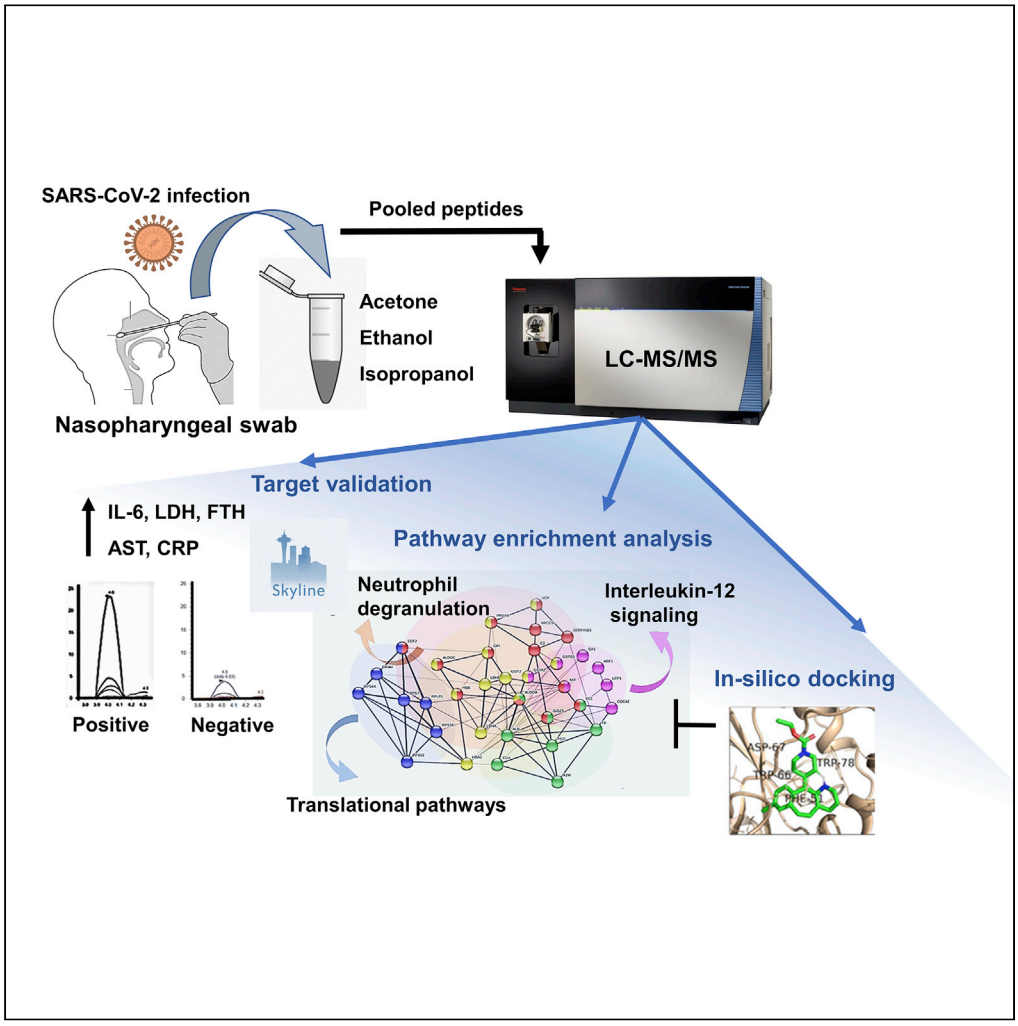
iScience

CellPress
OPEN ACCESS

Article

Proteomic investigation reveals dominant alterations of neutrophil degranulation and mRNA translation pathways in patients with COVID-19



Renuka Bankar, Kruthi Suvarna, Saicharan Ghantasala, ..., Om Shrivastav, Jayanthi Shastri, Sanjeeva Srivastava

sanjeeva@iitb.ac.in

HIGHLIGHTS

High-resolution mass spectrometry of swab identified 164 significant host proteins

Upregulation of LDH, Ferritin, and AST was validated by MRM assays

Significant alteration of immune response and translational pathways was observed

In silico docking identified Loratadine binding to interleukin signaling proteins

Bankar et al., iScience *24*, 102135
March 19, 2021
https://doi.org/10.1016/j.isci.2021.102135

iScience

CellPress
OPEN ACCESS

Article

Proteomic investigation reveals dominant alterations of neutrophil degranulation and mRNA translation pathways in patients with COVID-19

Renuka Bankar,[1,7] Kruthi Suvarna,[1,7] Saicharan Ghantasala,[2,7] Arghya Banerjee,[1,7] Deeptarup Biswas,[1,7] Manisha Choudhury,[1,7] Viswanthram Palanivel,[1] Akanksha Salkar,[1] Ayushi Verma,[1] Avinash Singh,[1] Amrita Mukherjee,[1] Medha Gayathri J. Pai,[1] Jyotirmoy Roy,[3] Alisha Srivastava,[1,4] Apoorva Badaya,[5] Sachee Agrawal,[6] Om Shrivastav,[6] Jayanthi Shastri,[6] and Sanjeeva Srivastava[1,8,*]

SUMMARY

The altered molecular proteins and pathways in response to COVID-19 infection are still unclear. Here, we performed a comprehensive proteomics-based investigation of nasopharyngeal swab samples from patients with COVID-19 to study the host response by employing simple extraction strategies. Few of the host proteins such as interleukin-6, L-lactate dehydrogenase, C-reactive protein, Ferritin, and aspartate aminotransferase were found to be upregulated only in COVID-19-positive patients using targeted multiple reaction monitoring studies. The most important pathways identified by enrichment analysis were neutrophil degranulation, interleukin-12 signaling pathways, and mRNA translation of proteins thus providing the detailed investigation of host response in COVID-19 infection. Thus, we conclude that mass spectrometry-detected host proteins have a potential for disease severity progression; however, suitable validation strategies should be deployed for the clinical translation. Furthermore, the *in silico* docking of potential drugs with host proteins involved in the interleukin-12 signaling pathway might aid in COVID-19 therapeutic interventions.

INTRODUCTION

The COVID-19 outbreak first identified in China in December 2019 has been officially declared a global pandemic by the World Health Organization (WHO). Last year, pneumonia-like symptoms caused by a new flu-like virus seemed to wipe out a part of the population in Wuhan, China, which was later named as 2019-novel coronavirus or severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by WHO. This viral disease manifests with mild fever, throat pain, dry cough, and severe respiratory difficulties to acute respiratory distress leading to breathlessness, pneumonia, and later multiple organ failure, which is the major cause of fatality rates (Chen et al., 2020). A positive viral infection is often diagnosed by either of the two ways; detection of viral load by real-time polymerase chain reaction (RT-PCR) and antibody-based tests (Chan et al., 2020). However, these techniques experience several drawbacks, such as RT-PCR faces several technical artifacts, including a high risk of cross-reaction with other potential coronaviruses and the amplicons to be contaminated; whereas, serological tests have very low-sensitivity in the early infection phase (Tang et al., 2020). Although it is important to screen individuals, mostly asymptomatic individuals, for a country with a large population such as India, it is crucial to use all its available infrastructure based on RT-PCR, next-generation sequencing (NGS), and mass spectrometry platforms, as well as antibody-based assays to expedite the testing. Further, for patients with COVID-19 with mild symptoms, identification of biological markers that define COVID-19 disease severity could help clinicians to provide better health-care facilities to patients in need. Therefore, exploring alternate and effective testing strategies, as well as identification of predictive biomarkers are very crucial to combat this pandemic.

One of the major concerns is the lack of any established therapy or treatment after COVID-19 infection and health-care systems around the world are relying on symptomatic care and supportive medicine. Currently, the disease is being treated by a combination of available antiviral drugs such as remdesivir, lopinavir, and ritonavir (Kim, 2020), and so on. Here, the scenario becomes burdensome when it comes to treating post-

[1]Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Powai, Mumbai, Maharashtra 400076, India

[2]Centre for Research in Nanotechnology and Sciences, Indian Institute of Technology Bombay, India

[3]Department of Chemical Engineering, Indian Institute of Technology Bombay, Powai, Mumbai, Maharashtra 400076, India

[4]University of Delhi, New Delhi, Delhi 110021, India

[5]Department of Chemistry, Indian Institute of Technology Bombay, Powai, Mumbai, Maharashtra 400076, India

[6]Kasturba Hospital for Infectious Diseases, Chinchpokli, Mumbai, Maharashtra 400034, India

[7]These authors contributed equally

[8]Lead contact

*Correspondence: sanjeeva@iitb.ac.in

https://doi.org/10.1016/j.isci.2021.102135

infection lung injury followed by multi-organ failure (Matthay et al., 2020). To target such key issues, proteomics-based investigations for the identification of biomarkers for COVID-19 has already begun among the scientific community all over the globe. Proteomics study on SARS-CoV-2-infected Caco-2 cells presented evidence of remodeling of cellular translation to overcome the changes occurring in host translation capacity as a response to the viral attack (Bojkova et al., 2020). Mass spectrometry-based proteomics from diluted gargle samples was able to identify the presence of a unique SARS-CoV-2 peptide, which belongs to viral nucleoprotein (Ihling et al., 2020).

Nasopharyngeal swab has a high viral titer, and it is one of the easiest, cheapest, and non-invasive ways of sample acquisition, currently being used for the RT-PCR-based detection of SARS-CoV-2 virus. Study conducted by Nikolaev et al. resulted in identification of SARS-CoV-2 nucleocapsid (N) protein, which pioneered the idea of developing COVID-19 biomarkers using nasopharyngeal swab samples (Nikolaev et al., 2020). It is evident that nasopharyngeal swab could be used to identify viral peptides for COVID-19 detection as well as for studying the host protein alteration using high-resolution mass spectrometry.

Host pathobiology plays a significant role in RNA virus infection. Pathomechanism of SARS-CoV-2 shows virus hijacks and modulates host cell machinery in multiple ways upon its entry. Till now, it is unclear how SARS-CoV-2 alters the host molecular pathways, although recent research has made some progress identifying and characterizing the biomolecular findings for disease onset and progression (Bock and Ortea, 2020). Previously, Bock et al. have re-analyzed the publicly available proteomic data generated by infecting SARS-CoV-2 in a cell line model and found several host proteins alteration, which were linked to various intracellular pathways (Bock and Ortea, 2020; Bojkova et al., 2020). In a study when nasopharyngeal swab samples from COVID-19-infected patients were analyzed, proteins belonging to the innate immune system were found to be significantly altered. Recent study on plasma proteomics revealed the potential clinical classifiers of COVID-19 infection (Messner et al., 2020). Such observations made at the macromolecular level may explain why a large part of innate immune machinery experiences the quintessential change. Till date, very few studies are available, which focuses on the unique biomolecular signature, that can play a crucial role in understanding the disease mechanism and clinical translation aiming at prognosis and cure.

In this study, we have used high-resolution mass spectrometry-based proteomics to identify viral peptides for detection and host proteins for the prognosis of COVID-19 disease severity. We have validated differentially expressed host proteins using multiple reaction monitoring (MRM)-based mass spectrometry assay. This is the first comprehensive proteomics study from an Indian hot spot COVID region, Mumbai, displaying the potential of proteins identified from mass spectrometry-based approaches for understanding the COVID-19 pathogenesis.

RESULTS

Detection of viral peptides from COVID-19 nasopharyngeal swab samples

The mass spectrometry-based label-free quantification is an efficient technique for a simple, rapid, and high-throughput analysis of the peptides from the clinical samples. One of the major challenges associated with the successful mass spectrometry-based detection of the unique peptides and proteins depends on the initial steps of protein precipitation and peptide enrichment. Here, we investigated the use of three different organic solvents (ethanol, acetone, and isopropanol) for the enrichment of peptides from a total of 101 nasopharyngeal swab samples collected in viral transport media (VTM). The swab samples in VTM tubes were heat-inactivated before addition of solvents. The detailed workflow for the mass spectrometry-based analysis of respiratory specimen is shown in Figure 1A. To study the disease progression, only those patients with COVID-19 who were confirmed by RT-PCR and presented non-severe-to-severe symptoms were selected. The clinical details of all the patient samples are shown in Table S1.

At the initial phase of the study, we performed shotgun mass spectrometry analysis for swab samples in three different solvent conditions. Comparison of the three solvents in detecting unique viral peptides showed that for spike glycoprotein and nucleoprotein similar number of peptides were obtained using each of the three solvents, whereas, in the case of replicase polyprotein 1ab protein, acetone exclusively yielded a higher number of enriched peptides (Figure 1B).

To further validate the viral peptides detected from the discovery study comprising a set of 13 patients, 12 peptides were selected, and MRM assays were performed on a set of 5 COVID-19-positive, 5 COVID-19-

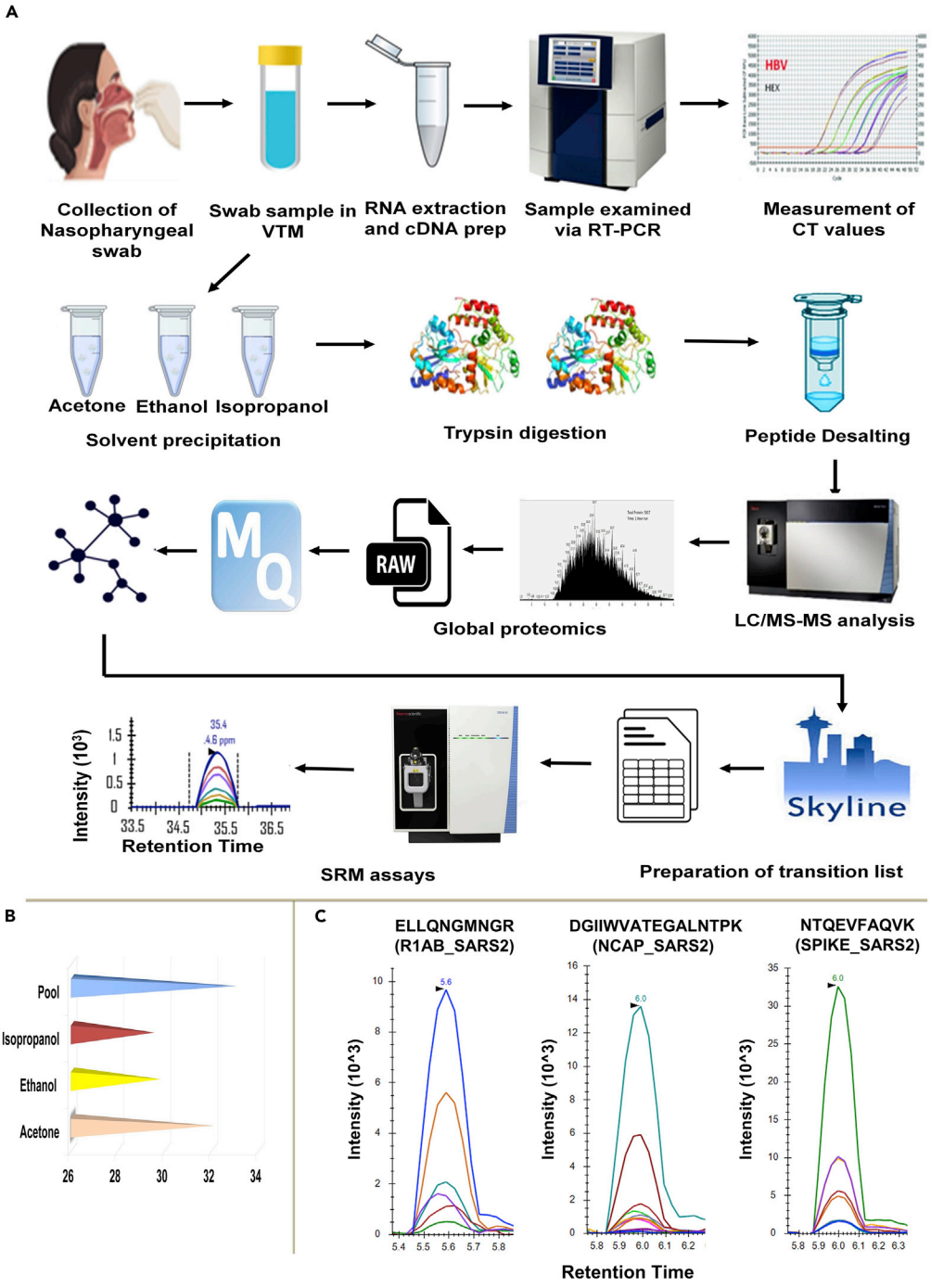


Figure 1. The workflow of sample preparation for detection of viral load and mass spectrometry analysis

Swab samples were collected in viral transfer media and RNA was extracted followed by cDNA preparation which was analyzed by RT-PCR for determining the viral load. For MS analysis, proteins in swab samples were precipitated using three organic solvents namely acetone, ethanol and isopropanol. The proteins were digested, desalted and subjected to mass spectrometry analysis. Raw files from Mass Spectrometry were analyzed using Maxquant and significant proteins were selected for MRM analysis.

(A) Schematic representation for RT-PCR and mass spectrometry-based detection of swab proteins.

(B) Comparison of various extraction methods. The sample pool used in this study was prepared from all three solvents, which yielded the maximum average number of unique viral peptides as compared to the other extraction methods.

