

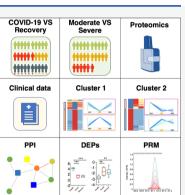
Article

# Proteomic Analysis Identifies Prolonged Disturbances in Pathways Related to Cholesterol Metabolism and Myocardium Function in the COVID-19 Recovery Stage

Yang Chen,<sup> $\nabla$ </sup> Hangping Yao,<sup> $\nabla$ </sup> Nan Zhang,<sup> $\nabla$ </sup> Jie Wu,<sup> $\nabla$ </sup> Shuaixin Gao, Jiangtao Guo, Xiangyun Lu, Linfang Cheng, Rui Luo, Xue Liang, Catherine C. L. Wong,<sup>\*</sup> and Min Zheng<sup>\*</sup>



attention has been given to the disease sequelae in patients recovering from COVID-19, and nothing is known about the mechanisms underlying these sequelae. Herein, we profiled the serum proteome of a cohort of COVID-19 patients in the disease onset and recovery stages. Based on the close integration of our proteomic analysis with clinical data, we propose that COVID-19 is associated with prolonged disorders in cholesterol metabolism and myocardium, even in the recovery stage. We identify potential biomarkers for these disorders. Moreover, severely affected patients presented more serious disturbances in these pathways. Our findings potentially support clinical decision-making to improve the prognosis and treatment of patients.



**KEYWORDS:** COVID-19, sequelae, quantitative proteomics, cholesterol metabolism, cardiomyopathy

## INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the disease it causes, coronavirus disease 2019 (COVID-19), first emerged in 2019 and have proceeded to cause a global health pandemic that continues on an unprecedented scale until today. The number of COVID-19 deaths around the world has surpassed 3.5 million, with more than 170 million people infected. Although first characterized as a primarily respiratory syndrome, COVID-19 is now known to involve multiple organ systems, including the heart, gastrointestinal tract, kidneys, liver, muscle, and central nervous system.<sup>1-3</sup> Several independent proteomic studies have revealed that SARS-CoV-2 infection triggers disruption of multiple biological pathways, including the immune response, coagulation, and metabolism.<sup>4-7</sup> Furthermore, studies found that in patients who had recovered from COVID-19, 87.4% reported persistence of fatigue and dyspnea. However, the underlying mechanisms are unknown and, therefore, no strategies exist to prevent these sequelae.<sup>8</sup> Thus, there is an urgent need for a systematic understanding of the disease mechanisms on a molecular level throughout the entire course of the disease including recovery.

Here, to have a complete picture of the disease onset, progression, and recovery, we established a long-term cohort of 10 patients with a 6 month follow-up. We carried out quantitative proteomic analysis of patient samples and we analyzed the data in conjunction with clinical results from regular blood tests. We found that the levels of a group of serum proteins related to infection changed after disease onset but were restored in the recovery stage, possibly representing disease stress reactions. Intriguingly, we found that another group of proteins changed during the disease stage but did not return to normal in the recovery stage, suggesting the presence of long-lasting disorders. Pathways functionally linked to the persistently changed proteins included blood coagulation, fibrin clot formation, cholesterol metabolism, and cardiopathy. These prolonged disturbances are more serious in severe than moderate patients, suggesting a possible correlation of these disorders with disease severity. We also identified potential biomarkers that could be used as indicators to prevent those disorders in advance before clinical symptoms emerge.

Our study brings attention to patient prognosis in the recovery stage. The combination of quantitative proteomic data and regular clinical serum tests shed light on the molecular changes that occur during both the disease and recovery stages. Our results potentially support clinical

Received: January 20, 2021



#### Journal of Proteome Research

decision-making to improve the prognosis and treatment of patients.

#### EXPERIMENTAL SECTION

#### Sample Collection

Serum samples were collected from The First Affiliated Hospital, School of Medicine, Zhejiang University. Blood samples for research purposes were collected in procoagulation vacuum tubes using standard venepuncture protocols. Serum was extracted by centrifugation for 10 min at 3000 rpm. The serum samples were inactivated at 65 °C for 1 h and subsequently stored at -80 °C before use. The patients with H7N9 avian influenza virus infection were recruited between 1 April and 10 June 2013. The H1N1 flu patients were retrospectively selected from the Inpatient Department of The First Affiliated Hospital of Zhejiang University School of Medicine. Patients were diagnosed according to their clinical symptoms and qPCR.

The study was approved by the Ethics Committee of The First Affiliated Hospital of Zhejiang University School of Medicine (ethical approval no. 2020IIT-255). All experiments with live virus were performed in a biosafety level 3 laboratory approved by The First Affiliated Hospital, School of Medicine Zhejiang University.

# Depletion of Highly Abundant Proteins from Serum Samples

High-Select Top14 Abundant Protein Depletion of serum samples was conducted according to the manufacturer instructions of "Multi Affinity Removal Column, Human-14" (Agilent Technologies). Briefly, 20  $\mu$ L of serum and 60  $\mu$ L of buffer A were mixed evenly, transferred into a Spin-X centrifuge tube filter 0.22  $\mu$ m cellulose acetate certified tube (COSTOR), and subsequently centrifuged at 16,000g for 1 min. The filtrate was removed into 29 × 5 mm preassembled plastic springs for High-Select Top14 Abundant Protein Depletion using a liquid chromatography system (Agilent Technologies 1290). The depleted samples were collected and the protein concentration was measured using a BCA Protein Assay Kit (Thermo Scientific) following the manufacturer's instructions.

#### Sample Preparation for DIA Analysis

The depleted serum sample (50  $\mu$ g) was precipitated by 1/3 volume of trichloroacetic acid (TCA) solution at 4 °C overnight and then centrifuged at 16,000g for 30 min at 4 °C. The precipitate was washed three times with acetone and then dried with a vacuum concentrator (Labconco, USA). The dried precipitate was dissolved in 40  $\mu$ L of 8 M urea (pH 8.5), incubated with 20 mM (2-carboxyethyl)phosphine hydrochloride (TCEP), and alkylated with 40 mM IAA. The mixture was diluted with 200  $\mu$ L of 100 mM Tris-HCl buffer (pH 8.5) to a final concentration of 1.3 M urea and then digested with 3  $\mu$ g of trypsin protease at 37 °C for 16 h. The sample was desalted using a Monospin C18 column (GL Science, Tokyo, Japan). The desalted peptides were vacuum-centrifuged to dryness and reconstituted in Milli-Q water with 0.1% formic acid (FA) for LC-MS analysis. Indexed retention time (IRT) calibration peptides were spiked into the samples before DIA analysis.

#### pubs.acs.org/jpr

#### Sample Preparation for Spectral Library Generation

The mixed protein  $(88 \ \mu g)$  from all samples was processed as above. Purified peptides were dissolved in 80  $\mu$ L of buffer (98% H<sub>2</sub>O and 2% acetonitrile).

