Clinical Proteomics

RESEARCH

Open Access

Serum proteomics for the identification of biomarkers to flag predilection of COVID19 patients to various organ morbidities

Madhan Vishal Rajan[1], Vipra Sharma[1], Neelam Upadhyay[1], Ananya Murali[2], Sabyasachi Bandyopadhyay[3] and Gururao Hariprasad[1*]

Abstract

Background COVID19 is a pandemic that has affected millions around the world since March 2020. While many patients recovered completely with mild illness, many patients succumbed to various organ morbidities. This heterogeneity in the clinical presentation of COVID19 infection has posed a challenge to clinicians around the world. It is therefore crucial to identify specific organ-related morbidity for effective treatment and better patient outcomes. We have carried out serum-based proteomic experiments to identify protein biomarkers that can flag organ dysfunctions in COVID19 patients.

Methods COVID19 patients were screened and tested at various hospitals across New Delhi, India. 114 serum samples from these patients, with and without organ morbidities were collected and annotated based on clinical presentation and treatment history. Of these, 29 samples comprising of heart, lung, kidney, gastrointestinal, liver, and neurological morbidities were considered for the discovery phase of the experiment. Proteins were isolated, quantified, trypsin digested, and the peptides were subjected to liquid chromatography assisted tandem mass spectrometry analysis. Data analysis was carried out using Proteome Discoverer software. Fold change analysis was carried out on MetaboAnalyst. KEGG, Reactome, and Wiki Pathway analysis of differentially expressed proteins were carried out using the STRING database. Potential biomarker candidates for various organ morbidities were validated using ELISA.

Results 254 unique proteins were identified from all the samples with a subset of 12–31 differentially expressed proteins in each of the clinical phenotypes. These proteins establish complement and coagulation cascade pathways in the pathogenesis of the organ morbidities. Validation experiments along with their diagnostic parameters confirm Secreted Protein Acidic and Rich in Cysteine, Cystatin C, and Catalase as potential biomarker candidates that can flag cardiovascular disease, renal disease, and respiratory disease, respectively.

Conclusions Label free serum proteomics shows differential protein expression in COVID19 patients with morbidity as compared to those without morbidity. Identified biomarker candidates hold promise to flag organ morbidities in COVID19 for efficient patient care.

*Correspondence:
Gururao Hariprasad
dr.hariprasadg@gmail.com

Full list of author information is available at the end of the article

Keywords COVID19, Organ morbidity, Clinical proteomics, Differential protein expression, Biomarkers, Complement factors, Coagulation factors

Background

COVID19 is an infection caused by a novel Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) that first emerged in Wuhan, China, in December 2019 [1]. WHO declared the disease a pandemic in March 2020 [2]. The COVID19 pandemic is devastating, and the long-term health consequences in the post-infection period remain unclear in terms of clinical manifestation and patient management [3–5].

While SARS-CoV-2 is known to cause substantial pulmonary diseases, including pneumonia and acute respiratory distress syndrome (ARDS), However, clinicians have observed many extra-pulmonary manifestations of COVID19 with organ morbidities beyond respiratory system in systems such as renal, myocardial, hepatic, gastrointestinal, nervous, and haematological systems [3–7]. These types of pathologies and prolonged illness after post-infection are described as 'Long COVID19' or 'post-COVID19 syndrome' [8]. Nalbandian and his group have made a comprehensive review on acute complicacies and organ-specific sequelae of COVID19 and have termed the post-COVID19 complicacies as 'Post-acute COVID19 syndrome' [4, 5]. In the post-COVID19 period 10–45% of patients experience respiratory sequelae, 10–22% of patients experience cardiovascular sequelae, 10–31% of patients experience neurological sequelae, and close to 20% of patients experience gastrointestinal sequelae [4]. This pattern of post-infection viral sequelae is not new and has been known for the SARS epidemic of 2003 and the Middle East respiratory syndrome (MERS) outbreak of 2012, where similar kinds of organotropism and persistent symptoms relating to organ morbidities [9–12]. Using various methods many studies have demonstrated the presence of SARS-like viral particles in epithelial cells of mucosa of the small and large intestines, of the renal distal tubules, neurons of the brain, and macrophages in different organs including the liver thereby establishing post viral sequelae [10, 13, 14].

Organ damage presents a significant challenge for clinicians when treating COVID19, as the clinical course can differ between patients [15, 16]. Some COVID19 survivors did not recover even after two years after acute infection, thereby necessitating a follow-up [17]. An objective tool to accurately flag organ morbidity in these patients is vital to understand the prognosis and treatment outcomes.

In the past years, OMICS-based technologies have helped to identify biomarkers candidates for diagnosis, prognosis, disease monitoring, disease recovery, and severity [18, 19]. We have recently carried out a repository-based proteomic analysis to delineate protein signatures in COVID19 related organ morbidity [20]. This has been encouraging enough to carry out proteomic analysis in COVID19 patients that can unravel molecular patterns that can flag organ dysfunction. Our group has been actively involved in carrying out clinical proteomics experiments for various health-related research questions relating to disease diagnosis and monitoring pharmacological responses in various clinical conditions [21–23]. We propose to carry out serum based proteomics to identify protein biomarkers that can flag various organ dysfunctions as post-COVID19 complication. The proposed study will help develop a translational proteomic platform to subsequently design diagnostics that will help clinicians to streamline COVID19 patient management.

Materials and methods

Clinical phenotyping and sample collection

Patients who tested PCR positive for SARS-CoV-2 were screened. Detailed clinical history was taken, examination was conducted, and pharmacological interventions were noted. Patients having single organ morbidity were recruited for the study, and their serum samples were collected and annotated. Those with multiple organ co-morbidities, having other co-existing infections or chronic ailments were excluded from the study. COVID19 patients who did not have any organ morbidity served as controls. The study overview is shown in Fig. 1. A total of 114 COVID19 serum samples of 200 µl each were collected. All samples were inactivated by sterilizing at 56°C for 30 min and stored at -80°C until further analysis [24]. While 29 samples (cardiovascular: 4, renal: 4, pulmonary: 3, gastrointestinal: 6, neurological: 4, hepatic: 4, and control: 4) were taken for the discovery phase of the proteomic experiment, 114 (cardiovascular: 20, Renal: 18, pulmonary: 20, gastrointestinal: 6, neurological: 13, hepatic: 17, and control: 20) were taken for the validation phase of the proteomic experiment.