(C) Representative MRM spectra for three viral peptides as observed in COVID-19 positive swab samples. The peptides belong to Replicase 1ab, Nucleocapsid and Spike proteins respectively.

negative, and 3 non-COVID/healthy controls (samples collected before COVID pandemic). The workflow and quality control checks for TSQ Altis mass spectrometer used for MRM assay is shown in Figures S2 and S3. The representative MRM spectra for three viral peptides (Replicase 1ab, Nucleocapsid and Spike proteins) as observed in COVID-19-positive swab samples are shown in Figure 1C. The final MRM list had 4 viral peptides (IQDSLSSTASALGK, NTQEVFAQVK, FLPFQQFGR, and QIAPGQTGK) belonging to SARS-CoV-2 Spike glycoprotein (P0DTC2); 4 peptides belonging to SARS-CoV-2 nucleoprotein (P0DTC9)(NSTPGSSR, GGSQASSR, DGIIWVATEGALNTPK, and RPQGLPNNTASWFTALTQHGK; 4 more peptides (NGSIHLYFDK, ELLQNGMNGR, DGHVETFYPK, DFMSLSEQLR) (P0DTD1) belonging to Replicase polyprotein 1ab (Table S5) and a SpikeTide (Schnatbaum et al., 2011) having sequence FEDGVLDPDYPR that was spiked into all these samples as an internal standard. Distinct MRM spectra were seen for all the selected viral peptides in the COVID samples but not in healthy controls (refer to the Skyline files in the data availability section). This study should be further validated on the larger cohort of COVID-19-positive and COVID-19-negative samples.

Alteration of host proteome in response to COVID-19 infection

To investigate the alteration in host proteome, mass spectrometry-based label-free quantification was performed on 68 patient's swab samples, extracted using three different solvents (ethanol, acetone, and isopropanol). These clinical samples were run in two batches, first batch consisted of 44 samples (26 COVID-19-positive, 11 recovered, and 7 negative) and the second batch consisted of 24 samples (11 severe and 13 non-severe). The mass spectrometry generated raw data sets were processed with MaxQuant software against the Human Swiss-Prot database, which includes a total of 20,353 proteins. The total number of proteins identified in all the groups is shown in Table S6. Around 164 common significant proteins were identified from the comparison between 3 groups (COVID-19 positive, negative, and recovered) (Figure S4). Interestingly, Partial least squares-discriminant analysis (PLSDA) indicated sample-ID 30 belonging to the recovered group, to be a part of cluster consisting of positive samples. This analysis indicates that the patient sample-ID 30 might have not recovered completely and thus was showing clustering with the COVID-19-positive samples (Figure 2A). The heatmap of top 25 proteins from 164 common significant proteins between 3 groups revealed that the protein angiotensinogen, complement component C3, T-complex protein 1 subunit delta and signal transducer and activator of transcription 1 (STAT1) was found to be highly upregulated in COVID-19-positive patients when compared with negative and recovered patients (Figure 2B). The two different comparisons of positive samples with recovered samples and true negative samples identified several significant proteins and they were also able to form a distinct cluster based on the dysregulated proteins (Figure S5). These results show that the panel of significant proteins was able to distinguish the three clusters even on different group comparisons.

Further, the batch of 24 COVID-19-positive samples which includes 11 non-severe and 13 severe patient samples were separately analyzed in MaxQuant using the same parameters discussed above. One sample of 24 samples was found to be an outlier due to its low correlation coefficient and hence removed. From the heatmap of top 25 proteins, 6 significant proteins were identified to be differently expressed in COVID-19 severe when compared to COVID-19 non-severe patient samples (Figure 2C). A few of the significant proteins identified were lipocalin-1, myeloid-derived growth factor, fibrinogen silencer-binding protein, and Cofilin-1. The partial least square-discriminant analysis of 23 samples was performed which found to show a clear distinction between the severe and non-severe cohorts (Figure 2D). VIP score based on PLSDA has been used to select top 15 features and 7 of these 15 features were found to be significantly differentially expressed between non-severe and severe cohort (Figure S6).

The proteomic analysis of COVID-19-positive and COVID-19-negative patients revealed few links with the clinical parameters. We found the protein lactate dehydrogenase A (LDH-A), lactate dehydrogenase B (LDH-B), STAT1 (a key regulator of IL-6), hemoglobin subunit alpha (HBA) and hemoglobin subunit beta (HBB) protein (hemoglobin subunits) to be upregulated in COVID-19-positive patients when compared to COVID-19-negative patients (Figure 2E). These results correlated with the symptoms of the patients and hence the significant proteins identified from positive and negative comparison were taken forward for the validation assay and pathway analysis.

Targeted MRM assay and clinical validation of the host proteins for COVID-19 prognosis

A few proteins such as L-lactate dehydrogenase A chain (P00338), L-lactate dehydrogenase B chain (P07195), Ferritin heavy chain (P02794), Ferritin light chain (P02792), aspartate aminotransferase,

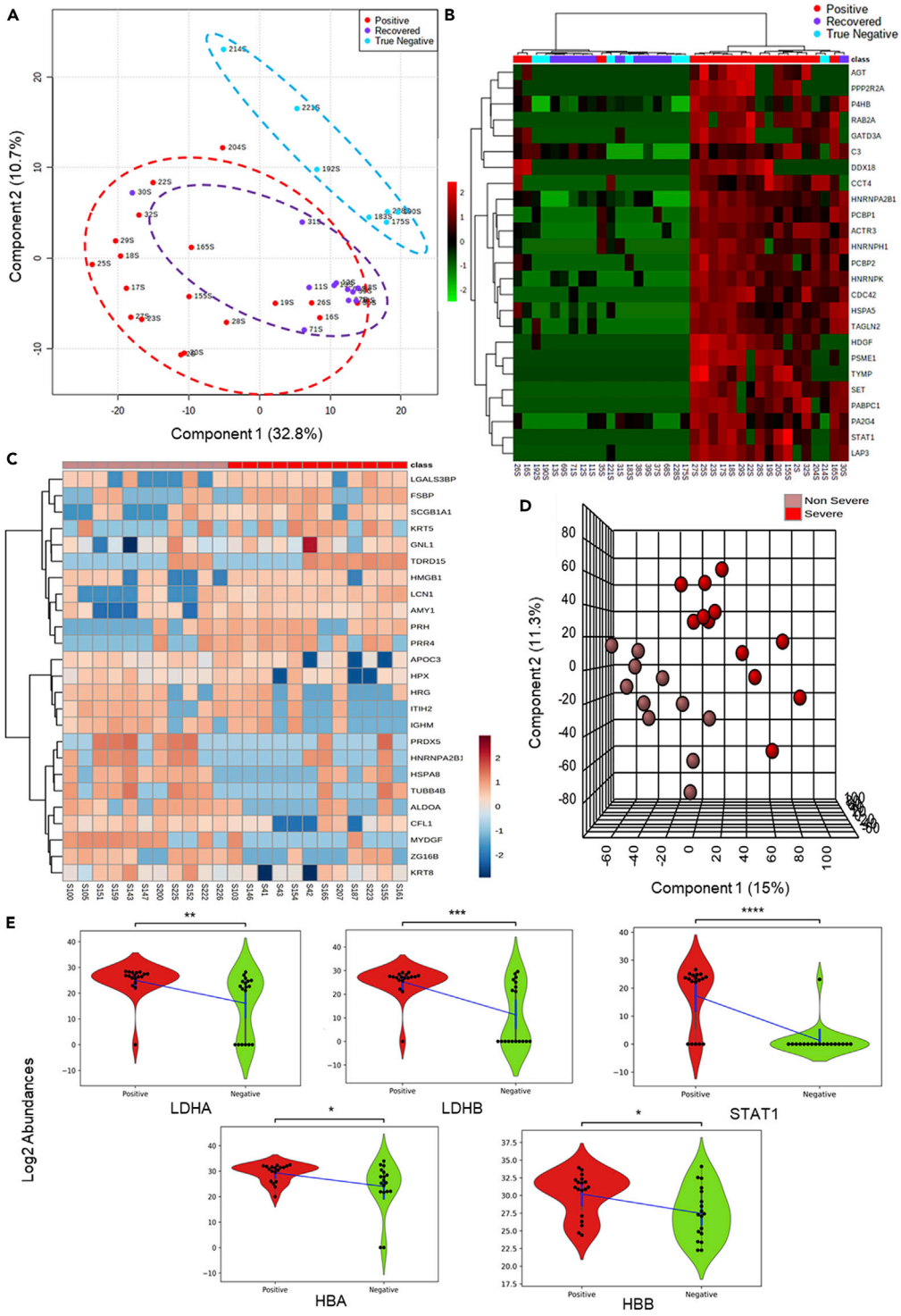


Figure 2. Alteration of host proteome in response to COVID-19 infection

To study the host proteome profile the mass spectrometry generated raw datasets were processed with MaxQuant software against the Human Swiss-Prot database.

(A) Map of the Segregation of Positive, True Negative, and Recovered samples using PLSDA. Analysis of 18 positive, 11 recovered and 7 true negative samples showed segregation into three clusters. True negative clusters distinctly classified from recovered and positive samples. Sample 30, a recovered sample found to be placed within the positive sample cluster.

Figure 2. *Continued*

(B) HeatMap of Positive, True Negative and Recovered samples. A list of 164 significant proteins found to be common between Positive vs Negative, Positive vs Recovered and Positive vs True Negative has been used to draw a hierarchical clustering based heatmap. The figure depicts the top 25 significant proteins found to segregate the groups using the ward clustering algorithm.

(C) The unsupervised heatmap of 25 significant protein shows a perturbation between the non-severe and severe group.

(D) Clustering analysis to segregate the Positive severe and non-severe samples. 12 severe positives and 11 non-severe positive samples were found to be segregated into two clusters in PLSDA.

(E) Violin plot showing the expressional difference of clinical protein markers of COVID diagnosis. Violin plot of LDH-A and LDH-B – the subunits of Lactate dehydrogenase found to be significantly upregulated in COVID-19-positive samples when compared with COVID-19-negative samples. STAT1, a key regulator of interleukin also found to be significantly upregulated in COVID-19-positive, having biological connections with D-dimers and creatine phosphokinase which are clinical markers. HBA and HBB subunits of hemoglobin are also found to be significantly upregulated in the COVID-19-positive samples. (Unpaired Welch's T test; ns: 5.00e-02 < p <= 1.00e+00; *: 1.00e-02 < p <= 5.00e-02; **: 1.00e-03 < p <= 1.00e-02; ***: 1.00e-04 <p <= 1.00e-03; ****: p <= 1.00e-04).

mitochondrial (P00505), aspartate aminotransferase, cytoplasmic (P17174) and albumin (P02768) which were found to be upregulated in COVID-19-positive patients when compared to COVID-19-negative patients in the LFQ data were selected and used for a targeted MRM study. Also, we checked for the clinical markers such as interleukin – 6 (IL-6), C-reactive protein, as well as proinflammatory cytokines such as tumor necrosis factor (TNF-α) which we could not identify in LFQ data but is clinically significant using targeted MRM assay. The following analysis was performed on the swab samples of 6 COVID-19-negative and 16 COVID-19-positive patients (consisting of 8 severe and 8 mild samples) were run on a triple quadrupole mass spectrometer coupled to an HPLC. The peak shapes of representative peptides for clinically significant proteins validated using MRM, as seen in Skyline are shown in Figures 3A and S7. Table S7 shows the list of MRM identified peptides between 16 COVID-19-positive and 6 COVID-19-negative samples, used for validation. Intensities of the peptide for proteins interleukin-6 (IL-6), L-lactate dehydrogenase A chain (LDH-A) and aspartate aminotransferase, and cytoplasmic (GOT1) were observed to be statistically significant between COVID-19-positive and COVID-19-negative patient samples ($p < 0.05$; Fold change > 1.5 at a confidence interval of 99%). A refined list of transitions that provided validated clinical marker proteins using MRM and gave statistically significant peptides (adjusted p value < 0.05; fold change > 1.5) between COVID-19-positive and COVID-19-negative swab samples are shown in Figure 3B and Table S8.

Molecular pathway perturbations with the progression of COVID-19 infection

A list of 452 significant proteins from COVID-19-positive samples, when compared to COVID-19-negative samples, were taken forward for Gene Ontology (GO) enrichment analysis using Metascape. The most prominent pathways identified were neutrophil degranulation, platelet degranulation, interleukin-12 signaling pathways, mRNA translation of proteins, and co-factor metabolomic process (Figures 4 and S8). The abundance of few proteins mapping into the prominent pathways such as peroxiredoxin-6 (PRDX6), macrophage migration inhibitory factor (MIF), glutathione S-transferase omega-1, 40S ribosomal protein S28 (RPS28), eukaryotic translation initiation factor 4H (EIF4H) and elongation factor 2 (EEF2) for COVID-19-positive and COVID-19-negative controls are shown in the violin plot (Figure 4).

Although we could not identify some of the proinflammatory cytokines in LFQ analysis, the GO enrichment analysis did reveal several proteins that might be upregulated in response to the stimulus from proinflammatory cytokines or might be inducing its release. For example, proteins such as RPS3 with the role in biological process response to TNF-α agonist, MIF protein involved in the interleukin-12 family signaling capable of activating T cells to produce proinflammatory cytokines such as TNF-α, interleukin (IL) 1β, IL-2, IL-6, IL-8, and interferon-γ, leading to inflammatory responses. Thus, the GO enrichment analysis showed several proteins belonging to the IL-6, IL-1, IL-23, IL-12, and IL-7 signaling pathway which plays an important role in the COVID-19 progression to severity (Table S9). We also identified ILK pathways such as interleukin-12 family and interleukin-1 signaling which belongs from cytokine signaling in the immune system pathway. We also found STAT1, a key regulator of IL-6, significantly upregulated in most of the COVID-19-positive when compared with the negative sample group (Figure S8).

Molecular docking of drugs with host proteins

We used the *in silico* approach of molecular docking to analyze the inhibition mechanism of few drugs on pathways related to immune response. In our study, we have docked the identified proteins from

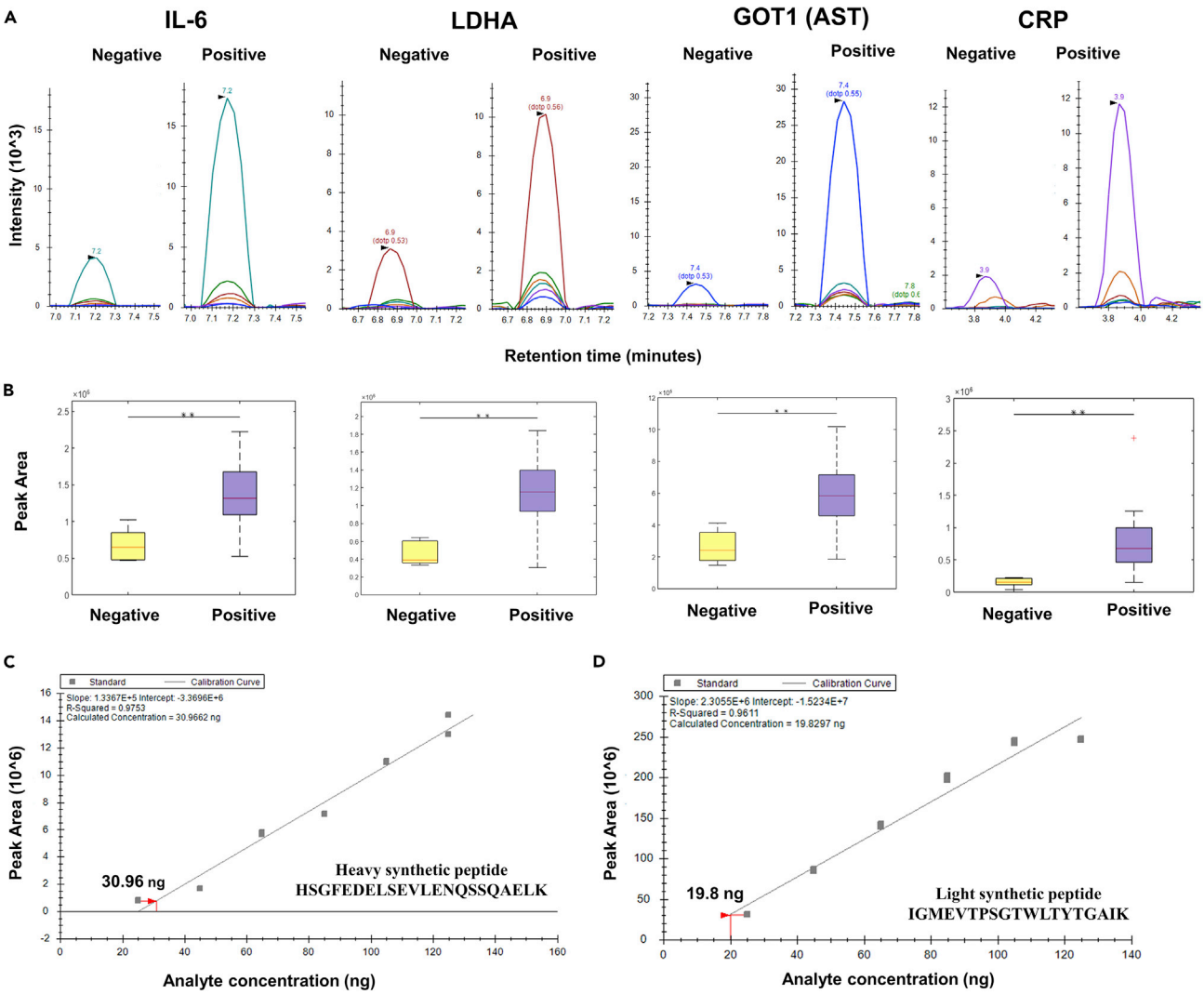


Figure 3. MRM validation of the host proteins upregulated in COVID-19-positive samples

The MRM analysis was performed on 6 COVID-negative and 16 COVID-positive patient swab samples. 1 μg peptide from each sample was injected into a Vanquish HPLC coupled to a TSQ Altis™ triple quadrupole and run against a transition list of peptides belonging to important clinical markers. The list of transitions was prepared for unique peptides of these selected proteins using Skyline (Version 20.2.1.286). This list included a spiked-in synthetic heavy peptide (THCLYTHVCDAIK) used for monitoring the consistency of the mass spectrometry runs. For identification of the sensitivity of the peptide detection, we performed serial dilution of two crude synthetic peptides (heavy and light). The concentration of peptide was calculated using the Scopes method from its O.D. value at 205 nm and 280 nm. Different concentrations of the peptides starting from 25 to 125 ng were run in TSQ Altis™ Triple Quadrupole Mass Spectrometer.

