Article

Proteo-Genomic Analysis of SARS-CoV-2: A Clinical Landscape of Single-Nucleotide Polymorphisms, COVID-19 Proteome, and Host Responses

Sheetal Tushir, Sathisha Kamanna, Sujith S. Nath, Aishwarya Bhat, Steffimol Rose, Advait R. Aithal, and Utpal Tatu*

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: A novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of coronavirus disease 2019 (COVID-19) and continues to be a global health challenge. To understand viral disease biology, we have carried out proteo-genomic analysis using next-generation sequencing (NGS) and mass spectrometry on nasopharyngeal swabs of COVID-19 patients to examine the clinical genome and proteome. Our study confirms the mutability of SARS-CoV-2 showing multiple single-nucleotide polymorphisms. NGS analysis detected 27 mutations, of which 14 are synonymous, 11 are missense, and 2 are extragenic in nature. Phylogenetic analysis of SARS-CoV-2 isolates indicated their close relation to a Bangladesh isolate and multiple origins of isolates within the country. Our proteomic analysis, for the first time, identified 13 different SARS-CoV-2 proteins from the clinical swabs. Of the total 41 peptides captured by high-resolution mass spectrometry, 8 matched to nucleocapsid protein, 2 to ORF9b, and 1 to spike glycoprotein and ORF3a, with remaining peptides mapping to ORF1ab polyprotein. Additionally, host proteome analysis revealed several key host proteins to be uniquely expressed in COVID-19 patients. Pathway analysis of these proteins points toward modulation in immune response, especially involving neutrophil and IL-12-mediated signaling. Besides revealing the aspects of host-virus pathogenesis, our study opens new avenues to develop better diagnostic markers and therapeutic approaches.

KEYWORDS: COVID-19, SARS-CoV-2, next-generation sequencing, mass spectrometry, COVID-19 proteomics, genomics, host proteome
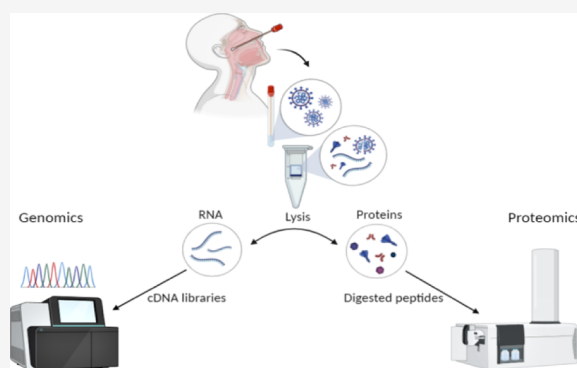
INTRODUCTION

Coronavirus disease 2019 (COVID-19) is the latest addition to the extensive list of infectious diseases caused by viruses that have jumped from animals to humans. The global outbreak of COVID-19 caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) originated from the city of Wuhan, China, in December 2019. After SARS-CoV in 2003[1,2] and Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012,[3] SARS-CoV-2 is the third of the coronavirus family to cross the species barrier and infect humans with severe respiratory disease. It has proven to be more dangerous than SARS-CoV and MERS-CoV due to its alarming transmission rate through respiratory droplets, encountering infected persons or contaminated surfaces.[4] SARS-CoV-2 was first identified by meta-transcriptomic sequencing from the bronchoalveolar lavage fluid of a patient in China, and the sequence was made available at the Global Initiative on Sharing All Influenza Data (GISAID) platform on 12th January 2020.[5] Phylogenetic analysis shows that about 30 kb of the genome of this new RNA virus is most closely related to a group of SARS-like coronaviruses (humans and bats) with an 89.1% similarity.[5] SARS-CoV-2 has 14 open-reading frames (ORFs), which codes for structural and non-structural (accessory and replication) proteins. ORF1ab is the largest of all comprising 21,291 nucleotides (about 2/3 of the genome) and codes for 15 non-structural proteins (replicase polyprotein) which comprise the viral replicase−transcriptase complex (RTC).[6,7] Structure protein genes located downstream to ORF1ab, aligned in the order spike (S), envelope (E), membrane (M), and nucleocapsid (N), with ORFs that code for accessory non-structural proteins, are located in between (Figure S1). The non-structural proteins are not part of the virion and hence expressed only during the actively
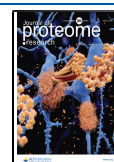
replicating phase in the infected host cell. The structural proteins are highly abundant as they are incorporated into the virion where N encapsulates the RNA genome and E, M, and S proteins form a capsid of lipid bilayer.[8]

Since the availability of the first SARS-CoV-2 sequence, more than 390,000 complete genome sequences have been added to the list by several laboratories across the world. Despite a large number of sequences being available, it is still not clear how fast the virus mutates and if the mutations impact its virulence in the context of the growing pandemic. In addition, very little information is available regarding the clinical proteome of the virus. Until recently, only a few proteins, mainly including structural proteins N and S, have been identified from clinical swabs.[9−11] Besides, host proteome studies from clinical samples are necessary to fill the void in understanding the host responses to viral infection.

