Complex changes in serum protein levels upon recovery from SARS-CoV2 infection

Smruti Pushalkar[1#*], Shaohuan Wu[1#], Shuvadeep Maity[1,2], Matthew Pressler[1], Justin Rendleman[1], Burcu Vitrinel[1], Ted Ross[3], Michael Carlock[3], Hyungwon Choi[4], and Christine Vogel[1*]

[1]Department of Biology, Center for Genomics and Systems Biology, New York University, New York, New York, USA; [2]Department of Biological Sciences, Birla Institute of Technology and Science-Pilani, Hyderabad, Telangana, India; [3]Department of Infectious Diseases, University of Georgia, Athens, Georgia, USA; and [4]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.

# equally contributing authors

*Correspondence to: sp117@nyu.edu, cvogel@nyu.edu

Abstract

The COVID-19 pandemic, triggered by severe acute respiratory syndrome coronavirus 2, has affected millions of people worldwide. Much research has been dedicated to our understanding of COVID-19 disease heterogeneity and severity, but less is known about recovery associated changes. To address this gap in knowledge, we quantified the proteome from serum samples from 29 recuperated COVID-19 patients and 29 age-, race-, and sex-matched healthy controls. Many proteins from pathways known to change upon acute COVID-19 illness, such as from the complement cascade, coagulation system, inflammation and adaptive immune system, had returned to levels seen in healthy controls. In comparison, we identified 22 and 15 proteins with significantly elevated and lowered levels, respectively, amongst recuperated COVID-19 cases compared to healthy controls. Some of the changes were similar to those observed for the acute phase of the disease, i.e. elevated levels of proteins from hemolysis, the adaptive immune systems, and inflammation. In contrast, some changes opposed those in the acute phase, e.g. elevated levels of CETP and APOA1 which function in lipid/cholesterol metabolism, and decreased levels of proteins from the complement cascade (e.g. C1R, C1, and VWF), the coagulation system (e.g. THBS1 and VWF), and the regulation of the actin cytoskeleton (e.g. PFN1 and CFL1) amongst recuperated COVID-19 cases. We speculate that some of these changes might originate from transient decreases in platelet counts upon recovery from the disease. Finally, we observed race-specific changes, e.g. with respect to immunoglobulins and cholesterol-metabolism-related proteins.

Introduction

The COVID-19 pandemic has affected >607 million people worldwide with ~6.5 million deaths (World Health Organization). Caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the disease is highly infective and exhibits an extensive clinical heterogeneity, from asymptomatic to symptomatic disease states [1]. While the primary manifestation of COVID-19 is in the respiratory tract, there is an increased risk of other life-threatening pathologies such as pulmonary embolism, myocardial infarction, and ischemic stroke with frequent venous and arterial thromboembolisms with the severity of disease [2]. Similarly, recovery from COVID-19 has displayed enormous heterogeneity, ranging from disappearance of symptoms within a few days to establishment of 'long COVID', also known as 'post-acute sequelae of SARS-CoV-2' (PASC) marked by a broad spectrum of the ongoing physiological changes [3,4].

Much work has been done to characterize the molecular changes during the acute phase of the disease, e.g. in patient plasma and serum samples [5–8]. Here, we define the acute phase as the first ten days after positive infection. Both untargeted and targeted proteomics approaches have identified dysregulation of various pathways including lipid homeostasis, immunoglobulins, inflammatory and antiviral cytokines, chemokines of innate and adaptive immunity, as well as complement and coagulation cascades [4,5,9–11]. These studies also observed platelet degranulation, lymphocyte apoptosis in serum, likely due to the viral mode of entry into the cells [4,5,9,10]. During acute infection, overproduction of proinflammatory cytokines (IL-6, IL-1β, and TNF-α) induces a 'cytokine storm' elevating the risk of clot formation, platelet activation, and ultimately hypoxia and multiorgan failure leading to high patient mortality [9,12]. Accordingly, serum proteomics using a highly sensitive targeted assay identified proteins of inflammation, cardiometabolic, and neurologic diseases to contribute to disease severity [13]. Other studies found an expansion in regulatory proteins of coagulation (APOH, FN1, HRG, KNG1, and PLG) and lipid homeostasis (APOA1, APOC1, APOC2, APOC3, and PON1) in serum as the disease progressed [5]. COVID-19 plasma samples demonstrated extensive changes in several key protein modifications, such as glycosylation, phosphorylation, citrullination and arginylation during the acute phase of the disease [14]. Even serum from COVID-19 infected asymptomatic individuals showed altered levels of coagulation and inflammation, such as fibrinogen, von willebrand factor (VWF), and thrombospondin-1 (TSP1)[3]. In comparison, long COVID/PASC patients appear to have altered levels of autoantibodies, localized inflammation, and reactivation of latent pathogens [4].

In comparison, much less information is available on the molecular impact of COVID-19 upon recovery. One study examining blood samples from severe COVID-19 patients upon recovery observed elevated erythrocyte sedimentation rates, increased levels of C-reactive protein (CRP), and reduced levels of serum albumin [15]. Another study showed that carbonic anhydrase 1 (CA1) was still bound to immunoglobulin IgA in COVID-19 patients within 2 weeks of recovery, unlike in any healthy, vaccinated healthy subjects or in COVID-19 patients after 6 months of recovery [16].

To understand the broader impact of COVID-19 after recovery, we profiled the serum proteome in 29 serum samples from recuperated COVID-19 cases and 29 samples from age-, sex-, and race-matched healthy controls. Recovery is defined here as the period following the first ten days after diagnosis (acute phase). We used mass spectrometry to quantify protein levels in the soluble fraction of the blood serum in an untargeted fashion and used linear regression models to deconvolute the effects of demographic parameters in differentially abundant serum proteins. We identified pathways whose member proteins had returned to the pre-infection levels, and pathways with member proteins that were still altered either consistent with or opposite to changes observed during the acute infection.

Results

Quantitation of 334 serum proteins

We quantified proteins from serum samples of 29 recuperated COVID-19 cases and age-, sex- and race-matched 29 healthy individuals. **Figure 1A** shows the experimental outline; **Supplementary Table S1** describes the cohort demographics. Participant age ranged from 22 to 61 years, with a median (interquartile range, IQR) of 44 (20) years. The male to female ratio was 1.2:1. About 55% of the individuals reported to have had symptoms at the time of the acute COVID-19 infection (during the first 10 days after diagnosis) while the remaining individuals reported no symptoms. The serum samples from recovered COVID-19 individuals were collected 9 to 70 days after diagnosis (Days since diagnosis); individuals showed no symptoms at the time of sample collection. Median (IQR) Days since diagnosis was 59 (15) for symptomatic and 32 (8) for asymptomatic individuals (**Supplementary Table S1**). Serum antibody titers were measured at the time of sample collection. About 24% of the patients showed no antibody response whereas 3%, 21%, 21% and 24% patients displayed low, moderate, high and very high antibody response respectively, as defined in **Supplementary Table S1**.

Figures 1B-D describe relationships between selected demographics. The age distributions between male and female participants were similar **(Figure 1B**). Individuals who reported symptoms during the acute phase of the infection had a significantly longer period until sample collection (**Figure 1C**) and significantly higher antibody titers (**Figure 1D**) than individuals without symptoms (p-value < 0.05). These relationships suggest that individuals participated in the study at a time based on severity of the infection: more severely ill individuals provided serum samples later than less ill individuals. This connection renders the interpretation of the impact of Days since diagnosis, Symptoms, and Titer levels on the serum proteome difficult.

We quantified a total of 334 proteins across all samples as depicted in **Figure 2**. About half of the proteins fall into three functional categories: immunoglobulins, complement cascade and high-density lipoproteins (32%, 9%, and 5%, respectively). The proteins were grouped into 20 clusters, based on their levels across the samples. Among them, 12 clusters contained one or more proteins with statistically significant differences between recuperated COVID-19 cases and healthy controls as discussed below.

To evaluate the quality of the proteomic measurements, we examined the temporal behavior of quality control (QC) samples which had been analyzed along with the cohort samples (**Supplementary Figure S1**). The QC samples consisted of pooled cohort samples and therefore contained serum peptides whose variability originated from technical sources, e.g. batch effects. The median coefficient of variance (CoV) for all proteins across the quality control samples was <30% in all four batches, which is consistent with what is expected from untargeted proteomics analyses. Only 40 proteins showed an average CoV>50% in the QC samples (**Supplementary Data File S1**); their quantification in the cohort samples is less reliable, and the proteins are marked in all figures. As additional quality controls, we examined several example proteins for their levels amongst demographic subsets (**Supplementary Figures S3 and S4**). The proteins showed the expected sex and body weight related differences.

Altered levels of components of the innate and adaptive immune system

First, we examined the overall differences in protein levels between samples from healthy controls and recuperated COVID-19 cases, regardless of the individuals' demographics. We conducted partial least squares discriminant analysis with the protein levels (**Figure 3A**). The major components explained 19% of

the variability and separated the data into two distinct clusters comprising the two cohorts. This separation confirmed that the proteomics data captured differences between recuperated COVID-19 cases and healthy controls.

The volcano plot in **Figure 3B** depicts the results from the overall comparison of protein levels between the two sample sets. Grey dots indicate proteins with similar levels in both cohorts: these proteins included the majority of complement (e.g., C1QB and C1RI subunits, C2 to C4, C8, and C9) and coagulation (e.g., coagulation factors; F5, and F10 to F13) cascades (**Supplementary Figure S6, Table 2**). As both pathways are heavily dysregulated during the acute phase of the infection [17–19], the results indicate that individuals returned to levels similar to those observed in healthy individuals.