#### High-pH Reversed-Phase Fractionation

The 80  $\mu$ L mixed peptides were fractioned into 62 fractions within 80 min on a chromatographic system (Waters Xevo ACQUITY UPLC, USA). The first fraction was mixed with the 62nd fraction, the second was mixed with the 61st fraction, and so on. All 31 mixed fractions were vacuum-centrifuged to dryness and dissolved in 10  $\mu$ L Milli-Q water with 0.1% formic acid (FA). IRT peptides were spiked before data-dependent acquisition analysis.

#### Liquid Chromatography

Peptide separation was carried out on a nanoElute liquid chromatography system (Bruker Daltonics). Digested peptides (200 ng) were separated within 90 min at a flow rate of 300 nL/min on a homemade column (25 cm  $\times$  75  $\mu$ m, 1.5  $\mu$ m C18-AQ particles (Dr. Maisch)). Mobile phases A and B were water and ACN with 0.1% formic acid, respectively. The %B was linearly increased from 2 to 22% within 70 min, then increased to 37% within 8 min, then further increased to 95% within 5 min, and finally maintained at 95% for the last 7 min.

## Mass Spectrometry

All 31 fraction samples were analyzed on a hybrid TIMS quadrupole time-of-flight mass spectrometer (Bruker timsTOF Pro) via a CaptiveSpray nanoelectrospray ion source. The mass spectrometer was operated in data-dependent mode for generation of the spectral library and in data-independent mode for sample analysis. The accumulation and ramp time were set at 100 ms each and the mass spectra were recorded in the range from m/z 100 to 1700 in positive electrospray mode. The ion mobility was scanned from 0.6 to 1.6 V s/cm<sup>2</sup>. The overall acquisition cycle of 1.16 s comprised 1 full TIMS-MS scan and 10 PASEF MS/MS scans. We define quadrupole isolation windows as a function of the TIMS scan time to achieve seamless and synchronous ramps for all applied voltages. Eight windows were defined for single 100 ms TIMS scans according to the m/z ion mobility plane. During scanning mode, the collision energy was ramped linearly as a function of the mobility from 59 eV at  $1/K_0 = 1.6 \text{ V s/cm}^2$  to 20 eV at  $1/K_0 = 0.6$  V s/cm<sup>2</sup>. For PRM confirmation, the peptide precursors of selected differential protein precursors were scheduled with 120 min retention time windows using a Q-Exactive HF mass spectrometer (Thermo Fisher Scientific, CA, USA). The PRM quantification was performed by Skyline software.

#### Generation of Spectral Libraries and DIA Data Analysis

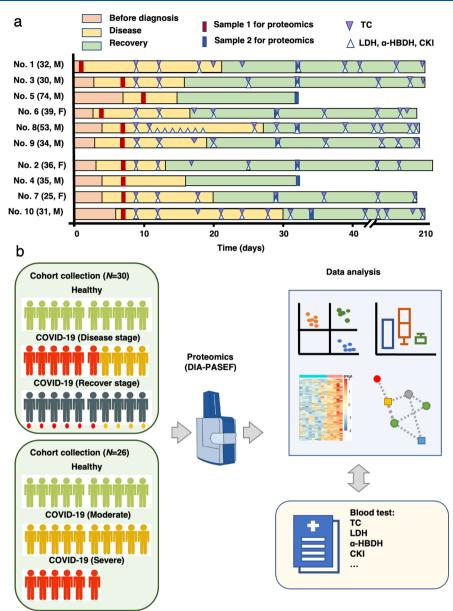
Spectral libraries were generated with Spectronaut version 14.2 (Biognosys) and all the parameters were default. The MS/MS spectra were matched against the human UNIPROT database (20,421 human entries, downloaded July 2019) and the SARS-CoV-2 UNIPROT database.

#### **Statistical Analysis**

OmicsBean software was used for data analysis including data imputation, normalization, and principal component analysis (PCA). Genefilter was used in calculation of the fold-change values of proteins. A fold change of 2, a fold change of <0.5, and a *P* value (*t* test) of 0.05 were used to filter differential expression proteins. Mfuzz was used to detect different

#### Journal of Proteome Research

pubs.acs.org/jpr

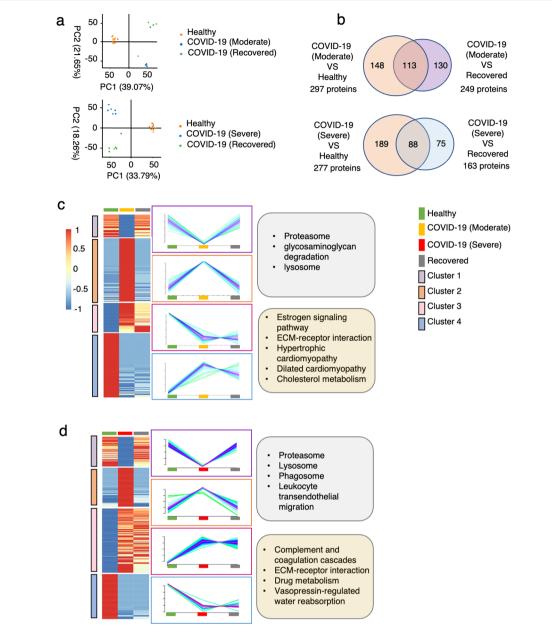


**Figure 1.** Summary of COVID-19 patient samples and experimental design. (a) Timeline of the disease course and sample acquisition in 10 COVID-19 patients including 6 severely affected patients (upper bars) and 4 moderately affected patients (lower bars). Further details are presented in Tables S1 and S2. For each patient, the disease stage (yellow bar) and recovery stage (green bar) are separated by a negative nasopharyngeal swab real-time PCR test for SARS-CoV-2. Samples were taken for proteomic studies during both the disease stage (thick red lines) and the recovery stage (thick blue lines). The triangles represent the time points for testing blood total cholesterol (purple triangles) and profiling myocardial enzymes (white triangles). (b) Experimental design for the quantitative proteomic analysis in this study. A total of 30 serum samples were taken from three groups (10 from healthy donors and 20 from patients in both the disease stage and the recovery stage as marked in A) and processed for quantitative proteomic analysis. Green, yellow, and red represent healthy donors, moderate patients, and severe patients, respectively. At the same time, by adding another six samples from moderate patients, a total of 26 samples were analyzed from three new groups: 10 samples from healthy donors, 10 samples from moderate patients, and 6 samples from severe patients.

subclustering models of gene expression among groups. Venn diagram, heatmap, and network visualization were performed using the ggplot228 package and Cytoscape v.3.5.129 implemented in the OmicsBean workbench. Ingenuity pathway analysis (IPA) was performed to explore the downstream effects in the dataset of significantly up- or downregulated proteins. The z-score algorithm was used to predict the activation state (either activated or inhibited) of each biological process. If the z-score  $\leq -2$ , then the process is predicted to be statistically significantly inhibited.

