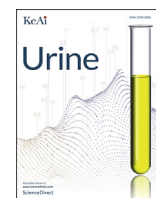




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Urine proteome of COVID-19 patients

Yanchang Li^{a,1}, Yihao Wang^{a,g,1}, Huiying Liu^{b,1}, Wei Sun^{a,1}, Baoqing Ding^c, Yinghua Zhao^a, Peiru Chen^a, Li Zhu^d, Zhaodi Li^a, Naikang Li^a, Lei Chang^a, Hengliang Wang^{d,**}, Changqing Bai^{b,***}, Ping Xu^{a,e,f,*}

^a State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences (Beijing), Research Unit of Proteomics & Research and Development of New Drug of Chinese Academy of Medical Sciences, Beijing Institute of Lifeomics, Beijing, 102206, China

^b The Fifth Medical Center of PLA General Hospital, Beijing, 100039, China

^c Department of Ecology and Evolutionary Biology, University of Connecticut, 75 North Eagleville Road, Unit 3043, Storrs, CT, 06269, USA

^d State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Biotechnology, Beijing, 100071, China

^e Key Laboratory of Combinatorial Biosynthesis and Drug Discovery of Ministry of Education, School of Pharmaceutical Sciences, Wuhan University, Wuhan, 430071, China

^f Guizhou University School of Medicine, Guiyang, 550025, China

^g Department of Pharmaceutical Sciences, Beijing Institute of Radiation Medicine, Beijing, 100850, China



ARTICLE INFO

Keywords:
COVID-19
Urine
Proteomics

ABSTRACT

The atypical pneumonia (COVID-19) caused by SARS-CoV-2 is a serious threat to global public health. However, early detection and effective prediction of patients with mild to severe symptoms remain challenging. The proteomic profiling of urine samples from healthy individuals, mild and severe COVID-19 positive patients with comorbidities can be clearly differentiated. Multiple pathways have been compromised after the COVID-19 infection, including the dysregulation of complement activation, platelet degranulation, lipoprotein metabolic process and response to hypoxia. This study demonstrates the COVID-19 pathophysiology related molecular alterations could be detected in the urine and the potential application in auxiliary diagnosis of COVID-19.

Introduction

Coronavirus Disease-2019 (COVID-19) is caused by a novel virus strain, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which is an unprecedented global health threat.^{1,2} By January 31, 2021, more than 100 million confirmed cases and 2.3 million fatalities, spreading in almost all countries and regions of the world, has been reported. Even worse, more than 500,000 new cases are being confirmed daily. However, no clinical drugs is available for highly infectious SARS-CoV-2, which further exacerbates the panic.

Tremendous efforts have been devoted to investigating the SARS-CoV-2, and its host response, epidemiological and clinical characteristics to mitigate the current pandemic.^{1,3–10} Statistics showed that the elderly have severe symptoms, especially those with comorbidities such as obesity, diabetes, heart and lung disease. The SARS-CoV-2 has been

reported to be harmful to lung, liver, heart, testis, bladder and kidney, where ACE2 are highly expressed.^{5,8,9,11–13} It has been estimated that about 80% of COVID-19 patients experiencing mild symptoms (M-COVID), recover with, or even without conventional medical treatment.¹⁰ However, the remaining 20% of patients with respiratory distress symptom may die rapidly without urgent and specialized intensive medical care, including immediate oxygen therapy, and mechanical ventilation.^{14,15} Disease stage significantly affects COVID-19 treatment and survivorship. The overall mortality rate for hospitalized patients varied from 2.3% for patients diagnosed at the early stage to 11% at the advanced stage.¹⁶ Unfortunately, the majority of cases are diagnosed at the advanced stage due to the lack of biomarkers and medical resources at the early stage. Therefore, it is critical to develop novel approaches to estimate the disease stages for patients in order to seek appropriate treatments and allocate scarce medical resources. In addition, novel

* Corresponding author. State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences (Beijing), Research Unit of Proteomics & Research and Development of New Drug of Chinese Academy of Medical Sciences, Beijing Institute of Lifeomics, Beijing, 102206, China.

** Corresponding author.

*** Corresponding author.

E-mail addresses: wanghl@bmi.ac.cn (H. Wang), baicq307@163.com (C. Bai), xuping@mail.ncpsb.org (P. Xu).

¹ These authors contributed equally.

<https://doi.org/10.1016/j.urine.2021.02.001>

Received 23 December 2020; Received in revised form 7 February 2021; Accepted 24 February 2021

Available online 5 March 2021

2590-2806/© 2021 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND

license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

detection methods that genuinely reflect the underlying changes of molecular and biological processes of COVID-19 patients would be favorable to understanding of SARS-CoV-2 pathogenesis.

Blood and urine are frequent biomaterials for discovery of biomarkers of human diseases because of their accessibility and non-invasiveness.^{17,18} The compositions of proteins detected in blood and urine samples can genuinely reflect the changes of the body health condition; thus, they are considered an important source for early warning and sensitive for disease detection.^{19,20} Recently, MS-based serum proteomics studies have been utilized to predict the severity of COVID-19 infection.^{18,21} Additionally, the urinary proteomics analysis showed the molecular changes of immunosuppression and tight junction impairment occurring in the early stage of COVID-19 infection.²² Therefore, a more detailed comprehensive profiling of the serum or urine proteome of COVID-19 patients will likely provide better diagnostics and clinical investigations of this disease.

In this study, we evaluated the diagnostic roles of urine samples on the progression of mild to severe type of COVID-19, and recovery state with cutting-edge urine proteomics.^{17,23} Six COVID-19 patients, comprised of 3 diagnosed as severe cases including one death and 3 mild patients, were investigated. To confirm the findings derived from the urine proteome, two recovery samples were further analyzed. We found that proteins related with complement activation and hypoxia were highly up-regulated, while proteins associated with platelet degranulation, and glucose and lipid metabolic process were especially down-regulated in the COVID-19 severe type patients. However, the changed proteins during the infectious phase recovered to normal in the recovery stage. We propose that urine proteome characterization can be potentially used to distinguish and predict the COVID-19 progression of

the mild to severe type. These urine proteome characteristics and changes may also shed light on the understanding of the COVID-19 pathogenesis.

Results

Characterization of urine proteomes in controls and six COVID-19 patients

In total, we assayed 40 urine specimens that passed quality check (QC), including 32 healthy controls, 6 COVID-19 patients and 2 corresponding recovery persons (Fig. 1 and S1). All patients were tested positive for the presence of SARS-CoV-2 nucleic acid. They all developed either fever or cough. Severe patients showed typical symptoms of fatigue and dyspnea (Fig. 1A). All patients had comorbidities, including 4 patients with essential hypertension, 1 patient with both essential hypertension and diabetes, and 1 patient with multiple metastases of colon cancer (dead on March 3, 2020) (Fig. 1A). According to the Diagnosis standards,¹⁰ these six patients were categorized into two disease types: three patients were defined as severe type acute respiratory syndrome (S-COVID) and the other three were diagnosed as mild type (M-COVID).