Dots with red and blue colors in **Figure 3B** mark the 22 and 15 proteins that were significantly elevated or decreased in the serum from recuperated COVID-19 cases and healthy controls, respectively (adjusted p-value < 0.05). **Supplementary Figure S4** shows the levels of these proteins measured in each sample. The recuperated COVID-19 patients had significantly elevated immunoglobulins, Orosomucoid 2 (ORM2), peroxiredoxin-2 (PRDX2), hemoglobin subunits (HBD, HBB and HBA1), as well as proteins involved in cholesterol transport such as cholesteryl ester transfer protein (CETP) and apolipoprotein A1 (APOA1). In comparison, healthy controls had significantly elevated levels of actin cytoskeleton signaling proteins, e.g. filamin (FLNA), profilin (PFN1), cofilin (CFL1), and actin beta (ACTB).

We also tested for differences in protein modifications between the two groups. While we did not enrich proteins with post-translational modifications, the type of mass spectrometry data we collected allowed for a retrospect analysis for modified peptides. To do so, we constructed computational libraries scanning all data for the occurrence of several frequently occurring modifications, i.e. mono- and di-hexose, and then used the library to analyze the cohort samples. We observed hexose addition (glycosylation) most frequently (**Supplementary Data File S2**): three of 55 hexose-modified peptides were significantly more abundant in recuperated COVID-19 cases than in healthy controls, when normalized for the abundance of the respective unmodified peptides (adjusted p-value < 0.05; **Supplementary Figure S5**). The three peptides originated from albumin (ALB) and immunoglobulin heavy constant alpha 1 (IGHA1); these two proteins did not show significantly differential levels across the two cohorts. Our results were consistent with extensive serum glycosylation observed in COVID-19 patients [20,21].

The impact of known demographics on serum protein signatures

Next, we tested for the impact of known demographics, i.e. Days since diagnosis, Symptoms, Age, Sex, and Race of the individuals on the serum protein levels, to identify the ‘interactions’ between demographics and respective COVID-19 states of the individuals. Due to its correlation with Symptoms and Days since diagnosis, we excluded Titer levels from the analyses. We tested i) protein levels in healthy individuals; ii) protein levels in recuperated COVID-19 cases; and iii) log base 2 ratios of the protein levels in the COVID-19 cases versus the matched healthy controls. **Figures 4** and **5** show the results involving the first two analyses. Results from all analyses are shown in **Supplementary Data File S1**.

Figure 4 shows five of the previously mentioned 12 clusters from significant overall differences in protein levels between the two cohorts as well as other proteins from the same clusters, but no significant interactions with any known demographics. The samples in **Figure 4** were ordered by their origin (healthy or recuperated cohort), as well as by sex and age. We observed three clusters (A to C) that contained proteins with lower levels in recuperated COVID-19 cases, including proteins functioning in cell adhesion and platelet degranulation, e.g. actin cytoskeleton (e.g. fibrinogen A (FGA), VWF), and innate immunity/the complement system (e.g. C1R, C1S, C1QA, C1QC, and GGH). Actin cytoskeleton proteins (FLNA, PFN1,

CFL1, ACTB, TAGLN2, and TSMB4X) showed the strongest bias both with respect to fold-change and significance, as also indicated in **Figure 3B**. The quantitation of TAGLN2 and TSMB4X quantitation was less reliable as marked in the figure (**Supplementary Data File S1**). In comparison, we observed two clusters (D and E) with higher protein levels in recuperated COVID-19 cases, e.g. immunoglobulins playing a role in antigen recognition (e.g. IGKV1-27, IGKV1-39, IGHV1-46, IGLV3-1, and PRDX2), hemoglobin subunits (e.g. HBB, HBA1, and HBD), and carbonic anhydrase (CA1).

Next, we examined clusters with differences between recuperated COVID-19 cases and healthy controls that were more complex, i.e. that involved interactions with demographic factors (**Figure 5**). We observed no significant interactions with Age, Sex, Days since diagnosis, or Symptoms. While this result may be in part due to the limited sample size, it might also be attributable to an intrinsic relationship between demographic variables (**Figure 1D**): individuals with symptoms and higher antibody titers during the acute phase tended to have serum samples collected at later time points than those without symptoms and lower antibody titers. We assume that Days since diagnosis and Symptoms/Titer have opposite effects on protein levels: early sample collection and a more severe acute phase of the disease should have similar effects on the serum. Therefore, the inverse relationship between these demographics across samples effectively eliminated any potential signal.

The only demographic variable with an interaction signal in several clusters was race (**Figure 5**). To illustrate the effect of race, samples in **Figure 5** were sorted into the healthy control and recuperated COVID-19 cohorts, and within the cohorts sorted according to race. The figure focussed on samples from only white or black individuals which comprised the majority in this study (n=26).

The first three clusters (F to H) in **Figure 5** comprised proteins with significant overall differences between recuperated COVID-19 cases and healthy controls. For some proteins, e.g. IGKV3-20, IGKV1-45, IGHG1, IGHV3-49, and IGLV3-27, this difference only existed for black, but not for white individuals. Most proteins from clusters F to H function were immunoglobulins, i.e. they functioned in adaptive immunity.

We observed the opposite race effect in proteins from clusters I to L: proteins exhibited overall lower levels in recuperated COVID-19 cases, but more so or only in samples from black individuals (**Figure 5**). Healthy individuals had no significant biases with respect to race. Proteins from these clusters included those from lipid transport (cluster J), e.g. apolipoproteins (APOB, APOC2, APOC3, APOC4, and APOA4), and cell maintenance proteins, e.g. selenoprotein P (SELENOP)(cluster K). Cluster I showed very weak interactions with race, with signatures very similar to those in clusters D and E (**Figure 4**), and it contained proteins from cholesterol transport, e.g. CETP and APOA1.

Discussion

Proteomic alterations in serum and plasma of mild, moderate, severe and critically ill COVID-19 patients have been studied extensively with respect to changes during the acute phase of the disease [11,22–27]. In comparison, much less is known about changes upon recovery. We present one of the few controlled studies investigating serum proteomic differences between recuperated COVID-19 cases and age-, sex-, and race-matched healthy controls.

Table 1 provides a qualitative summary of the results. While a total of 37 proteins (11%) showed significant differences in their levels (**Figure 3**), most of the observed proteome (89%) was similar between recuperated COVID-19 cases and healthy controls. Some of these proteins similar amongst the two cohorts were known to be altered during the acute phase of the infection, i.e. within 10 days since diagnosis, and their respective respective functions are listed in **Table 1A**: they included coagulation and the complement

system (**Supplementary Figure 6**), as well as some proteins involved in metabolism (e.g. LPA and PON1) and inflammation (e.g. ORM1, S100A8, and S100A9).

We found that the functions of the 37 proteins still altered amongst the recuperated COVID-19 cases belonged to two groups shown in **Table 1B** and **C**: pathways whose protein levels were altered in a manner consistent with changes during the acute phase, and pathways whose protein levels were inconsistent with these changes.

For example, we identified markers of acute inflammation (ORM2) and hemolysis (HBA1, HBB, HBD, and CA1) as well as the hemolytic anemia associated protein PRDX2 with significantly elevated levels in the recuperated COVID-19 cases compared to healthy controls (**Figure 3B, 4**), consistent with what had been found in acute COVID-19 patients with high IL-6 levels and severely ill patients [28](**Table 1B**). These findings indicated that elevated levels of proteins from inflammation and hemolysis were persistent for up to >2 months of recovery. CA1 is known to associate with the IgA-complex in acute COVID-19 patients but not in healthy individuals [16] - the elevated levels we observed indicated that this might still be the case.

Perhaps the most interesting changes were those opposite to those observed during the acute phase (**Table 1C**). Examples include proteins involved in coagulation (fibrinogen and VWF) whose levels were significantly reduced amongst recuperated COVID-19 cases (**Figure 4**). VWF mediates platelet attachment to damaged endothelium and acts as a carrier protein for coagulation factor VIII rendering protection from proteolytic degradation [29,30]. Its levels are known to be elevated in acute COVID-19 and long COVID/PASC patients signifying platelet activation and adhesion to endothelium [29,31] and COVID-19-associated endotheliopathy [32]. Fibrinogen regulates protective immune functions and clot formation [33]. During acute COVID-19, it is involved in thrombosis in lungs [34], and fibrinogen chains are strongly associated with COVID-19 fatalities [35]. The significantly decreased levels of VWF and fibrinogens in recuperated COVID-19 cases (**Table 1C**) contrasts the finding that most components of the coagulation and complement cascade had returned to pre-infection levels (**Table 1A, Supplementary Figure S6**). We speculate that these findings might be due to the temporary suppression of some pathways after recovery from the acute phase.

One such pathway relates to platelet counts in the blood. Platelets act as cellular immunomodulators interacting with endothelial cells and leukocytes in response to infections, and are therefore crucial during thrombosis and the host immune response [36]. During viral infections, low platelet counts, interactions with leukocytes, and platelet secretion may lead to protective or injurious immune effects [37]. Aberrant blot clot formation, such as thrombosis, is a known complication of COVID-19 infection [38,39]. As fibrinogen and VWF are engaged in platelet degranulation [40,41], we hypothesize that the observed decreased levels of proteins from the platelet degranulation pathway in samples from recovered COVID-19 cases (**Figure 4**) may be due to low platelet counts resulting from platelet consumption during COVID-19 infection. This hypothesis is supported by studies that suggest a 5% to 42% exhaustion in platelet counts for several months post infection (immune thrombocytopenia)[42–45] amongst survivors of both severe and non-severe COVID-19 patients [46–49]. While low platelet counts can occur any time during the acute phase of COVID-19, it has been frequently observed after clinical recovery [50], e.g. after three [51,52] or five weeks [53] after onset of symptoms, consistent with our findings.