#### **Data Availability**

The experimental data that support the findings of this study have been deposited in iProX (integrated proteome resources) of ProteomeXchange with the accession code PXD023305. Reference FASTA files contain human UNIPROT database (only reviewed entries) (20,421 human entries, downloaded July 2019) and SARS-CoV-2 UNIPROT database that combined with SARS protein database (38 viral entries, April 2020). The latest GO database (https://www.ebi.ac.uk/ QuickGO/) was used for pathway enrichment analysis. The source data underlying Figures 3c-h, 4c,d, and 5d, and Figures



**Figure 2.** Quantitative proteomic analysis of serum samples from the disease and recovery stages in moderate or severe patients. (a) Principal component analysis (PCA) showing the intergroup differences. Individuals in the healthy control group, the moderate or severe COVID-19 groups, and the recovered moderate or severe COVID-19 groups are indicated by colored dots in the figure. (b) Venn diagram of significantly changed proteins (cutoff value of fold change > 2 and fold change < 0.5; *P* (*t* test) < 0.05) in the moderate or severe COVID-19 group compared to the healthy control group and in the recovered moderate or recovered severe group compared to the healthy control group. Fold change indicates the protein level of the moderate or severe group to the healthy group. (*c*, d) Heatmap of proteomic changes in healthy controls, moderate patients, and recovered moderate patients (*c*) and in healthy controls, severe patients, and recovered severe patients (d). The color bar from red to blue represents the fold change from increasing to decreasing of all proteins identified in each group. Hierarchical clustering shows a clear group differentiation according to similarity. Intensity profiles and selected enriched KEGG pathways are indicated for the marked clusters. The color bar represents *z*-score change from -1 to 1.

S1, S2c, S3, S4, and S5b are provided as a source data file. Source data are provided with this paper.

#### RESULTS

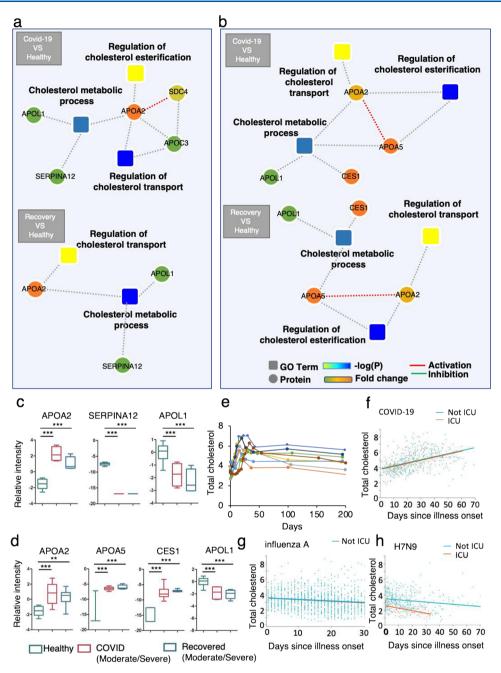
Journal of Proteome Research

#### Quantitative Proteomic Profiling of COVID-19 Serum in Both Disease and Recovery Stages

We collected serum samples from a cohort of 10 COVID-19 patients and 10 healthy donors. Four of the patients were diagnosed with moderate symptoms and six were diagnosed with severe symptoms according to the Chinese management guideline for COVID-19 (version 6.0) (http://www.nhc.gov. cn/yzygj/s7653p/202002/8334a8326dd94d329df351d 7da8aefc2/files/b218cfeb1bc54639af227f922bf6b817.pdf). Of note, patients were monitored by regular blood testing at multiple time points during the entire disease process from admission to recovery in a time range of about 6 months (Figure 1a and Tables S1 and S2). The samples for quantitative proteomics were taken from each patient at two time points (Figure 1a). The first sample was taken shortly after the patient was diagnosed with COVID-19, as confirmed

Article

Article

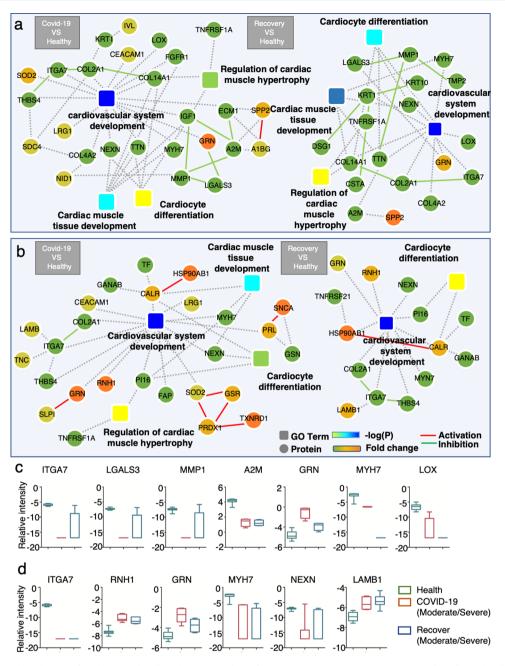


**Figure 3.** Long-term disturbance of pathways related to cholesterol metabolism in COVID-19 patients, extending from the disease stage to the recovery stage. (a, b) Interaction diagrams of the cholesterol metabolic process, regulation of cholesterol transport, and regulation of cholesterol esterification when moderate patients (a, top) or severe patients (b, top) are compared to healthy controls and when recovered moderate patients (a, bottom) or recovered severe patients (b, bottom) are compared to healthy controls. Proteins selected for analysis meet the conditions of fold change > 2 and fold change < 0.5; *P* (*t* test) < 0.05. The significance of the pathways represented by  $-\log(P \text{ value})$  (Fisher's exact test) is shown by the color scale at the bottom of panel (b), with dark blue representing the highest significance. The color bar from orange to green represents the fold change of the protein level from increasing to decreasing. Fold change indicates the protein level of the moderate or severe group to the healthy group. (*c*, d) Box-and-whisker plots showing the relative intensities of the indicated proteins in moderate (c) or severe (d) COVID-19 patients. These proteins are potential markers for long-term disorders of cholesterol metabolism. One-way ANOVA was used to analyze the data. For the healthy group, *n* = 6; for the recovered severe COVID-19 group, *n* = 6. Data are presented as median, upper and lower quartile, and min and max. (e) Long-term monitoring of total serum cholesterol levels in COVID-19 patients throughout the disease and recovery stages. (f–h) Comparison of serum total cholesterol levels of COVID-19 patients (f), influenza A patients (g), and avian influenza H7N9 patients (h) in days since disease onset. The orange lines indicate ICU admission; the blue lines indicate no ICU admission.