The severe COVID-19 patients showed ground-glass opacity in the lungs on Computed Tomography (CT) scanning (Fig. 1B). After treatment, the lung shadow disappeared and gradually recovered (Fig. 1B). Because the patient 4 (P4) had multiple metastases of colon cancer, only X-ray test was obtained (Figure S2). Interleukin-6 (IL-6) is an indicator of inflammatory storms.²⁴ We found the level of IL-6 in mild patients was 4.73 ± 2.03 pg/mL (mean \pm standard deviation), while the expression level of IL-6 in severe patients was significantly higher than the normal standard (≤ 7.0 pg/mL) and drastically fluctuated during the infection,

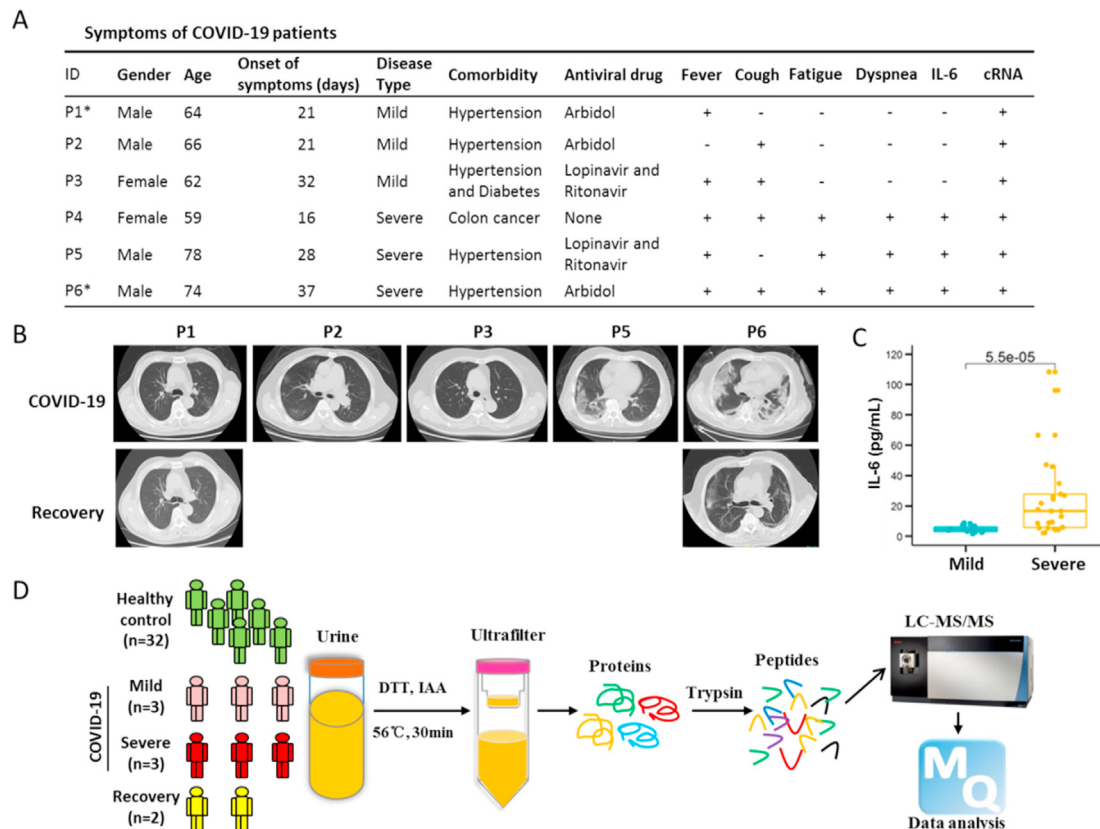


Fig. 1. Proteomics study on urine samples of COVID-19 patients. (A) Basic information and clinical symptoms of COVID-19 patients, including mild ($n = 3$) and severe ($n = 3$) patients. No.4 patient (P4) was with multiple metastases of colon cancer and died on March 3, 2020. No.1 (P1) and 6 (P6) patients labeled with asterisk indicated the persons providing the recovery urine samples. cRNA indicated that the SARS-CoV-2 nucleic acid. (B) Ground-glass opacity on Computed Tomography (CT) of COVID-19 patients. (C) The amounts of IL-6 between mild and severe COVID-19 patients. (D) Experimental design of urine proteomics for COVID-19 patients. Interleukin-6, IL-6.

indicating that the stress response to viral infection in S-COVID patients was more severe (Fig. 1C and S3).

The urine samples were collected after the diagnosis of the COVID-19. Four urine samples (H01–H04) of healthy controls were processed in parallel with the samples of COVID-19 patients (Fig. 1D), which were further compared with the other healthy sample datasets (H05–H32) generated in the laboratory following the same sample preparation processes and mass spectrometry analysis in order to detect sample heterogeneity. To confirm the proteome shift observed from the COVID-19 patients, we also collected urine samples from two recovered patients (P1 and P6) (Fig. 1A and S3).

As the sample size increased, the number of identified proteins in control group grew quickly, and gradually became saturated (Fig. 2A). The peptide over protein ratio was 6.0 (Table S1), indicating high quality and reliability of our protein identification. To improve the accuracy of COVID-19 and the corresponding recovery samples, 2 technical repeats were measured for each sample. A total number of 2656 proteins was identified from 32 healthy control samples (Fig. 2A and Table S1). We identified and quantified 1380 and 1641 proteins in urine samples from COVID-19 and two recovery persons in total, which were significantly lower than that of healthy controls (Fig. 2B and C, Tables S2 and S3).

There were 1008 proteins being commonly identified and quantified among the healthy controls, COVID-19 patients and recovered patients. However, 211 and 63 proteins were uniquely expressed in COVID-19 patients and recovery samples, respectively (Fig. 2D). The average abundance of the identified proteins for each group spanned about 6 orders of magnitude, with lower abundance for the COVID-19 samples compared with healthy and recovery ones (Fig. 2E). To check whether the SARS-CoV-2 proteins were present in the urine sample, we added

SARS-CoV-2 protein sequences to the human proteome database, and no related proteins were identified.