Protein isolation and trypsin digestion

20 µl of serum was diluted with 80 µl of 50 mM ammonium bicarbonate containing 8 M urea. The solution was vortexed at 1000 rpm for 5 min and sonicated for 1 min to achieve complete protein solubilization. Protein was estimated using the Bradford assay where 2 mg/ml Bovine Serum Albumin (BSA) was used as calibration standard. 100 µg of protein was taken and reduced with 10 mM dithiothreitol at 60 °C for 30 min, followed by alkylation with 50 mM iodoacetamide at room temperature in

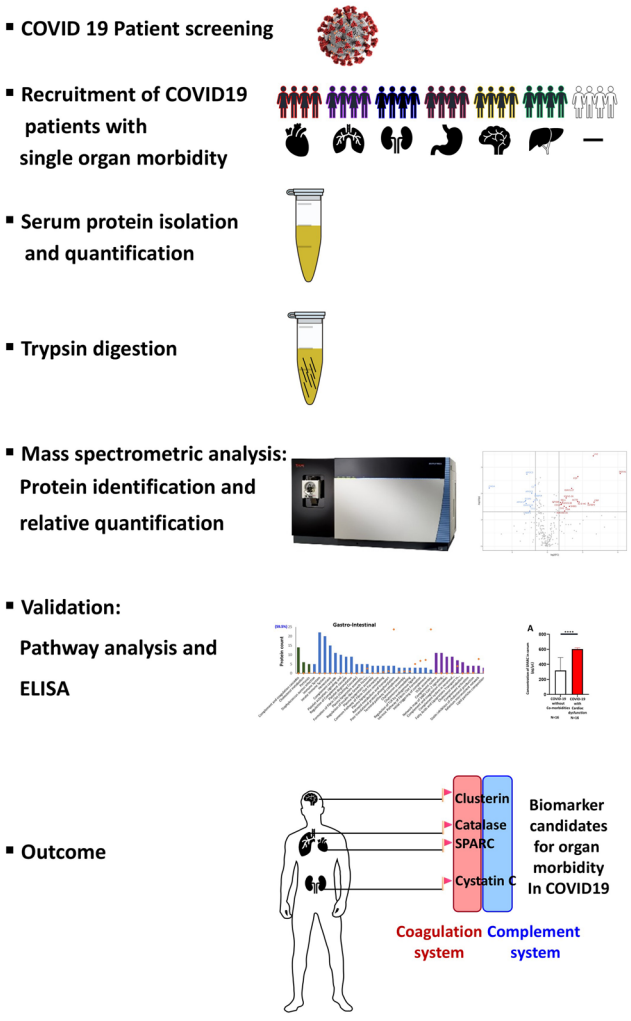

Fig. 1 Flowchart depicting the summary of methodology and outcome of this study

dark for 30 min. Final solution was diluted with 50 mM ammonium bicarbonate to reduce the urea concentration below 1 M. Proteomic grade Trypsin (Promega) was added at a 1:50 (w/v) enzyme to protein ratio and incubated at 37 °C for 12 to 18 h for protein digestion. Digestion was quenched by adding 1% of TFA and the digested peptides were purified using C18 reverse phase desalting columns. Purified peptides were lyophilized using SpeedVac vacuum concentrator (Thermo Fisher Scientific, Rockford, USA) and reconstituted in 0.1% formic acid containing LC-MS grade water (loading buffer) and concentrations were estimated using a NanoDrop Spectrophotometer (Thermo Fisher Scientific, Rockford, USA).

Liquid chromatography assisted mass spectrometry (LCMS/MS) analysis

LCMS/MS analysis was conducted on an Orbitrap Fusion Tribrid Mass Spectrometer coupled with an Easy-nLC1200 nano-flow LC system (Thermo Fisher Scientific, Rockford, USA). 1 µg of peptides from individual samples were loaded onto trap column (Acclaim PepMap 100, 3 µm, 100 Å, 75 µm x 3 cm; Thermo Fisher Scientific, Rockford, USA) and then resolved in an analytical column (Acclaim Pep-Map RSLC C18, 2 µm, 100 Å, 75 µm x 25 cm; Thermo Scientific, Rockford, USA) with a flow rate of 300 nL/min. Peptides were separated using a multi-step linear gradient of loading buffer and elution buffer (80% acetonitrile and 0.1% formic acid) at a flow rate of 300 nL/min. To elute the peptides, a 100 min gradient was used with a gradient composition of 5% elution buffer for 1 min, 8% for 10 min, 40% for 90 min, 95% for 10 min and 5% for 2 min. The mass spectra were acquired using Thermo Xcalibur (v.4.1) MS acquisition software (Thermo Fisher Scientific, Rockford, USA). For the analysis in data-dependent acquisition (DDA) mode, each scan cycle consisted of one full-scan mass spectrum (R=60 K, AGC=$5e^5$, max IT=50 ms, scan range=350–1700 m/z) followed by 20 MS/MS events in Linear Ion Trap (AGC=$1e^4$, max IT=35 ms). High energy collision dissociation (HCD) energy was set to 30%. Quadrupole isolation window was set to 1.2 Da and dynamic exclusion was set for 40 s.

Data analysis

Raw files obtained from the Orbitrap Fusion mass spectrometer were analyzed in Proteome Discoverer (v.2.4.1.15, PD 2.4, Thermo Fisher Scientific, Rockford, USA). Data are available via ProteomeXchange with identifier **PXD053440**. Human Swiss-Prot reviewed database from UniProtKB (https://www.uniprot.org/) containing 26,741 proteins was downloaded on 27th March 2023. Both canonical and isoform FASTA were taken. Label-Free Quantification (LFQ) approach was used in PD 2.4 and proteins were quantified using 'Minora Feature Detector' quantification node. Sequest algorithm was used to search peptides, where methionine oxidation and acetylation on protein N-terminus were set as variable modifications and Carbamido-methylation on cysteine was set as fixed modifications. MS1 match tolerance was set as 10 parts per million (ppm) and the MS2 tolerance was set as 0.6 Da. Searched peptides were validated using 'Percolator' node applying strict and relaxed FDR threshold of 0.01 and 0.05, respectively. Searched peptides were normalized against total peptide amount and only unique peptides were used for protein quantification. Proteins with abundance values in at least 50% of the samples were considered for statistical analysis. Data was further normalised by median-centred and log-transformed in MetaboAnalyst (Version 6.0; https://www.metaboanalyst.ca), and statistical significance analysis was done using Student's t-test, where a p-value<0.05 was considered for protein selection. Comparison of the sample groups from each phenotype for differential abundance of proteins

was done with criteria of $0.5 \geq$ Fold change ≥ 2, and statistical significance of p-value < 0.05. Partial Least-Squares Discriminant Analysis (PLS-DA) was used to analyse the categorization of clinical phenotypes and control groups based on complete protein expression.

Bioinformatics analysis

Differentially expressed genes were analysed using the STRING database v12 (Search Tool for the Retrieval of Interacting Genes/Proteins) [25]. Proteins that were $0.67 \geq$ Fold change ≥ 1.5 and statistically significant were taken for pathway analysis. Homo Sapiens was used as background species, and the enrichment analysis was run for Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathways, WikiPathways, and Reactome pathway. Results with FDR-adjusted p values < 0.01 were considered.

ELISA

Proteins SPARC, CST3, CLU, CAT, and DEFA1 were selected for the validation phase by Enzyme-Linked Immunosorbent Assay (ELISA). All protein concentrations were measured using the following commercially available ELISA kits: SPARC: Human ELISA kit (ab220654, Abcam, Cambridge, UK), CST3: Human ELISA kit (ab119589, Abcam, Cambridge, UK), CLU: Human ELISA kit (ab174447, Abcam, Cambridge, UK), CAT: Human ELISA kit (ab277396, Abcam, Cambridge, UK), and DEFA1: Human ELISA kit (CSB-E14155h, CUSABIO, Houston, TX, USA). All tests were conducted in duplicate as per the manufacturer's manual. ELISA plate readings were taken on SpectraMax i3x Multimode Microplate reader. Values below limit of detection were excluded from the analysis. Serum concentrations between the COVID19 morbidity and controls were compared using an independent Student t-test, and values of $p < 0.05$ were considered significant.

Statistical analysis

MetaboAnalyst (version 6.0) was used to normalise the protein abundance. ELISA data was analysed using GraphPad Prism 8 (GraphPad Software Inc., San Diego, CA, USA). To assess the diagnostic accuracy of each candidate biomarker marker, Receiver Operating Characteristic (ROC) curves of ELISA data were done in MetaboAnalyst (version 6.0). Area Under the Curve (AUC) was estimated with 95% confidence intervals. Optimum cut-off value was obtained at which the Yuden Index (sensitivity + specificity-1) was maximum. Likelihood ratio values were also computed. $p < 0.05$ was taken to test the significance of the AUC.