Further, we found significantly reduced levels of proteins of the actin cytoskeleton network amongst recuperated COVID-19 cases, e.g. PFN1 and CFL1 (**Figure 4**), partially contrasting what had been found for the acute phase (**Table 1C**). The actin cytoskeleton is critical in various pathways of the immune system, ranging from hematopoiesis and immune cell development, recruitment, migration, inter- and intra-cellular signaling, as well as response activation [54]. Many viruses interact with actin and actin-regulating signaling pathways within the host cell [55,56] and reprogram the cellular pathways [57]. Further, PFN1 is an important

player in activation of viral transcription and airway hyperresponsiveness [58,59], and it is known to be downregulated in non-severe COVID-19 patients [58]. CFL1 functions in T cell motility, T cell migration to lymphoid tissues, immune reconstitution, and immune control of viremia [60] and is dysregulated in HIV-infected patients [60,61]. Therefore, an additional interpretation of altered CFL1 levels observed in our data relates to possible changes in T cell mobility.

Further, we observed significant decreases of levels for some proteins of the complement cascade amongst the recuperated COVID-19 cases (**Figure 5**) contrasting observations from the acute phase [62,63](**Table 1C**). The complement cascade directly associates with altered blood coagulation in COVID-19 pathology [17,64,65], and blood coagulation, which in turn, involves platelet activation. We discussed the possible impact on platelet counts above which might also explain the temporary depletion of some proteins from the complement system.

Finally, by analyzing the impact of known demographics, we found no significant association of past COVID-19 infections with the individuals' Age, Sex, Days since diagnosis, and Symptoms. The lack of associations might be due to the small size of the cohort available and the heterogeneity amongst available samples. In comparison, we identified several proteins that were associated with Race which included mostly white and black individuals. Examples included many proteins of the adaptive immune response (**Figure 5**).

Other proteins with a race effect were from cholesterol metabolism and transport: their elevated levels amongst recuperated COVID-19 cases were observed more strongly amongst white individuals than black individuals (**Figure 5**). Apolipoproteins (APOA1 and APOB) are key regulators of cholesterol metabolism and transport [66,67] and can render protection against severe COVID-19 [68–70]. Other apolipoproteins such as APOCs are not known to play a critical role in COVID-19, but demonstrated a race effect in our data (**Figure 5**).

While grouped into a cluster with proteins with race effects (**Figure 5**, cluster I), CETP, APOA1, and SSA4 showed individually no significant race difference. CETP is linked to reverse cholesterol transport and associated with APOA1 and SSA4 [71]. The elevated SSA4 levels are consistent with findings from acute COVID-19 cases [26,66,72–74]; however, this is not true for CETP and APOA1 (**Table 1C**). APOA1 and APOA isoforms, which are also involved in the immune response and dyslipidemia, have been frequently observed in COVID-19 patients with acute inflammatory conditions [26,74,75]. The elevated APOA1 levels contrast our observations on other APOA isoforms which returned to levels similar to those in healthy controls (**Table 1A**) suggesting that further work will be necessary to decipher the complex role of cholesterol metabolism and transport.

Another protein with a race effect was SELENOP (**Figure 5**). SELENOP is expressed in the liver and secreted into plasma [76,77]. It has protective functions of host immune defense and tissue homeostasis [76,77]. SELENOP levels directly impact serum selenium levels [78], and higher serum selenium levels, in turn, have been associated with increased COVID-19 survival [26,78]. We observed elevated SELENOP levels only in the white population but not the African American individuals which suggest complex, possibly race dependent relationships. In general, while increasedCOVID-19 infection rates and deaths amongst African American, Hispanic, and Asian communities compared to the white population have been reported [79], the race-dependent changes that we observed will require further investigation prior to their interpretations.

In sum, our study provides insights into the proteomic landscape present at up to >two months after the infection (**Table 1**). While most of the proteome was similar to that found in healthy individuals, we identified several intriguing differences. The interpretation of these differences, e.g. with respect to a possible

temporary decrease in platelet counts, will have to be tested in future work through analysis of larger cohorts. Our findings might inspire some of these analyses to be conducted in a targeted fashion.

Methods

Sample collection

We conducted a retrospective case-control study of two cohorts with a total of 58 subjects, comprising healthy controls (n=29) and recuperated COVID-19 cases (n=29) which were age, sex and race matched. All participants had been recruited at the University of Georgia at Athens and provided written informed consent prior to participation. The study protocol was reviewed and approved by the University of Georgia Ethical Review Board. The participants' demographics were shown in **Supplementary Data File S1**. Antibody titers are the maximum dilutions at which antibodies were still detected. For visualization, titers were log transformed (base 10). Serum samples were heat inactivated at 56◦C for 30 min and stored at -80◦C.

Sample preparation

Serum samples including individual (n=58) and pooled samples were processed using a protocol described elsewhere [80]. In brief, 1 µl of serum sample (~70-80 µg protein) was lysed with 0.1% Rapigest (Waters, MA, USA) in 100 mM ammonium bicarbonate (Sigma, MO, USA) and denatured at 95◦C for 5 minutes. Further, the samples were reduced using 5 mM dithiothreitol (DTT, Sigma) at 60◦C for 30 minutes, followed by alkylation with 15 mM iodoacetamide (Sigma) at room temperature in the dark for 30 minutes. Subsequently, the samples were quenched with 10 mM DTT and digested overnight at 37◦C with Trypsin gold (Promega, WI, USA). The digestion was stopped and the surfactant was cleaved by treating samples with 200 mM HCl (Sigma) at 37◦C for 30 minutes. The samples were desalted on Hypersep C-18 spin tips (Thermo Fisher Scientific, MA, USA) and the peptides dried under vacuum at low heat (Eppendorf, CT, USA). The dried peptides were resuspended in 5% acetonitrile in 0.1% formic acid (Thermo Scientific) and quantified by fluorometric peptide assay kit (Thermo Fisher Scientific) prior to mass spectrometry analysis.

We analyzed the samples using an EASY-nLC 1200 (Thermo Fisher Scientific) connected to Q Exactive HF mass spectrometer (Thermo Fisher Scientific). We used an analytical column RSLC PepMan C-18 (Thermo Fisher Scientific, 2uM, 100Å, 75µm id x 50cm) at 55◦C with the mobile phase comprising buffer A (0.1% formic acid in water) and buffer B (90% acetonitrile in 0.1% formic acid), injecting approximately 400 ng peptides. The chromatographic gradient consisted of 155 minutes from buffer A to buffer B at a flow rate of 300 nl/min with the following steps: 2 to 5% buffer B for 5 minutes, 5 to 25% buffer B for 110 minutes, 25 to 40% buffer B for 25 minutes, 40 to 80% buffer B for 5 minutes, and 80 to 95% buffer B for 5 minutes and hold for additional 5 minutes at 95% for Buffer B.

The serum samples were analyzed using the data independent acquisition (DIA) mode with the following parameters: for full-scan MS acquisition in the Orbitrap, the resolution was set to 120,000, with scan range of 350 to 1650 m/z, the maximum injection time of 100 ms, and automatic gain control (AGC) target of 3e6. The data was acquired using 17 DIA variable windows in the Orbitrap with a resolution set at 60,000, AGC target of 1e6, and the maximum injection time in auto mode.

The run order was randomized, but paired COVID-19 cases and controls were analyzed in sequence, with a quality control (QC) sample run approximately every 6 samples (**Supplementary Data File S1**). The QC sample consisted of pooled serum samples that had been processed in a way identical to that of the

experimental samples. QC samples were analyzed for the variability of the quantitation of the 334 proteins through calculating the coefficient of variance within each batch (**Supplementary Data File S1**). Proteins with less reliable quantitation are marked in figures.

Data analysis

Data preprocessing. We used Spectronaut for all primary processing (v14, https://biognosys.com/software/spectronaut/). All 74 raw files were first converted to the HTRMS format with the HTRMS converter (centroid method). The converted files were then analyzed with the directDIA method using default settings. We exported intensity information at the fragment level for data preprocessing.

We used in-house R scripts to eliminate the batch effect arising from events such as a change of chromatographic columns between the sample runs. Initially, the duplicate ions were removed. Ions with >= 50% presence across the samples were retained. The data were first normalized by subtracting the median intensity within each sample from all ion intensities from that sample. We then $\log_2$-transformed intensity values of the ions and applied Gaussian kernel smoothing with a window of 5 samples. Later, the median intensities were equalized across different batches and the data was transformed back to linear scale. We further applied mapDIA [81] to select fragment ions with the best quality to estimate protein levels. The protein levels were $\log_{10}$-transformed and the values were used for statistical testing.

Processing for visualization. For data visualization in heatmaps, the data was standardized by subtracting the row median intensity (across samples) from the intensity value of each protein. The heatmaps were generated using R-scripts and Perseus (version 1.5.5.1)[82,83]. Hierarchical clustering was performed using Perseus setting the complete linkage method and'1 - Pearson correlation' as the distance metric.

Statistical testing

Overall comparison. To assess the overall abundance difference between recuperated COVID-19 cases and healthy controls, we used a two-tailed, paired t-test.

Linear regression models. To examine the effect of the demographic variables including age, sex, and race, we used three univariate and three multivariate linear regression models. Univariate models considered each variable separately; multivariate models considered each variable in the context of all other variables. All models were evaluated with respect to the variable's impact on predicting the abundance of a specific protein amongst the samples from i) the healthy control data set, ii) the recuperated COVID-19 data set, iii) the data set of paired [$\log_{10}$(COVID-19/healthy control)] values. Note that due to the age/sex/race matching, dataset iii) intrinsically controlled for some of the effects of age, sex, and race already. Further, we also considered for healthy controls (dataset i)) variables including Presence of symptoms and Days since diagnosis which were derived from the corresponding sample of the recuperated COVID-19 patient. We included the variables for control purposes: as their role in healthy individuals is meaningless, we expected no significant associations between the Presence of symptoms and Days since diagnosis when modeling healthy controls (dataset i)). Indeed, we found only a few and minor associations in the univariate models. These associations were likely due to additional links, such as between age and Presence of symptoms.

