The COVIDome Explorer researcher portal

Graphical Abstract

Highlights

- Matched multi-omics datasets are generated for COVID-19 patients and controls
- Transcriptome, proteome, metabolome, immune cell maps, and clinical data are linked
- An online researcher portal enabling data access and analysis is created
- Utility is illustrated by study of biosignatures associated with the biomarker CRP

Authors

Kelly Daniel Sullivan, Matthew Dominic Galbraith, Kohl Thomas Kinning, ..., Angelo D’Alessandro, Kirk Charles Hansen, Joaquin Maximiliano Espinosa

Correspondence
joaquin.espinosa@cuanschutz.edu

In brief

Sullivan et al. describe the development of a multidimensional dataset for the study of COVID-19 known as the COVIDome, which includes diverse clinical, transcriptome, proteome, metabolome, and immune cell datasets. A researcher portal known as the COVIDome Explorer was created to enable global data access and analysis.
COVID-19 pathology involves dysregulation of diverse molecular, cellular, and physiological processes. To expedite integrated and collaborative COVID-19 research, we completed multi-omics analysis of hospitalized COVID-19 patients, including matched analysis of the whole-blood transcriptome, plasma proteomics with two complementary platforms, cytokine profiling, plasma and red blood cell metabolomics, deep immune cell phenotyping by mass cytometry, and clinical data annotation. We refer to this multidimensional dataset as the COVIDome. We then created the COVIDome Explorer, an online researcher portal where the data can be analyzed and visualized in real time. We illustrate herein the use of the COVIDome dataset through a multi-omics analysis of biosignatures associated with C-reactive protein (CRP), an established marker of poor prognosis in COVID-19, revealing associations between CRP levels and damage-associated molecular patterns, depletion of protective serpins, and mitochondrial metabolism dysregulation. We expect that the COVIDome Explorer will rapidly accelerate data sharing, hypothesis testing, and discoveries worldwide.
and red blood cell metabolomes by MS, deep immune phenotyping by mass cytometry (MC) (measuring 100+ immune cell types), and seroconversion assays. All datasets are publicly accessible through a user-friendly, analysis-ready researcher portal dubbed the COVIDome Explorer (https://medschool.cuanschutz.edu/covidome/). Herein, we describe how the datasets were generated and analyzed, and explain how to use the COVIDome Explorer for rapid hypotheses testing, hypothesis generation, and real-time discoveries by experts and non-experts. We illustrate the prowess of the COVIDome dataset by completing a multi-omics analysis of biosignatures associated with varying levels of C-reactive protein (CRP), a clinical marker of poor prognosis in COVID-19 (Liu et al., 2020; Xu et al., 2020). This analysis revealed that high CRP levels associate with damage-associated molecular patterns (DAMPs), depletion of key members of the serpin family of serine protease inhibitors, and metabolic changes indicative of mitochondrial dysfunction.

RESULTS

The COVIDome: A multi-omics dataset for the study of COVID-19

In order to investigate variations in the endotype of COVID-19 patients, we completed a multi-omics assessment of 105 research participants, including 73 hospitalized COVID-19 patients versus 32 COVID-19-negative controls (Figure 1A). The demographics and clinical characteristics of this cohort are described in Table S1. All COVID-19-positive participants were hospitalized due to moderate symptoms, but none had developed severe clinical disease requiring intensive care unit (ICU) admission at the time of blood collection. Of note, some annotated comorbidities were more prevalent in the control group; however, the available clinical data do not distinguish between “history of” and “present at admission,” somewhat limiting the interpretation of these data. COVID-19 positivity was defined from results of PCR and/or antibody testing within 14 days of interpretation of these data. COVID-19 positivity was defined from results of PCR and/or antibody testing within 14 days of hospitalization due to moderate symptoms, but none had developed severe clinical disease requiring intensive care unit (ICU) admission at the time of blood collection.

To test the hypothesis-driven nature of the COVIDome datasets, we identified 2,299 differentially expressed genes (DEGs) in the bloodstream of the COVID-19 patients (Figure 1B). Examples of significantly upregulated DEGs include specific immunoglobulin sequences (e.g., IGHV1-24, IGLV3-1) (Figure 1D). Interactive volcano plots and box-and-whisker plots for this dataset can be generated in the Cytokine dashboard of the COVIDome Explorer at https://covidome.shinyapps.io/Cytokines/. Results of the linear model for MSD data can be found in Table S2D. To investigate metabolic dysregulation in COVID-19, we completed parallel targeted analyses of the RBC and plasma metabolomes using ultra-high-pressure liquid chromatography coupled to MS (UHPLC-MS) (see STAR Methods). RBC and plasma metabolomic signatures inform about different metabolic and physiological processes, with both common and unique metabolites measured in each matrix. Analysis of the RBC metabolome revealed 35 differentially abundant metabolites in COVID-19 patients, such as CXCL10 (C-X-C motif chemokine ligand 10, interferon-inducible protein 10 [IP10]) and interleukin (IL)-10 (Figure 1E). Interactive volcano plots and box-and-whisker plots for this dataset can be generated in the Cytokine dashboard of the COVIDome Explorer at https://covidome.shinyapps.io/Cytokines/. Results of the linear model for MS data can be found in Table S2C.

To generate the cytokine profile dataset, the same aliquot of plasma used for the proteomics analyses was employed to measure the levels of a selected list of immune modulatory factors via multiplexed immunoassays using MSD assays. A linear model adjusting for age and sex revealed many cytokines differentially abundant in the bloodstream of COVID-19 patients, such as CXCL10 (C-X-C motif chemokine ligand 10, interferon-inducible protein 10 [IP10]) and interleukin (IL)-10 (Figure 1E). Interactive volcano plots and box-and-whisker plots for this dataset can be generated in the Cytokine dashboard of the COVIDome Explorer at https://covidome.shinyapps.io/Cytokines/. Results of the linear model for MSD data can be found in Table S2B (SOMAmer proteomics) and Table S2C (MS proteomics). To generate the transcriptome dataset, whole blood was collected in PAXgene RNA tubes, and RNA was extracted and subjected to next-generation sequencing (see STAR Methods). Analysis of the transcriptome dataset using DESeq2 (Love et al., 2014) identified 2,299 differentially expressed genes (DEGs) in the bloodstream of the COVID-19 patients (Figure 1B). Examples of significantly upregulated DEGs include specific immunoglobulin sequences (e.g., IGHV1-24, IGLV3-1), indicative of seroconversion, as well as interferon-stimulated genes (ISGs) (e.g., MX1), indicative of an antiviral transcriptional response. An interactive volcano plot similar to that in Figure 1B enabling real-time data visualization can be found in the Transcriptome dashboard of the COVIDome Explorer at https://covidome.shinyapps.io/Transcriptome/. DESeq2 results can be found in Table S2A. To generate the SOMAmer proteomics dataset, plasma was analyzed with SOMAmer technology to measure the abundance of 4,800+ epitopes corresponding to 3,000+ unique proteins (see STAR Methods). Using a linear model adjusting for age and sex, we identified 970 differentially abundant epitopes in the plasma of COVID-19 patients (Figure 1C). Examples of significantly upregulated proteins include many ISGs, such as ISG15 and IFIT3 (interferon-induced protein with tetratricopeptide repeats 3) (Figure 1C). To generate the MS proteomics dataset, the same plasma aliquot used for SOMAmer proteomics was analyzed by MS (see STAR Methods). This approach enabled the quantification of 412 abundant proteins in plasma (Figure 1D). The MS proteomics dataset is highly complementary to the SOMAmer proteomics dataset, as it enables detection of many abundant proteins for which SOMAmer reagents are not available. For example, analysis of the MS proteomics dataset using a linear model adjusting for age and sex identified 74 differentially abundant proteins, including clear upregulation of immunoglobulin sequences not detected by the SOMAmer but which were also detected as upregulated in the transcriptome dataset (e.g., IGHV1-24, IGLV3-1) (Figure 1D). Interactive volcano plots and box-and-whisker plots for the two proteomics datasets can be generated in the Proteome dashboard of the COVIDome Explorer at https://covidome.shinyapps.io/Proteome/. Results of the linear models described herein can be found in Table S2B (SOMAmer proteomics) and Table S2C (MS proteomics).
metabolomics datasets can be generated in the Metabolome dashboard of the COVIDome Explorer at https://covidome.shinyapps.io/Metabolome/. Results of the linear models for metabolomics can be found in Table S2E (RBC metabolomics) and Table S2F (Plasma metabolomics).