In this study, genomic analysis of SARS-CoV-2 was performed by next-generation sequencing (NGS) on naso-pharyngeal swab samples of reverse transcription-polymerase chain reaction (RT-PCR)-positive individuals, collected from Bangalore, India. Variant analysis of these samples showed ≥11 mutations per sample. Phylogenetic analysis of these sequences with other variants from India as well as across the world revealed their close similarity to one of the Bangladesh isolates. SARS-CoV-2 phylogeny indicated the prevalence of isolates showing multiple origins within the country. Overall, through genomic analysis of SARS-CoV-2, our study highlighted increasing variations [single-nucleotide polymorphisms (SNPs)] in the viral genome and their role to understand its evolution and virulence.

In addition to sequencing the genome by NGS, our study also explored the clinical proteome and host-protein responses to SARS-CoV-2 infection by using high-resolution mass spectrometry (MS) (HRMS). We performed HRMS on nasopharyngeal swab samples of both RT-PCR-positive and RT-PCR-negative patients. In total, we identified 41 peptides matching to 13 different SARS-CoV-2 proteins, including proteins from ORF1ab polyprotein, spike glycoprotein, ORF3a, ORF9b, and nucleocapsid. Additionally, the host proteomic analysis revealed significant differences between RT-PCR-positive and RT-PCR-negative host proteomes. We found 441 host proteins uniquely present in positive samples. Most of these proteins are involved in neutrophil degranulation and activation pathways, indicating host immunological response to the virus. In conclusion, our proteomic analysis confirms the presence of SARS-CoV-2 proteins in nasal swab samples and also predicted host responses to viral infection, including identification of the neutrophil response as a key host response against SARS-CoV-2 infection.

■ MATERIALS AND METHODS

Sample Collection

Nasopharyngeal swab samples were collected from the diagnosed patients as a part of routine monitoring. A part of the samples after diagnosis was sent to the lab for the research purpose. Samples were classified into positive and negative based on the RT-PCR result targeting E and RNA-dependent RNA polymerase (RdRp) genes of the virus. Samples were collected only after the approved consent of the patients who were informed about the study. The study was conducted after approval of the Institutional Human Ethics Committee, IISc (19-01092020).

RNA Library Preparation

Total RNA from nasopharyngeal swabs of three positive patients (RT-PCR test) was extracted using Trizol-based extraction. RNA samples were quantified using the Qubit RNA Assay HS (Invitrogen). RNA purity was checked using a Nanodrop, and integrity was assessed on a TapeStation using RNA HS ScreenTapes (Agilent). The Qiagen SARS-CoV-2 Primer (Qiagen) was used to prepare libraries from RNA extracted from COVID-19-positive subjects. Viral RNA was converted to complementary DNA (cDNA) and used as a template for multiplex PCR with primers spanning the entire genome of the virus. The amplicons were then pooled and purified before proceeding for library preparation. During library preparations, the amplicons were subjected to a series of enzymatic steps that repair the ends, tails, and the 3′ end with a single "A" nucleotide, followed by ligation of the adapters. The adapter-ligated products were then purified and enriched using a limited cycle PCR. The final cDNA libraries were purified and checked for fragment size distribution on the TapeStation using D1000 DNA ScreenTapes (Agilent).

Next-Generation Sequencing of SARS-CoV-2

Prepared cDNA libraries were quantified using the Qubit High Sensitivity Assay (Invitrogen). Quantified libraries were pooled and diluted to final optimal loading concentrations for cluster amplification on an Illumina flow cell, followed by sequencing on an Illumina HiSeq X instrument to generate 150 bp paired-end reads.

Mutation Analysis

The quality of the reads was initially checked using FastQC v0.11.9.[12] Further, the sequencing adapters clipped at the 5′ and 3′ ends of reads were trimmed using Cutadapt v2.9.[13] The adapter-trimmed pair-end reads were then aligned to the Wuhan reference genome (accession no. NC_045512.2) downloaded from NCBI.[14] The fast and accurate read alignment was achieved by using a BWA v.0.7.12 aligner.[15] The aligned reads were sorted, soft-clippings were removed, and variant calling was performed using a GATK variant caller.[16] The reported variants were then annotated to study their effects in proteins and genes using the SNPEff tool.[17] The variant class, amino acid changes, and other relevant annotations were added to the variants.