We evaluated the following models: i) Univariate model: protein level (COVID) as a function of Age / Sex / Race / Presence of symptoms / Days since diagnosis; ii) Univariate model: protein level (CONTROL) as a function of Age / Sex / Race / Presence of symptoms / Days since diagnosis; iii) Univariate model: $\log_{10}$[protein level ratio (COVID/CONTROL)] as a function of Age / Sex / Race / Presence of symptoms /

Days since diagnosis; iv) Multivariate model: protein level (COVID) as a function of Age + Sex + Race + Presence of symptoms + Days since diagnosis; v) Multivariate model: protein level (CONTROL) as a function of Age + Sex + Race + Presence of symptoms + Days since diagnosis; and vi) Multivariate model: $\log_{10}$[protein level ratio (COVID/CONTROL)] as a function of Age + Sex + Race + Presence of symptoms + Days since diagnosis + Age*Sex. Due to the correlation between Titer and Symptoms as well as Days since diagnosis, we excluded Titer from the modeling to avoid overfitting. The main text focuses on the results of the multivariate models on CONTROL and COVID sets; the results from all models are presented in the **Supplementary Data File S1**, including corrections for multiple hypothesis testing (see below)**.**

Correction for multiple hypothesis testing. We corrected all P-values obtained from the overall comparison, the linear regression models, and ANOVA for multiple hypothesis testing using the Benjamini-Hochberg procedure [84]. To parse the results, we focussed on proteins with adjusted p-values < 0.05 as the primary set of proteins discussed. We considered an extended set with adjusted p-values < 0.20. All significance values are displayed in gray if not within these thresholds and in dark or light color if below the 0.05 or 0.20 threshold, respectively. Blue and red indicate the directionality of protein level difference. For visualization purposes only, we transformed adjusted p-values to derive a new value as (1-p) if the respective $\log_{10}$-transformed abundance fold change was positive, and as -(1-p) if the fold change was negative.

Post-translational modifications. To examine samples for changes in post-translational modifications, we constructed a Monohexose (+C6+H10+O5) and Dihexose (+C12+H20+O10) combined library with Spectronaut Pulsar (v14, https://biognosys.com/software/spectronaut/), using default settings except for the maximum number of variable modifications set to 3. Then we ran our DIA samples against this library, setting the minor grouping to "by modified sequence" and the differential abundance grouping to "minor" (peptide level). We exported peptide intensities from Spectronaut for further analysis. The data is available in Supplementary Data File S2.

We extracted peptides with Monohexose modifications as well as the corresponding unmodified peptides. We filtered out peptides with >20% missing values across the samples. We defined the modification level as $\log_2(\text{Intensity}_{\text{modified}}/(\text{Intensity}_{\text{unmodified}}+\text{Intensity}_{\text{modified}}))$ for each peptide. We calculated the $\log_2$ fold change between recuperated COVID-19 cases and healthy controls and used a paired t-test to determine the significance of the difference. We used the Benjamin-Hochberg procedure to adjust for multiple hypothesis testing. The Dihexose analysis resulted in only very few peptide hits.

Data availability

We deposited the raw files of the serum proteome in the PRIDE database [85] with the accession number PXD036597.

Author contributions

Christine Vogel conceptualized this project. Ted M. Ross and Michael A. Carlock provided samples. Shuvadeep Maity, Matthew Pressler, Justin Rendleman, and Burcu Vitrinel analyzed the samples. Smruti Pushalkar, Shaohuan Wu, Hyungwon Choi, and Christine Vogel performed the data analysis and wrote the manuscript. All authors read and approved of the submission of the manuscript.

Disclosure and competing interests statement

The authors declare that they have no conflict of interest.

Figures / Tables

Figure 1. Experimental design

A. Proteomic analysis of serum samples from 29 recuperated COVID-19 cases with age, sex, and race matched healthy controls. **B.-D.** Selection of demographics and their relationship with each other. **B.** Distribution of age between male and female cases. **C./D.** Distribution of Days since diagnosis and Antibody titer levels split by self-reported presence of symptoms during the acute phase. P-values were derived from t-tests. Symp. - symptoms.

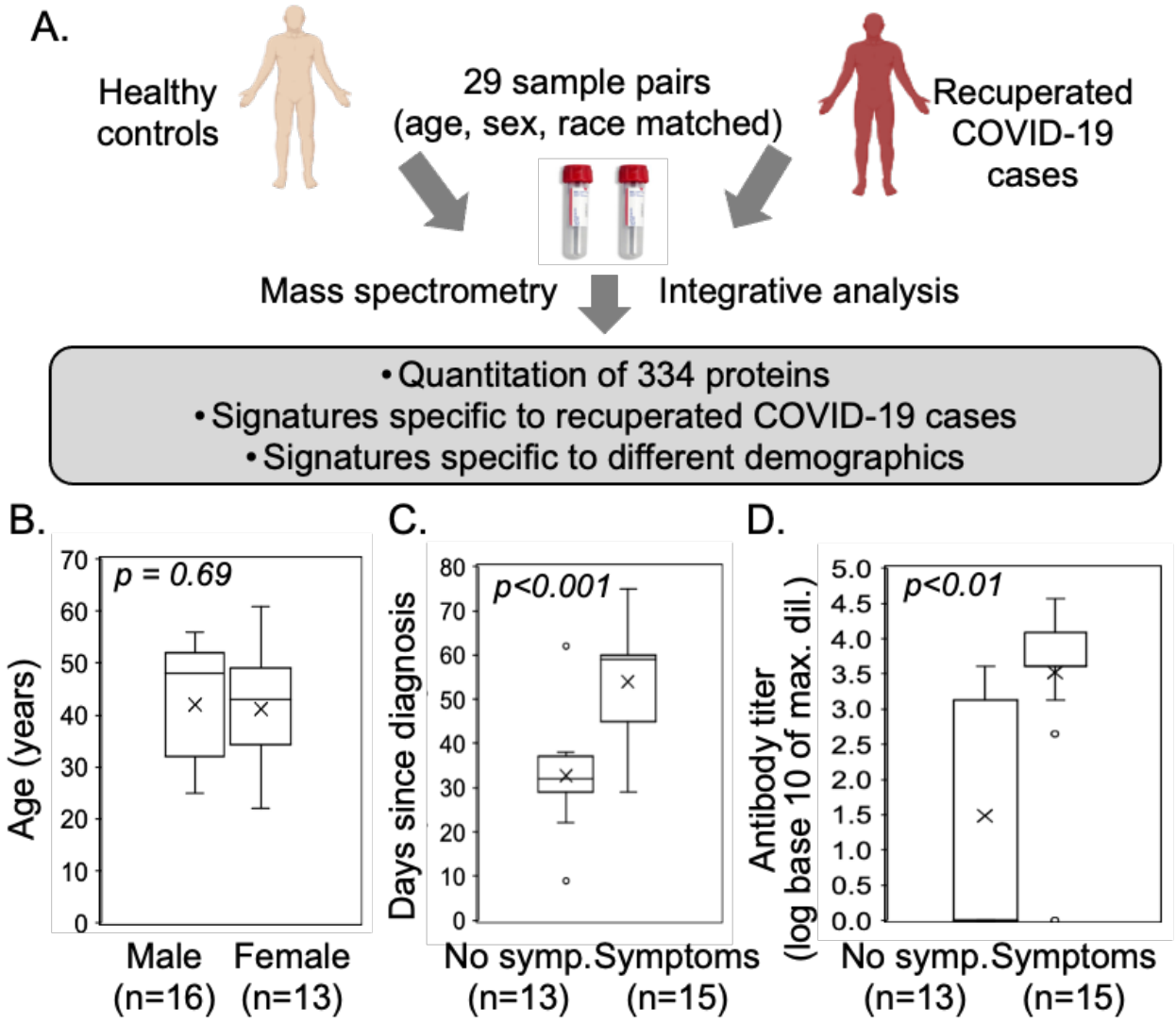

Figure 2. Serum levels of 334 serum proteins

Heatmap depicts normalized log_{10}-transformed levels for 334 proteins in sera from healthy controls and recuperated COVID-19 cases ordered with respect to sex and age. The upper panel provides additional sample information such as days passed since diagnosis of an acute COVID-19 infection and sample collection, self-reported presence of symptoms during the acute phase, antibody titer levels (log transformed), sex, age, and race of the individual. Hierarchical clustering identified 20 clusters with specific protein abundance patterns; letters indicate clusters discussed in detail. S.d. - since diagnosis