Results

Clinical profile

A total of 114 serum samples were collected from 94 COVID19 patients with various organ co-morbidities (64 men; age range: 8–78 years; median age: 43 years) and from 20 controls who had COVID19 but no co-morbidities (10 men; age range: 18–58 years; median age: 37.5 years). Distribution of patients across different clinical phenotypes is given in Table 1. The predominant age group of patients recruited in our study was 31–60 years, which comprised 72% of the participants. Of these, serum samples of 25 COVID19 patients with organ morbidities and 4 controls were taken for the discovery phase of the proteomic experiment. Distribution of patients across the phenotypes is as follows: Cardiovascular: 4 patients (4 men; median age: 53.5 years, age range: 38–60 years), Pulmonary: 3 patients (1 man; median age: 41 years, age range: 25–56 years), Neurological: 4 patients (3 men; median age: 41 years, age range: 30–51 years), Renal: 4 patients (3 men; median age: 48 years, age range: 44–57 years), Hepatic: 4 patients (4 men; median age: 43 years, age range: 31–61 years), Gastrointestinal: 6 patients (0 men; median age: 43 years, age range: 31–61 years years), and Control: 4 patients (2 men; median age: 43 years, age range: 34–57 years).

Table 1 Demography of COVID-19 patients with organ morbidity

S. No	Organ system affected	Number of samples	Male: Female	Median age with Range (years)	Clinical presentation	Phenotype
1	Cardiovascular (C)	20	16:4	50 (32–60)	Myocardial infarction on anticoagulants	COVID-19 with cardiovascular disease
2	Pulmonary (P)	20	10:10	41.5 (25–61)	Breathlessness on bronchodilators	COVID-19 with pulmonary disease
3	Neurological (N)	13	9:4	34 (8–66)	Epilepsy / migraine / depression	COVID-19 with neurological disease
4	Hepatic (L)	17	16:1	47 (18–71)	Mild abdominal distention with deranged Liver enzymes	COVID-19 with liver disease
5	Renal (R)	18	12:6	42 (19–78)	Kidney stone / Increased levels of urea and creatinine	COVID-19 with renal disease
6	Gastrointestinal (G)	6	0:6	40 (35–45)	Appendicitis / Gastritis	COVID-19 with gastrointestinal disease
7	Control (Con)	20	10:10	37.5 (18–58)	Fever/malaise	COVID-19 with no co-morbidities

Differential protein expression

A total of 919 proteins and 3425 peptides were identified from the raw files using Proteome Discoverer software. After applying a filter criterion for at least 2 unique peptides, 254 proteins were considered for further analysis. PLS-DA analysis of this subset of proteins clearly establishes a clear demarcation between organ morbidity phenotypes and control phenotypes in COVID19 (Fig. 2). Scores of the first two components were represented showing the ovals at 95% confidence intervals. With a criterion of a protein present in at least half of the biological replicates in each of the morbidity phenotypes, cardiovascular had 239 proteins, hepatic had 244 proteins, renal had 242 proteins, gastrointestinal had 243 proteins, neurological had 240 proteins, and pulmonary had 236 proteins. A two-fold differential expression with statistical significance criterion resulted in identification of 12–31 candidate biomarker proteins for each of the organ morbidities in COVID19. Hepatic phenotype had 15 upregulated and 16 downregulated proteins, gastrointestinal phenotype had 17 upregulated and 11 downregulated proteins, cardiovascular phenotype had 10 upregulated and 9 downregulated proteins, pulmonary phenotype had 10 upregulated and 8 downregulated proteins, renal phenotype had 10 upregulated and 5 downregulated proteins, and neurological phenotype had 4 upregulated and 8 downregulated proteins. These are represented as volcano plots in Fig. 3. Some of the differentially expressed proteins in the organ morbidity phenotype include: (1) acute phase reactants such as C-Reactive protein (CRP), Apolipoproteins (APOA2, APOC1, APOC2, APOC3, APOA5), haemopexin (HPX), and orosomucoid-2 (ORM2); (2) Immunoglobulins (IGKV1-27, IGKV3D-15, IGHV3-38, IGKV1-16, IGLV3-19, IGHV2-5); (3) Complement related proteins Ficolin-3 (FCN3) and, (4) coagulation related proteins such as Serine protease inhibitor C1 (SERPIC1), Factor 12 (F12), Factor 11 (F11), Factor 13 (F13A1), and Glycoprotein 1b α-chain (GP1BA). There are three proteins consistently upregulated in all the six organ morbidity phenotypes. They are: Defensin Alpha 1 (DEFA1): 27.2 fold; Lysozyme (LYZ): 10.3 fold; and Insulin-Like Growth Factor Binding Protein 2 (IGFBP2): 6.1 fold.

Pathway analysis

Functional enrichment analysis using KEGG, WikiPathways, and Reactome delineated 3–36 pathways that are implicated in the pathogenesis of the organ morbidities in COVID19. Out of these, around 60% involved complement and coagulation pathways. Percentage of complement and coagulation related pathways in each group were: 67% in the cardiovascular system, 40% in the neurological system, 86% in the renal system, 80% in the hepatic system, and 44% in the gastrointestinal system (Fig. 4). In addition, there are consistent overlaps of complement and coagulation pathways that are derived from