(A) Peak shapes of representative peptides for clinical marker proteins validated using MRM. The statistically significant change in the expression of the proteins was observed between COVID-19-positive and COVID-19-negative patient samples (T-test, $^{**}p < 0.05$; Fold change > 1.5 at a confidence interval of 99% - determined by Skyline).

(B) Box plots of representative peptides for clinical marker proteins validated using MRM.

(C) Standard curve of heavy synthetic peptide HSGFEDELSEVLENQSSQAELK. The crude heavy synthetic peptide was diluted in the range of 25 to 125 ng concentration. The standard curve for this peptide was plotted using the peak area against the concentration of the peptide. The intensity of the peak area was proportional to the amount of the synthetic peptide. The lowest amount of synthetic peptide detected was at 30.9 ng.

(D) Standard curve of light synthetic peptide HSGFEDELSEVLENQSSQAELK. The crude light synthetic peptide was diluted in the range of 25 to 125 ng concentration. The standard curve for this peptide was plotted using the peak area against the concentration of the peptide. The intensity of the peak area was proportional to the amount of the synthetic peptide. The lowest amount of synthetic peptide detected was at 19.8 ng.

neutrophil degranulation, translation and interleukin pathway against 29 FDA-approved, 9 clinical, and 20 pre-clinical trial drugs (Table S10). The inhibitors well established in the literature (Gordon et al., 2020) were used as a positive control against these target proteins. The control inhibitor gives us a possible cutoff for

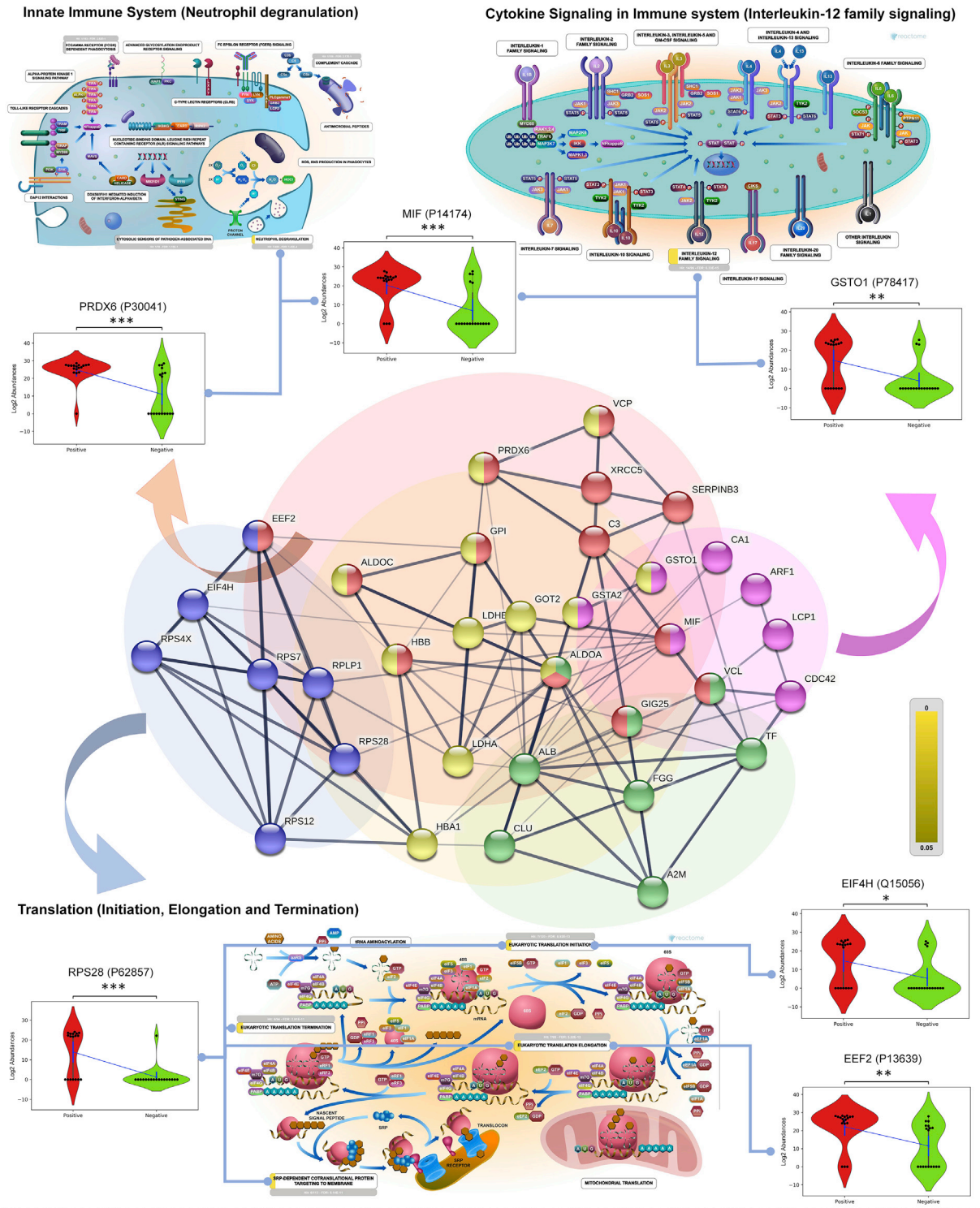

Innate Immune System (Neutrophil degranulation)
Cytokine Signaling in Immune system (Interleukin-12 family signaling)
MIF (P14174)
***
PRDX6 (P30041)
***
GSTO1 (P78417)
**
VCP
PRDX6
XRCC5
SERPINB3
EEF2
GPI
C3
CA1
ALDOC
GSTO1
ARF1
EIF4H
GOT2
GSTA2
LDHB
MIF
LCP1
RPS4X
HBB
RPS7
RPLP1
ALDOA
VCL
CDC42
GIG25
RPS28
LDHA
ALB
TF
FGG
RPS12
HBA1
CLU
A2M
0
0.05
Translation (Initiation, Elongation and Termination)
EIF4H (Q15056)
*
RPS28 (P62857)
***
EEF2 (P13639)
**
Log2 Abundances
Positive
Negative
Protein-Protein Interaction Network Annotation:
Neutrophil Degranulation
Cofactor Metabolic Process
Interleukin-12-mediated signaling Pathway
Platelet degranulation
Translation
COVID-19 Positive
COVID-19 Negative

Figure 4. Molecular pathway perturbations with the progression of COVID-19 infection

The figure represents the enriched GO biological processes with their co-expressed proteins in the form a bipartite network where few proteins has been shown in the form of violin plot for COVID-19-positive (high and low viral load patients) and COVID-19-negative controls. (Unpaired Welch's T-test, p-value annotation legends: ns: 5.00e-02 < p <= 1.00e+00; : 1.00e-02 < p <= 5.00e-02; *: 1.00e-03 < p <= 1.00e-02; *: 1.00e-04 < p <= 1.00e-03; **: p <= 1.00e-04). The representative images from the Reactome depicting the key pathways perturbed in the host are also shown.

the docking score. The selection of drugs for the protein was done by considering a couple of criteria. First, the binding energy of the drug was expected to be equal or higher than that of the control inhibitor. Second, the binding pocket of the drug was expected to be similar to the control drug. The proteins involved in the neutrophil degranulation, translation, and interleukin pathways are the mediators to distinguish between two cohorts (COVID-19-positive and COVID-19-negative) and hence were taken forward as potential drug targets to perform further computational studies. Based on our docking analysis, we observed that 1 pre-clinical trial drug binds to protein involved in the translation pathway (Figure S9) and 7 FDA-approved drugs to target neutrophil degranulation pathway (Figure S10); and 2 FDA-approved and 2 pre-clinical drugs for interleukin pathway (Figure 5). One FDA-approved drug loratadine (ZINC537931) and another, currently in pre-clinical studies, UCPH-101 (ZINC000040914195) was found to bind three major proteins (ADP-ribosylation factor 1, carbonic anhydrase 1 and MIF) from the interleukin pathway. Both, loratadine and UCPH-101 bind to the control ligand-binding pocket with a binding affinity more negative than control (<-8 Kcal/mol) (Figure 5). Similarly, few more drugs were found to bind proteins belonging to eukaryotic translation, as well as neutrophil degranulation pathway.

DISCUSSION

We have performed a comprehensive proteomics study from an Indian hot spot COVID region, Mumbai, displaying the potential of mass spectrometry-based approaches for severity progression of COVID-19 and clinical translation. To obtain highly efficient peptide enrichment from COVID-19-infected swab samples, we precipitated protein from swab using three different organic solvents namely, ethanol, isopropanol, and acetone, as well as prepared a pool of peptides from all three conditions. A few MS-based studies have relied on a complex sample preparation protocol, using TCA-acetone and in-gel digestion or a single organic solvent for the extraction of proteins identifying only a few unique peptides (Gouveia et al., 2020; Nikolaev et al., 2020). However, we observed the pool of peptides from three methods identified the maximum number of viral peptides using Orbitrap Fusion mass spectrometer. The MRM-based assay could specifically detect SARS-CoV-2 peptides from spike glycoprotein, nucleoprotein, and replicase polyprotein 1ab protein. These peptides showed a distinct elution profile and were eluted between 0.5 to 8 minutes. Thus, MRM-based assay confirmed the presence of SARS-CoV-2 peptides in COVID-19-positive samples as compared to healthy controls.

Currently, several researcher groups have studied the host proteome alteration in response to COVID-19 infection (Shen et al., 2020; Overmyer et al., 2020). However, still panel of host proteins to be routinely monitored in the patients or the pathways to be targeted are not yet identified and validated. Thus, to identify the potential prognostic markers and therapeutic targets, we performed a deep proteome profiling of COVID-19 respiratory samples based on label-free quantification using a high-resolution mass spectrometer. The nasopharynx being the primary site of the viral entry, the nasopharyngeal swab might be the better sample to study the complex molecular and immune response events in the infected host. From over 3749 host proteins, we could identify around 164 significant proteins in COVID-19-positive when compared to COVID-19-negative and recovered patients. Using targeted MRM assay, we could detect a statistically significant change in the intensity of the peptide of proteins such as interleukin-6, L-lactate dehydrogenase A chain, L-lactate dehydrogenase B chain, Ferritin heavy chain, Ferritin light chain, aspartate aminotransferase (mitochondrial), aspartate aminotransferase (cytoplasmic), C-reactive protein, and proinflammatory cytokines, TNF-α in COVID-19-positive patients when compared to COVID-19-negative patients. However, targeted assays based on these host proteins on a large cohort of patients using longitudinal samples are required for the successful clinical translation.

Currently, we have limited understanding of molecular pathways altered during the SARS-CoV-2 infection (Shen et al., 2020; Overmyer et al., 2020). Using enrichment analysis, we could identify proteins involved in neutrophil degranulation, platelet degranulation, interleukin-12 signaling related pathway, translational mechanism, and co-factor metabolomic process to be key GO enriched pathways altered in the COVID-19-positive patients. Recent studies state that there is an evidence of increased peripheral neutrophil-to-lymphocyte-ratio in the case of severe COVID-19 patients (Zheng et al., 2020), suggesting the fact

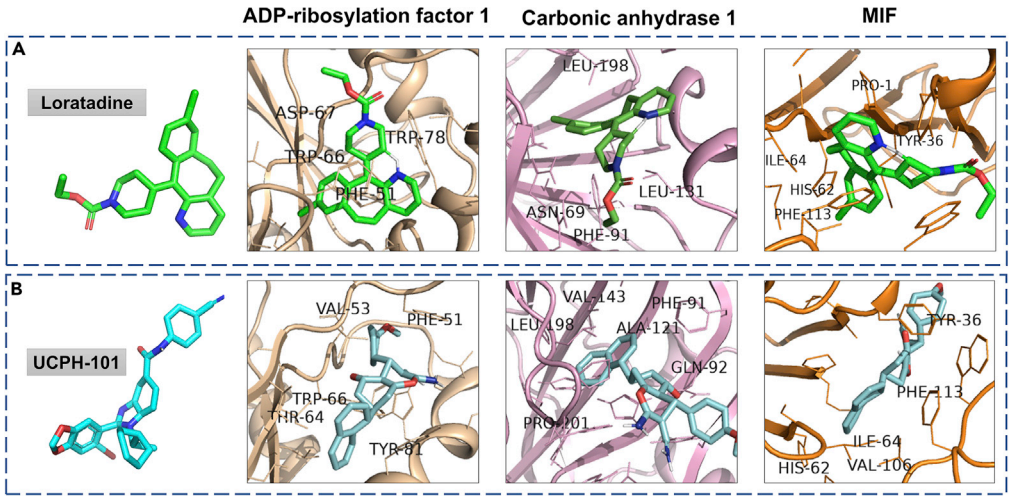


Figure 5. Molecular docking studies of drugs binding to the host proteins involved in interleukin-12 signaling pathways

Autodock Vina (1.1.2) was used to identify the drugs binding to the host proteins involved in the neutrophil degranulation, translation, and interleukin pathway. A screening of 29 FDA-approved, 9 clinical, and 20 pre-clinical trial drugs against the host protein identified several potential drug candidates targeting the interleukin-12 signaling pathway.

(A) shows the drug Loratadine (green) docked with three proteins from the interleukin pathway; ADP-ribosylation factor 1 (binding energy or BE -9.8 kcal/mol), carbonic anhydrase 1 (BE -8.3 kcal/mol) and macrophage migration inhibitory factor (MIF) (BE -8 kcal/mol).

(B) shows the same proteins docked with the drug UCPH-101 (cyan); ADP-ribosylation factor 1 (BE -9.9 Kcal/mol), carbonic anhydrase 1 (BE -8.5 kcal/mol) and MIF (BE -8.8 kcal/mol).

Both of the drugs bind to all three proteins with negative binding energy greater than their respective control inhibitor. The interacting amino acid residues, which are present on the ligand-binding pocket are labeled. Almost all of the interacting residues belong to hydrophobic amino acids.

that neutrophil degranulation is one of the major events that modulate the immune system post-infection. We identified several host proteins such as complement C3, glucose-6-phosphate isomerase, SERPINB3, peroxiredoxin-6, and transitional endoplasmic reticulum ATPase (VCP) mapping to neutrophil degranulation pathway. In recent studies, it has also been argued that neutrophils play a dual role during pathogen attack. They help our immune system to get rid of viruses or bacteria by releasing granule-derived mediators and activating complement, but accumulation or overreaction due to these mediators may lead to hyperinfection, tissue injury, or septic shock (Lacy, 2006). However, it is quite unclear whether SARS-CoV-2 directly targets the neutrophil degranulation pathway or is just a consequence of severe immune complications of SARS-CoV-2 infection. The interleukin IL-6 is a hallmark of the cytokine storm observed in the COVID-19 severe patients (Sun, 2020). We also identified several proteins such as STAT1 (a key regulator of IL-6) belonging to the interleukin signaling pathway upregulated in COVID-19-positive patients as compared to the COVID-19-negative patients (Matsuyama, 2020).

We also identified several host proteins belonging to platelet activation, aggregation, and degranulation pathways such as fibrinogen gamma chain, profilin-1, albumin, serotransferrin, and alpha-2-macroglobulin. Platelets function as exocytotic cells, secreting a plethora of effector molecules at sites of vascular injury. The increased fibrin formation and breakdown correlated with the high level of D-dimers observed in the COVID-19 patients with the worst outcomes (Tang, 2020). The increasing levels of FGG in severe cases might be due to liver injury, impairing hepatic fibrinogen secretion with acquired fibrinogen storage disease (Fraga, 2020).