Urine proteomics differentiates COVID-19 patients from healthy people

To assess the quantitative variation and accuracy of the MS datasets, each urine sample of COVID-19 patients and the respective recovered samples were technically repeated twice. The absolute quantitative information iBAQ value was used for further comparison and analysis. The correlation coefficient (R^2) of the two replicates for each sample was higher than 0.80 (Figure S4), indicating the MS data was acquired with high degree of consistency and reproducibility in this study.

Due to the differences in sample size and operation during the sample processing, we found significant quantitative variations among different samples (Figure S5A). Therefore, the median values of iBAQ for each sample dataset were normalized equally to reduce the potential biases before quantitatively comparing the samples under COVID-19 with healthy conditions (Figure S5B). We found that the correlation of samples within the health and recovery groups were higher than that between the healthy and patient groups (Figure S6).

We found that patients and healthy people can be divided into two categories based on our cluster analysis (Fig. 3), indicating the distinctive molecular characteristics between healthy and COVID-19 conditions. Interestingly, the urine samples of two recovery patients were clustered with healthy people (Fig. 3). We also found that normal control individual H5 and H6 were “incorrectly” clustered with the M-COVID patients. Nevertheless, the M-COVID and S-COVID samples were clustered into separated groups except for one M-COVID patient (P3) with both hypertension and diabetes. The differences between M-COVID and S-

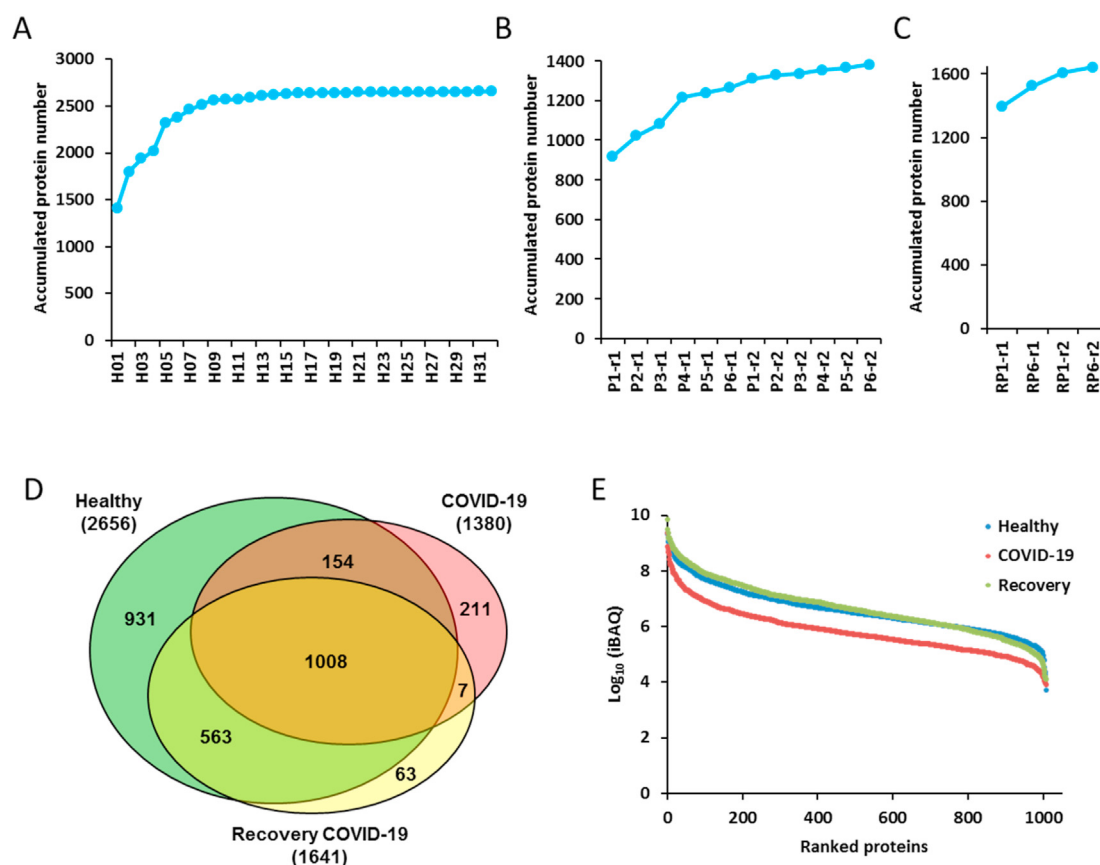


Fig. 2. Identification and quantification of urine samples from COVID-19 patients and healthy controls. (A&B&C) The accumulation curve of the quantified proteins from 32 healthy volunteers (A), 6 COVID-19 patients (B) and 2 recovery patients (C). (D) The Venn diagram for the identified urine proteins from the healthy volunteers, COVID-19 and recovery patients. (E) The dynamic range of the iBAQ abundance of identified proteins from healthy volunteers, COVID-19 patients and recovery ones. The average abundance for each group was calculated.

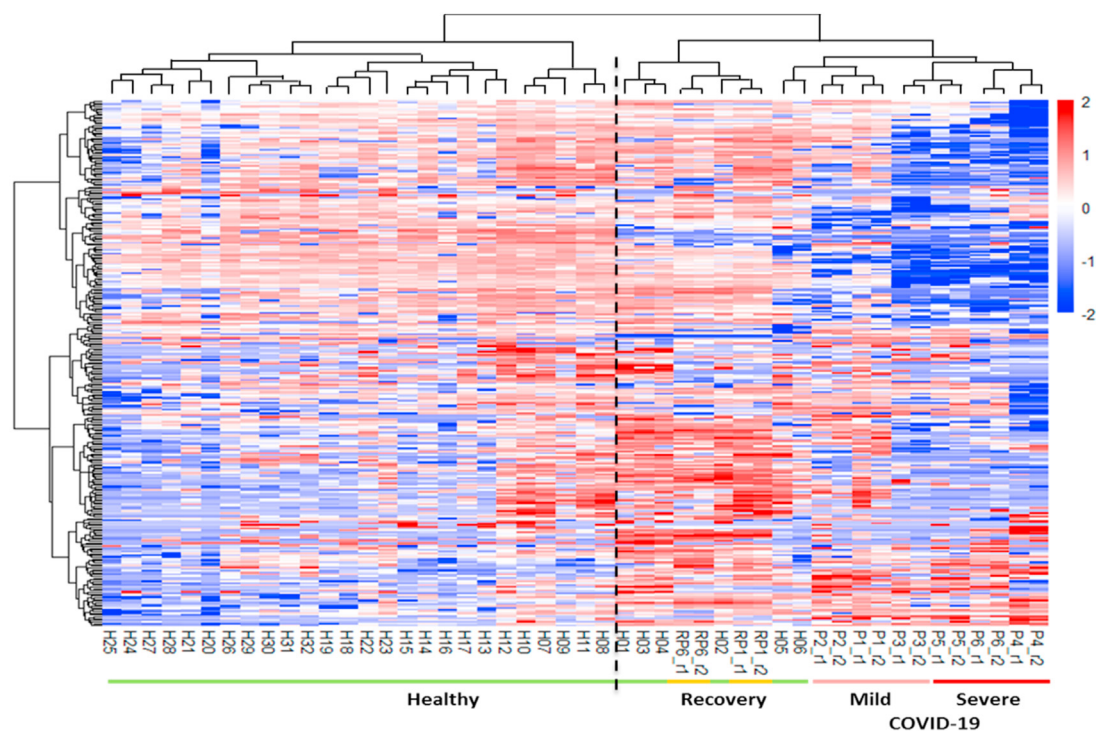


Fig. 3. Distinction of healthy volunteers, COVID-19 patients and recovery patients in proteomic features. The clustering heatmap analyses differentiates healthy volunteers from COVID-19 patients and recovery ones.