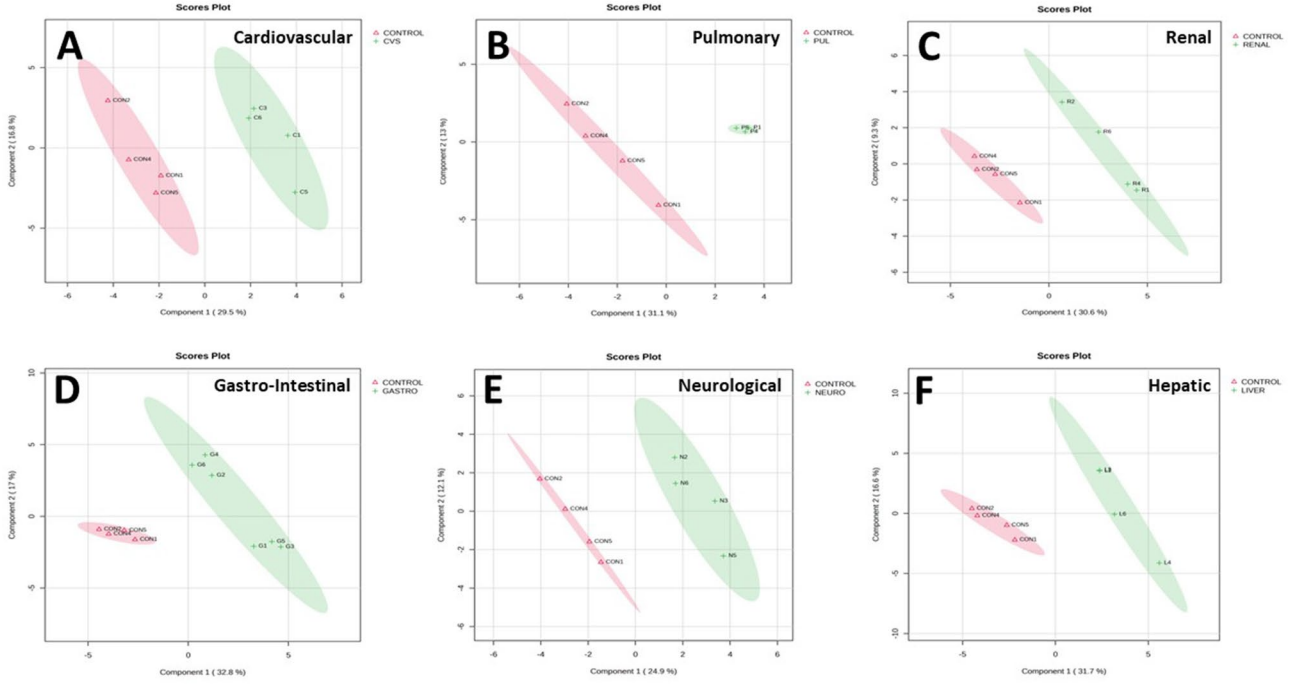

Fig. 2 Partial least-squares discriminant analysis (PLS-DA) based on abundance of proteins identified in each of the COVID19 biological replicates: (**A**) Control and Cardiovascular morbidity phenotype. (**B**) Control and Pulmonary morbidity phenotype. (**C**) Control and Renal morbidity phenotype. (**D**) Control and Gastrointestinal morbidity phenotype. (**E**) Control and Neurological morbidity phenotype. (**F**) Control and Hepatic morbidity phenotype

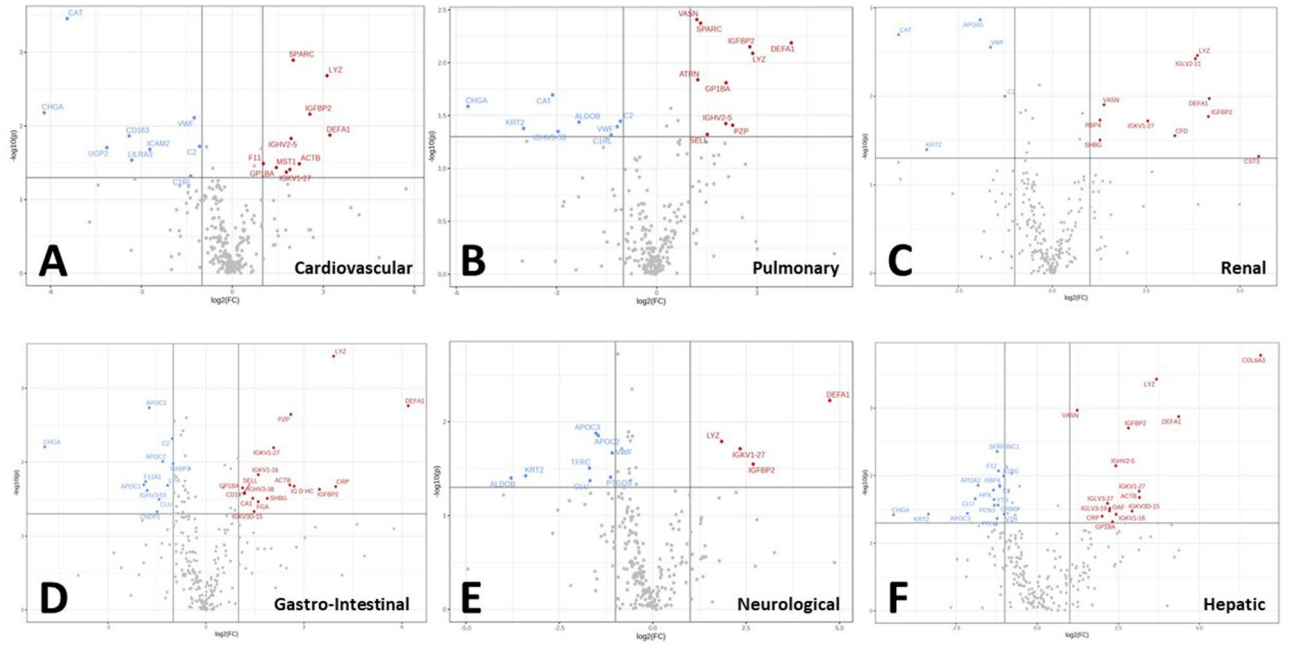

Fig. 3 Volcano plots showing differentially expressed proteins between control and organ morbidity phenotypes in COVID19. (**A**) Control and Cardiovascular morbidity phenotype. (**B**) Control and Pulmonary morbidity phenotype. (**C**) Control and Renal morbidity phenotype. (**D**) Control and Gastrointestinal morbidity phenotype. (**E**) Control and Neurological morbidity phenotype. (**F**) Control and Hepatic morbidity phenotype. With respect to organ morbidity phenotypes, Red dots represent proteins with increased fold change (> 2); blue dots represent proteins with decreased fold change (< 0.5); and grey dots represent proteins with no significant change

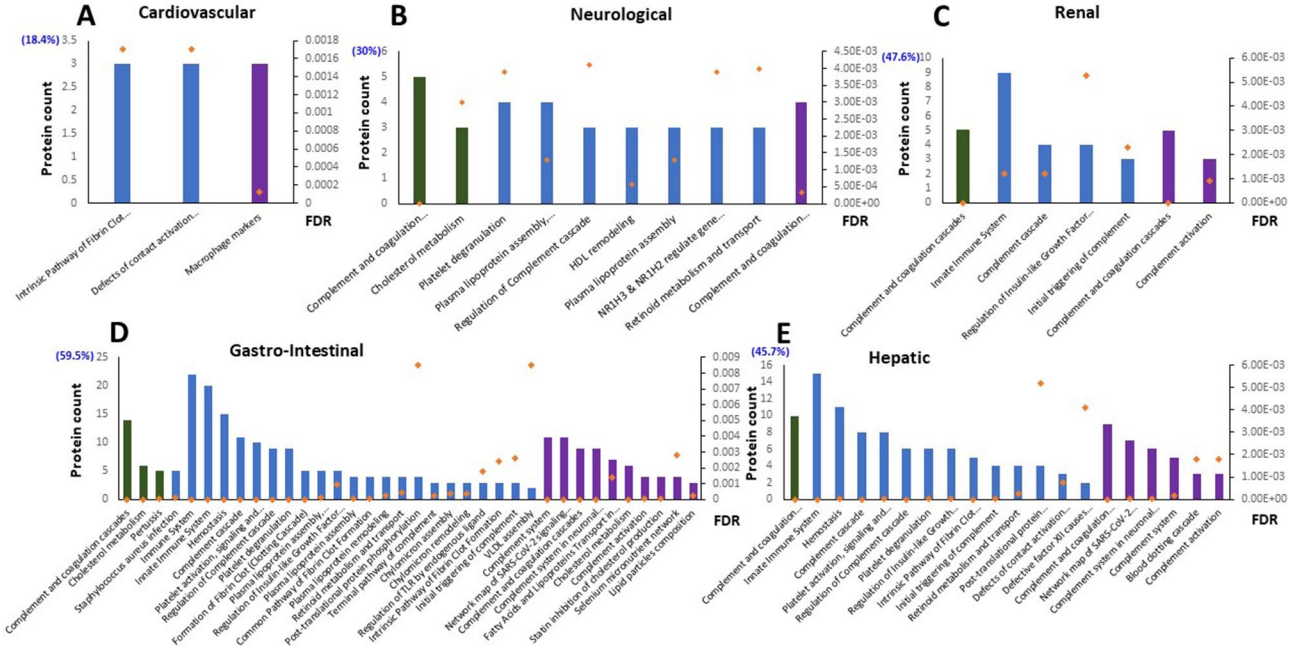

Fig. 4 Pathway analysis of differentially expressed proteins in organ morbidity phenotypes in COVID19. (**A**) Control and Cardiovascular morbidity phenotype. (**B**) Control and Neurological morbidity phenotype. (**C**) Control and Renal morbidity phenotype. (**D**) Control and Gastrointestinal morbidity phenotype. (**E**) Control and Hepatic morbidity phenotype. KEGG pathways are represented in green; Reactome pathways are represented in blue; and Wiki pathways are represented in purple. x-axis represents the different pathways. Left y-axis represents number of proteins annotated from previous studies in each of the pathways and percentage corresponding (blue) to total number of input proteins is shown for the top values in the axis. Right y-axis represents the FDR value