Translation event inside a eukaryote system requires several hundred proteins to act in sync. It becomes impossible for a virus to carry that much information in its genome thereby increasing its dependence on the host to make viral proteins. To achieve that, viral mRNAs compete with the host mRNAs to access limited cellular translational resources (Katze et al., 1985). Almost all coronavirus families initiate their translation in a 'cap-dependent pathway' with the help of SARS-CoV-2 N protein (Nakagawa et al., 2016). We identified several proteins such as EEF2 (P13639), ATP-dependent RNA helicase DDX18 (Q9NVP1), 60S

acidic RP P1 (P05386), 40S RP S4, X isoform (P62701), 40S RP S12 (P25398), elongation factor 1-gamma (Q15056), EIF4H (Q15056) belonging to the translational pathway. One of the proteins EIF4H was known to stimulate the RNA helicase activity of EIF4A in the translation initiation complex (Cencic et al., 2011). Krogan et al. reported that eIF4H, which interacts with NSP9 SARS-CoV-2 protein, is a partner of eIF4A, and observed a strong antiviral effect after treatment with the eIF4A inhibitor Zotatifin (Gordon et al., 2020). After infection by some viruses, the translation of the host mRNA is often suppressed, whereas translation initiation of RPs mRNA might increase and persist late (Shuo, 2019). We observed significant upregulation of 40S RP and 60S RP which might be a result of the virus-host interaction for enhancing the production of few RPs to maintain viral propagation. These results reveal that the composition of ribosome proteins may be different between COVID-19-infected hosts and in uninfected populations.

Interleukin-12 signaling pathway plays an important role in the coordination of the innate and adaptive immune response (Liu, 2005). We identified several host proteins such as plastin-2, MIF, glutathione S-transferase A2, carbonic anhydrase 1 (CA1), and ADP-ribosylation factor 1 (ARF1) involved in the interleukin-12 signaling pathways. The MIF secreted in response to viral infection can activate the T cells and macrophages to produce inflammatory cytokines, including TNF-α, IL-6 and others, leading to elevated inflammatory response also called "cytokine storm" (Vandenbark, 2020). In our study, we have detected high levels of MIF, TNF-α and IL-6 in COVID-19 patients which correlates with the symptoms of acute respiratory distress syndrome (ARDS) in severe patients. Thus, targeting these host proteins involved in the interleukin-12 signaling or inflammatory response might protect the patients from developing ARDS.

Further, we investigated the binding of the drug to the host proteins involved in the molecular pathways altered in the infected host. An extensive docking study of 29 FDA-approved, 9 clinical, and 20 pre-clinical trial drugs was performed to target neutrophil degranulation, translation, and interleukin-12 signaling pathways. We identified Loratadine and UCPH-101 as a potential drug to target the host proteins such as MIF, CA1, and ARF1 involved in the interleukin-12 signaling. Loratadine is an anti-histamine used for the treatment of allergies and is known for reducing the excessive cytokine proinflammatory storm (Canonica, 2011). UCPH-101 an inhibitor of excitatory amino acid transporter subtype 1 (EAAT1), is also shown to bind to these host proteins and thus might play an important role in blocking inflammatory response (Abrahamsen, 2013). We also identified several drugs and small molecules inhibiting the proteins involved in the translational and neutrophil degranulation pathways. These small molecules and drug candidates should further be validated using *in vitro* human cell line model.

In conclusion, our data emphasize that nasopharyngeal swab respiratory samples, which are routinely collected for the COVID-19 RT-PCR testing; could also be used for mass spectrometry-based detection of host proteins. Further, we have validated these proteins using MRM-based mass spectrometry assay. The significant proteins involved in the translational and neutrophil degranulation pathways might be potential targets for the COVID-19 therapeutics. This study opens up new opportunity for the researchers to understand the alteration of the host response to SARS-CoV-2 in the Indian population.

Limitations of the study

Considering the high viral count, swab samples are clinically used for the detection of COVID-19 by RT-PCR, however, using these samples for routine monitoring of biomarkers is a challenge. Swab samples being highly infectious, in the clinical settings most of the biochemical and immunoassay are designed for application using blood-based sample. The VTM components in the swab samples will also interfere with the immunoassay, thus might not be suitable for the clinical application. However, suitable validation strategies could be developed directly from swab samples and should be validated on large cohort of patients before the identified biomarkers could be taken forward for clinical translation.

Resource availability

Lead contact

Further information and request for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Sanjeeva Srivastava (sanjeeva@iitb.ac.in).

Materials availability

This study did not generate any new unique reagents and/or materials.

Data and code availability

All proteomics data associated with this study are present in the manuscript or the Supplementary information. The accession number for the raw MS data and search output files for proteomics data sets depositeded to the Proteome Xchange Consortium via the PRIDE partner repository is "PRIDE: PXD020580" and "PRIDE: PXD023016" (Link: https://www.ebi.ac.uk/pride/archive/loginreviewer20505@ebi.ac.uk. https://www.ebi.ac.uk/pride/archive/loginreviewer_pxd023016@ebi.ac.uk) The targeted proteomics data is deposited in the Panorama Public and can be accessed through this link: "Panorama Public: https://panoramaweb.org/COVID_Swab_MRM.url" panorama+srivastava@proteinms.net. The additional supplemental items are available from "Mendeley Data: https://doi.org/10.17632/rnfn3vhg63.1". The present research did not use any new codes.

METHODS

All methods can be found in the accompanying Transparent methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2021.102135.

ACKNOWLEDGMENTS

The active support from Prof. Ambarish Kunwar from the Department of Biosciences & Bioengineering to fabricate UV transport device for sample transport and Prof. Anirban Banerjee for the BSL-2 biosafety aspects is gratefully acknowledged. We are also grateful to Dr. Jaishree Garhyan, JNU, New Delhi for discussion on biosafety guidelines and Dr. Arup Acharjee, BHU, Varanasi for the critical discussion on the COVID project. The authors thank Kasturba Hospital for Infectious Diseases for sample collection and sharing sample related information. The study was supported through Science and Engineering Research Board (SERB), Department of Science & Technology, Ministry of Science and Technology, MinisGovernment of India (SB/S1/Covid-2/2020), and a special COVID seed grant (RD/0520-IRCCHC0-006) from IRCC, IIT Bombay to SS. MASSFIIT (Mass Spectrometry Facility, IIT Bombay) from the Department of Biotechnology (BT/PR13114/INF/22/206/2015) is gratefully acknowledged for MS-based proteomics work. The authors also acknowledge the Thermo Fisher Scientific engineers and application scientists for their support to our MASSFIIT facility during the extreme lock-down time. The authors thank the PRIDE team for helping with the mass spectrometric data downloading and deposit in ProteomeXChange/PRIDE and Reactome.org for pathway related figures. A.B acknowledge CSIR-SRF funding, M.C and K.S thank IIT Bombay-IPDF funding.

AUTHOR CONTRIBUTIONS

Concept and design of the study were carried out by S.S., J.S., and O.S. Sample collection was done by S.A., R.B, V.P, and A.S. and sample preparation was done by R.B., S.G., A.B., K.S., and M.C. Mass spectrometry analysis, and method optimization was done by A.B, S.G, S.S. MRM (multiple reaction monitoring) were done by S.G, S.S, M.G.J.P, Al. S, and J.R. Statistical Data Analysis, and Data Visualization were done by D.B., A.S., A.V., M.C., K.S., and S.S. Docking study was performed by A.B. Data submission was done by A.S. Manuscript writing and Review by S.S., K.S., M.C., A.B., S.G., R.B., D.B., A.V., A.S., and A.M.

DECLARATION OF INTERESTS

The authors have filed an Indian patent related to this work "Protein markers and method for prognosis of COVID-19 in individuals". (Application number: 202021034688). The authors declare no competing interests. "Proteomics-based method for detection of Coronavirus in a sample" (Application number 202021034687).

Received: September 22, 2020
Revised: December 17, 2020
Accepted: January 28, 2021
Published: March 19, 2021

REFERENCES

Abrahamsen, B., et al.; Schneider N, Erichsen MN (2013). Allosteric modulation of an excitatory amino acid transporter: the subtype-selective inhibitor UCPH-101 exerts sustained inhibition of EAAT1 through an intramonomeric site in the trimerization domain. J Neurosci *33* (3), 1068–1087.

Bock, J.O., and Ortea, I. (2020). Re-analysis of SARS-CoV-2-infected host cell proteomics time-course data by impact pathway analysis and network analysis: a potential link with inflammatory response. Aging (Albany. NY) *12*, 11277–11286.

Bojkova, D., Klann, K., Koch, B., Widera, M., Krause, D., Ciesek, S., Cinatl, J., and Münch, C. (2020). SARS-CoV-2 Infected Host Cell Proteomics Reveal Potential Therapy Targets. Nature *583* (7816), 469–472.

Canonica, G.W., et al.; Blaiss M (2011). Antihistaminic, Anti-Inflammatory, and Antiallergic Properties of the Nonsedating Second-Generation Antihistamine Desloratadine: A Review of the Evidence. World Allergy Organ J *4* (2), 47–53.

Cencic, R., Desforges, M., Hall, D.R., Kozakov, D., Du, Y., Min, J., Dingledine, R., Fu, H., Vajda, S., Talbot, P.J., et al. (2011). Blocking eIF4E-eIF4G interaction as a strategy to impair coronavirus replication. J. Virol. *85*, 6381–6389.

Chan, J.F.W., Yip, C.C.Y., To, K.K.W., Tang, T.H.C., Wong, S.C.Y., Leung, K.H., Fung, A.Y.F., Ng, A.C.K., Zou, Z., Tsoi, H.W., et al. (2020). Improved molecular diagnosis of COVID-19 by the novel, highly sensitive and specific COVID-19-RdRp/Hel real-time reverse transcription-PCR assay validated in vitro and with clinical specimens. J. Clin. Microbiol. *58*, e00310-20.

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., et al. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet *395*, 507–513.

Fraga, M.; Moradpour D, Artru F, Romailler E, Tschopp J, Schneider A, Chtioui H, Neerman-Arbez M, Casini A, Alberio L, Sempoux C (2020). Hepatocellular type II fibrinogen inclusions in a patient with severe COVID-19 and hepatitis. J Hepatol *73*, 967–970.

Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O'Meara, M.J., Rezelj, V.V., Guo, J.Z., Swaney, D.L., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature *583*, 459–468.

Gouveia, D., Miotello, G., Gallais, F., Gaillard, J.-C., Debroas, S., Bellanger, L., Lavigne, J.-P., Sotto, A., Grenga, L., Pible, O., et al. (2020). Proteotyping SARS-CoV-2 virus from nasopharyngeal swabs: a proof-of-concept focused on a 3 min mass spectrometry window. J. Proteome Res. *19*, 4407–4416.

Ihling, C., Taenzler, D., Hagemann, S., Kehlen, A., Huettelmaier, S., and Sinz, A. (2020). Mass spectrometric identification of SARS-CoV-2 proteins from gargle solution samples of COVID-19 patients. Journal of proteome research *19* (11), 4389–4392.

Katze, M.G., Detjen, B.M., Safer, B., and Krug, R.M. (1985). Translational control in influenza virus-infected cells. Virus Res. *3*, 39.

Kim, J.Y. (2020). Letter to the editor: case of the index patient who caused tertiary transmission of coronavirus disease 2019 in Korea: the application of lopinavir/ritonavir for the treatment of COVID-19 pneumonia monitored by quantitative RT-PCR. J. Korean Med. Sci. *35*, 1–4.

Lacy, P. (2006). Mechanisms of degranulation in neutrophils. Allergy Asthma Clin. Immunol. *2*, 98–108.

Liu, J., et al.; Cao S, Kim S (2005). Interleukin-12: an update on its immunological activities, signaling and regulation of gene expression. Curr Immunol Rev *1* (2), 119–137.

Matsuyama, Toshifumi, et al.; Kubli, S.P., Yoshinaga, S.K (2020). An aberrant STAT pathway is central to COVID-19. Cell Death Differ *27*, 3209–3225.

Matthay, M.A., Aldrich, J.M., and Gotts, J.E. (2020). Treatment for severe acute respiratory distress syndrome from COVID-19. Lancet Respir. Med. *8*, 433–434.

Messner, C.B., Demichev, V., Wendisch, D., Michalick, L., White, M., Freiwald, A., Textoris-Taube, K., Vernardis, S.I., Egger, A.S., Kreidl, M., et al. (2020). Ultra-high-throughput clinical proteomics reveals classifiers of COVID-19 infection. Cell Syst. 11–24.

Nakagawa, K., Lokugamage, K.G., and Makino, S. (2016). Viral and Cellular mRNA Translation in Coronavirus-Infected Cells (Elsevier Inc.).

Nikolaev, E.N., Indeykina, M.I., Brzhozovskiy, A.G., Bugrova, A.E., Kononikhin, A.S., Starodubtseva, N.L., Petrotchenko, E.V., Kovalev, G., Borchers, C.H., and Sukhikh, G.T. (2020). Mass Spectrometric detection of SARS-CoV-2 virus in scrapings of the epithelium of the nasopharynx of infected patients via Nucleocapsid N protein. J. Proteome Res. *19*, 4393–4397.

Overmyer, K.A., Shishkova, E., Miller, I.J., Balnis, J., Bernstein, M.N., Meyer, J.G., Quan, Q., Muehlbauer, L.K., Trujillo, E.A., He, Y., et al. (2020). Large-scale Multi-Omic Analysis of COVID-19 Severity. Cell Syst. *12*, 23–40.e7.

Schnatbaum, K., Zerweck, J., Nehmer, J., Wenschuh, H., Schutkowski, M., and Reimer, U. (2011). SpikeTidesTM—proteotypic peptides for large-scale MS-based proteomics. Nat. Methods *8*, i–ii.

Shen, B., Yi, X., Sun, Y., Bi, X., Du, J., Zhang, C., Quan, S., Zhang, F., Sun, R., Qian, L., et al. (2020). Proteomic and metabolomic characterization of COVID-19 patient sera. Cell, 59–72.

Shuo, Li, et al. (2019). Regulation of Ribosomal Proteins on Viral Infection. Cells *8(5)*, 508.

Sun, Xinjuan, et al.; Tianyuan Wang, Dayong Cai, Zhiwei Hu, Jin'an Chen, Hui Liao, Liming Zhi, Hongxia Wei, Zhihong Zhang, Yuying Qiu, Jing Wang, Aiping Wang (2020). Cytokine storm intervention in the early stages of COVID-19 pneumonia. Cytokine & Growth Factor Reviews *53*, 38–42.

Tang, N., Li, D., Wang, X., and Sun, Z. (2020). Abnormal coagulation parameters are associated with poor prognosis in patients with novel coronavirus pneumonia. J Thromb Haemost *18*, 844–847.

Tang, Y., Schmitz, J.E., Persing, D.H., and Stratton, C.W. (2020). Laboratory diagnosis of COVID-19: current issues and challenges. J. Clin. Microbiol. *58*, 1–9.

Vandenbark, A.A., et al.; Meza-Romero R, Offner H (2020). Surviving the storm: Dealing with COVID-19. Cellular immunology *354*, 104153.