by a positive test for SARS-CoV-2 based on real-time PCR (RT-PCR) analysis of a throat-swab specimen. The second sample was taken several days after the patient tested negative for SARS-CoV-2. To tackle the specific proteomic features of

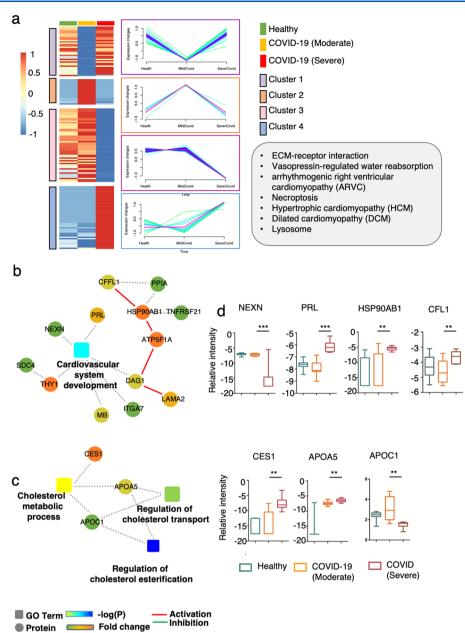
severe cases, we also obtained serum samples from a further six patients diagnosed with moderate COVID-19 to constitute a cohort containing 10 healthy controls, 10 moderate COVID-19 patients, and 6 severe COVID-19 patients (Figure 1b and

Е



**Figure 4.** Long-term disturbance of pathways related to the myocardium function in COVID-19 patients, extending from the disease stage to the recovery stage. (a, b) Interaction diagrams of the regulation of cardiac muscle hypertrophy, cardiac muscle tissue development, cardiocyte differentiation, and cardiovascular system development when moderate patients (a, left) or severe patients (b, left) are compared to healthy controls and when recovered moderate patients (a, right) or severe recovered patients (b, right) are compared to healthy controls. Proteins selected for analysis meet the conditions of fold change > 2 and fold change < 0.5; *P* (*t* test) < 0.05. The significance of the pathways represented by  $-\log(P value)$  (Fisher's exact test) is indicated by the color scale (b, bottom), with dark blue representing the highest significance. Fold change indicates the protein level of the moderate or severe group to the healthy group. (d) Box-and-whisker plots showing the relative intensities of the indicated proteins in moderate (c) or severe (d) COVID-19 patients. These proteins are potential markers for long-term disorders of the myocardium. One-way ANOVA was used to analyze the data. For the healthy group, *n* = 10; for the moderate COVID-19 group, *n* = 4; for the recovered moderate COVID-19 group, *n* = 6. Data are presented as median, upper and lower quartile, and min and max.

Table S2). All samples from the two cohorts were first subjected to depletion of high-abundance serum proteins and then processed by data-independent acquisition (DIA) mass spectrometry. The resulting quantitative proteomic data were subsequently analyzed for enriched KEGG pathways (Figure 1b). The proteomic data were correlated with clinical indicators to obtain more accurate dissection of the results. The raw data were first processed using a double boundary Bayes (DBB) imputation method and were standardized for further analysis. The data showed clear stratification of the different patient/control groups according to principal component analysis (PCA) (Figure 2a). A total of 1056 proteins were identified from all samples. The levels of 261 and 277 proteins significantly changed in moderate and severe



**Figure 5.** Quantitative proteomic analysis reveals disturbances in pathways related to cholesterol metabolism and myocardium function when severe COVID-19 patients are compared to moderate patients. (a) Proteomic changes in healthy controls, moderate patients, and severe patients. Heatmap of 157 proteins that are significantly changed when severe COVID-19 patients are compared to moderate COVID-19 patients. The color bar from red to blue represents the fold change from increasing to decreasing of all proteins identified in each group. Hierarchical clustering shows a clear group differentiation according to similarity. Intensity profiles and selected enriched KEGG pathways are indicated for the marked clusters. The color bar represents z-score change from -1 to 1. (b, c) Interaction diagrams of regulation of cardiovascular system development (b) or the cholesterol metabolic process, regulation of cholesterol transport, and regulation of cholesterol esterification (c) when severe COVID-19 patients are compared to moderate COVID-19 patients. Proteins selected for analysis meet the condition of fold change > 2 and fold change < 0.5; *P* (*t* test) < 0.05. The significance of the pathways represented by  $-\log(P \text{ value})$  (Fisher's exact test) is indicated by the color scale, with dark blue representing the most significant difference. Fold change indicates the protein level of the severe to moderate group. (d) Box-and-whisker plots showing the relative intensities of proteins related to cholesterol metabolism or myocardium function in healthy controls, moderate COVID-19 patients, or severe COVID-19 patients. One-way ANOVA was used to analyze the data. For the healthy group, *n* = 10; for the moderate COVID-19 group, *n* = 6. Data are presented as median, upper and lower quartile, and min and max.

COVID-19 patients, respectively, compared to healthy controls. Surprisingly, in recovered moderate and recovered severe patients, 243 and 163 proteins were significantly changed compared to healthy controls, and of these, 113 and 88 proteins overlapped with the disease-stage proteins (Figure 2b). Thus, a substantial number of proteins remain abnormal in the recovery stage. Next, subclustering analysis was performed on the significantly changed 261 or 277 proteins corresponding to moderate diseases to analyze their variation at health, disease, or recover stages. When moderate COVID-19 patients were compared to healthy controls, the significantly changed proteins showed four different subclustering patterns (Figure 2c). Subclusters 1 and 2 contain proteins that changed due to SARS-CoV-2 infection but were restored in the recovery stage. The enriched KEGG pathways include proteasome, glycosaminoglycan degradation, lysosome, and so forth, which partially represent the stress reaction to viral infection (Figure 2c). Among these proteins are GP70, CRP, SAA1, and SAA2, which are known to be biomarkers for acute phase response after infection and other stimuli<sup>9–11</sup> (Figure S1). However, more interestingly, subclusters 3 and 4 contain proteins that are significantly altered during infection but are not restored in the recovery stage. The highlighted pathways include ECM–receptor interaction, cholesterol metabolism,

dilated cardiomyopathy, and hypertrophic cardiomyopathy (Figure 2c). Similar patterns were found for the severe COVID-19 patients (Figure 2d). The proteomic analysis prompted us to look in more detail at the proteins and related pathways that showed long-term disturbance.