Lastly, we completed a comprehensive map of peripheral immune cell lineages using MC, which enabled the identification and curation of 100+ immune cell subsets (see STAR Methods). Toward this end, we utilized peripheral blood mononuclear cells (PBMCs) purified by Ficoll gradient from the same blood draw used for all other datasets and stained them with a panel of 40 metal-coupled antibodies designed to quantify many major and minor lymphoid and myeloid subsets (see STAR methods). In order to quantify differences in immune cell subsets within their parent lineage, we created seven different immune maps, stemming from (1) all live cells, (2) CD3+ T cells (all T cells), (3) CD4+ T cells, (4) CD8+ T cells, (5) CD19+ B cells, (6) CD11c+ monocytes (CD3 CD19 CD56-), and (7) CD1c+ myeloid dendritic cells.
(mDCs) (CD3⁺ CD19⁺ CD56⁺). Using a linear model adjusting for age and sex, we identified many immune cell types with significantly different frequencies among all live cells in COVID-19-positive patients, such as increased frequencies of plasmablasts and decreased frequencies of CD1c⁺ mDCs (Figure 1G). Interactive volcano plots and box-and-whisker plots for the seven immune maps can be generated in the Immune Maps dashboard of the COVIDome Explorer at https://covidome.shinyapps.io/ImmuneMaps/. Results of the linear models for each immune lineage can be found in Tables S2G–S2M.

In sum, the COVIDome dataset includes major data types for the study of diverse biological processes dysregulated in hospitalized COVID-19 patients.

The COVIDome Explorer: An online portal for real-time data analysis, visualization, and sharing

To facilitate quick and broad access to the COVIDome dataset, we created a user-friendly online portal, dubbed the COVIDome Explorer, which can be accessed online at covidome.org (see overview in Figure 2).

After data curation and quality control, each of the COVIDome datasets was linked at the sample level with a unique identifier, enabling cross-referencing among platforms. Then, each of the datasets was imported into applications developed using R, R Studio, and the R-based web application framework Shiny. Each application includes custom-developed features that enable rapid query, visualization, and download of data, in an interactive environment (see STAR Methods). The COVIDome Explorer hosts six dashboards: Cohort, Transcriptome, Proteome, Cytokines, Metabolome, and Immune Maps. Each dashboard runs within its own isolated and protected environment, hosted on the cloud-based Platform-as-a-Service (PaaS) environment “shinyapps.io.” When a user navigates to a specific dashboard via URL, individual instances of the Shiny application are instantly deployed to the shinyapps.io hosting platform, allowing for interaction and analysis throughout the duration of the user’s session. The Cohort dashboard is a simple description...
of the research cohort involved. The other dashboards are organized in a similar fashion and present similar options for analysis. This similarity allows users to become familiar with one dashboard, and then rapidly adapt to the use of the other dashboards. Each of the five analytical dashboards contains four tabs: Overview, Effect of COVID-19 status, Effect of Seroconversion, and Cross Omics Correlates. The Overview tab provides a summary of the approach, a brief explanation on how to use the dashboards, and, in some instances, links to data files that would guide users, such as catalogs of proteins, metabolites, cytokines, and immune cells present in each dataset. The Overview tab also points to publications that provide further detail about the methodology employed. The Effect of COVID-19 status tab enables users to investigate differences between the COVID-19-negative control cohort and COVID-19-positive patients. The Effect of Seroconversion tab enables users to investigate differences among COVID-19 patients with low versus high titers of anti-SARS-CoV-2 antibodies. A detailed description of the metrics of seroconversion employed and the definition of “sero-low” versus “sero-high” groups can be found in a recent publication illuminating how seroconversion stages COVID-19 into distinct pathophysiological states (Galbraith et al., 2021).

Upon entry into a given dashboard, users must select from a menu of options before data can be displayed. For the Proteome and Metabolome dashboards the first option is the choice of Platform: MS versus SOMAscan for the Proteome, Plasma versus Red Blood Cells for the Metabolome. For the Immune Maps dashboard, the first choice is to select one of seven parent lineages: Live Cells, CD3⁺ T, CD4⁺ T, CD8⁺ T, CD19⁺ B cells, Monocytes, Myeloid DCs. Next, users can choose a statistical test (linear model with age and sex adjustment; Kolmogorov-Smirnov test, Student’s t test, or Wilcoxon test) and an Adjustment Method for multiple hypotheses correction (None, Bonferroni, Benjamini-Hochberg [false discovery rate, FDR]). Users with a pre-formed hypothesis in mind interested in searching for a specific feature of interest (e.g., specific mRNA, protein, or immune cell type) may opt out of a multiple hypothesis adjustment method. In contrast, users exploring the data in an unbiased fashion should select an adjustment method to account for multiple hypotheses testing. Two other filters are sex (both, Male, Female) and Age (All, 21 & Over), which enable users to visualize all or a fraction of the dataset. At this stage, users can “Apply filters and generate plot,” which would then lead to the appearance of an interactive volcano plot displaying the results. Users can then “mouse over and click” individual features in the volcano plot to display a box-and-whisker plot for that specific feature. Alternatively, users can use the searchable menus to find a feature of interest. Once an individual feature has been selected, live links to external databases become available, including PubMed, GeneCards, GTEx, NCBI, and Wikipedia, thus allowing users to navigate away from the COVIDome Explorer and learn more about a gene, protein, cytokine, metabolite, or immune cell type of interest. Of note, both volcano plots and box-and-whisker plots can be downloaded as scalable vector graphics (.svg) files.

In each dashboard, the data being visualized can be accessed through the “Aggregated Data” or “Sample Level Data” tabs, two distinct interactive spreadsheets. In these tabs, users can filter by fold change and p value, sort by any of the columns visible (e.g., gene/protein name, fold change, p value), and search for individual features. Users can then download the data as a comma-separated values file (.csv), Microsoft Excel spreadsheet (.xlsx), or pdf files.

Altogether, the COVIDome Explorer dashboards enable data access and analysis by a broad range of users with different degrees of bioinformatics and biostatistics literacy, from those simply interested in a group comparison for a single protein, to those interested in sophisticated offline analyses of the downloaded datasets.

The COVIDome Explorer enables cross-omics integration and discoveries

To enable cross-platform discoveries, we created a “Cross Omics Correlates” tab in each dashboard that allows users to explore relationships among features both within and across COVIDome datasets for COVID-19-positive patients. To examine connections among features, the Cross Omics Correlates dashboard performs Spearman correlations using age- and sex-adjusted data. When a user accesses the Cross Omics Correlates tab, they navigate to a landing page with numbered instructions for use on the left as well as a “Take Tutorial” option. First, the user selects a query platform when applicable (e.g., MS Proteome versus SOMAscan Proteome); second, the user selects a query feature; third, the user selects a comparison platform; and fourth, the user clicks the “Generate Volcano Plot” button. This renders a volcano plot with correlation data for all features in the comparison platform versus the query feature, where Spearman rho values are represented along the x axis and Benjamini-Hochberg adjusted q values are represented along the y axis. An optional fifth step allows users to visualize XY scatterplots for analyte pairs of interest. For example, the mRNA encoding the viral restriction factor MX1, which is upregulated in the whole blood transcriptome of COVID-19-positive individuals (Figure 1B), is strongly correlated with many plasma proteins measured by SOMAscan (Figure 3A). Among the most positively correlated proteins are numerous other ISGs including CXCL10 and IFIT3 (Figure 3A). MX1 protein is also strongly correlated with MX1 mRNA levels (Figures 3A and 3B). This same method can be applied to any other pairwise cross-omics comparison. For instance, levels of the metabolite kynurenine in RBCs are positively correlated with plasma levels of WARs (tryptophanyl-tRNA synthetase 1) and NADK (nicotinamide adenine dinucleotide kinase), two enzymes involved in tryptophan metabolism (Figures 3C and 3D). WARS is an ISG that may be upregulated as a compensatory mechanism for kynurenine pathway-mediated tryptophan depletion (Sarkar et al., 2007; Adam et al., 2018), and NADK is a kinase that converts NAD⁺ produced by the kynurenine pathway into NAD phosphate (NADP) (Castro-Portuguez and Sutphin, 2020). The Cross Omics Correlates can also be used to compare immune cell subsets with gene expression data. For example, plasmablasts, which are among the most upregulated immune cells in COVID-19-positive individuals, are positively correlated with many immunoglobulin genes in the whole blood transcriptome, such as IGKV4-1, IGHV6-1, and IGLC2, indicative of B cell differentiation toward plasmablasts producing specific immunoglobulin sequences.
targeting SARS-CoV-2 (Figures 3E and 3F). These examples demonstrate the ability of the COVIDome dataset, in concert with the COVIDome Explorer, to reveal potentially biologically meaningful relationships among features across diverse -omics datasets.