Zheng, M., Gao, Y., Wang, G., Song, G., Liu, S., Sun, D., Xu, Y., and Tian, Z. (2020). Functional exhaustion of antiviral lymphocytes in COVID-19 patients. Cell. Mol. Immunol. *17*, 533–535.

iScience, Volume *24*

Supplemental Information

Proteomic investigation reveals dominant alterations of neutrophil degranulation and mRNA translation pathways in patients with COVID-19

Renuka Bankar, Kruthi Suvarna, Saicharan Ghantasala, Arghya Banerjee, Deeptarup Biswas, Manisha Choudhury, Viswanthram Palanivel, Akanksha Salkar, Ayushi Verma, Avinash Singh, Amrita Mukherjee, Medha Gayathri J. Pai, Jyotirmoy Roy, Alisha Srivastava, Apoorva Badaya, Sachee Agrawal, Om Shrivastav, Jayanthi Shastri, and Sanjeeva Srivastava

Supplemental Figures and Tables

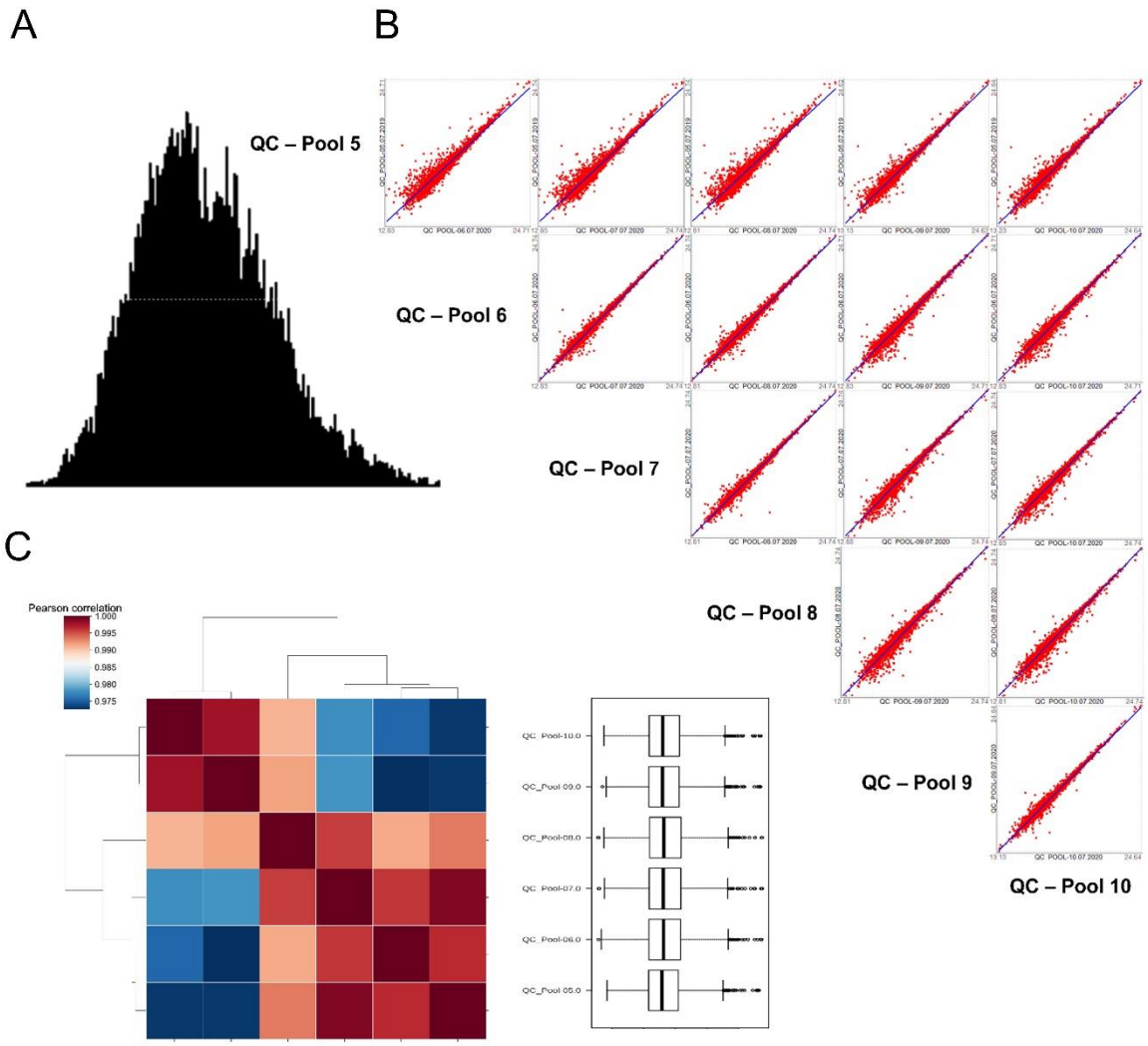


Figure S1. Quality check of the Mass Spectrometry run showing the distribution of all QC Pools, the pool-to-pool correlation and the Pearson r correlation between the pools and individual pool distribution respectively. Related to Figure 1 and transparent methods “sample preparation for LFQ”.

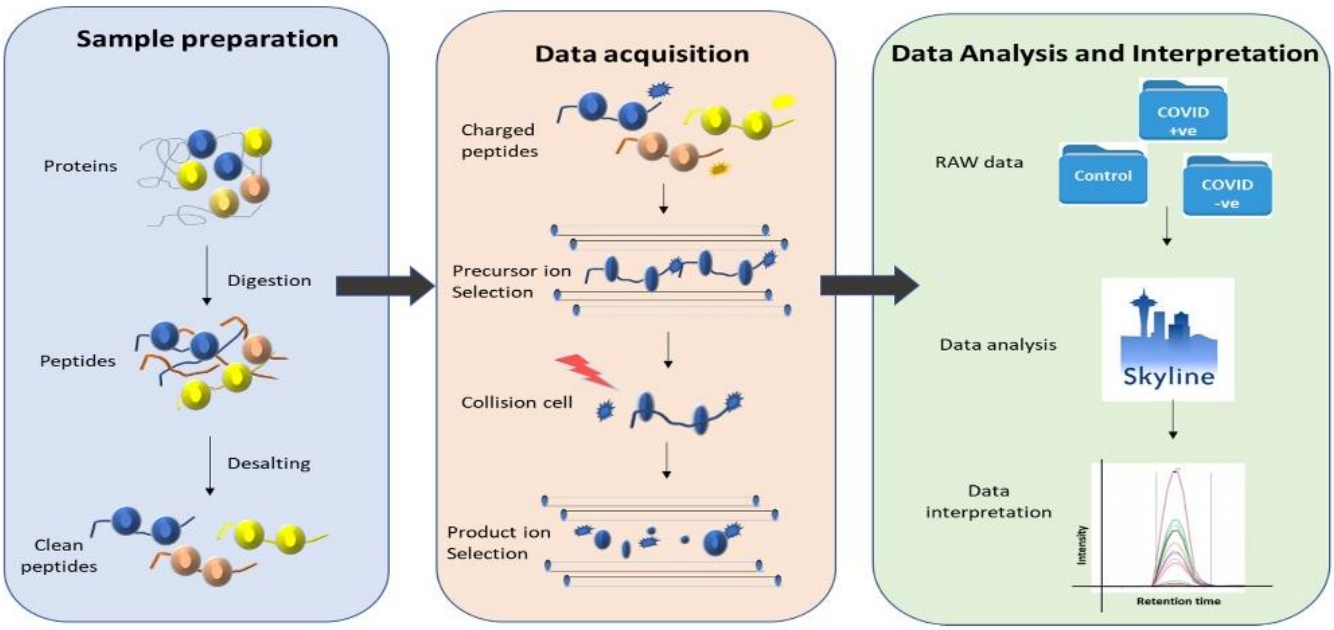


Figure S2. Workflow for Multiple Reaction Monitoring. Related to Figure 1C.

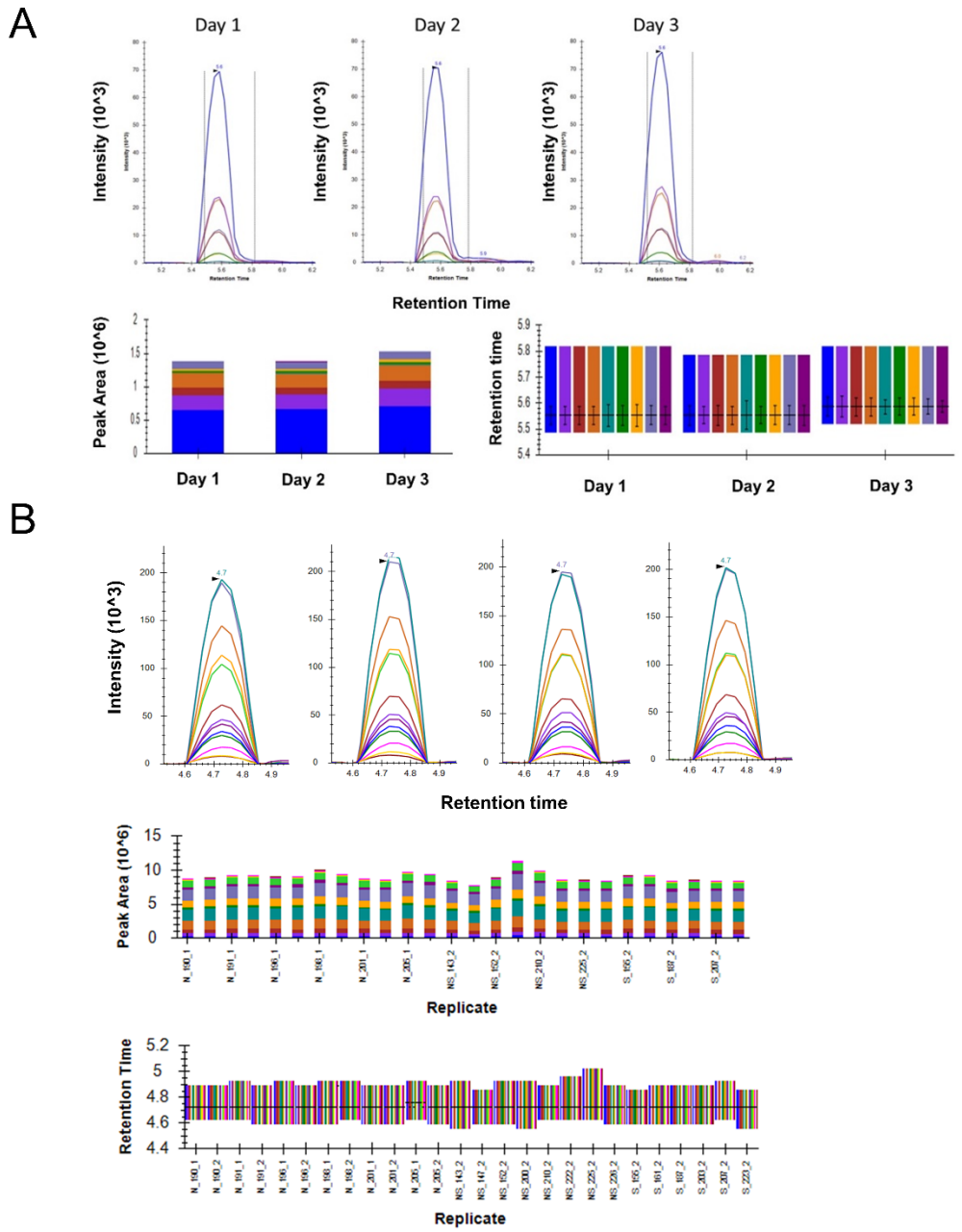


Figure S3. Quality check of TSQ Altis mass spectrometer used for MRM experiments showing the response of peak areas and retention times of the representative peptide for BSA and Spike-in peptide response. Related to Figure 1C.

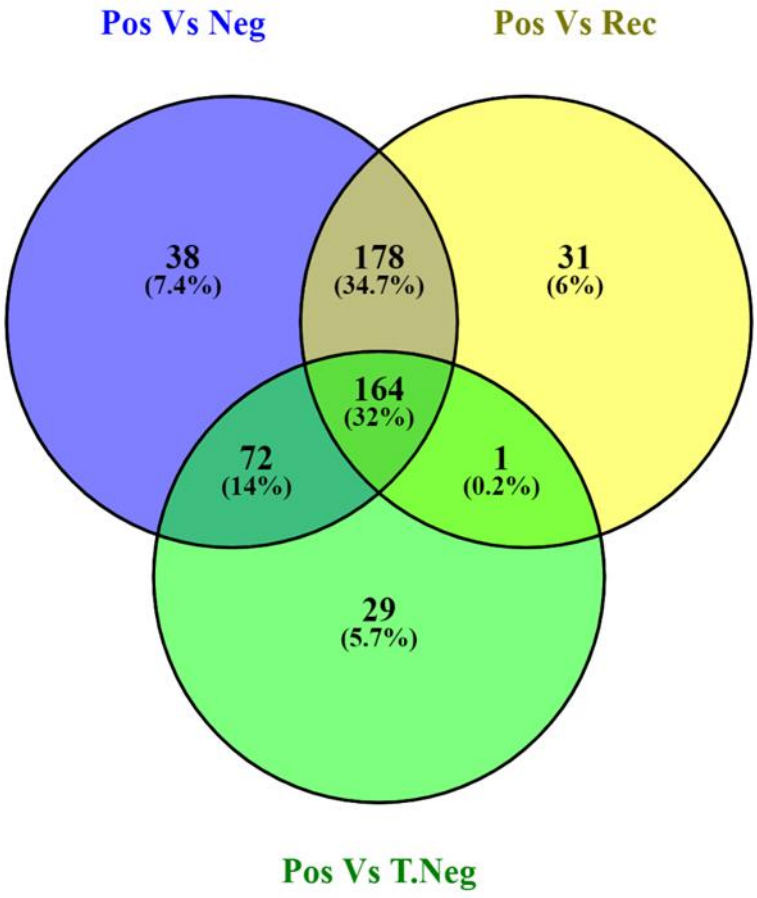


Figure S4. Venn diagram represents a list of 164 common significant proteins. Related to Figure 2A.

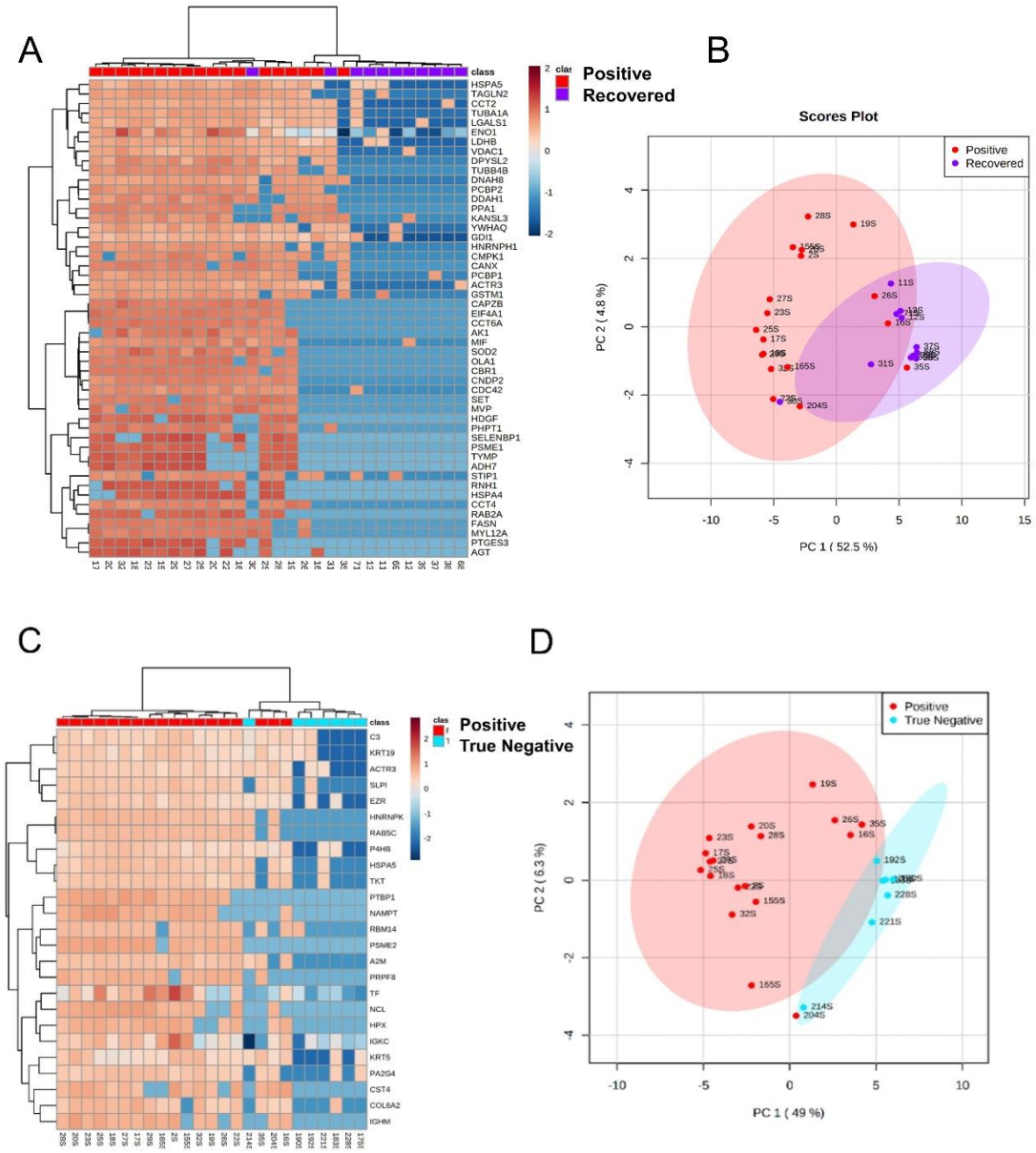


Figure S5. Unsupervised heat map and principal component analysis of differentially expressed proteins of COVID-19 Positive Vs True negative and recovered patients. Related to Figure 2B.

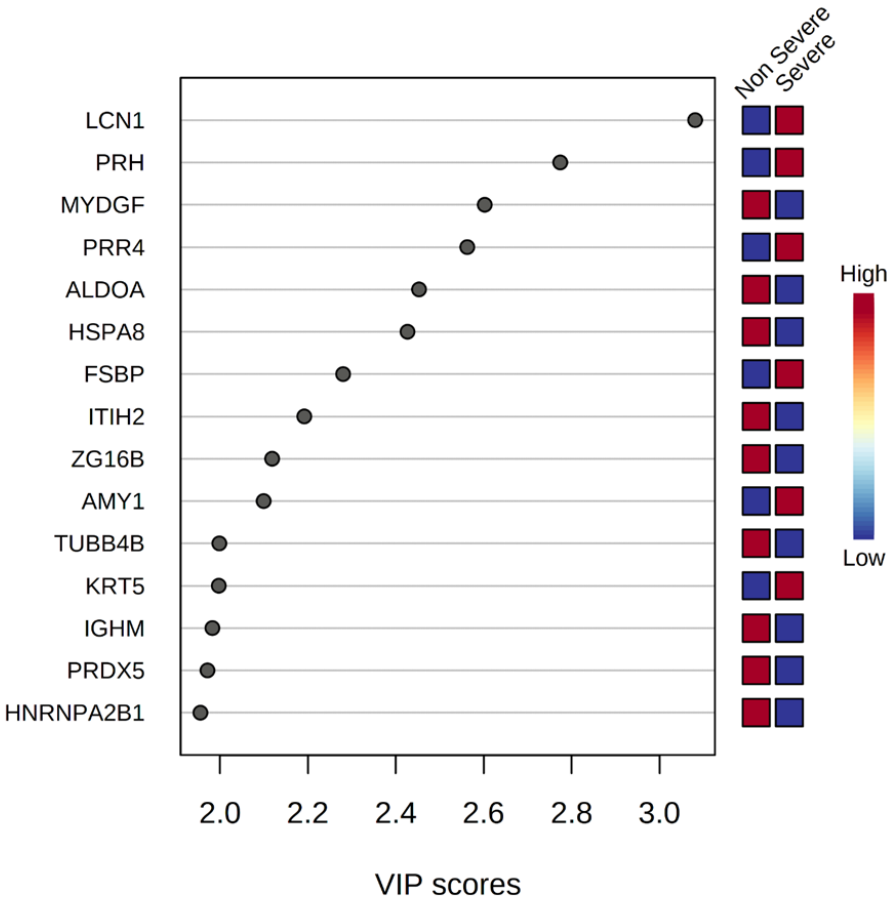


Figure S6. VIP score based PLSDA for COVID-19 severe when compared to COVID-19 non-severe. Related to Figure 2D.