COVID urine proteome samples might reflect different physiological responses of the COVID-19 infection at the proteome level.

To corroborate the identified distinct clusters in our cluster analysis, we also performed principal component analysis (PCA). The result showed that patients and healthy people were clearly divided into two groups (Figure S7). The two samples including one mild (Recovery P1, RP1) and one severe recovery patient (Recovery P6, RP6) were grouped with healthy samples. Interestingly, P1 and P2 COVID-19 patients belonged to M-COVID patient group with only hypertension complication were more closed to healthy control. We found that these two patients could be distinguished from S-COVID patients or P3 mild patient with hypertension and diabetes complications. The P3 mild patient was incorrectly classified as severe (Figure S7), possibly because this female patient has diabetes complication (Fig. 1A). These results imply that the urine proteomics analysis can be served as a potential auxiliary prediction tool to differentiate M-COVID and S-COVID patients.

Molecular features of urine proteome for the pathogenesis of severe COVID-19 patients

Despite only 6 COVID-19 urine samples were tested, our results showed a clear distinction between healthy control and COVID-19 patient urine proteomes. We used the fold change ≥ 2 and p -value < 0.05 as filters to find the differentially expressed proteins between mild and severe diseases compared to the healthy control, respectively. There were 86 and 83 significantly up-regulated proteins, and 100 and 172 significantly down-regulated proteins in mild and severe COVID-19 samples, respectively (Fig. 4A and B, Tables 4 and 5). To eliminate the confounding effects of the individual characteristics such as age or comorbidities, we also compared recovery samples with healthy controls and identified 278 differentially expressed proteins, including 152 up-regulated and 126 down-regulated proteins (Fig. 4C and Table S6), which were excluded in the further analysis. Finally, we identified 95 unique changed proteins for severe type disease, 44 for mild type and 75 overlapped ones for both types of COVID-19 (Fig. 4D). GO analysis of these changed urinary molecular features implied that the COVID-19

could result in the dysregulation of immune response, viral process, response to hypoxia, complement activation and platelet degranulation (Fig. 4E).

To identify specific proteins to distinguish the mild from severe type of COVID-19 patients, we clustered the commonly identified proteins for all of these four datasets into 16 significant discrete clusters with the quantified values (Figure S8) through mFuzz.²⁵ We chose the cluster 2 and 11 as severe COVID-19 up-changed from mild COVID-19 (Fig. 5A). Combined the filter results and significantly changed proteins from Fig. 4D, we identified 56 unique proteins conforming to the criteria. These proteins were highly associated with the complement activation, regulation of immune response, cellular oxidant detoxification, cellular response to hypoxia and oxidative stress-induced apoptosis, which might reflect the pathogenesis of the severe COVID-19. These results are consistent with the recent reported sera proteomics.^{21,22} We also chose the cluster 1 and 12 as the down-regulated filter of the severe COVID-19 from mild COVID-19 as well (Fig. 5C). These filtered proteins were highly associated with the platelet degranulation, glucose metabolic process, protein metabolic process and lipid metabolic and transport pathways.

The molecular features used to distinguish the patient type (M and S) in our classifier (Fig. 5B and D, Tables S4-5) contain several potential biomarkers which were highly associated with the clinical characteristics of mild and severe COVID-19. For example, the hypoxia up-regulated protein 1 (HYOU1) belonging to cluster 2 was more than three-fold higher in the severe COVID-19 (Fig. 5B). HYOU1 plays a pivotal role in cyto-protective cellular mechanisms triggered by oxygen deprivation and is highly expressed in tissues such as liver and pancreas that contain well-developed endoplasmic reticulum and also regulates large amounts of secretory proteins.^{26,27} Patients with hypoxia warrant more attention to their intravascular coagulation, such as the elevated levels of D-dimer, a blood marker of excess clotting. It was reported that the heparin could boost patients' low oxygen levels regardless of whether they were struggling to breathe.²⁸ In this study, we found that the heparin cofactor 2 (SERPIND1) belonging to cluster 10 was specially up-regulated more than four-fold higher in the mild and two-fold higher in the severe COVID-19 (Table S4). SERPIND1, also known as heparin cofactor II, is a