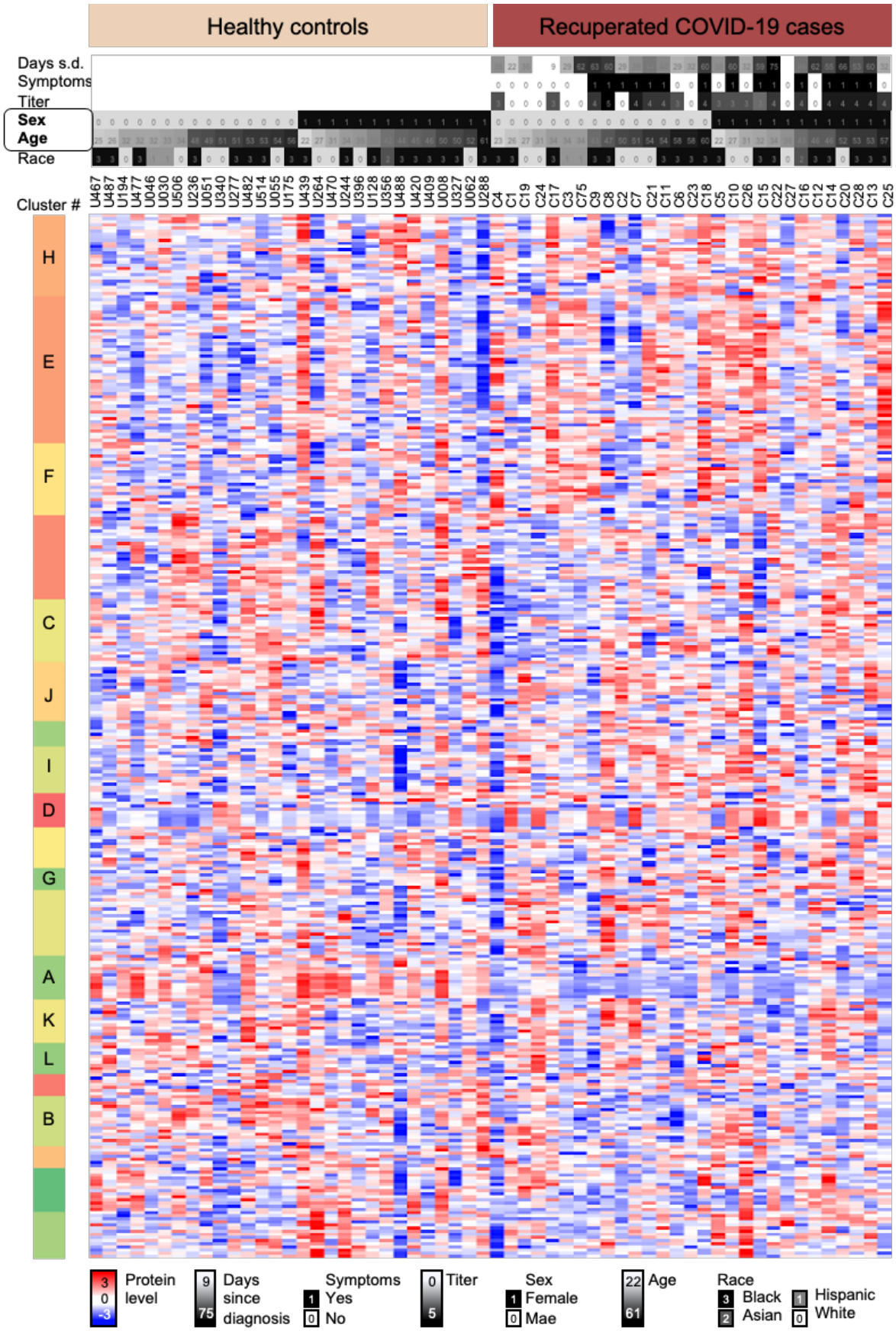
Healthy controls
Recuperated COVID-19 cases
Days s.d.
Symptoms
Titer
Sex
Age
Race
Cluster #
U467 U487 U194 U477 U046 U030 U506 U236 U051 U340 U277 U482 U514 U055 U175 U439 U264 U470 U244 U396 U128 U356 U488 U420 U409 U008 U327 U062 U288 C4 C1 C19 C24 C17 C3 C75 C9 C8 C2 C7 C21 C11 C6 C23 C18 C5 C10 C26 C15 C22 C27 C16 C12 C14 C20 C28 C13 C25
H
E
F
C
J
I
D
G
A
K
L
B
3 0 -3 Protein level
9 75 Days since diagnosis
Symptoms 1 Yes 0 No
0 5 Titer
Sex 1 Female 0 Mae
22 61 Age
Race 3 Black 2 Asian 1 Hispanic 0 White

Figure 3. Differential protein levels in recuperated COVID-19 cases and healthy controls

A. Partial least square discriminant analysis depicts distinct segregation between the two cohorts two major components. **B.** The volcano plot indicates fold changes and corresponding adjusted p-values of protein abundance between recuperated COVID-19 cases and healthy controls. Colored dots represent proteins with significantly higher abundance in the recuperated COVID-19 cohort (red) or healthy controls (blue), respectively (adjusted p-value <= 0.05). * Less reliable quantitation as determined by a high coefficient of variance across quality control samples (>50%)

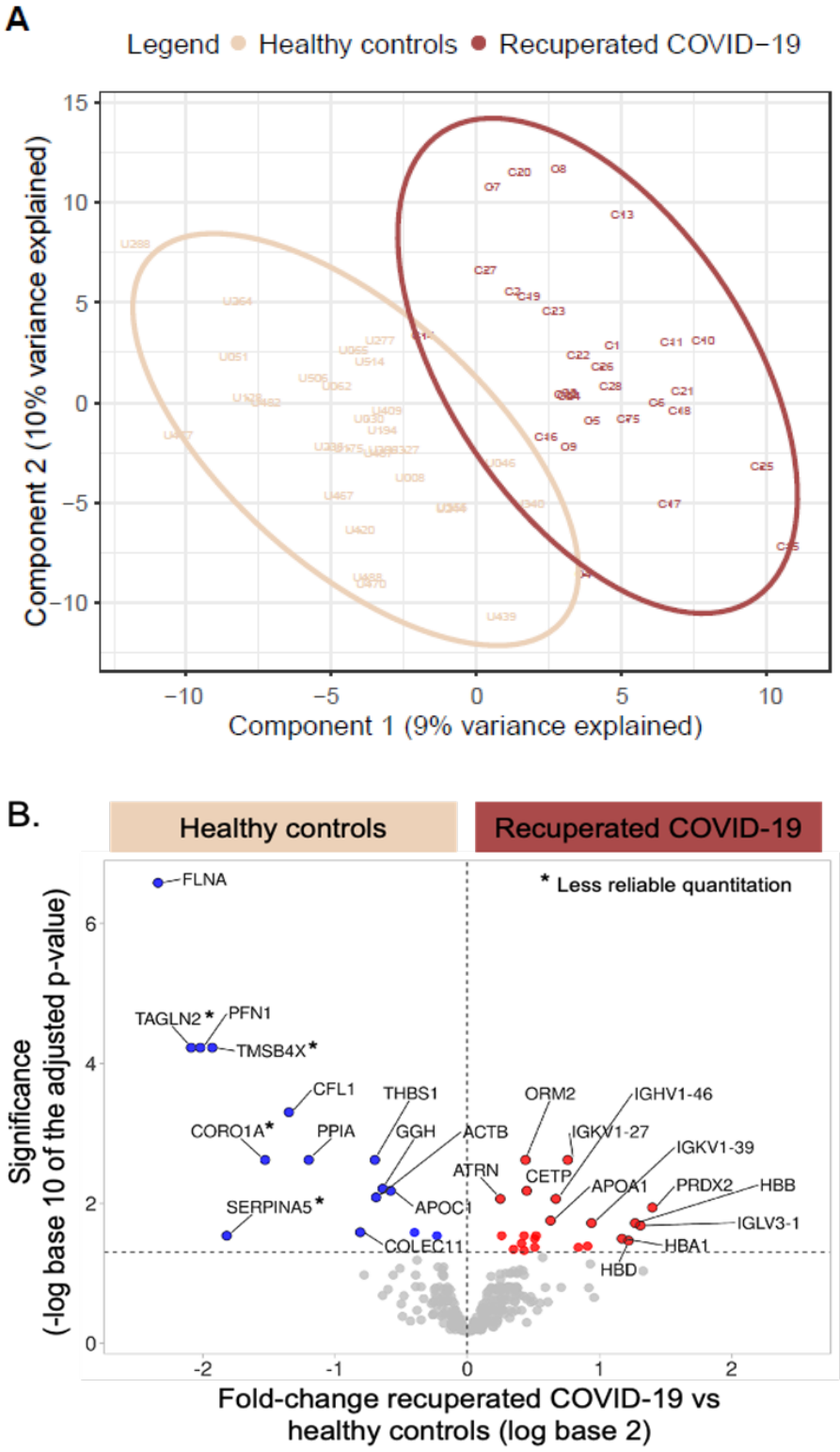
A
Legend
Healthy controls
Recuperated COVID−19
Component 2 (10% variance explained)
Component 1 (9% variance explained)
B.
Healthy controls
Recuperated COVID-19
* Less reliable quantitation
Significance
(-log base 10 of the adjusted p-value)
Fold-change recuperated COVID-19 vs
healthy controls (log base 2)
FLNA
TAGLN2*
PFN1
TMSB4X*
CFL1
THBS1
CORO1A*
PPIA
GGH
ACTB
ORM2
IGHV1-46
IGKV1-27
IGKV1-39
ATRN
CETP
APOA1
PRDX2
HBB
SERPINA5*
APOC1
IGLV3-1
COLEC11
HBA1
HBD

Figure 4. Protein abundance patterns marking differences between healthy controls and recuperated COVID-19 individuals independent of other factors

Heatmap showing example proteins from five clusters (marked in **Figure 2**). Panel **A**. shows the color-coded protein levels with samples sorted according to sex and age. Panel **B**. shows the significance values (associations) for different comparisons and different models. Significance values (adjusted p-values) were transformed as follows: if the observed $\log_{10}$-transformed abundance fold change was positive, we calculated 1-p; if negative, we calculated -(1-p). Dark colors indicate adjusted p-value <0.05; light colors adjusted p-value < 0.20; gray: no significance. Columns represent select comparisons: the Overall difference between recuperated COVID-19 cases and healthy controls; and the impact of Days since diagnosis, the presence of Symptoms, Sex, Age, and Race in a multivariate model using the healthy control (beige) or recuperated COVID-19 samples (brown). The complete results of the statistical testing are provided in Supplementary Data File S1. Example proteins were selected based on the statistical significance of the overall abundance difference between recuperated COVID-19 cases and healthy controls (with the adjusted p-value < 0.20 or < 0.05) and based on their relevance for the observed functional enrichment in each cluster. Functional enrichments were as follows: A: Cell adhesion, platelet degranulation; B and C: Innate immunity, complement system; D: Hemoglobin, adaptive immunity, immunoglobulins; and E: Adaptive immunity, immunoglobulins, complement system.