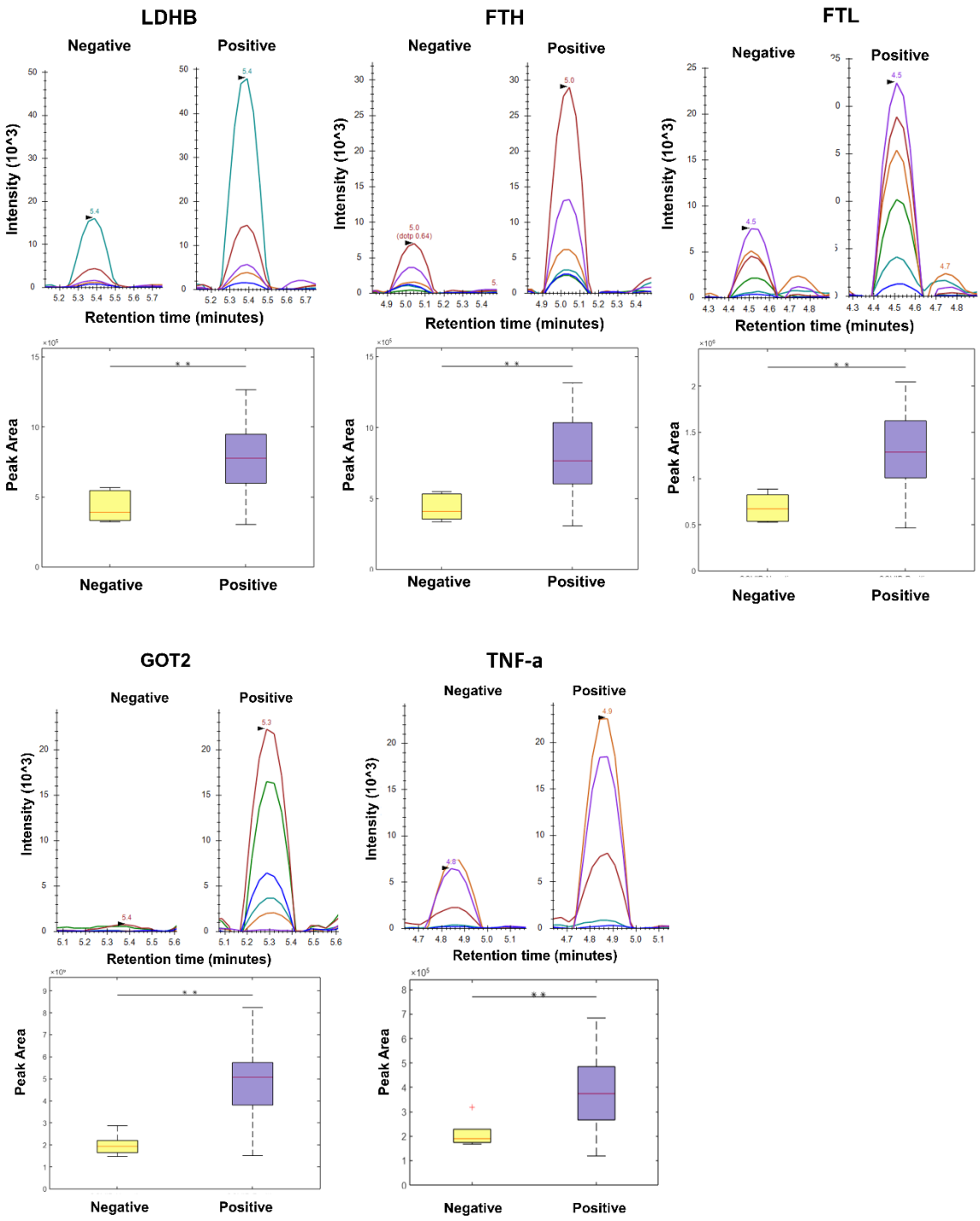


Figure S7. MRM analysis of clinically relevant markers performed on 6 COVID-19 negative and 16 COVID-19 positive samples. (T-test, $p < 0.05$; Fold change > 1.5 at a confidence interval of 99%). Related to Figure 3.

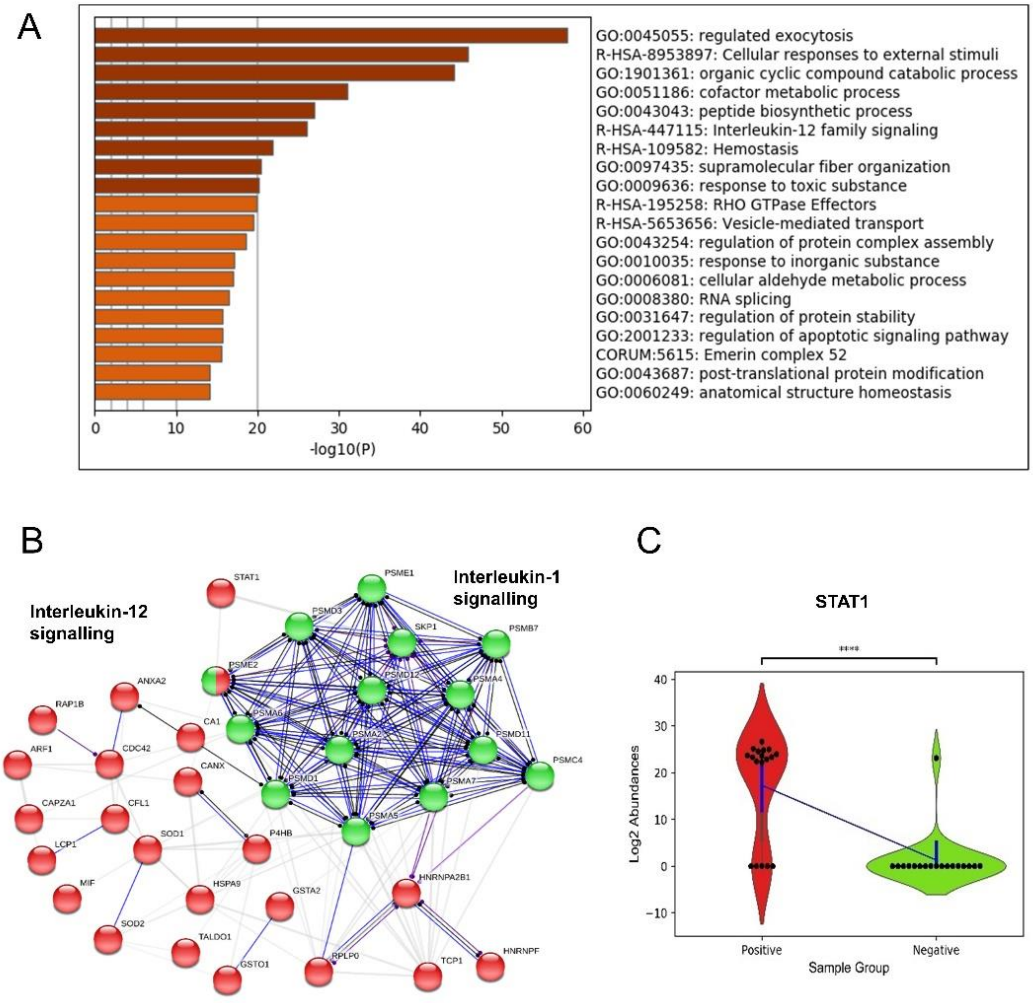


Figure S8. Gene ontology enrichment analysis of COVID-19 positive Vs negative using Metascape Pathway, protein-protein interaction network and STAT1 is shown in the form of violin plot (Unpaired Welch's T-test; ns: 5.00e-02 < p <= 1.00e+00; *: 1.00e-02 < p <= 5.00e-02; **: 1.00e-03 < p <= 1.00e-02; ***: 1.00e-04 <p <= 1.00e-03; ****: p <= 1.00e-04). Related to Figure 4.

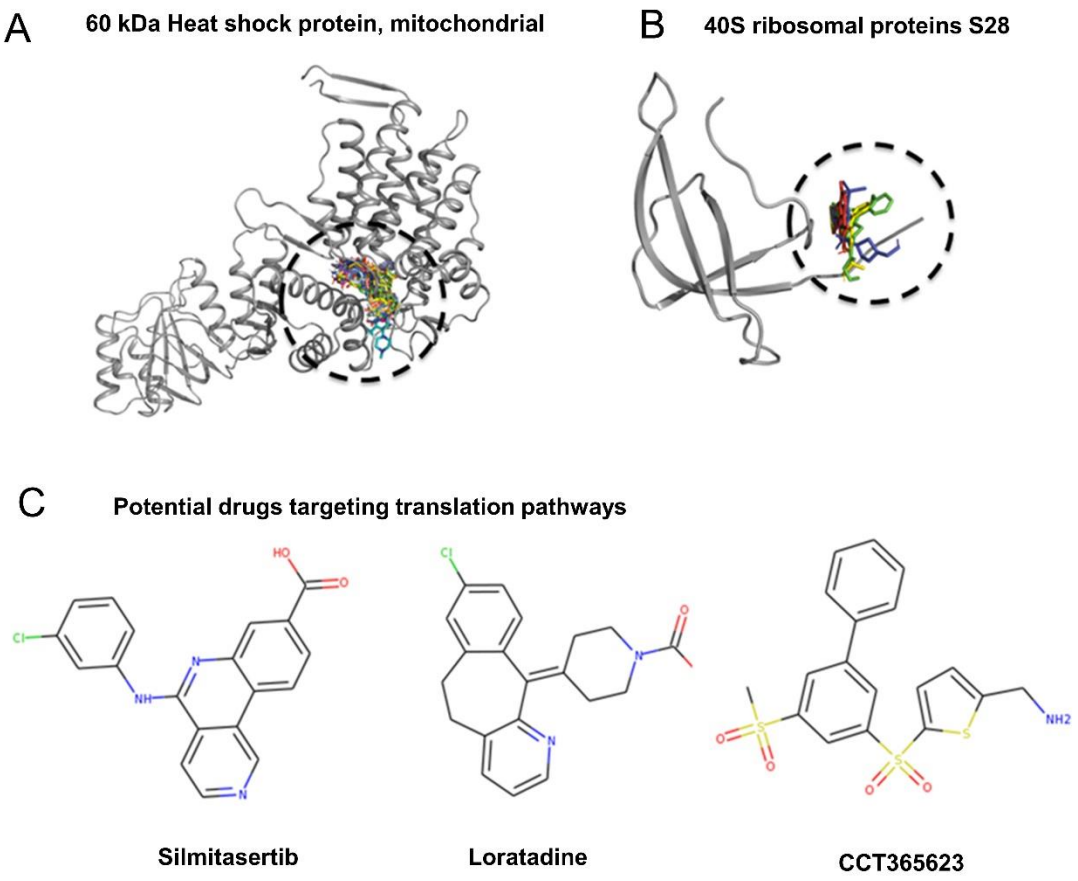


Figure S9. In-silico docking of host proteins involved in translational pathways. Related to Figure 5.

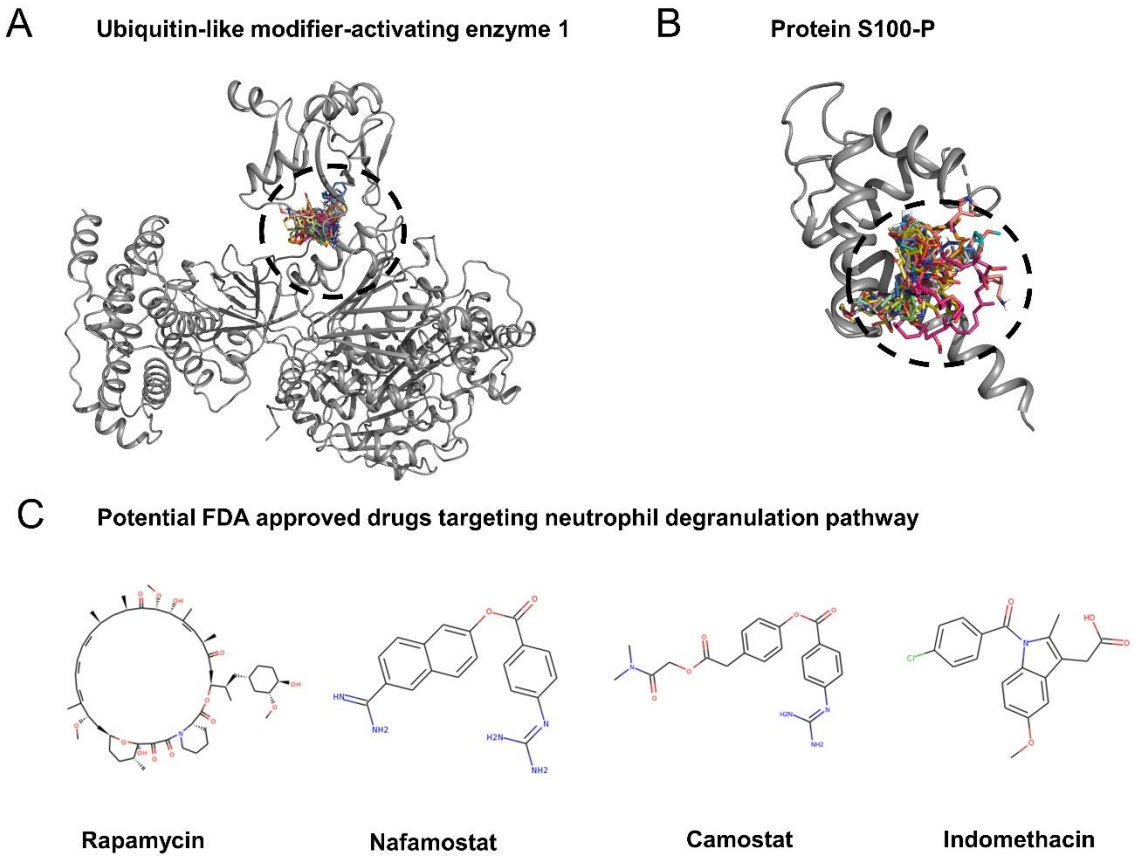
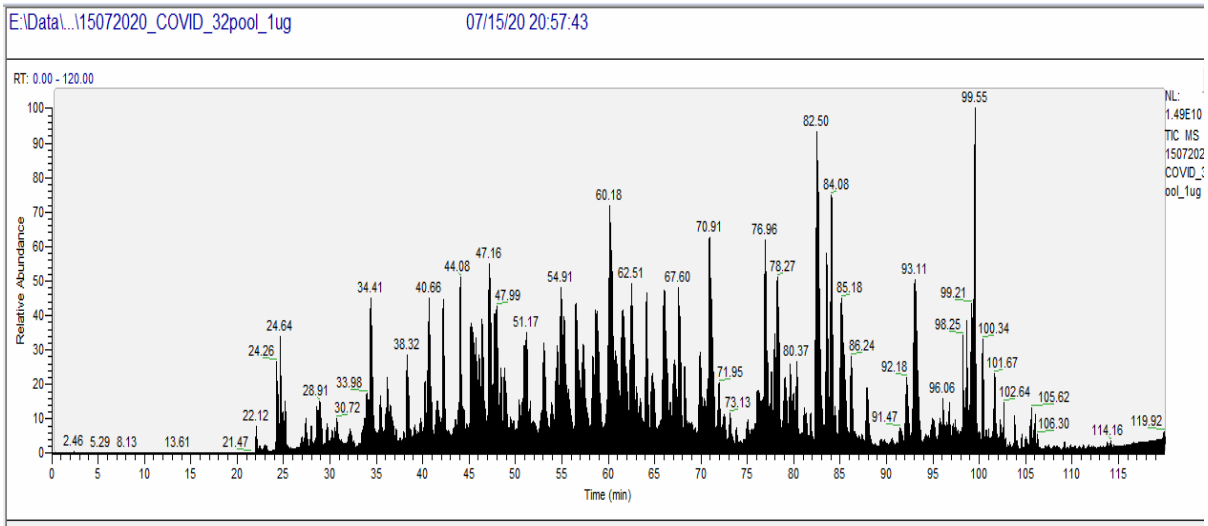


Figure S10. In-silico docking of host proteins involved in neutrophil degranulation pathways. Related to Figure 5.