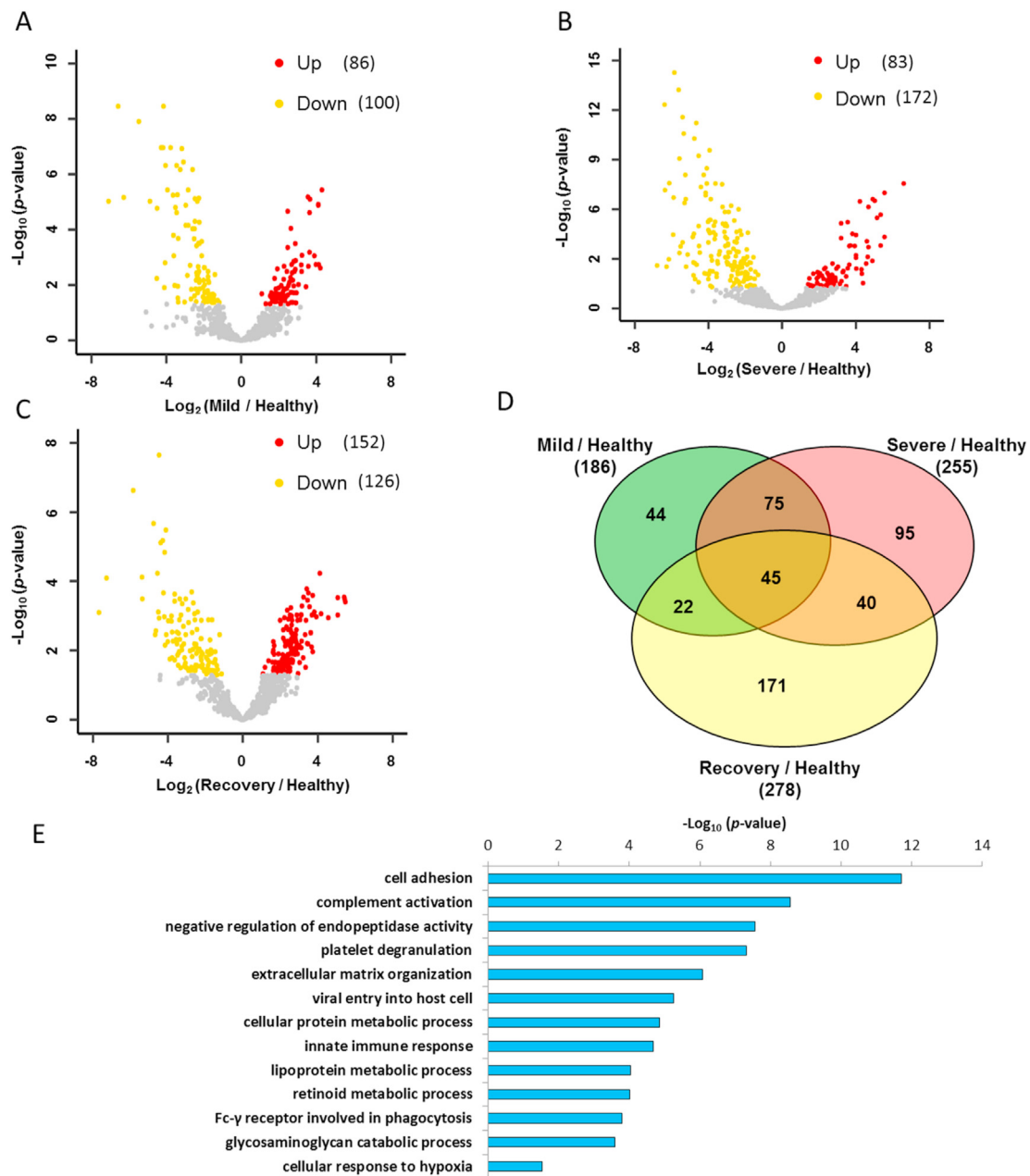


Fig. 4. Function distribution of dysregulated proteins in COVID-19 patients. (A&B&C) The volcano plots of the up-regulated and down-regulated proteins in different groups. Proteins with p-Value lower than 0.05 and fold change ≥ 2 were considered as significantly differential expression. (D) Venn diagrams of differential proteins in mild, severe COVID-19 patients and recovery patients compared with healthy volunteers. (E) The GO analysis of dysregulated proteins in the COVID-19 patients.

glycoprotein in human plasma that inhibits thrombin and chymotrypsin, and the rate of inhibition of thrombin is rapidly increased by Dermatan sulfate (DS), heparin (H) and glycosaminoglycans (GAG).^{29,30} We speculated that the SERPIND1 could be the protective response to reduce the risk of excess intravascular coagulation in the COVID-19 patients.

We also found that the cyclic AMP-responsive element-binding protein 3-like protein 3 (CREB3L3) belonging to cluster 10 (Figure S8) was specially up-regulated in the M-COVID (Table S4). In acute inflammatory response, CREB3L3 may activate expression of acute phase response (APR) genes, which was activated in response to cAMP stimulation.³¹ This might be the protective mechanism for body to fight against the virus.

For the down-regulated molecular clusters, the proteins related with platelet degranulation were also reported in the sera proteomics

recently.^{21,22} Additionally, the down-regulated pathways of lipid metabolic and transport in the COVID-19 patients caused our attention. The cholesterol homeostasis was reported to impact COVID-19 prognosis, virus entry and the antiviral therapies.³² In our data, the lipid metabolism and transporting, including the cholesterol homeostasis, were down-regulated in the S-COVID (Fig. 5D and Table S5). The proteins NPC intracellular cholesterol transporter 2 (NPC2), apolipoproteins A1 (APOA1), and Cubilin (CUBN) were changed with the similar trends (Fig. 5D). These results indicated that after the SARS-CoV-2 infection, the lipoprotein-mediated cholesterol uptake and transporting was disordered. Our study suggests that more characteristic molecular changes at protein levels can be used to build a predictive filter for the prospective identification of severe cases and shed light on the understanding of COVID-19 pathophysiology.