CRP levels associate with DAMPs

To illustrate the utility of the matched multidimensional COVIDome datasets, we analyzed multi-omics biosignatures associated with varying levels of CRP, an acute phase protein whose elevation in circulation has been consistently associated with poor prognosis in COVID-19. Repeatedly, higher CRP levels at the time of hospitalization and/or a rapid rise in CRP levels during hospitalization have been associated with increased probability of developing severe COVID-19 pathophysiology (Mousavi-Nasab et al., 2020; Mueller et al., 2020; Sharifpour et al., 2020). As expected, CRP is also elevated in our cohort of COVID-19 patients as measured by MS proteomics, as well as other acute phase proteins such as ferritin (FTL) (Figure 4A). To identify biosignatures associated with CRP levels among COVID-19-positive patients, we calculated Spearman correlations between CRP values measured by MS and all features in all COVIDome datasets, which revealed myriad mRNAs, proteins, and metabolites significantly associated with CRP levels (Figures S2A–S2F; Tables S3A–S3F). This analysis exercise confirmed known associations, such as positive correlations between CRP levels and the levels of serum amyloid proteins SAA1 and SAA2, the acute phase protein LBP (lipopolysaccharide binding protein), and the cytokines IL-6 and IL-10 (Jain et al., 2011). Notably, there were no significant associations between CRP levels and frequencies of immune cell types, neither among all live cells nor within major lymphoid and myeloid lineages, with the sole exception of increased frequencies of inflammatory subsets of monocytes (Figures S2G–S2H; Tables S3G–S3M).

In order to investigate associations between CRP levels and underlying pathophysiological processes, we first performed Metascape pathway enrichment analysis of the positively correlated proteins measured by SOMAscan. Somewhat unexpectedly, this analysis revealed enrichment of several groups of proteins associated with immune activation, such as signatures associated with systemic lupus erythematosus (SLE, e.g.,

Figure 3. Cross-omics correlations enabled by the COVIDome Explorer
(A, C, and E) (Top) Volcano plots for Spearman correlations between (A) MX1 mRNA levels and SOMAscan proteomics, (C) kynurenine levels from red blood cell (RBC) metabolomics and SOMAscan proteomics, and (E) plasmablast frequency and transcriptomics. The horizontal dashed lines indicate the statistical cutoff of q < 0.1 (FDR10). Numbers in the left and right quadrants indicate the number of features passing the statistical cutoff. (B, D, and F) (Bottom) Scatterplot for correlations of (B) MX1 mRNA levels with MX1 protein levels, (D) RBC kynurenine levels with WARS protein levels in plasma, and (F) plasmablast frequency with IGKV4-1 mRNA levels. Points are colored by density; lines represent linear model fit with 95% confidence interval. Sample sizes range from 65 to 71 depending on the platform.
Figure 4. CRP levels correlate with damage-associated molecular patterns

(A) Sina plots showing values for CRP and ferritin light chain (FTL) measured by MS proteomics comparing COVID-19-negative (−) to COVID-19-positive (+) patients. Data are presented as modified sina plots with boxes indicating median and interquartile range.

(B) Scatterplots displaying correlations between CRP levels versus SAA1, LBP, and IL-10. MSD, Meso Scale Discovery assay. Points are colored by density; lines represent linear model fit with 95% confidence interval.

(C) Metascape pathway enrichment analysis of proteins detected by SOMAscan proteomics that are significantly and positively correlated with CRP.

(D) Scatterplot displaying correlations between CRP levels and representative factors from systemic lupus erythematosus (CXCL10) and positive regulation of Th2 cytokine (IL-6) signatures. Points are colored by density as in (B); lines represent linear model fit with 95% confidence interval.

(E) Heatmap displaying changes in circulating levels of proteins in the DNA Methylation signature that are significantly positively correlated with CRP levels. The left column represents Spearman rho values for correlation with CRP levels, while the right columns display median Z scores for each feature for COVID-19-

(legend continued on next page)
CXCL10), positive regulation of T helper 2 (Th2) cytokines (e.g., IL-6), acute inflammatory response, cytolyis (composed mostly of complement subunits), and the IFN-γ-mediated signaling pathway (Figures 4C and 4D). Interestingly, this analysis also identified protein signatures associated with DNA methylation and response to heat. The DNA methylation group is comprised of 19 features including chromatin-associated factors (e.g., DNMT3L, DPY30, SUSD3, RBBP4, RBBPS) and several histones (Figure 4E). For example, H2AFZ (H2AZ) is significantly correlated with CRP and elevated in the plasma of COVID-19 patients (Figure 4F). The Response to Heat signature has 12 features, 4 of which are heat shock proteins (HSPs) (HSPA1A, HSPA1B, HSP90AA1, HSPH1) including HSPA1A (HSP72), which is both significantly correlated with CRP and elevated in COVID-19 (Figures 4G and 4H). Interestingly, both histones and HSPs can function as DAMP molecules whose presence in the bloodstream is consistently associated with tissue damage and trauma, and which in turn function as ligands for amplification of innate immune signaling (Huang et al., 2011). Circulating histones can be released from dying cells in the liver and they have been shown to drive downstream damage to both pulmonary and hepatic endothelial cells (Kawai et al., 2016). Histones in the bloodstream could also be interpreted as a sign of netosis and formation of neutrophil extracellular traps (NETs) (Papayannopoulos, 2008), one of the H2B-encoding genes, and the fifth most correlated mRNA is H2BC12, one of the H2B-encoding genes, and the fifth most correlated gene is H2AC19, one of the H2A-encoding genes, both of which are elevated in the whole blood transcriptome of COVID-19-positive patients (Figure S2I). Given that the mRNAs captured in this transcriptome analysis are derived from circulating immune cells and that histone mRNAs are transcribed during the S phase of the cell cycle, this could be interpreted as a sign of stronger immune cell activation and proliferation in patients with higher CRP levels.

Altogether, these results indicated that the prognostic value of high CRP levels in COVID-19 could be potentially tied to the accompanying depletion of important protective serpins and consequent dysregulation of the coagulation, fibrinolysis, and complement cascades, both of which have been involved in the etiology of severe COVID-19 pathology (Lo et al., 2020).
Figure 5. CRP levels correlate with depletion of protective serpins

(A and B) Correlation analysis of CRP with SERPINA5 (A) and SERPINA4 (B). (Left) Scatterplot for correlation of CRP with the indicated SOMAmer reagent. Points are colored by density; lines represent linear model fit with 95% confidence interval. (Right) Sina plot for indicated SOMAmer reagent with boxes indicating median and interquartile range.

(C) Heatmap displaying changes in circulating levels of complement and coagulation proteins significantly correlated with CRP levels with an absolute rho value greater than 0.3. The left column represents Spearman rho values, while the right columns display median Z scores for each feature for COVID-19-negative patients.

(legend continued on next page)
can be oxidized to pyruvate by lactate dehydrogenase, was also positively correlated with CRP (Figure 6A). Increases in lactate and pyruvate can be interpreted as increased glycolysis in patients with high CRP, perhaps driven by hypoxia leading to carbon flow from pyruvate to acetyl-coenzyme A (CoA). Increased glycolysis is a metabolic consequence of both immune cell activation and hypoxemia (Makowski et al., 2020; Frauwirth et al., 2002; Michalek et al., 2011; Jellusova, 2020; van Teijlingen Bakker and Pearce, 2020). Importantly, each of these three classes of metabolites represent entry points to the TCA cycle, and elevated levels of these features are consistent with mitochondrial dysfunction and decreased activity in the TCA cycle and the electron transport chain (ETC).

Given that these metabolic associations between CRP and plasma metabolites could be due to metabolic dysregulation in peripheral blood cells and/or various host tissues, we asked whether these associations could be explained by gene expression changes in circulating blood cells by analyzing the whole blood transcriptome dataset. We used the Ingenuity Pathway Analysis (IPA) software to identify gene sets enriched among the RNAs positively and negatively correlated with CRP, with a focus on metabolic pathways. Strikingly, the most significantly enriched metabolic pathway among negatively correlated mRNAs is oxidative phosphorylation (OXPHOS) (Figure 6B; see Figure S4A for positively correlated gene sets). The OXPHOS gene signature is comprised of 54 genes including many components of the ETC, such as reduced nicotinamide adenine dinucleotide (NADH):ubiquinone oxidoreductase subunits, cytochrome c complex subunits, and ATP synthase subunits, among others (Figure 6C). For example, expression of NDUVF3 (NADH:ubiquinone oxidoreductase subunit V3, complex I, mitochondrial respiratory chain, 10-kDa subunit) and COX411 (cytochrome c oxidase subunit 411) are both negatively correlated with CRP levels and significantly decreased in COVID-19 patients, as is the mitochondrial encoded cytochrome c oxidase III (MT-CO3) (Figure 6D). Therefore, accumulation of TCA carbon sources in plasma could be linked to decreased gene expression of ETC components in circulating blood cells. Interestingly, we noticed that the mRNAs encoding the glucose transporter SLC2A3 (GLUT3) and the monocarboxylate transporter SLC16A3 (monocarboxylate transporter 4 [MCT4]) were both positively and significantly correlated with CRP in the whole blood transcriptome of COVID-19 patients (Figure S4B). Increases in surface expression of SLC2A3 have been noted during activation of diverse lymphocytes, neutrophils, and platelets, and they are thought to mediate increased glucose uptake to fuel cell activation (Simpson et al., 2006). SLC16A3 catalyzes the bidirectional transport across the plasma membrane of many monocarboxylates such as lactate, pyruvate, as well as branched-chain oxo acids derived from leucine, valine, and isoleucine. In innate immune cells, lactate is produced and exported in large amounts via SLC16A3 during pro-inflammatory responses, and its expression is upregulated in activated macrophages (Weiss and Angiari, 2020). Furthermore, SLC16A3 expression is necessary for macrophage activation, as its deletion results in intracellular accumulation of lactate and decreased glycolysis (Weiss and Angiari, 2020).