Table S2. MS settings used for running the samples in Mass Spectrometry. Related to Figure 1A and Transparent Methods “MS analysis”.



15072020_COVID_32pool_1ug
Type: Unknown ID: 1 Row: 2

Sample Name:	
Study:	
Client:	
Laboratory:	
Company:	
Phone:	
Instrument Method:	E:\Data\Method\2019_Updated SOPs\LFQ and Label based\Label Free_120mins_OT_HCD_OT.meth
Processing Method:	
Vial:	B1
Injection Volume (µl):	4.00
Sample Weight:	0.00
Sample Volume (µl):	0.00
ISTD Amount:	0.00
Dil Factor:	1.00

15072020_COVID_32pool_1ug
Orbitrap Fusion
Orbitrap Fusion Method Summary

Creator: MININT-CVDF0SE\Fusion Last Modified: 11/2/2019 11:14:03 PM by MININT-CVDF0SE\Fusion

Global Settings
Use Ion Source Settings from Tune = True
Method Duration (min)= 120
Infusion Mode (LC)= False
FAIMS Mode = Not Installed
Internal Mass Calibration= User Defined Lock Mass
Application Mode = Peptide
Default Charge State = 1
Advanced Peak Determination = False
Xcalibur AcquireX enabled for method modifications = False
Internal Cal Positive
m/z
445.12003
Experiment 1
Experiment Name = MS
Start Time (min) = 0
End Time (min) = 120
Cycle Time (sec) = 3

Scan MasterScan
MSn Level = 1
Use Wide Quad Isolation = True
Detector Type = Orbitrap
Orbitrap Resolution = 60K
Mass Range = Normal
Scan Range (m/z) = 375-1700
Maximum Injection Time (ms) = 50
AGC Target = 400000
Microscans = 1
RF Lens (%) = 60
Use ETD Internal Calibration = False
DataType = Profile
Polarity = Positive
Source Fragmentation = False
Scan Description =

Filter MIPS
MIPS Mode = Peptide

Filter ChargeState
Include charge state(s) = 2-6
Include undetermined charge states = False
Include charge states 25 and higher = False

Filter DynamicExclusion
Exclude after n times = 1
Exclusion duration (s) = 40
Mass Tolerance = ppm
Mass tolerance low = 10
Mass tolerance high = 10

Table S3. MaxQuant analysis parameter file. Related to Transparent Methods. Related to Figure 1A and Transparent methods “LFQ data analysis”.

Parameter	**Value**
Version	1.6.6.0
User name	Admin
Machine name	ADMIN-PC
Include contaminants	TRUE
PSM FDR	0.01
PSM FDR Crosslink	0.01
Protein FDR	0.01
Site FDR	0.01
Min. peptide Length	6
Min. score for unmodified peptides	0
Min. score for modified peptides	40
Min. delta score for unmodified peptides	0
Min. delta score for modified peptides	6
Min. unique peptides	0
Min. razor peptides	1
Min. peptides	1
Use only unmodified peptides and	TRUE
Modifications included in protein quantification	Oxidation (M);Acetyl (Protein N-term)
Peptides used for protein quantification	Razor
Discard unmodified counterpart peptides	TRUE
Label min. ratio count	2
Use delta score	FALSE
iBAQ	FALSE
Instrument Type	Orbitrap
Match between runs	TRUE
Matching time window [min]	0.7
Match ion mobility window [indices]	0.05
Alignment time window [min]	10
Alignment ion mobility window [indices]	1
Find dependent peptides	FALSE
Decoy mode	revert
Include contaminants	TRUE
Advanced ratios	TRUE
Second peptides	TRUE
Stabilize large LFQ ratios	TRUE
Separate LFQ in parameter groups	FALSE
Require MS/MS for LFQ comparisons	TRUE
Main search max. combinations	200

Advanced site intensities	TRUE
Max. peptide mass [Da]	4600
Min. peptide length for unspecific search	8
Max. peptide length for unspecific search	25
Razor protein FDR	TRUE
Disable MD5	FALSE
Max mods in site table	3
Match unidentified features	FALSE
Evaluate variant peptides separately	TRUE
Variation mode	None
MS/MS tol. (FTMS)	20 ppm
Top MS/MS peaks per Da interval. (FTMS)	12
Da interval. (FTMS)	100
MS/MS deisotoping (FTMS)	TRUE
MS/MS deisotoping tolerance (FTMS)	7
MS/MS deisotoping tolerance unit (FTMS)	ppm
MS/MS tol. (ITMS)	0.5 Da
Top MS/MS peaks per Da interval. (ITMS)	8
Da interval. (ITMS)	100
MS/MS deisotoping (ITMS)	FALSE
MS/MS deisotoping tolerance (ITMS)	0.15
MS/MS deisotoping tolerance unit (ITMS)	Da
MS/MS ammonia loss (ITMS)	TRUE
MS/MS dependent losses (ITMS)	TRUE
MS/MS recalibration (ITMS)	FALSE
MS/MS tol. (Unknown)	0.5 Da
Top MS/MS peaks per Da interval. (Unknown)	8
Da interval. (Unknown)	100

Table S4. The COVID-19 database from Uniprot (SwissProt Database) with 13 proteins. Related to Figure 1C and Transparent Methods (LFQ data analysis).

Entry	Entry name	Protein names	Length
P0DTD1	R1AB_SARS2	Replicase polyprotein 1ab	7,096
P0DTC7	NS7A_SARS2	ORF7a protein	121
P0DTD2	ORF9B_SARS2	ORF9b protein	97
P0DTC9	NCAP_SARS2	Nucleoprotein	419
P0DTC3	AP3A_SARS2	ORF3a protein	275
P0DTD8	NS7B_SARS2	ORF7b protein	43
P0DTC8	NS8_SARS2	ORF8 protein	121
P0DTC6	NS6_SARS2	ORF6 protein	61
P0DTC1	R1A_SARS2	Replicase polyprotein 1a	4,405
P0DTD3	Y14_SARS2	Uncharacterized protein 14	73
P0DTC5	VME1_SARS2	Membrane protein	222
P0DTC2	SPIKE_SARS2	Spike glycoprotein	1,273
P0DTC4	VEMP_SARS2	Envelope small membrane protein	75

Table S5. Peptide sequences and MRM details of the SARS-CoV-2 peptides. Related to Figure 1C.

Protein Name	Compound	Polarity	Precursor (m/z)	Product (m/z)	Collision Energy (V)
Replicase	DFMSLSEQLR(+3)	Positive	409.2	632.336215	17.8
	DFMSLSEQLR(+3)	Positive	409.2	416.261593	17.8
	DFMSLSEQLR(+3)	Positive	409.2	316.671745	17.8
	DGHVETFYPK(+3)	Positive	398.192	244.165568	17.4
	DGHVETFYPK(+3)	Positive	398.192	442.231636	17.4
	DGHVETFYPK(+3)	Positive	398.192	392.697429	17.4
	DGHVETFYPK(+3)	Positive	398.192	328.176132	17.4
	DGHVETFYPK(+3)	Positive	398.192	340.843153	17.4
	DGHVETFYPK(+3)	Positive	398.192	295.156849	17.4
	ELLQNGMNGR(+2)	Positive	566.282	477.223828	19.9
	ELLQNGMNGR(+2)	Positive	566.282	232.140415	19.9
	ELLQNGMNGR(+2)	Positive	566.282	445.219068	19.9
	ELLQNGMNGR(+2)	Positive	566.282	324.647747	19.9
	ELLQNGMNGR(+2)	Positive	566.282	267.626284	19.9
	ELLQNGMNGR(+2)	Positive	566.282	239.115552	19.9
	NGSIHLYFDK(+3)	Positive	398.536	822.414465	17.4
	NGSIHLYFDK(+3)	Positive	398.536	409.208161	17.4
	NGSIHLYFDK(+3)	Positive	398.536	262.139747	17.4
	NGSIHLYFDK(+3)	Positive	398.536	468.252903	17.4
	NGSIHLYFDK(+3)	Positive	398.536	343.181415	17.4
	NGSIHLYFDK(+3)	Positive	398.536	229.123368	17.4
	NGSIHLYFDK(+3)	Positive	398.536	88.051433	17.4
Spike glycoprotein	QIAPGQTGK(+3)	Positive	300.503	658.351865	13.7
	QIAPGQTGK(+3)	Positive	300.503	490.261987	13.7
	QIAPGQTGK(+3)	Positive	300.503	204.134267	13.7
	QIAPGQTGK(+3)	Positive	300.503	329.67957	13.7
	QIAPGQTGK(+3)	Positive	300.503	245.634632	13.7
	QIAPGQTGK(+3)	Positive	300.503	102.570772	13.7
	QIAPGQTGK(+3)	Positive	300.503	145.085025	13.7
	QIAPGQTGK(+3)	Positive	300.503	102.398833	13.7
	QIAPGQTGK(+3)	Positive	300.503	68.716273	13.7
	FLPFQQFGR(+2)	Positive	570.303	507.267407	20
	FLPFQQFGR(+2)	Positive	570.303	379.208829	20
	FLPFQQFGR(+2)	Positive	570.303	232.140415	20
	FLPFQQFGR(+2)	Positive	570.303	391.700837	20
	FLPFQQFGR(+2)	Positive	570.303	318.16663	20
	FLPFQQFGR(+2)	Positive	570.303	254.137341	20
	NTQEVFAQVK(+2)	Positive	582.306	820.45633	20.4

	NTQEVFAQVK(+2)	Positive	582.306	691.413737	20.4
	NTQEVFAQVK(+2)	Positive	582.306	592.345323	20.4
	NTQEVFAQVK(+2)	Positive	582.306	445.276909	20.4
	NTQEVFAQVK(+2)	Positive	582.306	374.239795	20.4
	NTQEVFAQVK(+2)	Positive	582.306	410.731803	20.4
	NTQEVFAQVK(+2)	Positive	582.306	346.210506	20.4
	NTQEVFAQVK(+2)	Positive	582.306	296.676299	20.4
	IQDSLSSTASALGK(+4)	Positive	345.185	204.134267	15.4
	IQDSLSSTASALGK(+4)	Positive	345.185	367.705785	15.4
	IQDSLSSTASALGK(+4)	Positive	345.185	273.665932	15.4
	IQDSLSSTASALGK(+4)	Positive	345.185	159.112804	15.4
	IQDSLSSTASALGK(+4)	Positive	345.185	245.472949	15.4
	IQDSLSSTASALGK(+4)	Positive	345.185	159.100675	15.4
	IQDSLSSTASALGK(+4)	Positive	345.185	130.089999	15.4
Nucleoprotein (NCAP)	DGIIWVATEGALNTPK(+3)	Positive	562.301	700.398815	23.6
	DGIIWVATEGALNTPK(+3)	Positive	562.301	643.377351	23.6
	DGIIWVATEGALNTPK(+3)	Positive	562.301	572.340238	23.6
	DGIIWVATEGALNTPK(+3)	Positive	562.301	459.256174	23.6
	DGIIWVATEGALNTPK(+3)	Positive	562.301	244.165568	23.6
	DGIIWVATEGALNTPK(+3)	Positive	562.301	550.800945	23.6
	DGIIWVATEGALNTPK(+3)	Positive	562.301	504.952203	23.6
	DGIIWVATEGALNTPK(+3)	Positive	562.301	467.257515	23.6
	DGIIWVATEGALNTPK(+3)	Positive	562.301	429.562827	23.6
	DGIIWVATEGALNTPK(+3)	Positive	562.301	367.536389	23.6
	DGIIWVATEGALNTPK(+3).2	Positive	562.301	277.151987	23.6
	DGIIWVATEGALNTPK(+3).2	Positive	562.301	234.137789	23.6
	DGIIWVATEGALNTPK(+3).2	Positive	562.301	115.742599	23.6
	GGSQASSR(+3)	Positive	250.456	349.183009	11.8
	GGSQASSR(+3)	Positive	250.456	262.15098	11.8
	GGSQASSR(+3)	Positive	250.456	318.159002	11.8
	GGSQASSR(+3)	Positive	250.456	274.642988	11.8
	GGSQASSR(+3)	Positive	250.456	210.613699	11.8
	GGSQASSR(+3)	Positive	250.456	175.095142	11.8
	GGSQASSR(+3)	Positive	250.456	183.431084	11.8
	GGSQASSR(+3)	Positive	250.456	140.744891	11.8
	GGSQASSR(+3)	Positive	250.456	88.055177	11.8
	NSTPGSSR(+3)	Positive	269.131	262.15098	12.5
	NSTPGSSR(+3)	Positive	269.131	252.132256	12.5
	NSTPGSSR(+3)	Positive	269.131	203.605874	12.5
	NSTPGSSR(+3)	Positive	269.131	175.095142	12.5
	NSTPGSSR(+3)	Positive	269.131	131.579128	12.5

Table S6. The number of proteins and peptides identified in each dataset. Related to Figure 2A.

Subject	Number of Proteins
Total Protein Identified	2445
COVID-19 Positive	2203
COVID-19 Recovered	1540
COVID-19 True Negative	1288
COVID-19 Positive vs COVID-19 Negative (Significant protein)	452
COVID-19 Positive vs COVID-19 Recovered (Significant protein)	374
COVID-19 Positive vs COVID-19 True Negative (Significant protein)	266

Table S7. Table depicting the adjusted p value and fold change between 16 COVID-19 positive and 6 COVID-19 negative samples, used for validation. Related to Figure 3A.

Protein accession	Protein Gene name	Peptide	Isotope Label Type	Fold Change Result	Adjusted P-Value
P02794	FTH1	NVNQSLLELHK	light	1.89 (99% CI:1.07 to 3.35)	0.0116
P02794	FTH1	YFLHQSHEER	light	1.96 (99% CI:1.09 to 3.52)	0.0116
P02794	FTH1	MTTASTSQVR	light	1.51 (99% CI:0.89 to 2.55)	0.0375
P02792	FTL	LNQALLDLHALGSAR	light	2.06 (99% CI:1.14 to 3.72)	0.0103
P02792	FTL	KPAEDEWGK	light	1.65 (99% CI:0.93 to 2.9)	0.0242
P02792	FTL	ALFQDIK	light	1.61 (99% CI:0.9 to 2.89)	0.0326
P17174	GOT1	IVASTLSNPELFEEWTGNVK	light	2.03 (99% CI:1.08 to 3.81)	0.0116
P17174	GOT1	NLDYVATSIHEAVTK	light	3.11 (99% CI:1.31 to 7.38)	0.0077
P17174	GOT1	IANDNSLNHEYLPILGLAEFR	light	1.66 (99% CI:0.94 to 2.94)	0.0242
P00505	GOT2	FVTVQTISGTGALR	light	1.91 (99% CI:0.96 to 3.83)	0.021
P00505	GOT2	DAGMQLQGYR	light	1.61 (99% CI:0.93 to 2.81)	0.0252
P00505	GOT2	IGASFLQR	light	2.16 (99% CI:1.18 to 3.95)	0.0083
P00505	GOT2	ASAELALGENSEVLK	light	2.46 (99% CI:1.06 to 5.72)	0.0116
P00505	GOT2	DDNGKPYVLPSVR	light	13.68 (99% CI:2.54 to 73.65)	0.0038
P00338	LDHA	DQLIYNLLK	light	1.88 (99% CI:1.07 to 3.33)	0.0116
P00338	LDHA	GEMMDLQHGSLFLR	light	2.52 (99% CI:1.36 to 4.66)	0.0038
P00338	LDHA	SADTLWGIQK	light	1.96 (99% CI:1.18 to 3.27)	0.0077
P00338	LDHA	DLADELALVDVIEDK	light	1.99 (99% CI:1.09 to 3.63)	0.0116
P00338	LDHA	VHPVSTMIK	light	2.55 (99% CI:1.08 to 5.99)	0.0116
P07195	LDHB	IHPVSTMVK	light	1.75 (99% CI:1.03 to 3)	0.0116
P01375	TNF	ETPEGAEAKPWYEPIYLGGVFQLEK	light	1.61 (99% CI:0.96 to 2.72)	0.0217
P02741	CRP	QDNEILIFWSK	light	2.48 (99% CI:1.04 to 5.9)	0.0116

P02741	CRP	APLTKPLK	light	6.13 (99% CI:1.84 to 20.45)	0.0038
P05231	IL6	EALAENNLNLPK	light	2.07 (99% CI:1.04 to 4.1)	0.0116
P05231	IL6	YILDGISALR	light	1.82 (99% CI:1.04 to 3.18)	0.0116
P05231	IL6	EFLQSSLR	light	1.57 (99% CI:0.96 to 2.57)	0.0217
P05231	IL6	VLIQFLQK	light	1.75 (99% CI:0.96 to 3.17)	0.021
	Spike-in peptide	THCLYTHVCDAIK	heavy	0.96 (99% CI:0.85 to 1.08)	0.3093

Table S9. List of the human proteins found to be mapped with the Interleukin Signaling Pathway. Related to Figure 4.