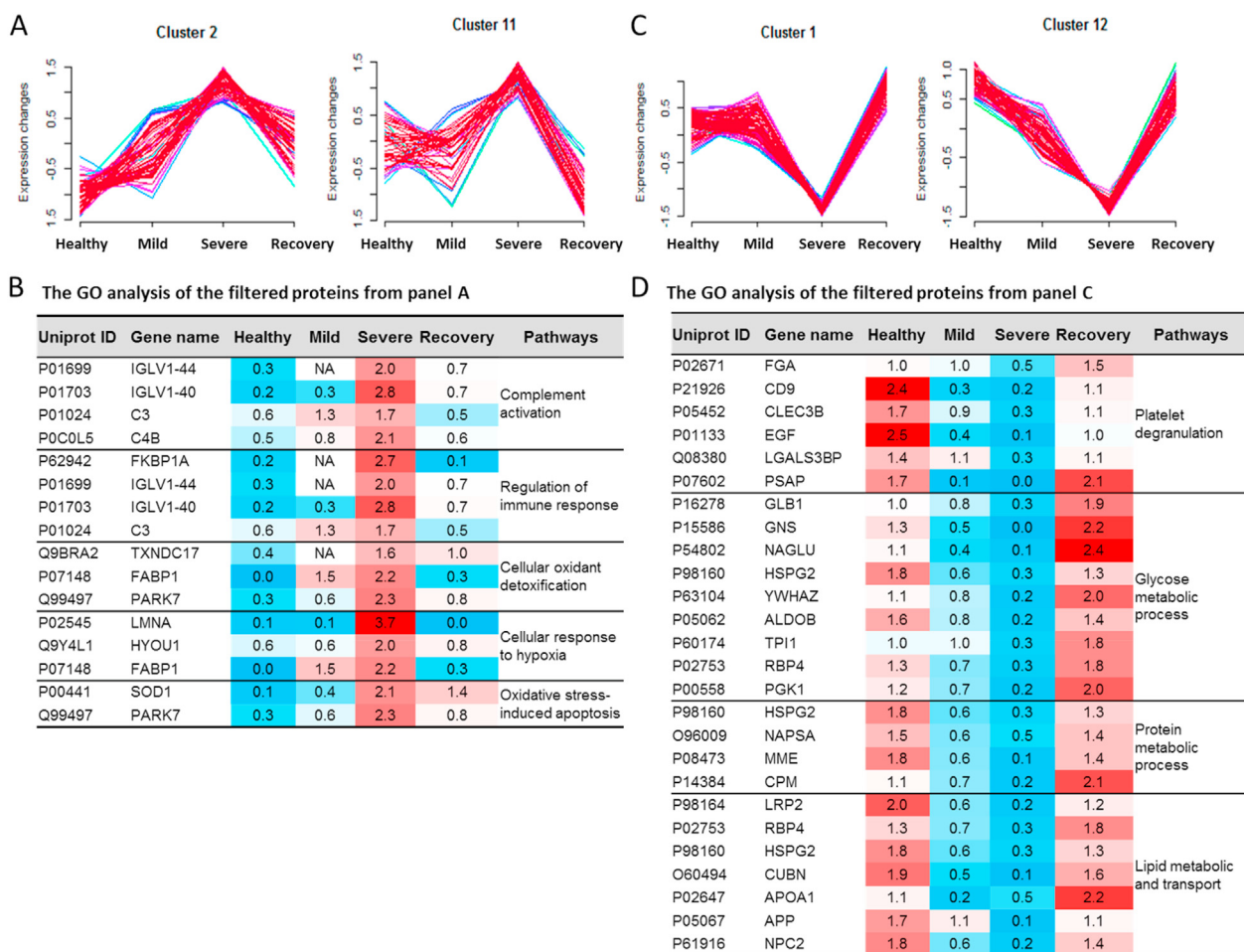


Fig. 5. Clustering of commonly identified proteins illustrated specific clusters of proteins in COVID-19 patients. The numbers 1–4 stands for the Health, Mild, Severe and Recovery, respectively. (A) The cluster 2 and 11 stands for the up-regulated trends uniquely in the severe type of COVID-19. (B) The GO analysis of the filtered proteins from panel A. (C) The cluster 1 and 12 stands for the down-regulated trends uniquely in the severe type of COVID-19. (D) The GO analysis of the filtered proteins from panel C.

Discussion

The COVID-19 pandemic caused by SARS-CoV-2 is not only putting huge pressure on global healthcare, but also having a devastating impact on the economy and society. Although much effort towards COVID-19 diagnostics and treatments has been made, the mortality of this infectious disease has not been significantly improved because of the limited mechanistic understanding of the pathogenesis.^{2,33} Patients progressing into the S-COVID often face very limited treatment options.^{8,9} Imaging technology, such as CT has been widely used to diagnose the COVID-19 patients, but suffers from high cost and demand for technical expertise. There is an urgent need for low-cost and reliable diagnostic techniques to estimate and predict the transition of severe COVID-19 patients from mild COVID-19 ones.

Urine is one of the most frequently studied biomaterials for biomarkers of human diseases in proteomics study because of its accessibility. It is less complex and has a relatively lower dynamic range with less technical challenges compared to blood.^{17,22,34–37} It is powerful to identify molecular groups to distinguish healthy controls, mild and serious COVID-19 patients through urinary proteomics. Then more robust and quick approaches could be developed for the targeted MS detection technology or multi-target microarray to improve the speed and throughput of sample detection. Our study demonstrated that urine profiling could separate the healthy control from COVID-19 patients and also tell recovery person from COVID-19. Specific proteome features for M-COVID and S-COVID patients were detected in the urine samples. This

is the first study to establish a link between the urine proteome and the understanding of the COVID-19 pathophysiology. Though the number of samples collected in this study was small, the obtained findings are consistent with the dysregulation of immune response, complement activation and platelet degranulation in previous blood and urine study,^{21,22} which supports the accuracy of our results. Our new findings of molecular dysregulation of lipoprotein metabolic process and response to hypoxia would have potential implications for clinical diagnosis and treatment. These significantly changed proteins and pathways were highly associated with the pathogenesis of the severe COVID-19. We concluded that urine proteome is an important source that warrants more attention for the understanding of COVID-19.

Given the possibility of using proteome as a diagnostic tool we have shown, we are also aware of the limitations in our COVID-19 urine proteomics study. First, only limited number of COVID-19 patient samples were included in our current proteomics study due to the lack of access to patient samples. More samples in future study will likely mitigate the possible sampling bias. Second, although the healthy controls and recovery samples were included, the coverage of population and representativeness was still limited when considering factors of genders and age.¹⁷ Most of test subjects¹⁷ were in the age group of more than 60, while younger patients were not included. Third, the current COVID-19 patients were only categorized as M-COVID and S-COVID. We could not obtain intermediate type of COVID-19 patient samples to improve the resolution to predict the trend and progress of atypical pneumonia caused by SARS-CoV-2 due to the limited number of patients. As a much

broader age spectrum of patient samples are included, the proteomic approach will be further validated and the signature proteins detected in specific patient groups can be further confirmed, which would allow fast developing and detecting the presence of specific antibodies in each group to predict the severity of COVID-19 patients. Forth, the batch effect of the sample processing and proteomics analysis may cause some deviations. Some healthy control datasets were generated before, though with the similar experiment processes. We recommended that the urine proteomics researches of the healthy, COVID-19 and their corresponding recovery samples were performed meanwhile if possible. Fifth, patients were subjected to different antiviral drug treatments, compounded with their age, preexisting health conditions as well as wide range of days of onset symptoms, it might cause some bias to the conclusion at this stage.

Altogether, our data demonstrate that a urine proteome-based proteomics study can reliably and sensitively differentiate COVID-19 patients from healthy people. It might be able to serve as a powerful tool to help scientists and clinicians fight the COVID-19 pandemic.

Materials and methods

Urine samples

All COVID-19 patients were diagnosed according to the Diagnosis and management plan of pneumonia with new coronavirus infection (Trial Version 6) in the Fifth Medical Center of Chinese PLA General Hospital, Beijing, China, between February 18 and March 3, 2020. According to the Diagnosis standards, those patients were classified as clinically severe type infection empirically based on a set of clinical characteristics, such as dyspnea, respiratory rate ($RR \geq 30$ times/min), mean oxygen saturation ($\leq 93\%$, resting state) or arterial blood oxygen partial pressure/oxygen concentration ($PaO_2/FiO_2 \leq 300$ mmHg), and/or lung infiltrates $> 50\%$ within 24–48 h. The patients classified as mild type infection were mainly manifested with the symptoms of fever, non-pneumonia or mild pneumonia cases. A total of 7 urine specimens from COVID-19 patients were collected. One of them was discarded because of severe renal failure. Patients with underlying diseases except renal dysfunction are indicated in Fig. 1A. The patients are aged from 59 to 78 years old (Fig. 1A). Among the six analyzed samples, two of them are female.

