20170658. Plasma from SJCI were obtained from patients consented under PH&S IRB approved protocol JWCI-18-0401, PH&S IRB # STUDY2018000254.

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

Flow Cytometry

Plots

Confirm that:

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

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

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

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

Methodology

Sample preparation	Cryopreserved peripheral blood mononuclear cells were thawed and washed once with RPMI (Gibco). Cells were treated with 1x RBC Lysis Buffer to eliminate red blood cells. Cells were treated with Fc blocker and monocyte blocker before being stained in Cell Staining Buffer with the following antibodies. T cell panel: BV421-CD3, PerCP/Cy5.5-CD4, and Alexa Fluor 488-CD8; monocyte panel: BV421-HLA-DR, PE-CD14, and APC/Cy7-CD16. After cell surface staining, cells were fixed in Fixation Buffer and washed with 1x Intracellular Staining Perm Wash Buffer twice before resuspension in 1x Intracellular Staining Perm Wash Buffer with the following antibodies. T cell panel: PE-Ki-67 and Alexa Fluor 647-HK2 (Abcam); Monocyte panel: FITC-S100A8/9 (Abcam) and Alexa Fluor 647-HK2 (Abcam). Cells were washed twice and analyzed on Attune NxT. All antibodies and reagents were purchased from BioLegend, unless otherwise specified.
Instrument	Cells were analyzed on Attune NxT (Thermo Fisher).
Software	Data were analyzed using FlowJo Software version 10.7.1.
Cell population abundance	Not applicable: cells were not sorted in this study.
Gating strategy	Lymphocytes and monocytes were gated on SSC-A/FSC-A, and then on FSC-H/FSC-W to gate on single cells. CD8+ Ki67+ T cells were identified using the single cell gate, and CD3-positive, CD8-positive, Ki67-positive gating strategy. S100A9high classical monocytes were identified using the single cell gate, CD14-positive, CD16-negative, S100A8/9-high gating strategy.

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