#### Long-Term Disruption of Proteins Related to Cholesterol Metabolism and Myocardium Function in COVID-19 Patients in the Disease and Recovery Stages

We found that proteins that changed significantly during infection and then remained changed in the recovery stage were enriched in functions related to blood coagulation, fibrin clot formation, cholesterol metabolism, and myocardium. Consistent with our proteomic data, thrombotic complications have been reported in COVID-19 patients.<sup>12</sup> We found that disturbed blood coagulation might be a long-lasting problem for patients (Figure S2). More intriguingly, we found that proteins in subclusters 3 and 4 from moderate COVID-19 patients were enriched in functions related to the cholesterol metabolic process, cholesterol transport and cholesterol esterification. Specifically, the level of APOL1, APOA2, and SERPINA12 were altered both in disease and recovery stages (Figure 3a,c). Proteins in subclusters 3 and 4 from severe COVID-19 patients showed similar results, with APOA5, APOL1, APOA2, and CSE1 identified (Figure 3b,d). Apolipoprotein A-II (ApoA2) and Apolipoprotein L1 (ApoL1) identified in this study are components of highdensity lipoproteins (HDLs), which are important for cholesterol reverse transport. The disturbance of serum HDLs may suggest defect in cholesterol reverse transport and cholesterol functions. HDLs also display major pleiotropic functions that could play a pivotal role in acute inflammatory conditions. Evidence suggests that HDLs are globally protective for the endothelial layer. Both HDL composition and functionality are profoundly modified under pathological conditions. One study showed that HDLs from COVID-19 patients were less protective for endothelial cells stimulated with TNF $\alpha$  than HDLs from healthy subjects.<sup>13</sup> The proteins identified may be potential indicators of a prolonged cholesterol metabolism disorder. From the clinical data obtained from patient blood samples (Figure 1a), we found that the serum total cholesterol level and the low-density lipoprotein level were increased at disease onset and did not fully return to the original values even at the 6 month followup, both in moderate and severe COVID-19 patients (Figure 3e and Figure S3). Notably, this phenotype was not found in patients with influenza A or H7N9 infection. According to the total cholesterol level of 600 COVID-19 patients, 1177 influenza A-infected patients and 522 H7N9-infected patients, the serum total cholesterol showed a tendency of increase over the disease onset only for COVID-19 patients (Figure 3f-h). It is worth noting that the serum total cholesterol level did not exceed the generally recognized normal range. Our study suggests that control of serum cholesterol needs to be taken into account when considering treatments to improve the prognosis.

We also observed long-term disturbances in the levels of proteins related to myocardium function. Most of the related proteins were downregulated in the COVID-19 disease stage and the downregulation persisted in the recovery stage both in moderate (Figure 4a) and severe COVID-19 patients (Figure 4b). Myocardial infarction and pressure overload induced a pathologic response of the myocardium termed cardiac remodeling. The cardiac myocyte growth (cardiac hypertrophy) and the deposition of the extracellular matrix (cardiac fibrosis) are hallmarks. Cardiac remodeling involves intense intercellular communication, wherein secreted factors such as proteases, protease inhibitors, cytokines, and chemokines serve a central role. Among identified proteins are TTN, MYH7, and NEXN, which function in the sarcomere, and COL14A1, COL2A1, ITGA7, MMP, THBS4, LGALS3, and LOX, which function in cell migration and cell-to-matrix adhesions (Figure 4c,d). It is worth mentioning that the levels of myocardial enzymes detected by clinical blood tests (Figure 1a) did not show a clear indication of cardiomyopathy in moderate or severe cases (Figure S4). Thus, the serum protein markers identified in this study may provide an early warning of cardiomyopathy before the severe heart function impairment detected by a regular blood test.

# Further Impairment of Cholesterol Metabolism and Myocardium Function in Severe COVID-19 Patients

The degree of severity of COVID-19 may be an important factor in poor prognosis. Thus, we aimed to focus on the specific features of severe patients by comparing them to moderate ones. We did quantitative proteomic analysis of the cohort containing 10 healthy controls, 10 moderate COVID-19 patients, and 6 severe patients (Figure 1b). Proteins that are significantly up- or downregulated when comparing severe to moderate patients fall into four specific subclusters (Figure 5a). Specifically, proteins with functions related to blood coagulation, cholesterol metabolism, and cardiovascular system development, which showed prolonged disruption in the previous analysis, were also enriched in subclusters 3 and 4 (Figure 5b,c and Figure S5). The levels of the identified proteins all showed significant changes to various degrees in severe patients compared to both moderate patients and healthy controls (Figure 5d). Among these proteins are NEXN, ITGA7, HSP90AB1, APOA5, CES1, and CFL1, which we identified above to be significantly changed in severe patients both in the disease and recovery stages. This result provides a good hint that severe patients may have more serious sequelae due to prolonged disturbance of the highlighted pathways.

The significance of some of the differentially expressed proteins identified in this study was independently validated by parallel reaction monitoring (PRM), a targeted quantitative MS method well suited for determining protein amounts in complex mixtures<sup>14,15</sup> (Figure S6). The validation results indicate that the intensities of APOA2, FGA, SOD2, APOL1, and CFL1 are significantly changed, consistent with the trends shown in Figures 3-5.

#### DISCUSSION

In this study, we applied mass spectrometry (MS)-based proteomics, which quickly delivered substantial amounts of

biological information from serum samples. Our approach provided a molecular atlas of serum proteins in an untargeted fashion, which does not depend on prior knowledge of the disease. This unbiased proteomic analysis, closely integrated with clinical symptoms, allows us to more accurately dissect the data. We obtained evidence for prolonged disturbances of pathways related to cholesterol metabolism and myocardium function, which provide insights into the mechanisms underlying the sequelae of COVID-19. This discovery enables us to identify clinically applicable markers and therapeutic targets and to design preventive strategies.

We found that cholesterol metabolism pathways were significantly affected in COVID-19 patients, both in moderate and severe cases, and were not restored during the recovery stage. Cholesterol is an abundant lipid in the membranes of eukaryotic cells. Cholesterol synthesis, uptake, or transport is fundamental for infection by viruses such as the flavivirus HCV or the retrovirus HIV to facilitate viral entry, formation of replicative complexes, and assembly or egress of new virus particles.<sup>16–18</sup> There are also reports of the role of cholesterol in the stability and infectivity of influenza A.<sup>19,20</sup> Likewise, it is possible that cholesterol supports viral infection through lipid subdomains on the plasma membrane that can harbor angiotensin-converting enzyme 2 (ACE2), which acts as the receptor for the S-protein of SARS-CoV-2.<sup>21,22</sup> Cellular cholesterol synthesis or transport may also accompany the other stage of the life cycle of SARS-CoV-2.<sup>23</sup> Pharmacological inhibition of cholesterol homeostasis reduces the replication of SARS-CoV-2.<sup>22</sup> Our proteomic data as well as the clinical data revealed that SARS-CoV-2 infection causes long-term disturbances, lasting for at least 6 months, of cholesterol metabolism. Thus, more attention should be given to monitoring cholesterol metabolism in COVID-19 patients, and cholesterol-lowering therapies might be useful for good prognosis.