A total of 32 urine specimens from healthy controls (CTL, $n = 32$) were collected at Beijing Proteome Research Center, Beijing, China. The midstream of the morning urine was obtained for this study. Healthy controls are aged from 22 to 39 years old without any underlying disease. Among them, 11 are female (Figure S1).

All participants have provided signed informed consent and samples were collected with ethics approval from institutional review board (IRB) from the Fifth Medical Center of Chinese PLA General Hospital and Beijing Proteome Research Center. Our research strictly followed the standards indicated by the Declaration of Helsinki.

Proteomics sample preparation and LC-MS/MS analysis

Human urine proteomics samples were prepared as described previously with slight modification.^{23,34–37} Briefly, 1 mL urine samples were centrifuged at 2000 g for 4 min to remove cell debris before reduced with 5 mM dithiothreitol (DTT) at 56 °C for 30 min, which could also inactivate the virus. The treated samples were alkylated with 10 mM iodoacetamide in dark at room temperature for 30 min. The supernatant was loaded into a 10 kDa ultrafiltration tube and the larger molecular weight proteins (proteome) were separated from the endogenous peptides (peptidome) by centrifugation. Proteome samples were digested with trypsin at 37 °C for 14 h then the digestion reaction was terminated by 1% formic acid (FA). The digested peptides were desalted through a StageTip^{38,39} and dried before LC-MS/MS analysis.

The dried peptides were dissolved with 20 μ L loading buffer (1% formic acid, FA; 1% acetonitrile, ACN). 6 μ L sample was taken for LC-MS/MS analysis on an Orbitrap Fusion Lumos coupled with EASY-nLC 1200

(Thermo Fisher Scientific, Waltham, MA, USA).

The samples were loaded onto a self-packed trap column (2 cm \times 150 μ m) and then separated by a capillary column (15 cm \times 150 μ m), both packed with C18 reverse phase particle (1.9 μ m, Phenomenex, Torrance, California, USA). The peptides were eluted with a 120 min nonlinear gradient: 6% B for 10 min, 9–14% B for 15 min, 14–30% B for 50 min, 30–40% B for 30 min, 40–95% B for 3 min, 95% B for 7 min, 95–6% B for 1 min, 6% B for 4 min (Buffer A, 0.1% FA in ddH₂O; Buffer B, 0.1% FA and 80% ACN in ddH₂O; flow rate, ~ 600 nL/min).

The parameters for MS detecting were as follows: The full MS survey scans were performed in the ultra-high-field Orbitrap analyzer at a resolution of 120,000 and trap size of 500,000 ions over a mass range from 300 to 1400 m/z. MS/MS scan were detected in IonTrap and the 20 most intense peptide ions with charge states 2 to 7 were subjected to fragmentation via higher energy collision-induced dissociation (1×10^4 AGC target, 35 ms maximum ion time).

Data processing and label-free quantification

The raw files were searched with MaxQuant software (v1.5.3.0) against the database composed of Human fasta downloaded from Swiss-Prot (version released in 2020.02) and the SARS-CoV-2 virus fasta downloaded from NCBI (RefSeq GCF_009858895.2). Mass tolerance of the first search for precursor ions was set to 20 ppm. Full cleavage by trypsin was set and a maximum of two missed cleavages was allowed. The protein identification must met the following criteria: (1) the peptide length ≥ 7 amino acids; (2) the FDR $\leq 1\%$ at the PSM, peptide and protein levels.

The peptides were quantified by the peak area derived from their MS1 intensity with MaxQuant software.⁴⁰ The intensity of unique and razor peptides was used to calculate the protein intensity. The intensity based absolute quantification (iBAQ) algorithm was used as protein quantification value.⁴¹ In order to exclude the influence of differences in sample sizes and loading amounts for MS analysis, we used median value of each sample to normalize protein iBAQ values.⁴² All missing values were substituted with the minimal value.

Statistical analyses

Overlapped 1008 proteins were used for the subsequent statistical analysis. Pearson correlation analysis of all datasets was realized by Perseus.⁴³ Differential proteins were filtered using R package limma (version 3.34.9). The significantly differentially expressed proteins were selected using the criteria of adjusted p value less than 0.05 and \log_2 FC larger than 1. Proteins were clustered using R package mFuzz (version 2.46.0) into 16 significant discrete clusters.

Pathway analyses

The function of differential proteins was analyzed in David Bioinformatics Resources (<https://david.ncicrf.gov/>) and Human Protein Atlas (<http://www.proteinatlas.org/>) platforms including tissue-specific enrichment, molecular function, biological process, cellular component, etc.

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the iProX partner repository.⁴⁴ The accession numbers for the mass spectrometry proteomics data reported in this paper are the iProX (<https://www.iprox.org/>) dataset identifier: IPX0002166000. All the data will be publicly released upon publication.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgements

This study was funded by the MOST (2017YFC0906600, 2020YFE0202200 & 2017YFA0505002), the National Natural Science Foundation of China (31700723, 31670834, 31870824, 91839302 & 31901037), the Innovation Foundation of Medicine (16CXZ027, 20SWAQX34, BWS17J032, AWS17J008 & 19SWAQ17), National Megaprojects for Key Infectious Diseases (2018ZX10302302 & 2018ZX10732202-003), the Foundation of State Key Lab of Proteomics (SKLP-K201704 & SKLP-K201901), and the grant for Research Unit of Proteomics & Research and the Beijing-Tianjin-Hebei Basic Research Cooperation Project (J200001), and CAMS Innovation Fund for Medical Sciences (2019-I2M-5-017).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.urine.2021.02.001>.

Author contributions

P.X., C.B., and H.W. designed the experiment and supervised this project. Y.L., Y.W. and W.S. performed the proteomic experiments and data analysis with the help of Y.Z., P.C., and L.C.. H.L. and C.B. collected the samples and clinical data with the help of L.Z., Z.L. and N.L.. Data were interpreted and presented by all co-authors. P.X., Y.L., B.D., and Y.W. wrote the manuscript with inputs from co-authors.