D.s.d. - Days since diagnosis; Titer - log base 10 transformed antibody titer as defined in methods; * Less reliable quantitation as determined by a high coefficient of variance across quality control samples (>50%)

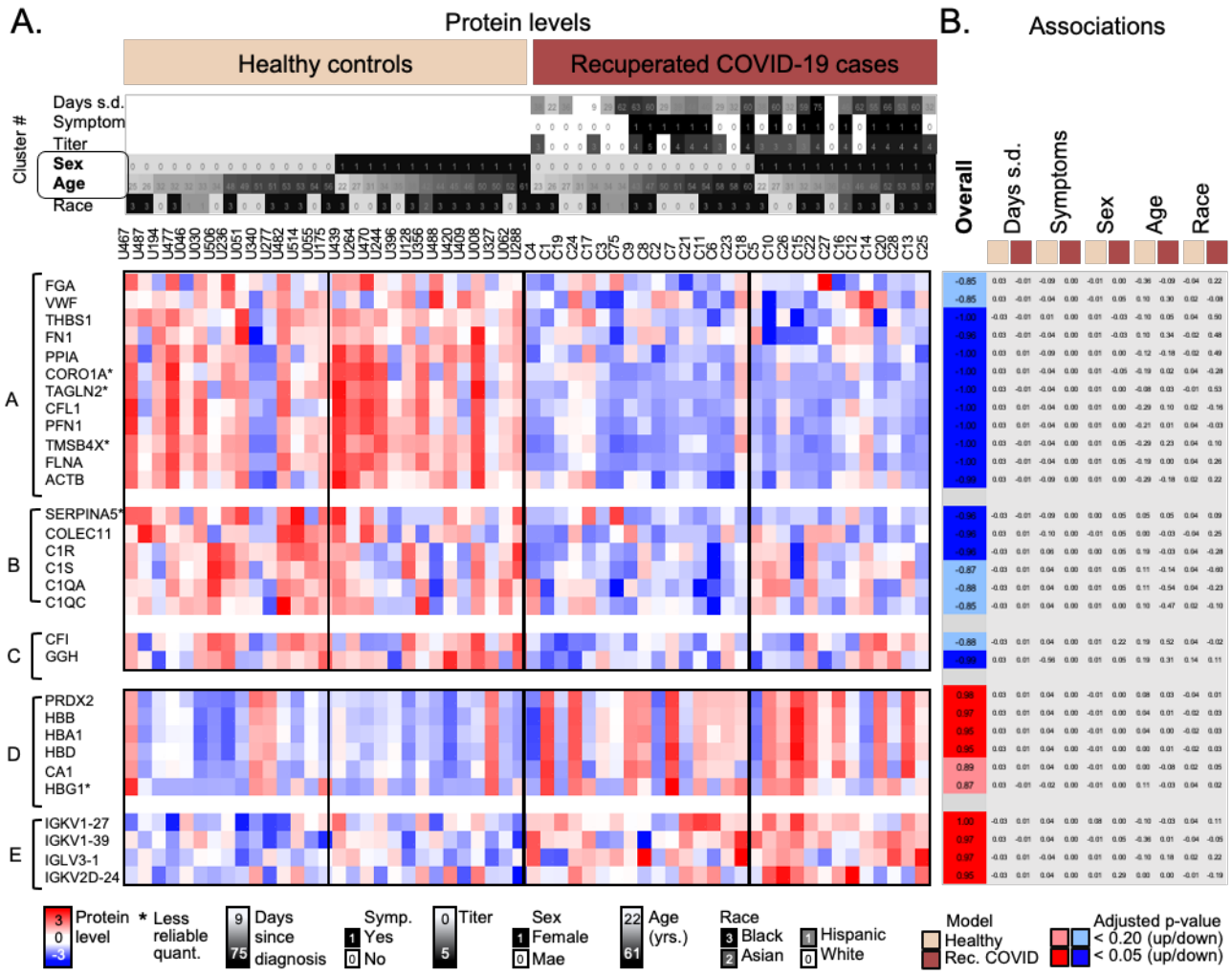
A.
Protein levels
Healthy controls
Recuperated COVID-19 cases
Cluster #
Days s.d.
Symptom
Titer
Sex
Age
Race
A
FGA
VWF
THBS1
FN1
PPIA
CORO1A*
TAGLN2*
CFL1
PFN1
TMSB4X*
FLNA
ACTB
B
SERPINA5*
COLEC11
C1R
C1S
C1QA
C1QC
C
CFI
GGH
D
PRDX2
HBB
HBA1
HBD
CA1
HBG1*
E
IGKV1-27
IGKV1-39
IGLV3-1
IGKV2D-24
B.
Associations
Overall
Days s.d.
Symptoms
Sex
Age
Race
Protein level
* Less reliable quant.
Days since diagnosis
Symp. Yes No
Titer
Sex Female Mae
Age (yrs.)
Race Black Asian Hispanic White
Model Healthy Rec. COVID
Adjusted p-value < 0.20 (up/down) < 0.05 (up/down)

Figure 5. Protein abundance differences between recuperated COVID-19 cases and healthy controls with interactions with other factors

Heatmap showing example proteins from seven clusters (marked in **Figure 2**). Panel **A**. shows the color-coded protein levels with samples sorted according to race and the presence of symptoms. Panel **B**. shows the significance values (associations) for different comparisons and different models. Significance values (adjusted p-values) were transformed as follows: if the observed $\log_{10}$-transformed abundance fold change was positive, we calculated 1-p; if negative, we calculated -(1-p). Dark colors indicate adjusted p-value <0.05; light colors adjusted p-value < 0.20; grey: no significance. Columns represent select comparisons: the Overall difference between recuperated COVID-19 cases and healthy controls; and the impact of Days since diagnosis, the presence of Symptoms, Sex, Age, and Race in a multivariate model using the healthy control (beige) or recuperated COVID-19 samples (brown). The complete results of the statistical testing are provided in Supplementary Data File S1. Example proteins were selected based on the statistical significance of the overall abundance difference between recuperated COVID-19 cases and healthy controls (with the adjusted p-value < 0.20 or < 0.05) and based on their relevance for the observed functional enrichment in each cluster. Function enrichments were as follows: F: Adaptive immunity, immunoglobulins; H: Adaptive immunity, immunoglobulins; I: Cholesterol transport; J: Lipid metabolism, cholesterol transport; K: Cell adhesion; G and L: n/a.

D.s.d. - Days since diagnosis; Titer - log base 10 transformed antibody titer as defined in methods; * Less reliable quantitation as determined by a high coefficient of variance across quality control samples (>50%)

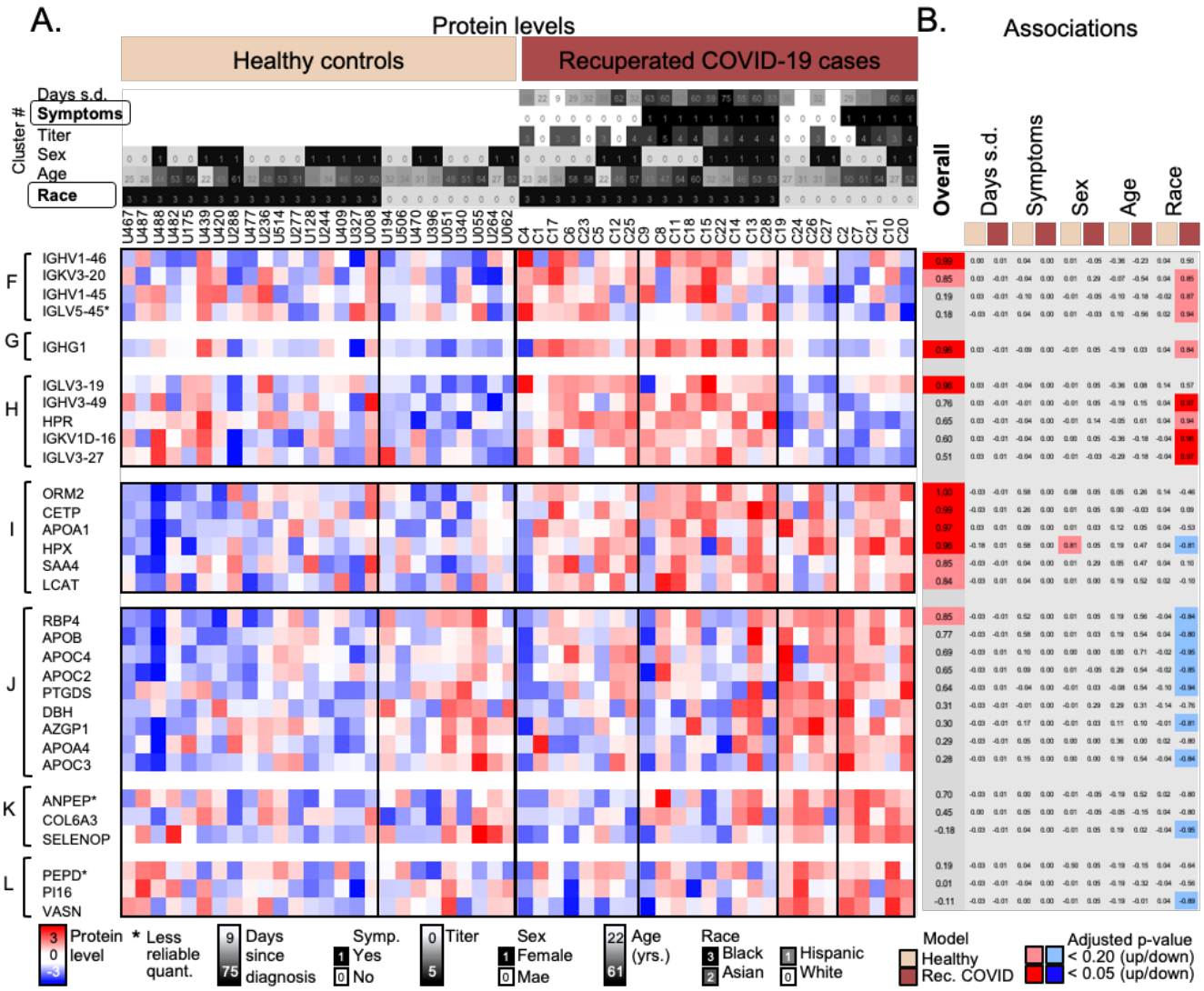
A.
Protein levels
Healthy controls
Recuperated COVID-19 cases
Cluster #
Days s.d.
Symptoms
Titer
Sex
Age
Race
F
IGHV1-46
IGKV3-20
IGHV1-45
IGLV5-45*
G
IGHG1
H
IGLV3-19
IGHV3-49
HPR
IGKV1D-16
IGLV3-27
I
ORM2
CETP
APOA1
HPX
SAA4
LCAT
J
RBP4
APOB
APOC4
APOC2
PTGDS
DBH
AZGP1
APOA4
APOC3
K
ANPEP*
COL6A3
SELENOP
L
PEPD*
PI16
VASN
B.
Associations
Overall
Days s.d.
Symptoms
Sex
Age
Race
Protein level
* Less reliable quant.
Days since diagnosis
Symp. Yes No
Titer
Sex Female Mae
Age (yrs.)
Race Black Asian
Hispanic White
Model Healthy Rec. COVID
Adjusted p-value < 0.20 (up/down) < 0.05 (up/down)