Altogether, the metabolic changes associated with CRP could be understood, in part, as the byproduct of metabolic remodeling of circulating blood cells, whereby decreased expression of OXPHOS genes and increased expression of glucose and monocarboxylate transporters would lead to increased glucose uptake, decreased OXPHOS, and consequent accumulation of glycolysis end products (lactate, pyruvate) and other carbon sources for the TCA cycle (branched-chain amino acids, acyl carnitines).

**DISCUSSION**

The global health crisis imposed by the COVID-19 pandemic has inspired new approaches for rapid collaboration, open access to manuscripts under review, and data sharing. Herein, we describe the rapid creation of a user-friendly researcher portal enabling easy access and real-time analysis of matched multi-omics datasets for COVID-19. The first batch of biospecimens for the COVIDome project was received by this team in July 2020, and the COVIDome Explorer was publicly launched in November 2020, thus spanning only 5 months from sample processing to portal launch. Between November 2020 and June 2021, more than 800 unique users from 36 countries had utilized the portal according to session data gathered from Google Analytics. Currently, the second batch of samples is being subjected to identical multi-omics analyses. Importantly, the COVIDome Explorer can easily ingest datasets from other teams to be displayed in its dashboards, which would then enable the comparison of results across different studies.

With the advent of multi-omics platforms, it is now possible to rapidly investigate hundreds of molecular, cellular, and physiological processes from a single biospecimen. Such a systems biology approach enables the integration of findings across different methodologies and layers of biological information to expedite the pace of discovery into the etiology of a medical condition. In this report, we illustrate the power of this approach by exploring biosignatures associated with CRP, a well-characterized marker of inflammation across numerous medical conditions, including COVID-19. Although it is well established that CRP levels and trajectory have prognostic value in COVID-19 (Mousavi-Nasab et al., 2020; Mueller et al., 2020; Sharifpour et al., 2020), the exact pathophysiological processes associated with this clinical biomarker of inflammation remain to be fully elucidated. What exactly is being revealed by high baseline levels and/or rapid elevation of CRP in COVID-19? Our analysis demonstrates that, in addition to the well-established links between CRP and other markers of inflammation and immune activity, CRP levels associate with DAMPs, depletion of protective controls (–) versus COVID-19-positive patients (+). Z scores were calculated from the adjusted values for each SOMAmer in each sample, based on the mean and standard deviation of COVID-19-negative samples. Asterisks indicate a significant difference between COVID-19 patients and the control group. (D–G) Scatterplots and sina plots as in (A) for KLKB1, KLK13, C9, and C3, respectively. q values in each are derived from linear models. Sample sizes range from 30 to 31 for COVID-19-negative controls and from 69 to 71 for COVID-19-positive patients, depending on the platform.
Figure 6. CRP levels correlate with dysregulated mitochondrial metabolism in blood cells

(A) Scatterplot displaying correlations between CRP levels and indicated metabolites. Points are colored by density; lines represent linear model fit with 95% confidence interval.

(B) Histogram displaying the results of Ingenuity Pathway Analysis (IPA) of metabolic pathways for mRNAs measured in the whole blood transcriptome analysis that are significantly and negatively correlated with CRP.

(C) Heatmap displaying expression changes in mRNAs in the oxidative phosphorylation (OXPHOS) IPA signature from (B). The left column represents Spearman rho values for correlations with CRP, while the right columns display median Z scores for each feature for COVID-19-negative controls (+/C0) versus COVID-19-positive patients (+). Z scores were calculated from the adjusted RPKM (reads per kilobase transcript per million mapped reads) values for each mRNA in each sample, based on the mean and standard deviation of COVID-19-negative samples. Asterisks indicate a significant difference between COVID-19 patients and the control group.

(legend continued on next page)
serpins, and dysregulation of mitochondrial metabolism in blood cells in COVID-19.

The association between CRP and DAMPs reveals that CRP levels inform about the extent of tissue damage in COVID-19. High levels of CRP associate with increased circulating levels of intracellular proteins released into the bloodstream during organ damage (e.g., histones, HSPs), which can further exacerbate the inflammatory phenotype. In turn, increased tissue damage could be conceptually tied to the clear depletion of the protective serpins SERPINA4 and SERPINA5, the most anti-correlated proteins with CRP in our proteomics datasets. SERPINA4/5 depletion could lead to exacerbated, harmful levels of activity within the coagulation system, kallikrein-kinin system, and complement cascade, all of which can contribute to COVID-19 pathology. Depletion of SERPINA5 could unleash a protease storm within the coagulation cascade, leading to coagulopathies and thromboembolism in COVID-19 (Becker, 2020). Given that both SERPINA4 and SERPINA5 inhibit KLKs, the serine proteases driving production of the vasoactive peptide bradykinin, their depletion could contribute to the so-called “bradykinin storm” in COVID-19 linked to accumulation of fluids in the lungs and respiratory failure (Garvin et al., 2020).

Lastly, since KLKs also activate the complement cascade, SERPINA4/5 depletion could lead to harmfully high levels of complement activity and consequent tissue damage by the membrane attack complex (MAC). Of note, all of these processes are suitable to pharmacological modulation and are the focus of many ongoing clinical trials testing the efficacy of blood thinners (Rentsch et al., 2021), kinin receptor antagonists (van de Veerdonk et al., 2020), and complement inhibitors (Mastellos et al., 2021) in COVID-19. Therefore, we posit that CRP could serve as a biomarker to stratify the patient cohorts in these clinical trials to assess potential differences between individuals with varying CRP levels. We also hypothesize that SERPINA4 and/or SERPINA5 administration could be a valid therapeutic strategy in COVID-19 to reduce organ damage, especially in patients with high CRP levels (Rau et al., 2007; Suzuki, 2008).

Interpretation of the association between CRP levels and markers of dysregulated mitochondrial metabolism must consider a combination of metabolic effects on circulating blood cells and host tissues. Plasma metabolomics can inform about metabolic alterations in the peripheral immune cell repertoire, platelets and RBCs, but also about dysregulated metabolism in various organs. CRP levels correlated with increased levels of three different carbon sources for the TCA cycle: branched chain amino acids (leucine, isoleucine), end products of glycolysis (lactate, pyruvate), and acyl carnitines, all of which could be explained by decreased activity in the TCA cycle and ETC. Indeed, when analyzing the transcriptome of circulating immune cells, the top gene signature negatively associated with CRP was OXPHOS, with expression of many components of the ETC being downregulated in patients with high CRP. Furthermore, these changes were accompanied by increased mRNA expression of the glucose transporter SLC2A3 and the monocarboxylate transporter SLC16A3, which can be associated with activation of different immune cell subsets (Weiss and Angiari, 2020; Simpson et al., 2008). Dysregulation of mitochondrial metabolism is increasingly appreciated in COVID-19 (Bertscher et al., 2020). Notably, disruption of the TCA cycle has been reported downstream of inflammatory stimulation via a mechanism shunting citrate to succinate, driving additional inflammation, largely in myeloid cells (Tannahill et al., 2013; Mills et al., 2016; Makowski et al., 2020). Importantly, downregulation of OXPHOS and ETC genes has been demonstrated in the liver in the case of hepatitis C infection (Gerresheim et al., 2019), and in the diaphragm, liver, and peripheral blood during sepsis (Callahan and Supinski, 2005; Weiss et al., 2014; Eyenga et al., 2014).

In sum, the COVIDome datasets and the COVIDome Explorer facilitate rapid hypothesis generation and testing, revealing unexpected associations between diverse molecular, cellular, and pathophysiological processes in COVID-19, even for well-studied factors such as CRP.