Gene Symbol	Description	Biological Process (GO)
HDGF	heparin binding growth factor	GO:0098760 response to interleukin-7;GO:0098761 cellular response to interleukin-7;GO:0036498 IRE1-mediated unfolded protein response
PSME2	proteasome activator subunit 2	GO:0035722 interleukin-12-mediated signaling pathway;GO:0071349 cellular response to interleukin-12;GO:0070671 response to interleukin-12
CANX	calnexin	GO:0034975 protein folding in endoplasmic reticulum;GO:0070106 interleukin-27-mediated signaling pathway;GO:0070757 interleukin-35-mediated signaling pathway
PDIA3	protein disulfide isomerase family A member 3	GO:0034975 protein folding in endoplasmic reticulum;GO:0098760 response to interleukin-7;GO:0098761 cellular response to interleukin-7
STIP1	stress induced phosphoprotein 1	GO:0098760 response to interleukin-7;GO:0098761 cellular response to interleukin-7;GO:0071345 cellular response to cytokine stimulus
GSTA2	glutathione S-transferase alpha 2	GO:1901685 glutathione derivative metabolic process;GO:1901687 glutathione derivative biosynthetic process;GO:0035722 interleukin-12-mediated signaling pathway
RPLP0	ribosomal protein lateral stalk subunit P0	GO:0000027 ribosomal large subunit assembly;GO:0035722 interleukin-12-mediated signaling pathway;GO:0071349 cellular response to interleukin-12
P4HB	prolyl 4-hydroxylase subunit beta	GO:0038155 interleukin-23-mediated signaling pathway;GO:0046598 positive regulation of viral entry into host cell;GO:0034378 chylomicron assembly

H3C14	H3 clustered histone 14	GO:0038111 interleukin-7-mediated signaling pathway;GO:0000183 chromatin silencing at rDNA;GO:0098760 response to interleukin-7
H3C1	H3 clustered histone 1	GO:0038111 interleukin-7-mediated signaling pathway;GO:0006335 DNA replication-dependent nucleosome assembly;GO:0034723 DNA replication-dependent nucleosome organization
CA1	carbonic anhydrase 1	GO:0006730 one-carbon metabolic process;GO:0015701 bicarbonate transport;GO:0035722 interleukin-12-mediated signaling pathway
TUBA1B	tubulin alpha 1b	GO:0071353 cellular response to interleukin-4;GO:0070670 response to interleukin-4;GO:0030705 cytoskeleton-dependent intracellular transport
RAC1	Rac family small GTPase 1	GO:0032707 negative regulation of interleukin-23 production;GO:0032627 interleukin-23 production;GO:0032667 regulation of interleukin-23 production
HNRNPF	heterogeneous nuclear ribonucleoprotein F	GO:0035722 interleukin-12-mediated signaling pathway;GO:0071349 cellular response to interleukin-12;GO:0070671 response to interleukin-12

Transparent Methods

Subject recruitment and swab collection

For this study, we obtained nasopharyngeal swab samples from 101 adult patients who visited Kasturba hospital, Mumbai. The study was approved by the Indian Institute of Technology (IIT), Bombay ethical committee, and Kasturba hospital for infectious diseases, Institutional Review Board. Based on the RT-PCR results and clinical information these patients were grouped into positive (n = 70), true negative (n = 21) and recovered (n = 10). The positive patients were further categorized into non-severe (n = 41) and severe (n = 23) based on the clinical symptoms and oxygen supplementation (**Table S1**). A nasopharyngeal swab was collected by a medical practitioner maintaining proper infection control. A sterile cotton swab was used to collect a clinical specimen and store it in a tube containing viral transport media (VTM) at 4 °C. Sample processing was performed in the BSL-2 facility as per WHO and ICMR guidelines. Around 800 µL of the sample from the stored tube was dispensed into a sterile tube. The tube was incubated at 65 °C for 45 minutes for heat inactivation of the virus. Further, 200 µL of swab samples collected in VTM was added into the Eppendorf tube containing 600 µL of acetone, ethanol, and isopropanol (1:3) each. All the tubes were kept at -20 °C for 4 hours. Next, the Eppendorf tubes were centrifuged at 15,000 g for 20 minutes at 4 °C. Proteins present in the samples precipitated in the form of pellet and supernatant was discarded. Dried pellets in the Eppendorf tubes were stored at -80 °C till further processing.

Sample preparation for Label-Free-Quantification (LFQ)-based mass spectrometry analysis

The Eppendorf tubes containing precipitated proteins obtained using different solvent conditions were further processed using in-solution digestion. Around 75 µL of freshly prepared urea lysis buffer (8 M urea, 50 mM Tris pH 8.0, 75 mM NaCl, 1mM $MgCl_2$) was added to each protein pellet. After ensuring pellets are completely dissolved, the sample was stored at 4°C until use. The protein quantification was performed by Bradford protein assay using standard concentrations of Bovine Serum Albumin (BSA). The distribution of all sample (human tissue samples) pools (quality control (QC)) and the Pearson correlation between the pools and individual pool distribution respectively are shown in **Figure S1**. From each sample, ~ 30 µg of protein was taken in a fresh Axygen 1.5 ml tube for enzymatic digestion. Further to reduce the disulfide bonds in the proteins, around 20 mM (final concentration) of tris (2-carboxyethyl) phosphine (TCEP) was

added and the tubes were incubated at 37 °C for 60 minutes. After reduction, Iodoacetamide (IAA) was added to a final concentration of 37.5 mM alkylate and incubated for 20 minutes at room temperature in the dark. The sample was diluted with a buffer containing 25 mM Tris pH 8.0 and 1 mM $CaCl_2$. This step was necessary to dilute the urea concentration to less than 1 M. Trypsin (Pierce) was added at an enzyme/substrate ratio of 1:30 and tubes were incubated at 37°C in a shaking dry bath for 16 hours. The samples were dried completely in a vacuum centrifuge. Once digested, peptides were desalted using C18 spin columns (Pierce, catalog number 89870) (Venkatesh et al., 2020). The desalting protocol was followed as per the manufacturer's instructions.

Quantification of desalted peptides and MS analysis

Desalted peptides were reconstituted in 0.1 % (v/v) formic acid (FA). Absorbance was measured using MultiSkan Go (Thermo Fischer Scientific). Peptide concentrations were quantified using the Scopes method. 0.1 % FA was used to equilibrate the pre-column (Thermo Fisher Scientific, P/N 164564, S/N 10694527) and analytical column (Thermo Fisher Scientific, P/N ES803A, S/N 10918620). 1 µg of peptides were injected into the liquid chromatography (LC) column at a flow rate of 300 nl/min, and the LC gradient was set to 120 minutes. The peptides were then analyzed on Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Fischer Scientific) with easy nano-LC 1200 system. DDA (Data Dependent Acquisition) mode was used along with a scan range and mass resolution of 375-1700 m/z and 60,000 respectively. A mass tolerance window was kept to 10 ppm and High energy Collision Dissociation method (HCD) of fragmentation was used to collect all MS/MS data. The specification of the fusion pre column used was Acclaim PepMap 100, 100um*2cm, nanoviper C18, 5um ,100 A and for the fusion Analytical column was PepMap RSLC C18 2um,100A, 75um*50cm, 75um*50cm. An LC gradient of 120 minutes was used to separate peptides on the nano-LC column. Solvent A comprised of 0.1% FA in water and solvent B comprised of 80% ACN in 0.1% FA water. Solvent B was increased from 5% to 30% from the 5th minute to 80th minute of the run followed by an increase of solvent B to 60% and then to 90% till the 115th minute. Solvent B was maintained at 90% for the last 5 minutes. Before processing the samples quality check of the instrument was also done by monitoring some peptides of the BSA sample via Panorama software.

Thermo Xcalibur software version 4.0 was used for data acquisition. **Table S2** shows the LC-MS/MS settings of label-free quantification performed using high-resolution mass spectrometry.

LFQ Data analysis and statistical data analysis

The MaxQuant analysis parameter file is shown in **Table S3**. The raw datasets were processed with MaxQuant (Tyanova et al., 2016) (v1.6.6.0) against the Human Swiss-Prot database (Version: 2020_04) which includes a total of 20,353 proteins and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Swiss-Prot database of 13 proteins (**Table S4**), searched with the built-in Andromeda Search Engine of MaxQuant (Tyanova et al., 2016). All these datasets were downloaded between 09.07.2020 to 13.07.2020. Raw files were processed using Label-Free-Quantification (LFQ) parameters and setting label-type as "standard" with a multiplicity of 1. The Orbitrap was set to Orbitrap Fusion mode. Trypsin was used for digestion with a maximum missed cleavage of 2 for Human and COVID. Carbamido-methylation of Cysteine (+57.021464 Da) was set as the fixed modification, whereas oxidation of Methionine (+15.994915 Da) was set as the variable modification. The False-Discovery-Rate (FDR) was set to 1% for the protein and peptide levels to ensure high reliability of the protein detection. Decoy mode was set to "revert", and the type of identified peptides was set to "unique+razor". The unique peptides were filtered using the Human Swiss-Prot database in Skyline (Version 20.1).

The protein output file of MaxQuant was taken forward, removing contaminants, and proteins identified with q-value greater than 0.05.3749 proteins were identified from 68 samples. The chromatograms of these samples were checked and the coefficient of Correlation analysis (Spearman Rank correlation) was taken forward to understand the quality of the datasets. 8 samples out of 44 samples were found to have quality issues and have been excluded from the study before any statistical analysis. LFQ intensities of 36 samples provided 2445 proteins, which includes 18 COVID-19 Positive samples, 11 recovered samples, and 7 True negative samples. The analysis was performed without any missing value imputation, log 2-fold change intensities were taken forward and a Welch's t-test was performed. The p-value cut-off of less than 0.05 was determined for the significance level. A list of 164 common significant proteins was found between Positive vs Negative, Positive vs Recovered and Positive vs True Negative.

The second batch of 24 COVID positive samples which includes 11 non-severe and 13 severe patient samples were separately analyzed in MaxQuant using the same parameters. 1 sample out of 24 samples was having the least correlation coefficient and was coming as an outlier were removed. The partial least square-discriminant analysis was performed using 23 samples and proteins were selected based on variable importance in projection (VIP) score in Metaboanalyst software (Xia et al., 2015).

The violin plots were drawn with the Log2 transformed data, where the significance level was calculated based on t-test independent samples with Bonferroni correction, (p-value annotation legend: ns: 5.00e-02 < p <= 1.00e+00; *: 1.00e-02 < p <= 5.00e-02; **: 1.00e-03 < p <= 1.00e-02; ***: 1.00e-04 < p <= 1.00e-03; ****: p <= 1.00e-04. The heatmap has been plotted using Metaboanalyst where Distance measure parameter were set to Euclidean and the clustering algorithm were set to Ward clustering. All the statistical data analysis were done using Microsoft excel, R (Version 4.0.2), Python (Version 3.7.6), Metaboanalyst (Xia et al., 2015), and Hierarchical Clustering Explorer (Version 3.0) (Seo and Shneiderman, 2002).

Multiple Reaction Monitoring (MRM) assays

For all the MRM experiments, a TSQ Altis mass spectrometer (ThermoFisher Scientific, USA) coupled to a Vanquish UHPLC system (ThermoFisher Scientific, USA) was used. The peptides were separated using a Hypersil Gold C18 column 1.9 µm, 100 X 2.1 mm (ThermoFisher Scientific, USA) at a flow rate of .45 ml/min for a total time of 10 minutes. The binary buffer system used was 0.1% FA as buffer A and 80% ACN in 0.1% FA as the buffer B. The gradient used for chromatographic separation of the peptides was as follows - 2-45% buffer B for 6 minutes, 45-95% buffer B for 0.5 minutes, 95% buffer B for 0.5 minutes, 95%-2% buffer B for 0.5 minutes, and 2% buffer B for 2.5 minutes. The transition lists were prepared on Skyline (MacLean et al., 2010). The missed cleavage criterion was 0, precursor charges +2, +3, and product charges +1, +2 with y ion transitions (from ion 2 to last ion -1) were included.

For experiments dealing with viral peptide detection, 16 peptides corresponding to 3 proteins were monitored resulting in a final list consisting of 12 peptides we used for MRM assays (**Table S5**). All the collision energy values used in the experiment were as determined by Skyline software (Ver 20.1.1.196) (MacLean et al., 2010). For each of the 10 COVID patient samples and 3 non-COVID/healthy controls the peptides from all three extraction solvents were pooled and used for the MRM study. The non-COVID

samples were used to establish baseline cut-off for the peaks being detected. A SpikeTide (Schnatbaum et al., 2011), having sequence FEDGVLDPDYPR was spiked into the sample to monitor the consistency of the mass spectrometry runs. This synthetic peptide has C-terminal Arginine (R) heavy labelled. 1 µg of peptides from each sample were injected and run against the list. After the data acquisition of the samples, all the downstream data analysis was performed in Skyline (MacLean et al., 2010).

Besides, a few host proteins such as L-lactate dehydrogenase A chain (P00338), L-lactate dehydrogenase B chain (P07195), Ferritin heavy chain (P02794), Ferritin light chain (P02792), Aspartate aminotransferase, mitochondrial (P00505), Aspartate aminotransferase, cytoplasmic (P17174) and Albumin (P02768) which were found to be upregulated in COVID-19 positive when compared to COVID-19 negative in the LFQ data were selected and used for a targeted MRM study. We have also checked for the clinical markers such as IL-6 – Interleukin – 6, CRP – C reactive protein as well as proinflammatory cytokines TNF-α – Tumor necrosis factor which we could not identify in LFQ analysis but are clinically significant in literature. The following analysis was performed on swab samples of 6 COVID-19 negative and 14 COVID-19 positive patients (consisting of 6 severe and 8 mild samples). Around 30 µg of protein was digested using trypsin, which resulted in approximately 15 µg peptides. The concentration of peptide was calculated using the Scopes method from its OD value at 205 nm and 280 nm. 1µg peptide of each sample was injected in the Vanquish UHPLC system (ThermoFisher Scientific, USA) linked to TSQ Altis™ Triple Quadrupole Mass Spectrometer. The peptides were separated and analyzed under the experimental conditions explained above. The list of transitions was prepared for unique peptides of these selected proteins using Skyline (Ver 20.2.1.286) and SRM Atlas. The list had 27 peptides from 9 host proteins mentioned above. This list also included a spiked-in heavy synthetic peptide (THCLYTHVCDAIK) essential for monitoring the consistency of the mass spectrometry runs. Also, the standard curve for the heavy and light synthetic peptide was monitored and the lowest range of detection for both the synthetic peptides was found to 25 ng (**Figure 3C-D**). Equal amounts of BSA were also run daily in sequence with the samples to check the uniformity of the instrument response. This was to monitor any day-wise fluctuation in the response of the instrument (**Figure S3**).

Pathways and Protein Interaction Network data analysis

The GO enrichment analysis was performed in Metascape and the biological pathway mapped proteins were taken forward for visualization. STRING (Szklarczyk et al., 2019) (Version 11) and Reactome.org (Fabregat et al., 2018) (Version 73) were used to perform protein-protein interaction analysis and pathway mapping respectively.

Molecular docking

The docking experiment was performed using AutoDock Vina 1.1.2 (Trott, O., Olson, 2019). The 3D structure of the drugs was obtained from the PubChem database (Kim et al., 2019) and the ZINC15 database. The drugs were converted from SDF format to PDB format using PyMOL software (Rigsby and Parker, 2016). The coordinates for the proteins were retrieved from the Protein Data Bank (PDB) (Berman et al., 2002). Using PyRx software, the PDB format of the protein and the drugs were converted to a PDBQT format, which is a readable file format for AutoDock Vina. The docking was done with the exhaustiveness value set to 50 using the blind docking method. In this method, the grid box is large enough to select the entire protein. This increases the chances of obtaining all possible ligand-receptor complexes. The docked complexes were visualized using PyMOL software. The binding interactions between the drug and the protein were calculated using the protein-ligand interaction profiler (PLIP) server (Salentin et al., 2015).

Supplemental References

Berman, H.M., Battistuz, T., Bhat, T.N., Bluhm, W.F., Bourne, P.E., Burkhardt, K., Feng, Z., Gilliland, G.L., Iype, L., Jain, S., et al. (2002). The protein data bank. Acta Crystallogr. Sect. D Biol. Crystallogr. 58, 899–907.

Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., Haw, R., Jassal, B., Korninger, F., May, B., et al. (2018). The Reactome Pathway Knowledgebase. Nucleic Acids Res. 46, D649–D655.

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B.A., Thiessen, P.A., Yu, B., et al. (2019). PubChem 2019 update: Improved access to chemical data. Nucleic Acids Res. 47, D1102–D1109.

Rigsby, R.E., and Parker, A.B. (2016). Using the PyMOL application to reinforce visual understanding of protein structure. Biochem. Mol. Biol. Educ. 44, 433–437.

Salentin, S., Schreiber, S., Haupt, V.J., Adasme, M.F., and Schroeder, M. (2015). PLIP: Fully automated protein-ligand interaction profiler. Nucleic Acids Res. 43, W443–W447.

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 47, D607–D613.

Trott,O., Olson, A.J. (2019). Autodock vina: improving the speed and accuracy of docking. J. Comput. Chem. 31, 455–461.

Tyanova S., Temu T., Cox J., The MaxQuant computational platform for mass spectrometry-based shotgun proteomics Nat. Protoc. 2016 2301-2319

Xia J., Sinelnikov I.V., Han B., Wishart D.S., MetaboAnalyst 3.0-making metabolomics more meaningful Nucleic Acids Res. 2015 W251-W257