Table 1. Qualitative summary of the results

A. Proteins with altered levels during COVID-19 infection compared to healthy individuals as reported in literature, but with levels similar to those in healthy individuals (this study). **B.** Proteins with altered levels during acute COVID-19 infection and consistent, significant changes observed in this study (recuperated cases). **C.** Proteins with altered levels during acute COVID-19 infection and inconsistent, significant changes observed in this study (recuperated cases). Recuperation was defined as >ten days after diagnosis, with no symptoms present at the day of sample collection. Elevated, Lower - observed levels compared to healthy controls; Healthy - observed levels similar to healthy controls.

A.

Pathway	**Example proteins**	**Known:**	**This Study:**
		Acute phase COVID-19	**Recuperated COVID-19**
Complement cascade	C2, C3, C4A, C4B, C8, C9, MASP1, MBL2, FCN2	Elevated	Healthy
Coagulation system/thrombosis	F5, F10, FGB, FGG, KNG1	Elevated	Healthy
	F13B	Lower	
Adaptive immune system	IGLV3-25, IGKV1-5	Elevated	Healthy
Metabolism	LPA, PON1	Elevated	Healthy
	APOA2, APOA4	Lower	
Inflammation/Immune response	CRP, LDH, ORM1, S100A8, S100A9	Elevated	Healthy

B.

Pathway	**Example proteins**	**Known:**	**This Study:**
		Acute phase COVID-19	**Recuperated COVID-19**
Hemolysis	HBA1, HBB, HBD, PRDX2, CA1	Elevated	Elevated
Adaptive immune system	IGHG1, IGHV1-24, IGHV1-46, IGHV1-69, IGKV3D-15, IGLV3-19	Elevated	Elevated
Inflammation/Immune response	ORM2, SAA4	Elevated	Elevated

C.

Pathway	**Example proteins**	**Known:**	**This Study:**
		Acute phase COVID-19	**Recuperated COVID-19**
Complement cascade	C1R, C1S, CFI, COLEC11, VWF	Elevated	Lower
Coagulation system/thrombosis	VWF, FN1, THBS1	Elevated	Lower
Actin cytoskeleton/Cell adhesion/Platelet degranulation	PFN1, CFL1, FLNA, VWF	Elevated	Lower
Metabolism	CETP, APOA1	Lower	Elevated

References

1. Tzilas, V., Manali, E., Papiris, S. & Bouros, D. Intravenous Immunoglobulin for the Treatment of COVID-19: A Promising Tool. *Respiration; international review of thoracic diseases* vol. 99 1087–1089 (2020).
2. Laudanski, K. *et al.* Unbiased Analysis of Temporal Changes in Immune Serum Markers in Acute COVID-19 Infection With Emphasis on Organ Failure, Anti-Viral Treatment, and Demographic Characteristics. *Front. Immunol.* **12**, 650465 (2021).
3. Beltrán-Camacho, L. *et al.* The serum of COVID-19 asymptomatic patients up-regulates proteins related to endothelial dysfunction and viral response in circulating angiogenic cells ex-vivo. *Mol. Med.* **28**, 40 (2022).
4. Ahamed, J. & Laurence, J. Long COVID endotheliopathy: hypothesized mechanisms and potential therapeutic approaches. *J. Clin. Invest.* **132**, (2022).
5. Geyer, P. E. *et al.* High-resolution serum proteome trajectories in COVID-19 reveal patient-specific seroconversion. *EMBO Mol. Med.* **13**, e14167 (2021).
6. Park, J. *et al.* In-depth blood proteome profiling analysis revealed distinct functional characteristics of plasma proteins between severe and non-severe COVID-19 patients. *Sci. Rep.* **10**, 22418 (2020).
7. Messner, C. B. *et al.* Ultra-High-Throughput Clinical Proteomics Reveals Classifiers of COVID-19 Infection. *Cell Syst* **11**, 11–24.e4 (2020).
8. Fontana, S. *et al.* Comparative Proteome Profiling and Functional Analysis of Chronic Myelogenous Leukemia Cell Lines. *Journal of Proteome Research* vol. 6 4330–4342 Preprint at https://doi.org/10.1021/pr0704128 (2007).
9. Eslamifar, Z., Behzadifard, M., Soleimani, M. & Behzadifard, S. Coagulation abnormalities in SARS-CoV-2 infection: overexpression tissue factor. *Thromb. J.* **18**, 38 (2020).
10. Zhong, W. *et al.* Next generation plasma proteome profiling of COVID-19 patients with mild to moderate symptoms. *EBioMedicine* **74**, 103723 (2021).
11. Overmyer, K. A. *et al.* Large-Scale Multi-omic Analysis of COVID-19 Severity. *Cell Syst* **12**, 23–40.e7 (2021).
12. Broman, N., Feuth, T., Oksi, J. & COVIDSTORM study group. 'Early administration of tocilizumab in hospitalized COVID-19 patients with elevated inflammatory markers; COVIDSTORM' - Author's reply. *Clin. Microbiol. Infect.* (2022) doi:10.1016/j.cmi.2022.06.028.
13. Beltrami, A. P. *et al.* Combining Deep Phenotyping of Serum Proteomics and Clinical Data via Machine Learning for COVID-19 Biomarker Discovery. *Int. J. Mol. Sci.* **23**, (2022).
14. Vedula, P., Tang, H.-Y., Speicher, D. W., Kashina, A. & UPenn COVID Processing Unit. Protein Posttranslational Signatures Identified in COVID-19 Patient Plasma. *Front Cell Dev Biol* **10**, 807149 (2022).
15. Ali, K. M., Ali, A. M., Tawfeeq, H. M., Figueredo, G. P. & Rostam, H. M. Hypoalbuminemia in patients following their recovery from severe coronavirus disease 2019. *J. Med. Virol.* **93**, 4532–4536 (2021).
16. Duan, F. *et al.* A novel strategy for identifying biomarker in serum of patient with COVID-19 using immune complex. *Signal Transduct Target Ther* **7**, 63 (2022).
17. Java, A. *et al.* The complement system in COVID-19: friend and foe? *JCI Insight* **5**, (2020).
18. Lo, M. W., Kemper, C. & Woodruff, T. M. COVID-19: Complement, Coagulation, and Collateral Damage. *J. Immunol.* **205**, 1488–1495 (2020).
19. Afzali, B., Noris, M., Lambrecht, B. N. & Kemper, C. The state of complement in COVID-19. *Nat. Rev. Immunol.* **22**, 77–84 (2022).
20. Xie, Y. & Butler, M. Serum N-glycomic profiling may provide potential signatures for surveillance of COVID-19. *Glycobiology* **32**, 871–885 (2022).
21. Vicente, M. M. *et al.* Altered IgG glycosylation at COVID-19 diagnosis predicts disease severity. *European Journal of Immunology* vol. 52 946–957 Preprint at https://doi.org/10.1002/eji.202149491