**STAR METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
  - Blood processing
  - Whole blood transcriptome
  - Plasma proteomics by mass spectrometry
  - Plasma proteomics by SOMAscan assays
  - Cytokine profiling and seroconversion by multiplex immunoassay
  - Mass cytometry analysis of immune cell types
  - Mass spectrometry-based metabolomics of plasma and red blood cells
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Analysis of transcriptome data
  - Analysis of MS-proteomic data
  - SOMAscan data
  - Analysis of MSD cytokine profiling data
  - Analysis of LCMS-metabolomics data
  - Analysis of mass cytometry data
  - Differential abundance analysis

(D) (Left) Scatterplots for correlations between CRP and the indicated mRNAs. Points are colored by density as in (A); lines represent linear model fit with 95% confidence interval. (Right) Sina plots for indicated mRNAs with boxes indicating median and interquartile range. q values in each sina plot are from DESeq2. Sample size is 30 for COVID-19-negative controls and from 65 to 71 for COVID-19-positive patients, depending on the platform.

(E) Summary of findings indicating dysregulation of mitochondrial metabolism in the bloodstream of COVID-19-positive patients with CRP levels. Created with graphic elements from BioRender.com.
SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2021.109527.

ACKNOWLEDGMENTS

We are grateful to Dr. Thomas Flaig and the Office of the Vice Chancellor for Research at the University of Colorado Anschutz Medical Campus for their leadership in setting up the COVID-19 Biobank at the University of Colorado and also to the COVID-19 Biobank Steering Committee for overall support of this project. We thank members of the Biorepository Shared Resource, especially Dr. Adrie Van Bokhoven, Zachary Grasmick, and Hannah Schumman; members of the Human Immune Monitoring Shared Resource, especially Dr. Jill Slansky, Jodi Livesay, Troy Schedlin, and Dr. Jennifer McWilliams; members of the Flow Cytometry Shared Resource of the University of Colorado Cancer Center, especially Dr. Eric Clambey, Alistair Acosta, Christine Childs, and Kristina Terrell; as well as Dr. Aaron Issaian for assistance with MS proteomics data analysis. We also thank the SomaLogic team for their support and the Meso Scale Discovery team for generous support with seroconversion assays. We are grateful to Dr. Ian Brooks, Dr. Michelle Edelmann, and the Health Data Compass Data Warehouse project (https://www.healthdatacompass.org) for the clinical data. This work was supported by NIH grants R01AI150305, 3R01AI150305-01S1, UL1TR002535, 3UL1TR002535-03S2, R01HL148442, R01HL149714, R01HL148151, R21HL150032, P30CA046934, R35GM124939, RM1GM131968, and K23AR070897, as well as grants from the Boettcher Foundation and Fast Grants. Additional generous support was received from the Chancellor’s Discovery Innovation Fund at the CU Anschutz Medical Center, the Global Down Syndrome Foundation, the Anna and John J. Sie Foundation, the LGA Foundation, and Lyda Hill Philanthropies.

AUTHOR CONTRIBUTIONS

J.M.E. designed the project and organized the multiple collaborations required for creation of the COVIDome dataset project and the COVIDome Explorer. K.W.B., N.C.L., and M.G.M. built the COVIDome database and Researcher Portal. K.D.S., M.D.G., K.T.K., P.A., K.P.S., R.E.G., R.M.B., K.R.J., S.A.R., M.E.D., J.A.R., R.E.G., T.G., A.A.M., T.D.B., E.W.-Y.H., A.D., K.C.H., and J.M.E. designed experiments and analyzed data. K.D.S. and J.M.E. wrote the manuscript. All authors reviewed the manuscript.

DECLARATION OF INTERESTS

J.M.E. serves in the COVID Development Advisory Board for Elly Lilly and has provided consulting services to Gilead Sciences Inc. J.M.E. also serves on the Cell Reports Advisory Board. The remaining authors declare no competing interests.

INCLUSION AND DIVERSITY

We worked to ensure gender balance in the recruitment of human subjects. One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in science. One or more of the authors of this paper self-identifies as a member of the LGBTQ+ community. One or more of the authors of this paper received support from a program designed to increase minority representation in science.

Received: March 4, 2021
Revised: June 30, 2021
Accepted: July 22, 2021
Published: July 28, 2021

REFERENCES

Adam, I., Dewi, D.L., Mooiweer, J., Sadik, A., Mohapatra, S.R., Berdel, B., Keil, M., Sonner, J.K., Thedieck, K., Rose, A.J., et al. (2018). Upregulation of tryptophan-tRNA synthetase adapts human cancer cells to nutritional stress caused by tryptophan degradation. Oncolmunology 7, e1486535.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—A Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169.

Buckner, R.C. (2020). COVID-19 update: Covid-19-associated coagulopathy. J. Thromb. Thrombolysis 50, 54–67.

Burtscher, J., Cappellano, G., Omori, A., Koshiba, T., and Millet, G.P. (2020). Mitochondria: In the Cross Fire of SARS-CoV-2 and Immunity. iScience 23, 101631.

Bushnell, B., Rood, J., and Singer, E. (2017). BBMerge—Accurate paired shotgun read merging via overlap. PLoS ONE 12, e0185056.

Callahan, L.A., and Supinski, G.S. (2005). Downregulation of diaphragm electron transport chain and glycolytic enzyme gene expression in sepsis. J. Appl. Physiol. (1985) 99, 1120–1126.

Castro-Portuguez, R., and Sutphin, G.L. (2020). Kynurenine pathway, NAD+ synthesis, and mitochondrial function: Targeting tryptophan metabolism to promote longevity and healthspan. Exp. Gerontol. 132, 110841.

Chao, J., Biedsoe, G., and Chao, L. (2016). Protective role of kallistatin in vascular and organ injury. Hypertension 68, 533–541.

Chong, J., Wishart, D.S., and Xia, J. (2019). Using MetaboAnalyst 4.0 for comprehensive and integrative metabolomics data analysis. Curr. Protoc. Bioinformatics 68, 486.

De Livera, A.M., Dias, D.A., De Souza, D., Rupasinghe, T., Pyke, J., Tull, D., Roessner, U., McConville, M., and Speed, T.P. (2012). Normalizing and integrating metabolomics data. Anal. Chem. 84, 10768–10776.

Ecke, S., Geiger, M., Resch, I., Jerabeck, I., Sting, L., Maier, M., and Binder, B.R. (1992). Inhibition of tissue kallikrein by protein C inhibitor. Evidence for identity of protein C inhibitor with the kallikrein binding protein. J. Biol. Chem. 267, 7048–7052.

Eyenga, P., Rousset, D., Morel, J., Rey, B., Romestaint, C., Teulier, L., Sheu, S.S., Goudable, J., Négrier, C., and Vale, J.P. (2014). Early septic shock induces loss of oxidative phosphorylation yield plasticity in liver mitochondria. J. Physiol. Biochem. 70, 285–296.

Finck, R., Simonds, E.F., Jager, A., Krishnaswamy, S., Sachs, K., Fanti, W., Pe‘er, D., Nolan, G.P., and Bendall, S.C. (2013). Normalization of mass cytometry data with bead standards. Cytometry A 83, 483–494.

Fauwirth, K.A., Riley, J.L., Harris, M.H., Parry, R.V., Rathmell, J.C., Plas, D.R., Finck, R., Simonds, E.F., Jager, A., Krishnaswamy, S., Sachs, K., Fanti, W., Pe‘er, D., Nolan, G.P., and Bendall, S.C. (2013). Normalization of mass cytometry data with bead standards. Cytometry A 83, 483–494.

Fauwirth, K.A., Riley, J.L., Harris, M.H., Parry, R.V., Rathmell, J.C., Plas, D.R., Finck, R., Simonds, E.F., Jager, A., Krishnaswamy, S., Sachs, K., Fanti, W., Pe‘er, D., Nolan, G.P., and Bendall, S.C. (2013). Normalization of mass cytometry data with bead standards. Cytometry A 83, 483–494.

Gafbraith, M.D., Kinning, K.T., Sullivan, K.D., Baxter, R., Araya, P., Jordan, K.R., Russell, S., Smith, K.P., Granath, R.E., Shaw, J.R., et al. (2021). Seroconversion stages COVID19 into distinct pathophysiological states. eLife 10, e65508.

Garvin, M.R., Alvarez, C., Miller, J.I., Prates, E.T., Walker, A.M., Aronow, B., Mast, A.E., Justice, A., Aronow, B., and Jacobson, D. (2020). A mechanistic model and therapeutic interventions for COVID-19 involving a RAS-mediated bradykinin storm. eLife 9, e59177.

Gershem, G.K., Roeb, E., Michel, M.A., and Niepman, M. (2019). Hepatitis C virus downregulates core subunits of oxidative phosphorylation, reminiscent of the Warburg effect in cancer cells. Cells 8, 1410.

Gold, L., Walker, J.J., Wilcox, S.K., and Williams, S. (2012). Advances in human proteomics at high scale with the SOMaScan proteomics platform. N. Biotechnol. 29, 543–549.