References

- Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579(7798):270–273.
- Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med*. 2020;382(8):727–733.
- Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020;579(7798):265–269.
- Kim D, Lee JY, Yang JS, et al. The architecture of SARS-CoV-2 transcriptome. *Bioinformatics*. 2020;29:15–21.
- Bojkova D, Klann K, Koch B, et al. SARS-CoV-2 Infected Host Cell Proteomics Reveal Potential Therapy Targets. Preprint available at Research Square; 2020.
- Davidson AD, Williamson MK, Lewis S, et al. Characterisation of the Transcriptome and Proteome of SARS-CoV-2 Using Direct RNA Sequencing and Tandem Mass Spectrometry Reveals Evidence for a Cell Passage Induced In-Frame Deletion in the Spike Glycoprotein that Removes the Furin-like Cleavage Site. bioRxiv; 2020.
- Ghinai I, McPherson TD, Hunter JC, et al. First known person-to-person transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in the USA. *Lancet*. 2020;395:1137–1144, 10230.
- Guan WJ, Ni ZY, Hu Y, et al. *Clinical Characteristics of Coronavirus Disease 2019 in China*. 18th 382. *N Engl J Med*; 2020:1708–1720.
- Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med*. 2020;382(12):1199–1207.
- Tncpere T. *The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19)*. China CDC Weekly; 2020.
- Zhang C, Shi L, Wang F. Liver injury in COVID-19: management and challenges. *Lancet Gastroenterol. Hepatol*. 2020;5(5):428–430.
- Hanley B, Lucas SB, Youd E, et al. Autopsy in suspected COVID-19 cases. *J Clin Pathol*. 2020;73(5):239–242.
- Nie X, Qian L, Sun R, et al. Multi-organ Proteomic Landscape of COVID-19 Autopsies. *Cell*; 2021:775–794.
- Murthy S, Gomersall CD, Fowler RA. Care for critically ill patients with COVID-19. *J Am Med Assoc*. 2020;323(15):1499–1500.
- Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese center for disease control and prevention. *J Am Med Assoc*. 2020;323(13):1239–1242.
- Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020;395:507–513, 10223.
- Leng W, Ni X, Sun C, et al. Proof-of-Concept workflow for establishing reference intervals of human urine proteome for monitoring physiological and pathological changes. *EBioMedicine*. 2017;18:300–310.
- Su Y, Chen D, Yuan D, et al. Multi-Omics resolves a sharp disease-state shift between mild and moderate COVID-19. *Cell*. 2020;183(6):1479–1495 e20.
- Kentsis A, Shulman A, Ahmed S, et al. Urine proteomics for discovery of improved diagnostic markers of Kawasaki disease. *EMBO Mol Med*. 2013;5(2):210–220.
- Zhao M, Li M, Yang Y, et al. A comprehensive analysis and annotation of human normal urinary proteome. *Sci Rep*. 2017;7(1):3024.
- Shen B, Yi X, Sun Y, et al. *Proteomic and Metabolomic Characterization of COVID-19 Patient Sera*. medRxiv; 2020.
- Tian W, Zhang N, Jin R, et al. Immune suppression in the early stage of COVID-19 disease. *Nat Commun*. 2020;11(1):5859.
- Zhang C, Leng W, Sun C, et al. Urine proteome profiling predicts lung cancer from control cases and other tumors. *EBioMedicine*. 2018;30:120–128.
- Del Giudice M, Gangestad SW. Rethinking IL-6 and CRP: why they are more than inflammatory biomarkers, and why it matters. *Brain Behav Immun*. 2018;70:61–75.
- Kumar L, E Futschik M. A software package for soft clustering of microarray data. *Bioinformatics*. 2007;2(1):5–7.
- Cechowska-Pasko M, Bankowski E, Chene P. The effect of hypoxia on the expression of 150 kDa oxygen-regulated protein (ORP 150) in HeLa cells. *Cell Physiol Biochem*. 2006;17(1–2):89–96.
- Lewandrowski U, Moebius J, Walter U, et al. Elucidation of N-glycosylation sites on human platelet proteins: a glycoproteomic approach. *Mol Cell Proteomics*. 2006;5(2):226–233.
- Negri EM, Piloto BM, Morinaga LK, et al. *Heparin Therapy Improving Hypoxia in COVID-19 Patients - a Case Series*. medRxiv; 2020:2020, 04.15.20067017.
- Rau JC, Deans C, Hoffman MR, et al. Heparin cofactor II in atherosclerotic lesions from the pathological determinants of atherosclerosis in youth (PDAY) study. *Exp Mol Pathol*. 2009;87(3):178–183.
- Aihara K, Azuma H, Akaike M, et al. Heparin cofactor II as a novel vascular protective factor against atherosclerosis. *J Atherosclerosis Thromb*. 2009;16(5):523–531.
- Luebke-Wheeler J, Zhang K, Battle M, et al. Hepatocyte nuclear factor 4alpha is implicated in endoplasmic reticulum stress-induced acute phase response by regulating expression of cyclic adenosine monophosphate responsive element binding protein H. *Hepatology*. 2008;48(4):1242–1250.
- Wei C, Wan L, Zhang Y, et al. *Cholesterol Metabolism—Impact for SARS-CoV-2 Infection Prognosis, Entry, and Antiviral Therapies*. medRxiv; 2020:2020, 04.16.20068528.
- Wang D, Hu C, Zhu F, et al. *Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China*. JAMA; 2020.
- Jakobsen KR, Paulsen BS, Bæk R, et al. Exosomal proteins as potential diagnostic markers in advanced non-small cell lung carcinoma. *J Extracell Vesicles*. 2015;4, 26659.
- Hoom EJ, Pisitkun T, Zietse R, et al. Prospects for urinary proteomics: exosomes as a source of urinary biomarkers. *Nephrology (Carlton)*. 2005;10(3):283–290.
- Barratt J, Topham P. Urine proteomics: the present and future of measuring urinary protein components in disease. *CMAJ (Can Med Assoc J)*. 2007;177(4):361–368.
- Shao C, Li M, Li X, et al. A tool for biomarker discovery in the urinary proteome: a manually curated human and animal urine protein biomarker database. *MCP*. 2011; 10(11).
- Zhai L, Chang C, Li N, et al. Systematic research on the pretreatment of peptides for quantitative proteomics using a C 18 microcolumn. *Proteomics*. 2013;13(15): 2229–2237.
- Li Y, Wang Z, Zhou W, et al. A rapid and easy protein N-terminal profiling strategy using (N-Succinimidyl)oxycarbonylmethyl)tris(2,4,6-trimethoxyphenyl)phosphonium bromide (TMPP) labeling and StageTip. *Proteomics*. 2017;17(13–14).
- Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008;26(12):1367–1372.
- Schwahnhauser B, Busse D, Li N, et al. Global quantification of mammalian gene expression control. *Nature*. 2011;473(7347):337.
- Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*. 2016;11(12):2301–2319.
- Tyanova S, Temu T, Sinitcyn P, et al. The Perseus computational platform for comprehensive analysis of (prote) omics data. *Nat Methods*. 2016;13(9):731.
- Ma J, Chen T, Wu S, et al. iProX: an integrated proteome resource. *Nucleic Acids Res*. 2019;47(D1):D1211–D1217.