(2022).
22. Patel, H. *et al.* Proteomic blood profiling in mild, severe and critical COVID-19 patients. *Sci. Rep.* **11**, 6357 (2021).
23. Al-Nesf, M. A. Y. *et al.* Prognostic tools and candidate drugs based on plasma proteomics of patients with severe COVID-19 complications. *Nat. Commun.* **13**, 946 (2022).
24. Nuñez, E. *et al.* Mapping the Serum Proteome of COVID-19 Patients; Guidance for Severity Assessment. *Biomedicines* **10**, (2022).
25. Wu, S. *et al.* Longitudinal Serum Proteome Characterization of COVID-19 Patients With Different Severities Revealed Potential Therapeutic Strategies. *Front. Immunol.* **13**, 893943 (2022).
26. Villar, M. *et al.* Characterization by Quantitative Serum Proteomics of Immune-Related Prognostic Biomarkers for COVID-19 Symptomatology. *Front. Immunol.* **12**, 730710 (2021).
27. Demichev, V. *et al.* A proteomic survival predictor for COVID-19 patients in intensive care. *PLOS Digital Health* **1**, e0000007 (2022).
28. D'Alessandro, A. *et al.* Serum Proteomics in COVID-19 Patients: Altered Coagulation and Complement Status as a Function of IL-6 Level. *J. Proteome Res.* **19**, 4417–4427 (2020).
29. Ladikou, E. E. *et al.* Von Willebrand factor (vWF): marker of endothelial damage and thrombotic risk in COVID-19? *Clin. Med.* **20**, e178–e182 (2020).
30. Haberichter, S. L. von Willebrand factor propeptide: biology and clinical utility. *Blood* **126**, 1753–1761 (2015).
31. Chen, J. & López, J. A. Interactions of platelets with subendothelium and endothelium. *Microcirculation* **12**, 235–246 (2005).
32. Goshua, G. *et al.* Endotheliopathy in COVID-19-associated coagulopathy: evidence from a single-centre, cross-sectional study. *Lancet Haematol* **7**, e575–e582 (2020).
33. Ko, Y.-P. & Flick, M. J. Fibrinogen Is at the Interface of Host Defense and Pathogen Virulence in Staphylococcus aureus Infection. *Semin. Thromb. Hemost.* **42**, 408–421 (2016).
34. Dolhnikoff, M. *et al.* Pathological evidence of pulmonary thrombotic phenomena in severe COVID-19. *Journal of thrombosis and haemostasis: JTH* vol. 18 1517–1519 (2020).
35. Rezaei-Tavirani, M. *et al.* Fibrinogen Dysregulation is a Prominent Process in Fatal Conditions of COVID-19 Infection; a Proteomic Analysis. *Archives of Academic Emergency Medicine* **9**, e26–e26 (2021).
36. Fard, M. B., Fard, S. B., Ramazi, S., Atashi, A. & Eslamifar, Z. Thrombosis in COVID-19 infection: Role of platelet activation-mediated immunity. *Thromb. J.* **19**, 59 (2021).
37. Hottz, E. D., Bozza, F. A. & Bozza, P. T. Platelets in Immune Response to Virus and Immunopathology of Viral Infections. *Front. Med.* **5**, 121 (2018).
38. Seyoum, M., Enawgaw, B. & Melku, M. Human blood platelets and viruses: defense mechanism and role in the removal of viral pathogens. *Thromb. J.* **16**, 16 (2018).
39. Biswas, S. *et al.* Blood clots in COVID-19 patients: Simplifying the curious mystery. *Med. Hypotheses* **146**, 110371 (2021).
40. Ruggeri, Z. M. The role of von Willebrand factor in thrombus formation. *Thromb. Res.* **120 Suppl 1**, S5–9 (2007).
41. Ruggeri, Z. M. Mechanisms initiating platelet thrombus formation. *Thromb. Haemost.* **78**, 611–616 (1997).
42. Liu, Y. *et al.* Association between platelet parameters and mortality in coronavirus disease 2019: Retrospective cohort study. *Platelets* **31**, 490–496 (2020).
43. Zhang, Y. *et al.* Mechanisms involved in the development of thrombocytopenia in patients with COVID-19. *Thromb. Res.* **193**, 110–115 (2020).
44. Guan, W.-J. *et al.* Clinical Characteristics of Coronavirus Disease 2019 in China. *N. Engl. J. Med.* **382**, 1708–1720 (2020).
45. Yang, X. *et al.* Thrombocytopenia and its association with mortality in patients with COVID-19. *J. Thromb. Haemost.* **18**, 1469–1472 (2020).
46. Li, Q. *et al.* Hematological features of persons with COVID-19. *Leukemia* **34**, 2163–2172 (2020).
47. Jiang, S.-Q., Huang, Q.-F., Xie, W.-M., Lv, C. & Quan, X.-Q. The association between severe COVID-19 and low platelet count: evidence from 31 observational studies involving 7613 participants. *Br. J.*

Haematol. **190**, e29–e33 (2020).
48. Asrie, F. *et al.* Baseline Thrombocytopenia and Disease Severity Among COVID-19 Patients, Tibebe Ghion Specialized Hospital COVID-19 Treatment Center, Northwest Ethiopia. *J. Blood Med.* **13**, 315–325 (2022).
49. Khave, L. J. *et al.* Association between thrombocytopenia and platelet profile with morbidity/mortality of severe and non-severe COVID-19 patients. *Rev. Assoc. Med. Bras.* **67**, 1670–1675 (2021).
50. Alonso-Beato, R. *et al.* Immune thrombocytopenia and COVID-19: Case report and review of literature. *Lupus* **30**, 1515–1521 (2021).
51. Bhattacharjee, S. & Banerjee, M. Immune Thrombocytopenia Secondary to COVID-19: a Systematic Review. *SN Compr Clin Med* **2**, 2048–2058 (2020).
52. Alharbi, M. G. *et al.* COVID-19 associated with immune thrombocytopenia: a systematic review and meta-analysis. *Expert Rev. Hematol.* **15**, 157–166 (2022).
53. Davoodian, A. *et al.* Severe Immune Thrombocytopenia Post-COVID-19: A Case Report. *Cureus* **13**, e19544 (2021).
54. Moulding, D. A., Record, J., Malinova, D. & Thrasher, A. J. Actin cytoskeletal defects in immunodeficiency. *Immunol. Rev.* **256**, 282–299 (2013).
55. Lyman, M. G. & Enquist, L. W. Herpesvirus interactions with the host cytoskeleton. *J. Virol.* **83**, 2058–2066 (2009).
56. Taylor, M. P., Koyuncu, O. O. & Enquist, L. W. Subversion of the actin cytoskeleton during viral infection. *Nat. Rev. Microbiol.* **9**, 427–439 (2011).
57. Nawaz-ul-Rehman, M. S. *et al.* Viral Replication Protein Inhibits Cellular Cofilin Actin Depolymerization Factor to Regulate the Actin Network and Promote Viral Replicase Assembly. *PLoS Pathog.* **12**, e1005440 (2016).
58. Shen, B. *et al.* Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell* **182**, 59–72.e15 (2020).
59. Wang, R., Cleary, R. A., Wang, T., Li, J. & Tang, D. D. The association of cortactin with profilin-1 is critical for smooth muscle contraction. *J. Biol. Chem.* **289**, 14157–14169 (2014).
60. He, S. *et al.* Cofilin hyperactivation in HIV infection and targeting the cofilin pathway using an anti-αβ integrin antibody. *Sci Adv* **5**, eaat7911 (2019).
61. Wu, Y. *et al.* Cofilin activation in peripheral CD4 T cells of HIV-1 infected patients: a pilot study. *Retrovirology* **5**, 95 (2008).
62. Heinrich, P. C., Castell, J. V. & Andus, T. Interleukin-6 and the acute phase response. *Biochem. J* **265**, 621–636 (1990).
63. Castell, J. V. *et al.* Interleukin-6 is the major regulator of acute phase protein synthesis in adult human hepatocytes. *FEBS Lett.* **242**, 237–239 (1989).
64. Conway, E. M. & Pryzdial, E. L. G. Is the COVID-19 thrombotic catastrophe complement-connected? *J. Thromb. Haemost.* **18**, 2812–2822 (2020).
65. Perico, L. *et al.* Immunity, endothelial injury and complement-induced coagulopathy in COVID-19. *Nat. Rev. Nephrol.* **17**, 46–64 (2021).
66. Li, H. *et al.* Serum Amyloid A is a biomarker of severe Coronavirus Disease and poor prognosis. *J. Infect.* **80**, 646–655 (2020).
67. Basavaraju, P. *et al.* Genetic Regulatory Networks of Apolipoproteins and Associated Medical Risks. *Front Cardiovasc Med* **8**, 788852 (2021).
68. Feingold, K. R. The bidirectional link between HDL and COVID-19 infections. *Journal of lipid research* vol. 62 100067 (2021).
69. Catapano, A. L., Pirillo, A., Bonacina, F. & Norata, G. D. HDL in innate and adaptive immunity. *Cardiovasc. Res.* **103**, 372–383 (2014).
70. AbdelHafez, M. A. Protective and therapeutic potentials of HDL and ApoA1 in COVID-19 elderly and chronic illness patients. *Bull. Natl. Salmon Resour. Cent.* **46**, 222 (2022).
71. McPherson, R. & Marcel, Y. Role of cholesteryl ester transfer protein in reverse cholesterol transport. *Clin. Cardiol.* **14**, I31–4 (1991).
72. Leng, L. *et al.* Sera proteomic features of active and recovered COVID-19 patients: potential diagnostic and prognostic biomarkers. *Signal transduction and targeted therapy* vol. 6 216 (2021).

73. Zinellu, A., Paliogiannis, P., Carru, C. & Mangoni, A. A. Serum amyloid A concentrations, COVID-19 severity and mortality: An updated systematic review and meta-analysis. *Int. J. Infect. Dis.* **105**, 668–674 (2021).
74. Zhu, Z. *et al.* Low serum level of apolipoprotein A1 may predict the severity of COVID-19: A retrospective study. *J. Clin. Lab. Anal.* **35**, e23911 (2021).
75. Ulloque-Badaracco, J. R., Hernandez-Bustamante, E. A., Herrera-Añazco, P. & Benites-Zapata, V. A. Prognostic value of apolipoproteins in COVID-19 patients: A systematic review and meta-analysis. *Travel Med. Infect. Dis.* **44**, 102200 (2021).
76. Carlson, B. A. *et al.* Selenoproteins regulate macrophage invasiveness and extracellular matrix-related gene expression. *BMC Immunol.* **10**, 57 (2009).
77. Huang, Z., Rose, A. H. & Hoffmann, P. R. The role of selenium in inflammation and immunity: from molecular mechanisms to therapeutic opportunities. *Antioxid. Redox Signal.* **16**, 705–743 (2012).
78. Zhang, J., Saad, R., Taylor, E. W. & Rayman, M. P. Selenium and selenoproteins in viral infection with potential relevance to COVID-19. *Redox Biol* **37**, 101715 (2020).
79. Hill, L. & Artiga, S. COVID-19 Cases and Deaths by Race/Ethnicity: Current Data and Changes Over Time. *KFF* https://www.kff.org/coronavirus-covid-19/issue-brief/covid-19-cases-and-deaths-by-race-ethnicity-current-data-and-changes-over-time/ (2022).
80. Allgoewer, K. *et al.* New Proteomic Signatures to Distinguish Between Zika and Dengue Infections. *Mol. Cell. Proteomics* **20**, 100052 (2021).
81. Teo, G. *et al.* mapDIA: Preprocessing and statistical analysis of quantitative proteomics data from data independent acquisition mass spectrometry. *Journal of Proteomics* vol. 129 108–120 Preprint at https://doi.org/10.1016/j.jprot.2015.09.013 (2015).
82. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
83. The R Project for Statistical Computing. https://www.R-project.org/.
84. Website. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x doi:10.1111/j.2517-6161.1995.tb02031.x.
85. Perez-Riverol, Y. *et al.* The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).