Gu, Z., Eills, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32, 2847–2849.

Huang, H., Evankovich, J., Yan, W., Nace, G., Zhang, L., Ross, M., Liao, X., Bil- liar, T., Xu, J., Esmon, C.T., and Tsung, A. (2011). Endogenous histones
function as alarmins in sterile inflammatory liver injury through Toll-like receptor 9 in mice. Hepatology 54, 999–1008.

Huber, W., Carey, V.J., Gentleman, R., Anders, S.,Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., et al. (2015). Orchestration high-throughput genomic analysis with Bioconductor. Nat. Methods 12, 115–121.

Irmscher, S., Döring, N., Haider, L.D., Jo, E.A.H., Kopka, I., Dunker, C., Jacobsen, I.D., Luo, S., Sleevogt, H., Lorkowski, S., et al. (2018). Kallikrein cleaves C3 and activates complement. J. Innate Immun. 10, 94–105.

Jain, S., Gautam, V., and Naseem, S. (2011). Acute-phase proteins: As diagnostic tool. J. Pharm. Bioallied Sci. 3, 118–127.

Jellusova, J. (2020). The role of metabolic checkpoint regulators in B cell survival and transformation. Immunol. Rev. 295, 39–53.

Johnson, M., Wagstaffe, H.R., Gilmour, K.C., Mai, A.L., Lewis, J., Hunt, A., Sirr, J., Bengt, C., Grandjean, L., and Goldblatt, D. (2020). Evaluation of a novel multiplexed assay for determining IgG levels and functional activity to SARS-CoV-2. J. Clin. Virol. 130, 104572.

Kawai, C., Kotani, H., Miyao, M., Ishida, T., Jemail, L., Abiru, H., and Tamaki, K. (2016). Circulating extracellular histones are clinically relevant mediators of multiple organ injury. Am. J. Pathol. 186, 829–843.

Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. 37, 907–915.

Li, H., Handsaker, B., Wysocki, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.: Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079.

Liu, F., Li, L., Xu, M., Wu, J., Luo, D., Zhu, Y., Li, B., Song, X., and Zhou, Y. (2020). Prognostic value of interleukin-6, C-reactive protein, and procalcitonin in patients with COVID-19. J. Clin. Virol. 127, 104370.

Lo, M.W., Kemper, C., and Wooddruff, T.M. (2020). COVID-19: Complement, coagulation, and collateral damage. J. Immunol. 205, 1488–1495.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550.

Makowski, L., Chab, M., and Rathmell, J.C. (2020). Immunometabolism: From basic mechanisms to translation. Immunol. Rev. 295, 5–14.

Martin-Murphy, B.V., Holt, M.P., and Ju, C. (2010). The role of damage associated molecular pattern molecules in acetaminophen-induced liver injury in mice. Toxicol. Lett. 192, 387–394.

Mastellos, D.C., Skendros, P., Calado, R.T., Risitano, A.M., and Lambris, J.D. (2011). Cutting edge: Distinct glycolytic and lipid oxidative metabolic programs are essential for effector and regulatory CD4+ T cell subsets. J. Immunol. 186, 3293–3299.

Michalek, R.D., Gerriets, V.A., Jacobs, S.R., Macintyre, A.N., MacIver, N.J., Michalek, R.D., Gerriets, V.A., Jacobs, S.R., Macintyre, A.N., MacIver, N.J., Alcorn, C., and Hsieh, E.W.Y. (2019). Minimizing batch effects in mass cytometry data. Front. Immunol. 10, 2367.

Nemkov, T., Reisz, J.A., Gehrke, S., Hansen, K.C., and D’Alessandro, A. (2019). High-throughput metabolomics: Isotopic and gradient mass spectrometry-based methods. Methods Mol. Biol. 1978, 13–26.

O’Connor, N.P., Abramson, S., Nadeau, J.H., and Cresswell, P. (2013). Succinate is an inflammatory signal that induces IL-1β production in human islets. Diabetes 62, 238–242.

O’Kearney, R.C., Mulligan, A., and Vannucci, S.J. (2008). The facilitative glucose transporter GLUT3: 20 years of distinction. Am. J. Physiol. Endocrinol. Metab. 295, E242–E253.

Suzuki, K. (2008). The multi-functional serpin, protein C inhibitor: Beyond thrombosis and hemostasis. J. Thromb. Haemost. 6, 2017–2026.

Sarkar, S.A., Wong, R., Hackl, S.I., Moua, O., Gill, R.W., Wiseman, A., Davidson, H.W., and Hutton, J.C. (2007). Induction of indoleamine 2,3-dioxygenase by interfere-γ in human islets. Diabetes 56, 72–79.

Schuler, R.P., Jackson, C., Garcia-Perez, J.E., Baxter, R.M., Ogolla, S., Robichaud, R., Ghosh, D., Rudra, P., and Hsieh, E.W.Y. (2019). Minimizing batch effects in mass cytometry data. Front. Immunol. 10, 2367.

Rau, J.C., Beaulieu, L.M., Huntington, J.A., and Church, F.C. (2007). Serpins in thrombosis, hemostasis and fibrolysis. J. Thromb. Haemost. 5 (Suppl 1), 102–115.

Rentsch, C.T., Beckman, J.A., Tomlinson, L., Gellad, W.F., Alcorn, C., Kidwai-Khan, F., Skanderson, M., Brittain, E., King, J.T., Jr., Ho, Y.L., et al. (2021). Early initiation of prophylactic anticoagulation for prevention of coronavirus disease 2019 mortality in patients admitted to hospital in the United States: Cohort study. BMJ 372, n311.
Differential expression of the nuclear-encoded mitochondrial transcriptome in pediatric septic shock. Crit. Care 18, 623.

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis (Springer-Verlag).

Xu, J.B., Xu, C., Zhang, R.B., Wu, M., Pan, C.K., Li, X.J., Wang, Q., Zeng, F.F., and Zhu, S. (2020). Associations of procalcitonin, C-reaction protein and neutrophil-to-lymphocyte ratio with mortality in hospitalized COVID-19 patients in China. Sci. Rep. 10, 15058.

Zunder, E.R., Finck, R., Behbehani, G.K., Amir, A.D., Krishnaswamy, S., Gonzalez, V.D., Lorang, C.G., Bjornson, Z., Spitzer, M.H., Bodenmiller, B., et al. (2015). Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm. Nat. Protoc. 10, 316–333.
### STAR★METHODS

**KEY RESOURCES TABLE**

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| **Antibodies**      |        |            |
| Anti-Human CD45     | Fluidigm | Cat# 3089003B; RRID: AB_2661851 |
| Anti-Human CD57     | Biolegend | Cat# 322302; RRID: AB_2661815 |
| Anti-Human CD11c    | BD bioscience | Cat# 555390; RRID: AB_395791 |
| Anti-Human CD16     | eBioscience | Cat# 16-0167-85; RRID: AB_11040983 |
| Anti-Human CD198 (CCR6) | Biolegend | Cat# 353402; RRID: AB_10918625 |
| Anti-Human CD19     | Fluidigm | Cat# 3142001B; RRID: AB_2651155 |
| Anti-Human CD123    | Fluidigm | Cat# 3143014B; RRID: AB_2811081 |
| Anti-Human CCR5     | Fluidigm | Cat# 3144007A; RRID: AB_2892770 |
| Anti-Human IgD      | Fluidigm | Cat# 3146005B; RRID: AB_2811082 |
| Anti-Human CD1c     | Miltenyi | Cat# 130-108-032; RRID: AB_2661165 |
| Anti-Human CD38     | Biolegend | Cat# 303502; RRID: AB_314354 |
| Anti-Human CD127    | Fluidigm | Cat# 3149011B; RRID: AB_2661792 |
| Anti-Human CD86     | Fluidigm | Cat# 3150020B; RRID: AB_2687852 |
| Anti-Human ICOS     | Biolegend | Cat# 313502; RRID: AB_416326 |
| Anti-Human CD141    | Biolegend | Cat# 344102; RRID: AB_2201808 |
| Anti-Human Tim3     | Fluidigm | Cat# 3153008B; RRID: AB_2687644 |
| Anti-Human TIGIT    | Fluidigm | Cat# 3154016B; RRID: AB_2888926 |
| Anti-Human CD27     | Fluidigm | Cat# 3155001B; RRID: AB_2687645 |
| Anti-Human CXCR3    | Fluidigm | Cat# 3156004B; RRID: AB_2687646 |
| Anti-Human CD45RA   | Biolegend | Cat# 304102; RRID: AB_314406 |
| Anti-Human PD-1     | Biolegend | Cat# 329941; RRID: AB_2563734 |
| Anti-Human PDL1     | Fluidigm | Cat# 3159029B; RRID: AB_2861413 |
| Anti-Human CD14     | Fluidigm | Cat# 3160001B; RRID: AB_2687634 |
| Anti-Human Tbet     | Fluidigm | Cat# 3161014B; RRID: AB_2858233 |
| Anti-Human Ki67     | Fluidigm | Cat# 3162012B; RRID: AB_2888928 |
| Anti-Human CD33     | Fluidigm | Cat# 3163023B; RRID: AB_2687857 |
| Anti-Human CD95     | Fluidigm | Cat# 3164008B; RRID: AB_2858235 |
| Anti-Human Foxp3    | Biolegend | Cat# 14-4774-82; RRID: AB_467552 |
| Anti-Human Eomes    | Biolegend | Cat# 14-4877-82; RRID: AB_2572882 |
| Anti-Human CCR7     | Fluidigm | Cat# 3167009A; RRID: AB_2858236 |
| Anti-Human CD8a     | Fluidigm | Cat# 3168002B; RRID: AB_2892771 |
| Anti-Human CD25     | Fluidigm | Cat# 3169003B; RRID: AB_2661806 |
| Anti-Human CD3      | Fluidigm | Cat# 3170001B; RRID: AB_2811085 |
| Anti-Human CXCR5    | Fluidigm | Cat# 3171014B; RRID: AB_2858239 |
| Anti-Human IgM      | Fluidigm | Cat# 3172004B; RRID: AB_2810858 |
| Anti-Human HLA-DR   | Fluidigm | Cat# 3173005B; RRID: AB_2810248 |
| Anti-Human CD4      | Fluidigm | Cat# 3174004B; RRID: AB_2687862 |
| Anti-Human CCR4     | R&D | Cat# MAB1567-500; RRID: AB_2892772 |
| Anti-Human CD56     | Miltenyi | Cat# 130-113-012; RRID: AB_2728090 |
| Anti-Human CD11b    | Fluidigm | Cat# 3209003B; RRID: AB_2687854 |