Table 2 Potential biomarker proteins that can flag different organ morbidities in COVID-19 patients

S. No	Protein	Accession No.	Organ system involved	Fold change	Clinical phenotypes in which the protein is present		Relevance in this study	Ref
					Controls	Morbidity		
1	SPRC_HUMAN	P09486	Cardiovascular System	4.0	4/4	4/4	SPARC reduces cardiac inflammation and mortality by improving endothelial barrier function during viral myocarditis.	[39]
							SPARC expression increases following myocardial injury.	[41]
2	CYTC_HUMAN	P01034	Renal System	44.1	3/4	4/4	CST3 is a marker for detecting changes in GFR and has high predictive value for COVID-19-related AKI	[45, 46]
3	CLUS_HUMAN	P10909	Neurological System	-3.2	4/4	4/4	CLU is a complement cytolysis inhibitor, and protects nerve cells by reducing inflammation.	[42]
							Low levels of CLU correlate with COVID19 severity	[43]
4	CATA_HUMAN	P04040	Pulmonary System	-8.7	4/4	2/3	CAT regulates production of cytokines preventing oxidative injury and suppressing SARS-CoV-2 replication.	[48, 51]
							protects against pulmonary fibrosis and prevents lung epithelial cells from hydrogen peroxide-induced apoptosis.	[56, 57]
5	DEF1_HUMAN	P59665	Common protein	27.2	2/4	25/25	DEFA1 is secreted upon neutrophil activation which is the first line of defence in viral infection	[54]
							DEFA1 correlates with COVID-19 severity; High levels are associated with fatal outcomes in patients	[55]

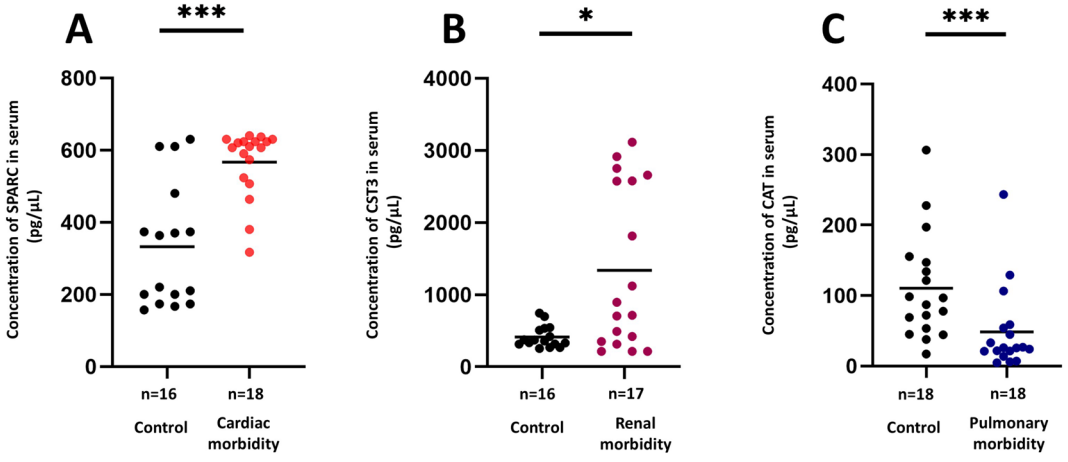

Fig. 5 ELISA of candidate biomarker proteins. (**A**) SPARC. (**B**) Cystatin C. (**C**) Catalase. The Scatter plots represent the individual concentrations of all the samples in the group as dots. Mean of the values is indicated by horizontal lines. * indicates $p < 0.05$; and *** indicates $p < 0.001$

three different data bases. In the neurological system and gastrointestinal system, which have less than half the pathways accounting for complement and coagulation, pathways relating to lipid metabolism and transport were a standout feature.

Potential biomarker candidates for organ morbidity in COVID19

All the dysregulated proteins were screened based on their function and relevance to this study to identify potential biomarker candidates for organ morbidities in COVID19 (Table 2). They are: Secreted Protein Acidic and Rich in Cysteine (SPARC) for cardiovascular disease; Clusterin (CLU) for neurological disease; Cystatin C (CST3) for renal disease; and Catalase (CAT) for respiratory disease. Of the three proteins that are upregulated in all morbidity phenotypes, Defensin Alpha 1 (DEFA1) is a neutrophil generated response to viral infection, while lysozyme and Insulin like growth factor binding protein 2 are non-specific mediators of acute inflammation. While ELISA was carried out for all five potential biomarker candidate proteins, only SPARC, CST3, and CAT showed results that validated the differential expression noted in the discovery phase (Fig. 5). Average Coefficient of variation for SPARC, CST3, and CAT are 9.3%, 5.8%, and 8.7% respectively. The corresponding ROC that was plotted is shown in Fig. 6, and the estimated diagnostic parameters are listed in Table 3. Sensitivity values of 64.7 -94.0%, and specificity values of 75.0 −94.4% are reasonable scores that reflect the usefulness of these biomarker candidates to flag organ morbidity in COVID19.

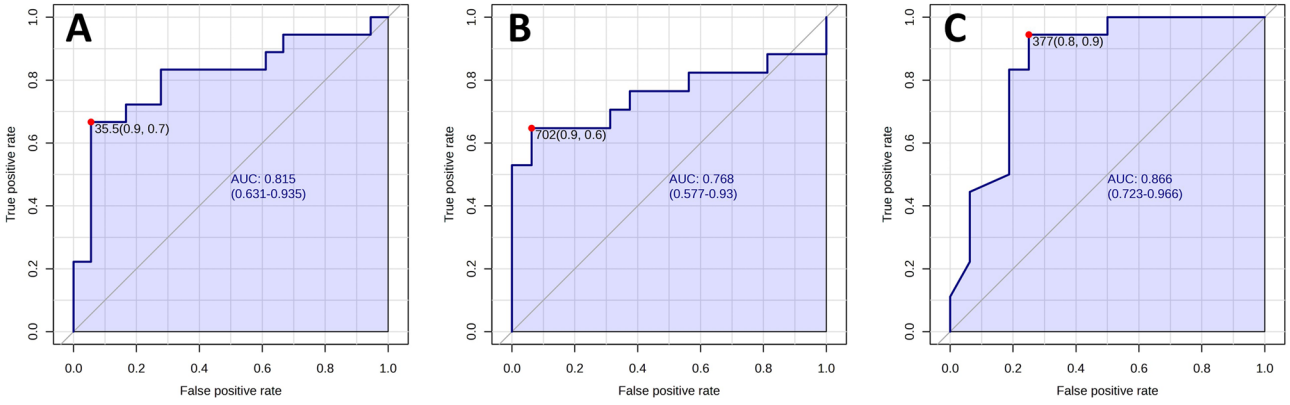

Fig. 6 ROC of candidate biomarker proteins. (**A**) SPARC. (**B**) Cystatin C. (**C**) Catalase. Area under the curve (AUC) and their 95% confidence intervals (blue area) are depicted. Red dot indicates optimal cut-off for the best sensitivity and specificity values