We found that proteins enriched in regulation of cardiac muscle hypertrophy, cardiac muscle tissue development, cardiocyte differentiation, and cardiovascular system development were significantly changed in COVID-19 disease and were not restored during recovery. Our finding is in accordance with multiple reports of major cardiac complications in COVID-19 patients and of COVID-19-related myocarditis. Moreover, a recent study observed that the majority of recovered patients presented impaired cardiac function, which indicates that COVID-19 results in long-term heart sequelae.<sup>24</sup> We speculate that there are potentially various degrees of COVID-19-associated cardiac function impairments due to single factors or combinations of multiple factors, including direct viral myocardial damage, hypoxia, hypotension, cholesterol metabolism disorder, or enhanced inflammatory status, and they last for a long time.<sup>25,26</sup> Some impairments are too preliminary to be detected from clinical indicators in patients' serum but could be identified by proteomic analysis, which is a much more sensitive approach. However, so far, cardiac dysfunction only comes to attention when there are organ-level impairments, which are relatively late for treatment. We identified potential protein signatures, the level of which changed in advance of regular serum cardiomyopathy indicators, and thus could be used as a warning of the long-term disorder. Among the candidate protein markers are MYH7 and TTN, the transcriptional level of which is significantly changed in SARS-CoV-2-infected iPSC-derived cardiac cells.<sup>27</sup> This may provide mechanistic

support of our findings. The potential markers identified in this study need further validation in large cohorts and could further be used as clinical indicators to prevent cardiomyopathy in COVID-19 patients, thus improving prognosis.

In addition, when severe patients were compared to moderate patients, the cholesterol metabolism and cardiomyopathy pathways were more severely affected. Thus, we speculate that there is a correlation between these pathways and disease severity. This prediction is supported by a systematic review of 54 published articles that observed that cholesterol levels are apparently related to COVID-19 severity.<sup>28</sup> Furthermore, it has been reported that elevated levels of high-sensitivity troponin-I or natriuretic peptides, which are biomarkers of cardiac damage or dysfunction, are the strongest predictor of mortality in hospitalized patients.<sup>29-32</sup> In our study, pathways related to cholesterol metabolism and myocardium function show prolonged disturbance and are also related to disease severity. Thus, we reasoned that monitoring and prevention of these disorders are required to improve the prognosis of severely affected COVID-19 patients.

Serum protein analysis has been published by multiple groups including unbiased proteomic studies on COVID-19 patient cohort, proteomic analysis of a certain class of proteins enriched from serum and targeted group of proteins by an antibody array or aptamer. They also reported significantly altered pathways as host-cell immune response, coagulation, and cholesterol metabolism pathways. However, the individual proteins involved in these pathways are not exactly the same. We compared all relevant proteins quantified in other studies with data in our paper and found that most of the published differentially expressed proteins show a similar trend in our studies, for example, SAA1, SAA2, CRP, C1R, SERPINA3, APOC1, APOC2, APOC3, APOA4, APOE, GRN, ORM1, and SAP.<sup>4–6,13,33,34</sup> However, there are other proteins that are not significantly changed or are not detected in our study. Notably, FETUB and ApoA2 showed an increased level in COVID-19 compared to healthy controls, while other studies showed a decreased level in severe COVID-19 patients. This is possibly due to the regional difference and the different races of the patient cohort, as well as the exact time point from disease onset. The different detection method might be another explanation.

Limitations in this study lie in that changes in the serum level of proteins could not be simply corresponded to the tissue level of the proteins and the exact mechanisms of these proteins. The serum proteomic data only provided information on enriched pathways composed of differentially expressed proteins in serum based on the available database. Furthermore, the quantitation of these potential markers in tissue will provide more hints on their functions. To go another step further, more rigorous studies tracing the origin of these proteins, how these proteins are released, and how these proteins functions in disease conditions by cells or a mouse model are required in the future mechanism studies.

#### CONCLUSIONS

In conclusion, our study closely integrates quantitative proteomics and regular clinical serum tests in a well-established disease cohort throughout the disease and recovery stages. We highlight long-term disturbances in pathways related to cholesterol metabolism and cardiopathy, which could not be diagnosed based on regular clinical indicators. We also provide potential protein markers for these impairments. Our findings may provide evidence and indications to guide clinical treatment to improve prognosis.

#### ASSOCIATED CONTENT

#### **Supporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.1c00054.

Figure S1: relative intensity of proteins indicating viral infection, box-and-whisker plots showing the relative intensities of the indicated proteins when comparing healthy donors, COVID-19 patients, and the same patients recovered from COVID-19; Figure S2: longterm disturbance of blood coagulation in COVID-19 patients in the disease and recovery stages; Figure S3: long-term monitoring of the serum level of low density lipoprotein-C in COVID-19 patients throughout the disease and recovery stages; Figure S4: long-term monitoring of serum myocardial enzyme profiles in COVID-19 patients throughout the disease and recovery stages; Figure S5: disturbance of blood coagulation pathways in severe COVID-19 patients compared to moderate COVID-19 patients; Figure S6: PRM validation results for differentially expressed proteins in COVID-19 patients; Table S1: demographic information of the patients enrolled in the study; Table S2: clinical characteristics of the COVID-19 patients (PDF)

#### AUTHOR INFORMATION

#### **Corresponding Authors**

- Catherine C. L. Wong Center for Precision Medicine Multi-Omics Research, Peking University Health Science Center, Peking University, Beijing 100191, China; School of Basic Medical Sciences, Peking University Health Science Center, Beijing 100191, China; Peking University First Hospital, Beijing 100034, China; Peking-Tsinghua Center for Life Sciences, Beijing 100871, China; Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing 100069, China; orcid.org/0000-0001-7270-980X; Email: catherine wong@bjmu.edu.cn
- Min Zheng State Key Laboratory for Diagnosis and Treatment of Infectious Disease, National Clinical Research Center for Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China; Email: minzheng@zju.edu.cn

#### **Authors**

- Yang Chen Center for Precision Medicine Multi-Omics Research, Peking University Health Science Center, Peking University, Beijing 100191, China; School of Basic Medical Sciences, Peking University Health Science Center, Beijing 100191, China
- Hangping Yao State Key Laboratory for Diagnosis and Treatment of Infectious Disease, National Clinical Research Center for Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China; ⊚ orcid.org/0000-0001-6742-7074
- Nan Zhang Center for Precision Medicine Multi-Omics Research, Peking University Health Science Center, Peking University, Beijing 100191, China; School of Basic Medical Sciences, Peking University Health Science Center, Beijing 100191, China