**Critical commercial assays**

| **U-PLEX Biomarker Group 1 (hu) 71-Plex** | Meso Scale Discovery (MSD) | Cat# K15081K |
| **V-PLEX Vascular Injury Panel 2 Human Kit** | Meso Scale Discovery (MSD) | Cat# K15198D |

(Continued on next page)
RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed and will be fulfilled by the Lead Contact, Joaquin Espinosa (joaquin.espinosa@cuanschutz.edu).

Materials availability
This study did not generate new unique reagents.
**Data and code availability**

All data generated for this manuscript is made available through the online researcher gateway of the COVIDome Project, known as the COVIDome Explorer, which can be accessed at covidome.org. The RNaseq data have been deposited in NCBI Gene Expression Omnibus, with series accession number GSE167000. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Perez-Riverol et al., 2019) with the dataset identifier ProteomeXchange: PXD022817. The mass cytometry data has been deposited in Flow Repository: FR-FCM-Z367. The metabolomics data have been deposited in the Metabolomics Workbench: PR001110. All code required to run the COVIDome Explorer applications can be found at https://github.com/cusom/CUSOM.COVIDome.Shiny-Apps (Zenodo https://doi.org/10.5281/zenodo.5081091) and https://github.com/cusom/CUSOM.ShinyHelpers (Zenodo https://doi.org/10.5281/zenodo.5081093). Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

**EXPERIMENTAL MODEL AND SUBJECT DETAILS**

Research participants were recruited and consented for participation in the COVID Biobank of the University of Colorado Anschutz Medical Campus [Colorado Multiple Institutional Review Board (COMIRB) Protocol # 20-0685]. Data was generated from deidentified biospecimens and linked to demographics and clinical metadata procured through the Health Data Compass of the University of Colorado under COMIRB Protocol # 20-1700. Participants were hospitalized either at Children’s Hospital Colorado or the University of Colorado Hospital. Blood samples were taken at variable times during hospital stays, with 75% of samples obtained within 4 days of admission and 90% within 10 days. COVID-19 status was defined by a positive PCR reaction using samples obtained from nasal swabs in most (~95%) cases and/or antibody test. Of the COVID-19-positive patients, 88% were SARS-CoV-2-positive by PCR, and the remaining 12% by antibody test. Cohort characteristics can be found in Table S1.

**METHOD DETAILS**

**Blood processing**

Blood samples were collected into EDTA tubes, PAXgene RNA, and sodium heparin tubes. After centrifugation, EDTA plasma was used for MS proteomics, SOMAscan® proteomics, as well as multiplex immunoassays using MSD technology for both cytokine profiles and seroconversion assays. From sodium heparin tubes, PBMCs were obtained by the Ficoll gradient method before cryopreservation and assembly of batches for MC analysis (see below).

**Whole blood transcriptome**

RNA was purified from PAXgene Blood RNA Tubes (PreAnalytiX/QIAGEN) using a PAXgene Blood RNA Kit (QIAGEN), according to the manufacturer’s instructions. RNA quality was assessed using an Agilent 2200 TapeStation and quantified by Qubit (Life Technologies). Globin RNA depletion, poly-A(+) RNA enrichment, and strand-specific library preparation were carried out using a Universal Plus mRNA-Seq with NuQuant, Human Globin AnyDeplete (Tecan). Paired-end 150 bp sequencing was carried out on an Illumina NovaSeq 6000 instrument by the Genomics Shared Resource at the University of Colorado Anschutz Medical Campus.

**Plasma proteomics by mass spectrometry**

Plasma samples were digested in S-Trap filters (Protifi, Huntington, NY) according to the manufacturer’s procedure. Briefly, a dried protein pellet prepared from organic extraction of patient plasma was solubilized in 400 μL of 5% (w/v) SDS. Samples were reduced with 10 mM DTT at 55 °C for 30 min, cooled to room temperature, and then alkylated with 25 mM iodoacetamide in the dark for 30 min. Next, a final concentration of 1.2% phosphoric acid and then six volumes of binding buffer [90% methanol; 100 mM triethylammonium bicarbonate (TEAB); pH 7.1] were added to each sample. After gentle mixing, the protein solution was loaded into an S-Trap filter, spun at 2000 rpm for 1 min, and the flow-through collected and reloaded onto the filter. This step was repeated three times, and then the filter was washed with 200 μL of binding buffer 3 times. Finally, 1 μg of sequencing-grade trypsin (Promega) and 150 μL of digestion buffer (50 mM TEAB) were added onto the filter and digestion carried out at 47 °C for 1 h. To elute peptides, three stepwise buffers were applied, 200 μL of each with one more repeat, including 50 mM TEAB, 0.2% formic acid in H2O, and 50% acetonitrile and 0.2% formic acid in H2O. The peptide solutions were pooled, lyophilized and resuspended in 1 mL of 0.1% FA. 20 μL of each sample was loaded onto individual Evotips for desalting and then washed with 20 μL 0.1% FA followed by the addition of 100 μL storage solvent (0.1% FA) to keep the Evotips wet until analysis. The Evasor One system (Evasor, Odense, Denmark) was used to separate peptides on a Pepsee column, (150 μm internal diameter, 15 cm) packed with ReproSil C18 1.9 μm, 120A resin. The system was coupled to a timsTOF Pro mass spectrometer (Bruker Daltonics, Bremen, Germany) via a nano-electrospray ion source (Captive Spray, Bruker Daltonics). The mass spectrometer was operated in PASEF mode. The ramp time was set to 100 ms and 10 PASEF MS/MS scans per topN acquisition cycle were acquired. MS and MS/MS spectra were recorded from m/z 100 to 1700. The ion mobility was scanned from 0.7 to 1.50 Vs/cm. Precursors for data-dependent acquisition were isolated within ± 1 Th and fragmented with an ion mobility-dependent collision energy, which was linearly increased from 20 to 59 eV in positive mode. Low-abundance precursor ions with an intensity above a threshold of 500 counts but below a target value of 20000 counts were repeatedly scheduled and otherwise dynamically excluded for 0.4 min. Raw data file conversion to peak lists in the MGF format, downstream
identification, validation, filtering and quantification were managed using FragPipe version 13.0. MSFragger version 3.0 was used for database searches against a Human isoform-containing UniProt fasta file (version 08/11/2020) with decoys and common contaminants added. The identification settings were as follows: Trypsin, Specific, with a maximum of 2 missed cleavages, up to 2 isotope errors in precursor selection allowed for, 10.0 ppm as MS1 and 20.0 ppm as MS2 tolerances; fixed modifications: Carbamidomethylation of C (+57.021464 Da), variable modifications: Oxidation of M (+15.994915 Da), Acetylation of protein N-term (+42.010565 Da), Pyrrolidine from peptide N-term Q or C (−17.026549 Da). The Philosopher toolkit version 3.2.9 (build 1593192429) was used for filtering of results at the peptide and protein level at 0.01 FDR. Label-free quantification was performed by AUC integration with matching between all runs using IonQuant.