Table 3 Diagnostic parameters of validated proteins

Protein	Cut-off values (pg/μL of serum)	Statistical Parameters				
		AUC (95% CI: LL- UL)	Sensitivity (%)	Specificity (%)	Likelihood Ratio +	Likelihood Ratio -
SPRC_HUMAN (Cardiac)	≥ 377.0	0.868 (0.703–0.978)	94.0	75.0	3.7	0.07
CYTC_HUMAN (Renal)	≥ 702	0.746 (0.572–0.923)	64.7	93.8	10.3	0.4
CATA_HUMAN (Pulmonary)	≤ 35.5	0.815 (0.651–0.946)	66.7	94.4	12.0	0.35

Discussion

In our study, we observed that males are more commonly affected by COVID19 than females, which is in line with other COVID19-related studies [26, 27]. Presenting clinical features and lab tests such as myocardial ischemia, breathlessness, depression, deterioration in kidney function, elevated liver enzyme levels seen in patients recruited in this study are very similar to the those observed in previous studies [4, 28]. It may be noted that majority of the patients with cardiovascular disorder suffered from myocardial infarction, a hyper coagulable-thrombotic disease.

Differential protein expression shows a lot of interesting outcomes. Discriminative analysis shows intra-group homogeneity and inter-group variability among the clusters of biological replicate phenotypes. This strengthens the causal association between protein expression profiling and clinical outcomes in COVID19 patients recruited in our study. While the homogeneity exists for all the organ morbidity phenotypes in the study, it is more enhanced in the case of pulmonary morbidity. This is probably due to the fact that lungs are primary site of infection for SARS-CoV-2. Differentially expressed proteins are a cumulative effect of body response to SARS-CoV-2 virus. While identification of acute phase reactants is indicative of systemic inflammatory response syndrome, identification of different immunoglobulins is a clear indicative of the humoral response mediated against the viral antigens. Identification of high number of coagulation factors and complement factors as dysregulated proteins provides ample evidence of their role in causation across the various organ morbidities in COVID19.

Mapping of their respective pathways implicates thrombo-inflammation as one of the key pathogenic mechanisms in COVID19. Possible reasons for association of 'complement-coagulation system' with COVID19 are: (1) SARS-CoV-2 virus by causing endothelial damage and thrombo-inflammation activating complement system [29]. (2) Spike protein and nucleo-capsid proteins of the virus are recognised by lectin pathway leading to complement activation [30]. (3) Immunoglobulins directed against receptor binding domain of the spike protein initiates classical complement pathway [31]; virus attack on the immune system causes a cytokine storm that triggers coagulation complications [32]. (4) spike protein of SARS-CoV-2 by binding to heparan sulfate and competing with factor H disrupts alternative pathway of complement activation [33, 34]. (5) Neutrophils which are major cellular component of innate immune system against viral infection releases various products including coagulants and complement factors [35]. It is therefore clear that 'complement-coagulation' activation in COVID19 is indeed a crucial factor in the pathogenesis of organ morbidity in COVID19. This is supported by immunofluorescence detection of complement deposits in the lungs, kidneys, and liver tissues of COVID19 patients [36]. Also, COVID19-associated coagulopathy

causes significant damage to multiple organs, including the lungs, heart, kidneys, and brain, contributing to the high mortality in severe COVID19 [35, 37].

It is very evident from the experiments and validation studies that there is a good possibility of candidate biomarker proteins for organ morbidity in COVID19. Some of these proteins and their functional roles in certain organ systems are discussed here. (**1**) Secreted Protein Acidic and Rich in Cysteine (SPARC) plays a significant role in regulating cellular interactions with the extracellular matrix (ECM) [38]. In the cardiac tissue it: (a) Reduces inflammation by preserving endothelial glycocalyx integrity in viral myocarditis [39]; (b) in response to myocardial injury, it aids in tissue regeneration [40, 41]; and (c) rescue myocytes that are compromised by viral infections [41]; (**2**) Clusterin is a complement cytolysis inhibitor which is expressed by nerve cells as a defence mechanism against endogenous complement attack, thereby reducing inflammation and cerebral edema [42]. Decreased levels of Clusterin seen in the neurological patients recruited in this study possibly indicates disease severity highly predictive of poor outcomes [43]. (**3**) Cystatin C is an endogenous cysteine proteinase inhibitor produced by all nucleated cells and is a sensitive marker for detecting changes in glomerular filtration rate (GFR) [44, 45]. Serum Cystatin C has demonstrated a high predictive value for COVID19-related Acute Kidney Injury and was also associated with COVID-19 severity and mortality [46, 47]. (**4**) Catalase is an intracellular antioxidant enzyme present in alveolar epithelial cells that helps in protecting lung tissue from oxidative stress in COVID19 [48–50]. In addition to, it is also involved in regulating cytokines [51]. (**5**) DEFA-1 is produced by the innate immune system and epithelial cells and mainly stored in neutrophilic granules [52, 53]. In response to viral infection, including COVID19, neutrophils migrate to the site of infection and, upon activation, release multiple molecules, including alpha-defensins such as DEFA1 [54]. The accumulation of neutrophils is observed in severe COVID19 patients compared to non-severe patients, and DEFA1 levels have been found to be significantly higher in patients with poor outcomes [55]. Excellent diagnostic parameters estimated for some of these proteins establish possible usefulness of these biomarker candidates to flag organ morbidities in COVID19.

Conclusions

Serum of COVID19 patients with various co-morbidities exhibit differentially expressed proteins. Complement and coagulation are the main pathways implicated in the pathogenesis of organ morbidity in COVID19. Potential biomarkers that could possibly flag various organ morbidities in COVID19: Secreted Protein Acidic And Rich in Cysteine for cardiac pathology, Catalase for the Pulmonary pathology, and Cystatin C for the Renal pathology. Diagnostic values with a minimum of 65% sensitivity and 75% specificity offers a potential platform for development of diagnostics for organ morbidity in COVID19.

Abbreviations

AGC	Automatic gain control
ARDS	Acute respiratory distress syndrome
AUC	Area under the curve
BSA	Bovine serum albumin
CAT	Catalase
CLU	Clusterin
CRP	C-reactive protein
CST3	Cystatin C
DDA	Data dependent acquisition
DEFA1	Defensin Alpha 1
ECM	Extracellular matrix
ELISA	Enzyme linked immunosorbent assay
FDR	False discovery rate
GFR	Glomerular filtration rate
HCD	High energy collision dissociation
IGFBP2	Insulin like growth factor binding protein 2
IL-6	Interleukin-6
KEGG	Kyoto encyclopaedia of genes and genomes
LCMS	Liquid chromatography mass spectrometry
LDH	Lactate dehydrogenase
LFQ	Label free quantification
LYZ	Lysozyme
MAC	Membrane attack complex
MERS	Middle east respiratory syndrome
MS	Mass spectrometry
m/z	Mass to charge ratio
NETs	Neutrophil extracellular traps
PCR	Polymerase chain reaction
PD	Proteome discoverer
PLSDA	Partial least squares discriminant analysis
PCT	Procalcitonin
PMN	Polymorphonuclear leukocytes
ppm	Parts per million
ROC	Receiver operating characteristic
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SPARC	Secreted protein acidic and rich in cysteine
STRING	Search tool for the retrieval of interacting genes/proteins

Acknowledgements

Proteomics lab at Central Core Research Facility, AIIMS for the assistance, and Translational Health Science and Technology Institute (THSTI), Faridabad for serum sample sourcing are acknowledged.