- Jie Wu State Key Laboratory for Diagnosis and Treatment of Infectious Disease, National Clinical Research Center for Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China
- Shuaixin Gao Center for Precision Medicine Multi-Omics Research, Peking University Health Science Center, Peking University, Beijing 100191, China
- Jiangtao Guo Center for Precision Medicine Multi-Omics Research, Peking University Health Science Center, Peking University, Beijing 100191, China
- Xiangyun Lu State Key Laboratory for Diagnosis and Treatment of Infectious Disease, National Clinical Research Center for Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China
- Linfang Cheng State Key Laboratory for Diagnosis and Treatment of Infectious Disease, National Clinical Research Center for Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China
- Rui Luo State Key Laboratory for Diagnosis and Treatment of Infectious Disease, National Clinical Research Center for Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China
- Xue Liang State Key Laboratory for Diagnosis and Treatment of Infectious Disease, National Clinical Research Center for Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.1c00054

#### **Author Contributions**

 $^{\nabla}$ Y.C., H.Y., N.Z., and J.W. contributed equally to this work. Author Contributions

C.C.W. and M.Z. conceived the project. H.Y., J.W., X.L., L.C., R.L., and X.L. collected clinical samples and clinical test data. C.C.W. supervised the mass spectrometry proteomic experiments. N.Z. and S.G. performed the proteomic experiments. Y.C., C.C.W., M.Z., and H.Y. analyzed the data and wrote the manuscript.

#### Notes

The authors declare no competing financial interest.

#### ACKNOWLEDGMENTS

The authors thank Bruker Daltonics Inc. for their support in proteomic analysis. The authors are grateful to OmicsBean (Gene For Health Inc.) for their assistance in data analysis. The authors thank Spectronaut (Biognosys Inc.) for their support in database searching. This work is supported by Big Science Strategy Plan (2020YFE0202200), the National Natural Science Foundation of China (91754108) and the Ministry of Science and Technology of the People's Republic of China (2018YFA0507102), the Major Project of Zhejiang Provincial Science and Technology Department (2021C03043 and 2021C03039), the National Science and Technology Major Project for the Control and Prevention of Major Infectious Diseases in China (2018ZX10711001, 2018ZX10102001, and 2018ZX10302206), the Fundamental Research Funds for the Central Universities (BMU2017YJ003 and BMU2018XTZ002), and Research Funds from the

Health@InnoHK Program launched by Innovation Technology Commission of the Hong Kong Special Administrative Region and the PKU-Baidu Fund (2019BD007).

#### REFERENCES

(1) De Felice, F. G.; Tovar-Moll, F.; Moll, J.; Munoz, D. P.; Ferreira, S. T. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) and the Central Nervous System. *Trends Neurosci.* **2020**, *43*, 355–357.

(2) Varga, Z.; Flammer, A. J.; Steiger, P.; Haberecker, M.; Andermatt, R.; Zinkernagel, A. S.; Mehra, M. R.; Schuepbach, R. A.; Ruschitzka, F.; Moch, H. Endothelial cell infection and endotheliitis in COVID-19. *Lancet* **2020**, 395, 1417–1418.

(3) Zhang, C.; Shi, L.; Wang, F. S. Liver injury in COVID-19: management and challenges. *Lancet Gastroenterol. Hepatol.* **2020**, *5*, 428–430.

(4) Shu, T.; Ning, W.; Wu, D.; Xu, J.; Han, Q.; Huang, M.; Zou, X.; Yang, Q.; Yuan, Y.; Bie, Y.; Pan, S.; Mu, J.; Han, Y.; Yang, X.; Zhou, H.; Li, R.; Ren, Y.; Chen, X.; Yao, S.; Qiu, Y.; Zhang, D. Y.; Xue, Y.; Shang, Y.; Zhou, X. Plasma Proteomics Identify Biomarkers and Pathogenesis of COVID-19. *Immunity* **2020**, 1108.

(5) Hou, X.; Zhang, X.; Wu, X.; Lu, M.; Wang, D.; Xu, M.; Wang, H.; Liang, T.; Dai, J.; Duan, H.; Xu, Y.; Yu, X.; Li, Y. Serum Protein Profiling Reveals a Landscape of Inflammation and Immune Signaling in Early-stage COVID-19 Infection. *Mol. Cell. Proteomics* **2020**, *19*, 1749–1759.

(6) Shen, B.; Yi, X.; Sun, Y.; Bi, X.; Du, J.; Zhang, C.; Quan, S.; Zhang, F.; Sun, R.; Qian, L.; Ge, W.; Liu, W.; Liang, S.; Chen, H.; Zhang, Y.; Li, J.; Xu, J.; He, Z.; Chen, B.; Wang, J.; Yan, H.; Zheng, Y.; Wang, D.; Zhu, J.; Kong, Z.; Kang, Z.; Liang, X.; Ding, X.; Ruan, G.; Xiang, N.; Cai, X.; Gao, H.; Li, L.; Li, S.; Xiao, Q.; Lu, T.; Zhu, Y.; Liu, H.; Chen, H.; Guo, T. Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell* **2020**, *182*, 59–72.e15 e15.

(7) Tian, W.; Zhang, N.; Jin, R.; Feng, Y.; Wang, S.; Gao, S.; Gao, R.; Wu, G.; Tian, D.; Tan, W.; Chen, Y.; Gao, G. F.; Wong, C. C. L. Immune suppression in the early stage of COVID-19 disease. *Nat. Commun.* **2020**, *11*, 5859.

(8) Carfi, A.; Bernabei, R.; Landi, F.; Gemelli Against, C.-P.-A. C. S. G. Persistent Symptoms in Patients After Acute COVID-19. *JAMA* **2020**, 324, 603–605.

(9) Kushner, I. The phenomenon of the acute phase response. *Ann. N. Y. Acad. Sci.* **1982**, 389, 39–48.

(10) Liu, L.; Zhu, J.; Yang, J.; Li, X.; Yuan, J.; Wu, J.; Liu, Z. GP73 facilitates hepatitis B virus replication by repressing the NF-  $\kappa$ B signaling pathway. *J. Med. Virol.* **2020**, 3327.

(11) Sack, G. H., Jr. Serum Amyloid A (SAA) Proteins. Subcell Biochem. 2020, 94, 421–436.

(12) Giannis, D.; Ziogas, I. A.; Gianni, P. Coagulation disorders in coronavirus infected patients: COVID-19, SARS-CoV-1, MERS-CoV and lessons from the past. *J. Clin. Virol.* **2020**, *127*, 104362.

(13) Begue, F.; Tanaka, S.; Mouktadi, Z.; Rondeau, P.; Veeren, B.; Diotel, N.; Tran-Dinh, A.; Robert, T.; Velia, E.; Mavingui, P.; Lagrange-Xelot, M.; Montravers, P.; Couret, D.; Meilhac, O. Altered high-density lipoprotein composition and functions during severe COVID-19. *Sci. Rep.* **2021**, *11*, 2291.

(14) Hoofnagle, A. N.; Wener, M. H. The fundamental flaws of immunoassays and potential solutions using tandem mass spectrometry. *J. Immunol. Methods* **2009**, 347, 3–11.