**Plasma proteomics by SOMAscan assays**

125 μL EDTA plasma was analyzed by SOMAscan® assays using previously established protocols (Gold et al., 2012). Briefly, each of the 4800+ SOMAmer® reagents binds a target peptide and is quantified on a custom Agilent hybridization chip. Normalization and calibration were performed according to SOMAscan® Data Standardization and File Specification Technical Note (SSM-020) (Gold et al., 2012). The output of the SOMAscan® assay is reported in relative fluorescent units (RFU).

**Cytokine profiling and seroconversion by multiplex immunoassay**

Multiplex immunoassays MSD assays were performed on EDTA plasma aliquots following manufacturer’s instructions (Meso Scale Discovery, MSD). A list of immune factors measured by MSD can be found in Table S2D. Values were extrapolated against a standard curve using provided calibrators. Seroconversion assays against SARS-CoV-2 proteins and the control protein from the Flu A Hong Kong H3 virus were performed in a multiplex immunoassay using the IgG detection readout according to manufacturer’s instructions (MSD). Relative values were extrapolated against a standardized curve consisting of pooled COVID-19 positive reference plasma (Johnson et al., 2020).

**Mass cytometry analysis of immune cell types**

Cryopreserved PBMCs were thawed, washed twice with Cell Staining Buffer (CSB) (Fluidigm), and counted with an automated cell counter (Countess II - Thermo Fisher Scientific). Extracellular staining on live cells was done in CSB for 30 min at room temperature, in 3-5 × 10⁶ cells per sample. Cells were washed with 1X PBS (Fluidigm) and stained with 1 mL of 0.25 mM cisplatin (Fluidigm) for 1 min at room temperature for exclusion of dead cells. Samples were then washed with CSB and incubated with 1.6% PFA (Electron Microscopy Sciences) during 10 min at room temperature. Surface staining with antibodies that work on fixed epitopes was performed in CSB for 30 min at room temperature (see Table S4 and Key resources table for antibody information). Cells were washed twice with CSB and fixed in Fix/Perm buffer (eBioscience) for 30 min, washed in permeabilization buffer (eBioscience) twice, then intracellular factors were stained in permeabilization buffer for 45 min at 4°C. Cells were washed twice with Fix/Perm Buffer and were labeled overnight at 4°C with Cell-ID Intercalator-Ir (Fluidigm) for DNA staining. Cells were then analyzed on a Helios instrument (Fluidigm). To make all samples comparable, pre-processing of mass cytometry data included normalization within and between batches via polystyrene beads embedded with lanthanides as previously described (Finck et al., 2013). Files were debarcoded using the MATLAB DebarcodeTool (Zunder et al., 2015). Then normalization again between batches relative to a reference batch based on technical replicates (Schuyler et al., 2019). Gating was performed using CellEngine (Primitybio) as previously described (Galbraith et al., 2021).

**Mass spectrometry-based metabolomics of plasma and red blood cells**

Samples were thawed on ice and extracted via a modified Folch method (chloroform/methanol/water 8:4:3), which completely inactivates other coronaviruses, such as MERS-CoV. Briefly, 20 μL of sample was diluted in 130 μL of LC-MS grade water, 600 μL of ice-cold chloroform/methanol (2:1) was added, and the samples were vortexed for 10 s. Samples were then incubated at 4°C for 5 minutes, quickly vortexed (5 s), and centrifuged at 14,000 g for 10 minutes at 4°C. The top (i.e., aqueous) phase was transferred to a new tube for metabolomics analysis and flash frozen. The bottom (i.e., organic) phase was transferred to a new tube for lipidomics analysis, then dried under N₂ flow.

Analyses were performed using a Vanquish UHPLC coupled online to a Q Exactive high resolution mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). Samples (10 μL per injection) were randomized and analyzed in positive and negative electrospray ionization modes (separate runs) using a 5-minute C18 gradient on a Kinetex C18 column (Phenomenex) as described (Nemkov et al., 2019). Data were analyzed using Maven (Princeton University, Princeton, NJ, USA) in conjunction with the KEGG database and an in-house standard library.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

Preprocessing, statistical analysis, and plot generation for all datasets was carried out using R (R 4.0.1 / Rstudio 1.3.959 / Bioconductor v 3.11) (Huber et al., 2015, R Core Team, 2020, RStudio Team, 2020), as detailed below.
Analysis of transcriptome data
RNaseq data yield was ~40-80 x 10^6 raw reads and ~32-71 x 10^6 final mapped reads per sample. Reads were demultiplexed and converted to fastq format using bcl2fastq (bcl2fastq v2.20.0.422). Data quality was assessed using FASTQC (v0.11.5) (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and FastQ Screen (v0.11.0, https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/). Trimming and filtering of low-quality reads was performed using bbdruk from BBTools (v37.99) (Bushnell et al., 2017) and fastq-mcf. from ea-utils (v1.05, https://expressionanalysis.github.io/ea-utils/). Alignment to the human reference genome (GRCh38) was carried out using HISAT2 (v2.1.0) (Kim et al., 2019) in paired, spliced-alignment mode with a GRCh38 index with a Gencode v33 annotation GTF, and alignments were sorted and filtered for mapping quality (MAPQ > 10) using Samtools (v1.5) (Li et al., 2009). Gene-level count data were quantified using HTSeq-count (v0.6.1) (Anders et al., 2015) with the following options (–strand=reverse –minqual=10 –type=exon–mode=intersection-nonempty) using a Gencode v33 GTF annotation file. Differential gene expression in COVID-19-positive versus COVID-19-negative was evaluated using DESeq2 (version 1.28.1) (Love et al., 2014) in R (version 4.0.1), using q < 0.1 (FDR < 10%) as the threshold for differentially expressed genes.

Analysis of MS-proteomic data
Raw Razor intensity data were filtered for high abundance proteins by removing those with > 70% zero values in both COVID-19-negative and COVID-19-positive groups. For the remaining 407 abundant proteins, 0 values (8,363 missing values of 44,363 total measurements) were replaced with a random value sampled from between 0 and 0.5x the minimum non-zero intensity value for that protein. Data was then normalized using a scaling factor derived from the global median intensity value across all proteins / sample median intensity across all proteins (De Livera et al., 2012).

SOMAscan data
Normalized data (RFU) was imported and converted from a SOMAscan®.dat file using a custom R package (SomaDataIO) for use in all subsequent analysis.

Analysis of MSD cytokine profiling data
Plasma concentration values (pg/mL) for each of the cytokines and related immune factors measured across multiple MSD assay plates was imported to R, combined, and analytes with > 10% of values outside of detection or fit curve range flagged. For each analyte, missing values were replaced with either the minimum (if below fit curve range) or maximum (if above fit curve range) calculated concentration and means of duplicate wells used in all further analysis.

Analysis of LCMS-metabolomics data
Peak intensity data was imported to R. Across the 171 metabolites, 0 values (486 missing values of 21,033 total measurements) were replaced with a random value sampled from between 0 and 0.5x the minimum non-zero intensity value for that metabolite. Data was then normalized using a scaling factor derived from the global median intensity value across all proteins / sample median intensity and used for downstream analysis. This normalization method has been widely employed for MS data, including on the MetaBAnalyst platform (Chong et al., 2019), and performs comparably to a number of other normalization methods (Vallikangas et al., 2018).

Analysis of mass cytometry data
Cell population frequencies were exported from CellEngine as percentages of various parental lineages and used for subsequent analysis.

Differential abundance analysis
Differential abundance analysis for MS proteomics, SOMAscan® proteomics, MSD cytokine profiling, LCMS metabolomics, and mass cytometry data was performed using linear models with log2 concentration as the outcome variable and age, sex, and COVID-19 status as independent variables. Multiple hypothesis correction was performed with the Benjamini-Hochberg method using a false discovery rate (FDR) threshold of 10% (q < 0.1).

Correlation analysis
To identify features in each dataset that correlate with CRP levels in COVID-19-positive samples, Spearman rho values and p values were calculated against values adjusted for Sex and Age using the removeBatchEffect function from the limma package (v 3.44.3) from each dataset using the rcorr function from the Hmisc package (v 4.4-0), with Benjamini-Hochberg correction of p values and an estimated false discovery fate threshold of 0.1. For visualization, XY scatterplots with points colored by local density were generated using a custom density function and the ggplot2 (v3.3.1) package (Wickham, 2016).

Data visualization
To visualize differences between COVID-19-negative samples and COVID-19-positive samples, Z-scores were calculated for each feature based on the mean and standard deviation of COVID-19-negative samples, and visualized as heatmaps and/or modified sina plots using the ComplexHeatmap (v2.4.2) (Gu et al., 2016), ggplot2 (v3.3.1), and ggforce (v0.3.1) packages (Pedersen, 2019).