Author contributions

GH conceptualized and designed the experimental flow. MVR and SB collected the sample tissues. MVR, VS, NU, SB, GH carried out the experiments. AM assisted GH in reviewing the literature. MVR and GH carried out the analysis and drafted the manuscript. GH edited and revised the manuscript. All authors reviewed the manuscript.

Funding

This work was financially supported by the Indian Council of Medical Research, New Delhi (BMI/2021/6364).

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD053440 [58].

Declarations

Ethics approval and consent to participate

This study was conducted after approval was obtained from the Institute Ethics Committee at All India Institute of Medical Sciences New Delhi, India (IECPG-374/26.05.2022) and is in accordance with the Declaration of Helsinki. Informed consent was obtained from all subjects and/or their legal guardian(s).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Biophysics, All India Institute of Medical Sciences, New Delhi, India
[2]Subbaiah Institute of Medical Sciences, New Delhi, Karnataka, India
[3]Proteomics Sub-Facility, Centralized Core Research Facility, All India Institute of Medical Sciences, New Delhi, India

Received: 29 June 2024 / Accepted: 23 October 2024
Published online: 01 November 2024

References

1. Gallo Marin B, Aghagoli G, Lavine K, Yang L, Siff EJ, Chiang SS, et al. Predictors of COVID-19 severity: a literature review. Rev Med Virol. 2021;31:e2146.
2. Ciotti M, Ciccozzi M, Terrinoni A, Jiang W-C, Wang C-B, Bernardini S. The COVID-19 pandemic. Crit Rev Clin Lab Sci. 2020;57:365–88.
3. Huang C, Huang L, Wang Y, Li X, Ren L, Gu X, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. Lancet. 2023;401:e21–33.
4. Nalbandian A, Sehgal K, Gupta A, Madhavan MV, McGroder C, Stevens JS, et al. Post-acute COVID-19 syndrome. Nat Med. 2021;27:601–15.
5. Gupta A, Madhavan MV, Sehgal K, Nair N, Mahajan S, Sehrawat TS, et al. Extrapulmonary manifestations of COVID-19. Nat Med. 2020;26:1017–32.
6. Attaway AH, Scheraga RG, Bhimraj A, Biehl M, Hatipoğlu U. Severe covid-19 pneumonia: pathogenesis and clinical management. BMJ. 2021;372:n436.
7. Shakaib B, Zohra T, Ikram A, Shakaib MB, Ali A, Bashir A, et al. A comprehensive review on clinical and mechanistic pathophysiological aspects of COVID-19 Malady: How far have we come? Virol J. 2021;18:120.
8. Najafi MB, Javanmard SH. Post-COVID-19 Syndrome Mechanisms, Prevention and Management. Int J Prev Med. 2023;14:59.
9. Gerges Harb J, Noureldine HA, Chedid G, Eldine MN, Abdallah DA, Chedid NF, et al. SARS, MERS and COVID-19: clinical manifestations and organ-system complications: a mini review. Pathog Dis. 2020;78:ftaa033.
10. Gu J, Gong E, Zhang B, Zheng J, Gao Z, Zhong Y, et al. Multiple organ infection and the pathogenesis of SARS. J Exp Med. 2005;202:415–24.
11. Ngai JC, Ko FW, Ng SS, To K-W, Tong M, Hui DS. The long-term impact of severe acute respiratory syndrome on pulmonary function, exercise capacity and health status. Respirology. 2010;15:543–50.
12. Saad M, Omrani AS, Baig K, Bahloul A, Elzein F, Matin MA, et al. Clinical aspects and outcomes of 70 patients with Middle East respiratory syndrome coronavirus infection: a single-center experience in Saudi Arabia. Int J Infect Dis. 2014;29:301–6.
13. Humar A, McGilvray I, Phillips MJ, Levy GA. Severe acute respiratory syndrome and the liver. Hepatology. 2004;39:291–4.
14. Yang Z, Xu M, Yi J-Q, Jia W-D. Clinical characteristics and mechanism of liver damage in patients with severe acute respiratory syndrome. Hepatobiliary Pancreat Dis Int. 2005;4:60–3.
15. Vosko I, Zirlik A, Bugger H. Impact of COVID-19 on Cardiovascular Disease. Viruses. 2023;15:508.
16. Zuccon W, Comassi P, Adriani L, Bergamaschini G, Bertin E, Borromeo R, et al. Intensive care for seriously ill patients affected by novel coronavirus sars - CoV – 2: Experience of the Crema Hospital, Italy. Am J Emerg Med. 2021;45:156–61.
17. Huang L, Li X, Gu X, Zhang H, Ren L, Guo L, et al. Health outcomes in people 2 years after surviving hospitalisation with COVID-19: a longitudinal cohort study. Lancet Respir Med. 2022;10:863–76.
18. Battaglini D, Lopes-Pacheco M, Castro-Faria-Neto HC, Pelosi P, Rocco PRM. Laboratory Biomarkers for Diagnosis and Prognosis in COVID-19. Front Immunol. 2022;13:857573.
19. Villar M, Urra JM, Rodríguez-del-Río FJ, Artigas-Jerónimo S, Jiménez-Collados N, Ferreras-Colino E et al. Characterization by Quantitative Serum Proteomics of Immune-Related Prognostic Biomarkers for COVID-19 Symptomatology. Front Immunol [Internet]. 2021 [cited 2024 Jun 6];12. https://www.frontiersin.org/journals/immunology/articles/https://doi.org/10.3389/fimmu.2021.730710/full
20. Bandyopadhyay S, Rajan MV, Kaur P, Hariprasad G. Identification of potential biomarkers to predict organ morbidity in COVID-19: A repository based proteomics perspective. Biochem Biophys Rep. 2023;35:101493.
21. Sehrawat U, Pokhriyal R, Gupta AK, Hariprasad R, Khan MI, Gupta D, et al. Comparative Proteomic Analysis of Advanced Ovarian Cancer Tissue to Identify Potential Biomarkers of Responders and Nonresponders to First-Line Chemotherapy of Carboplatin and Paclitaxel. Biomark Cancer. 2016;8:43–56.
22. Pathania S, Khan MI, Kumar A, Gupta AK, Rani K, Ramesh Parashar T, et al. Proteomics of Sentinel Lymph Nodes in Early Breast Cancer for Identification of Thymidylate Synthase as a Potential Biomarker to Flag Metastasis: A Preliminary Study. Cancer Manag Res. 2020;12:4841–54.
23. Gupta AK, Pokhriyal R, Khan MI, Kumar DR, Gupta R, Chadda RK, et al. Cerebrospinal Fluid Proteomics For Identification Of α2-Macroglobulin As A Potential Biomarker To Monitor Pharmacological Therapeutic Efficacy In Dopamine Dictated Disease States Of Parkinson's Disease And Schizophrenia. Neuropsychiatr Dis Treat. 2019;15:2853–67.
24. Shen B, Yi X, Sun Y, Bi X, Du J, Zhang C, et al. Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. Cell. 2020;182:59–e7215.
25. von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, et al. STRING: known and predicted protein–protein associations, integrated and transferred across organisms. Nucleic Acids Res. 2005;33:D433–7.
26. Swain CK, Rout HS. Gender and age group-wise inequality in health burden and value of premature death from COVID-19 in India. Aging Health Res. 2023;3:100151.
27. Mazumder A, Arora M, Bharadiya V, Berry P, Agarwal M, Behera P et al. SARS-CoV-2 epidemic in India: epidemiological features and *in silico* analysis of the effect of interventions [Internet]. F1000Research; 2020 [cited 2024 Jun 21]. https://f1000research.com/articles/9-315
28. Parotto M, Gyöngyösi M, Howe K, Myatra SN, Ranzani O, Shankar-Hari M, et al. Post-acute sequelae of COVID-19: understanding and addressing the burden of multisystem manifestations. Lancet Respiratory Med. 2023;11:739–54.
29. Afzali B, Noris M, Lambrecht BN, Kemper C. The state of complement in COVID-19. Nat Rev Immunol. 2022;22:77–84.
30. Ali YM, Ferrari M, Lynch NJ, Yaseen S, Dudler T, Gragerov S, et al. Lectin Pathway Mediates Complement Activation by SARS-CoV-2 Proteins. Front Immunol. 2021;12:714511.
31. Holter JC, Pischke SE, de Boer E, Lind A, Jenum S, Holten AR, et al. Systemic complement activation is associated with respiratory failure in COVID-19 hospitalized patients. Proc Natl Acad Sci U S A. 2020;117:25018–25.
32. Savla SR, Prabhavalkar KS, Bhatt LK. Cytokine storm associated coagulation complications in COVID-19 patients: Pathogenesis and Management. Expert Rev Anti Infect Ther. 2021;19:1397–413.
33. Yu J, Gerber GF, Chen H, Yuan X, Chaturvedi S, Braunstein EM, et al. Complement dysregulation is associated with severe COVID-19 illness. Haematologica. 2022;107:1095–105.
34. Yu J, Yuan X, Chen H, Chaturvedi S, Braunstein EM, Brodsky RA. Direct activation of the alternative complement pathway by SARS-CoV-2 spike proteins is blocked by factor D inhibition. Blood. 2020;136:2080–9.
35. Conway EM, Mackman N, Warren RQ, Wolberg AS, Mosnier LO, Campbell RA, et al. Understanding COVID-19-associated coagulopathy. Nat Rev Immunol. 2022;22:639–49.
36. Macor P, Durigutto P, Mangogna A, Bussani R, De Maso L, D'Errico S, et al. Multiple-Organ Complement Deposition on Vascular Endothelium in COVID-19 Patients. Biomedicines. 2021;9:1003.
37. Iba T, Warkentin TE, Thachil J, Levi M, Levy JH. Proposal of the Definition for COVID-19-Associated Coagulopathy. J Clin Med. 2021;10:191.
38. Rosset EM, Bradshaw AD. SPARC/osteonectin in mineralized tissue. Matrix Biol. 2016;52–54:78–87.
39. Rienks M, Carai P, van Teeffelen J, Eskens B, Verhesen W, Hemmeryckx B, et al. SPARC preserves endothelial glycocalyx integrity, and protects against adverse cardiac inflammation and injury during viral myocarditis. Matrix Biol. 2018;74:21–34.