(15) Ronsein, G. E.; Pamir, N.; von Haller, P. D.; Kim, D. S.; Oda, M. N.; Jarvik, G. P.; Vaisar, T.; Heinecke, J. W. Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics. *J. Proteomics* **2015**, *113*, 388–399.

(16) Osuna-Ramos, J. F.; Reyes-Ruiz, J. M.; Del Angel, R. M. The Role of Host Cholesterol During Flavivirus Infection. *Front. Cell Infect. Microbiol.* **2018**, *8*, 388.

(17) Gonzalez-Aldaco, K.; Torres-Reyes, L. A.; Ojeda-Granados, C.; Jose-Abrego, A.; Fierro, N. A.; Roman, S. Immunometabolic Effect of Cholesterol in Hepatitis C Infection: Implications in Clinical Management and Antiviral Therapy. *Ann. Hepatol.* **2018**, *17*, 908–919.

pubs.acs.org/jpr

(18) Bukrinsky, M.; Sviridov, D. Human immunodeficiency virus infection and macrophage cholesterol metabolism. *J. Leukocyte Biol.* **2006**, *80*, 1044–1051.

(19) Bajimaya, S.; Frankl, T.; Hayashi, T.; Takimoto, T. Cholesterol is required for stability and infectivity of influenza A and respiratory syncytial viruses. *Virology* **2017**, *510*, 234–241.

(20) Bajimaya, S.; Hayashi, T.; Frankl, T.; Bryk, P.; Ward, B.; Takimoto, T. Cholesterol reducing agents inhibit assembly of type I parainfluenza viruses. *Virology* **2017**, *501*, 127–135.

(21) Lu, Y.; Liu, D. X.; Tam, J. P. Lipid rafts are involved in SARS-CoV entry into Vero E6 cells. *Biochem. Biophys. Res. Commun.* 2008, 369, 344–349.

(22) Wang, R.; Simoneau, C. R.; Kulsuptrakul, J.; Bouhaddou, M.; Travisano, K. A.; Hayashi, J. M.; Carlson-Stevermer, J.; Zengel, J. R.; Richards, C. M.; Fozouni, P.; Oki, J.; Rodriguez, L.; Joehnk, B.; Walcott, K.; Holden, K.; Sil, A.; Carette, J. E.; Krogan, N. J.; Ott, M.; Puschnik, A. S. Genetic Screens Identify Host Factors for SARS-CoV-2 and Common Cold Coronaviruses. *Cell* **2021**, 106.

(23) Radenkovic, D.; Chawla, S.; Pirro, M.; Sahebkar, A.; Banach, M. Cholesterol in Relation to COVID-19: Should We Care about It? *J. Clin. Med.* **2020**, *9*, 1909.

(24) Puntmann, V. O.; Carerj, M. L.; Wieters, I.; Fahim, M.; Arendt, C.; Hoffmann, J.; Shchendrygina, A.; Escher, F.; Vasa-Nicotera, M.; Zeiher, A. M.; Vehreschild, M.; Nagel, E. Outcomes of Cardiovascular Magnetic Resonance Imaging in Patients Recently Recovered From Coronavirus Disease 2019 (COVID-19). *JAMA Cardiol.* 2020, 1265. (25) Kochi, A. N.; Tagliari, A. P.; Forleo, G. B.; Fassini, G. M.; Tondo, C. Cardiac and arrhythmic complications in patients with COVID-19. *J. Cardiovasc. Electrophysiol.* 2020, 31, 1003–1008.

(26) Siripanthong, B.; Nazarian, S.; Muser, D.; Deo, R.; Santangeli, P.; Khanji, M. Y.; Cooper, L. T., Jr.; Chahal, C. A. A. Recognizing COVID-19-related myocarditis: The possible pathophysiology and proposed guideline for diagnosis and management. *Heart Rhythm* **2020**, *17*, 1463–1471.

(27) Perez-Bermejo, J. A.; Kang, S.; Rockwood, S. J.; Simoneau, C. R.; Joy, D. A.; Ramadoss, G. N.; Silva, A. C.; Flanigan, W. R.; Li, H.; Nakamura, K.; Whitman, J. D.; Ott, M.; Conklin, B. R.; McDevitt, T. C. SARS-CoV-2 infection of human iPSC-derived cardiac cells predicts novel cytopathic features in hearts of COVID-19 patients. *Sci. Transl. Med.* **2020**, DOI: 10.1101/2020.08.25.265561.

(28) Zaki, N.; Alashwal, H.; Ibrahim, S. Association of hypertension, diabetes, stroke, cancer, kidney disease, and high-cholesterol with COVID-19 disease severity and fatality: A systematic review. *Diabetes. Metab. Syndr.* **2020**, *14*, 1133–1142.

(29) Shao, M. J.; Shang, L. X.; Luo, J. Y.; Shi, J.; Zhao, Y.; Li, X. M.; Yang, Y. N. Myocardial injury is associated with higher mortality in patients with coronavirus disease 2019: a meta-analysis. *J. Geriatr. Cardiol.* **2020**, 17, 224–228.

(30) Zhou, F.; Yu, T.; Du, R.; Fan, G.; Liu, Y.; Liu, Z.; Xiang, J.; Wang, Y.; Song, B.; Gu, X.; Guan, L.; Wei, Y.; Li, H.; Wu, X.; Xu, J.; Tu, S.; Zhang, Y.; Chen, H.; Cao, B. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* **2020**, *395*, 1054–1062.

(31) Ruan, Q.; Yang, K.; Wang, W.; Jiang, L.; Song, J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. *Intensive Care Med.* **2020**, *46*, 846–848.

(32) Wang, D.; Hu, B.; Hu, C.; Zhu, F.; Liu, X.; Zhang, J.; Wang, B.; Xiang, H.; Cheng, Z.; Xiong, Y.; Zhao, Y.; Li, Y.; Wang, X.; Peng, Z. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA* **2020**, *323*, 1061–1069.

(33) Di, B.; Jia, H.; Luo, O. J.; Lin, F.; Li, K.; Zhang, Y.; Wang, H.; Liang, H.; Fan, J.; Yang, Z. Identification and validation of predictive factors for progression to severe COVID-19 pneumonia by proteomics. *Signal Transduction Targeted Ther.* **2020**, *5*, 217.

## Journal of Proteome Research

pubs.acs.org/jpr

Article

(34) Rieder, M.; Wirth, L.; Pollmeier, L.; Jeserich, M.; Goller, I.; Baldus, N.; Schmid, B.; Busch, H. J.; Hofmann, M.; Thimme, R.; Rieg, S.; Kern, W.; Bode, C.; Duerschmied, D.; Lother, A. Serum Protein Profiling Reveals a Specific Upregulation of the Immunomodulatory Protein Progranulin in Coronavirus Disease 2019. J Infect. Dis. 2021, 223, 775–784.