40. Ghanemi A, Yoshioka M, St-Amand J. Secreted Protein Acidic and Rich in Cysteine as A Regeneration Factor: Beyond the Tissue Repair. Life (Basel). 2021;11:38.
41. Deckx S, Johnson DM, Rienks M, Carai P, Deel EV, der Velden JV, et al. Extracellular SPARC increases cardiomyocyte contraction during health and disease. PLoS ONE. 2019;14:e0209534.
42. Charnay Y, Imhof A, Vallet PG, Kovari E, Bouras C, Giannakopoulos P. Clusterin in neurological disorders: Molecular perspectives and clinical relevance. Brain Res Bull. 2012;88:434–43.
43. Siggins MK, Davies K, Fellows R, Thwaites RS, Baillie JK, Semple MG, et al. Alternative pathway dysregulation in tissues drives sustained complement activation and predicts outcome across the disease course in COVID-19. Immunology. 2023;168:473–92.
44. Jacobsson B, Lignelid H, Bergerheim U. s. r. Transthyretin and cystatin C are catabolized in proximal tubular epithelial cells and the proteins are not useful as markers for renal cell carcinomas. Histopathology. 1995;26:559–64.
45. Bagshaw SM, Bellomo R. Cystatin C in acute kidney injury. Curr Opin Crit Care. 2010;16:533–9.
46. Yildirim C, Ozger HS, Yasar E, Tombul N, Gulbahar O, Yildiz M, et al. Early predictors of acute kidney injury in COVID-19 patients. Nephrol (Carlton). 2021;26:513–21.
47. Zinellu A, Mangoni AA, Cystatin C. COVID-19 severity and mortality: a systematic review and meta-analysis. J Nephrol. 2022;35:59–68.
48. Nishikawa M, Hashida M, Takakura Y. Catalase delivery for inhibiting ROS-mediated tissue injury and tumor metastasis. Adv Drug Deliv Rev. 2009;61:319–26.
49. Tsermpini EE, Glamočlija U, Ulucan-Karnak F, Redenšek Trampuž S, Dolžan V. Molecular Mechanisms Related to Responses to Oxidative Stress and Antioxidative Therapies in COVID-19: A Systematic Review. Antioxidants. 2022;11:1609.
50. Lage SL, Amaral EP, Hilligan KL, Laidlaw E, Rupert A, Namasivayan S et al. Persistent Oxidative Stress and Inflammasome Activation in CD14highCD16 – Monocytes From COVID-19 Patients. Front Immunol [Internet]. 2022 [cited 2024 Jun 21];12. https://www.frontiersin.org/journals/immunology/articles/https://doi.org/10.3389/fimmu.2021.799558/full
51. Qin M, Cao Z, Wen J, Yu Q, Liu C, Wang F, et al. An Antioxidant Enzyme Therapeutic for COVID-19. Adv Mater. 2020;32:2004901.
52. Lehrer RI, Lu W. α-Defensins in human innate immunity. Immunol Rev. 2012;245:84–112.
53. Ganz T. Defensins: antimicrobial peptides of innate immunity. Nat Rev Immunol. 2003;3:710–20.
54. Faurschou M, Borregaard N. Neutrophil granules and secretory vesicles in inflammation. Microbes Infect. 2003;5:1317–27.
55. Shrivastava S, Chelluboina S, Jedge P, Doke P, Palkar S, Mishra AC, et al. Elevated Levels of Neutrophil Activated Proteins, Alpha-Defensins (DEFA1), Calprotectin (S100A8/A9) and Myeloperoxidase (MPO) Are Associated With Disease Severity in COVID-19 Patients. Front Cell Infect Microbiol. 2021;11:751232.
56. Odajima N, Betsuyaku T, Nagai K, Moriyama C, Wang D-H, Takigawa T, et al. The Role of Catalase in Pulmonary Fibrosis. Respir Res. 2010;11:183.
57. Arita Y, Harkness SH, Kazzaz JA, Koo H, Joseph A, Melendez JA, et al. Mitochondrial localization of catalase provides optimal protection from H2O2-induced cell death in lung epithelial cells. Am J Physiol Lung Cell Mol Physiol. 2006;290:L978–986.
58. Perez-Riverol Y, Bai J, Bandla C, García-Seisdedos D, Hewapathirana S, Kamatchinathan S, et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. 2022;50:D543–52.

Publisher's note