Nucleotide Sequence Accession Number

The SARS-CoV-2 whole-genome sequences have been submitted to NCBI under accession number PRJNA668889.

Phylogenetic Analysis

The evolutionary history was inferred by using the Maximum Likelihood method and the Tamura−Nei model.[18] The tree with the highest log likelihood (−525441.09) was generated. The initial tree for the heuristic search was obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura−Nei model and then selecting the topology with the superior log likelihood value. The tree was drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 40 nucleotide sequences with a total of 29,945 positions in the final data set. Codon positions included were 1st + 2nd + 3rd + Noncoding. Evolutionary analysis was conducted in MEGA X.[19]

Nucleotide Variations showing SNPs
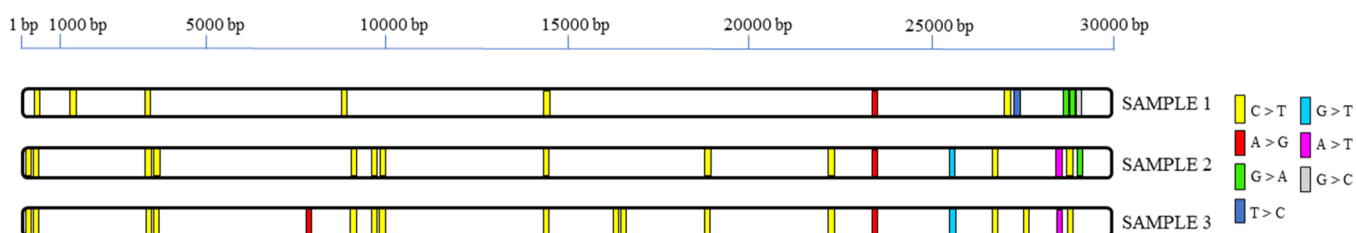
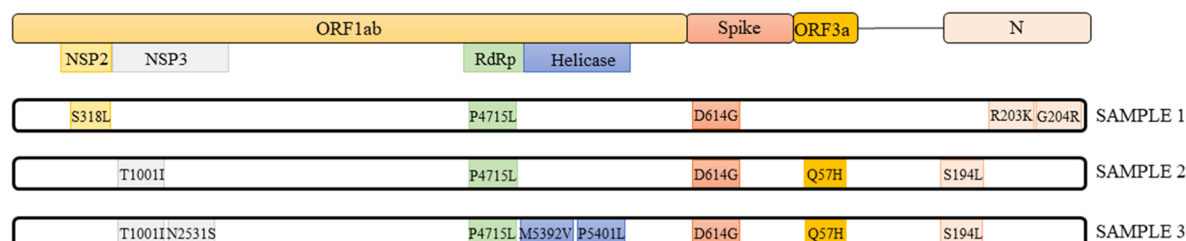**Missense Variations showing amino acid substitution**

Figure 1. Variant analysis in Indian isolates of the SARS-CoV-2 genome. The top panel of the figure indicates nucleotide variations (SNPs) in the genome sequence of the SARS-CoV-2 Bangalore isolates against the reference Wuhan-Hu-1 isolate complete genome sequence. Specific transitions and transversions with their color coding are mentioned on the right side of the panel. The bottom panel indicates the position of the missense mutations against the reference Wuhan-Hu-1 SARS-CoV-2 isolate.

Sample Preparation for Mass Spectrometry

In-solution trypsin digestion was performed to extract the peptides from the protein solution. Briefly, samples collected in viral transport medium (VTM) were centrifuged at 14,000 rpm for 15 min at 4 °C. The supernatant was collected in a separate microcentrifuge tube, and the pellet was washed with 1× PBS. Further, the pellet containing epithelial cells was lysed by the 1× Triton buffer (chilled). The lysate in the supernatant was collected after centrifuging at 14,000 rpm for 20 min at 4 °C. Proteins in both the lysate and supernatant were precipitated by the addition of chilled acetone and incubated at −80 °C for 2 h. Precipitated proteins were washed with chilled acetone and dissolved in 50 mM ammonium bicarbonate. Proteins were reduced using 10 mM dithiothreitol (Sigma-Aldrich) in 50 mM ammonium bicarbonate at 56 °C for 45 min, followed by alkylation with 55 mM iodoacetamide in 50 mM ammonium bicarbonate at 37 °C for 30 min in the dark. In-solution digestion was carried out by adding trypsin (Promega) 1 μg/μL to a final protease-to-protein ratio of 1:50 (w/w) and incubated at 37 °C for 16 h, with frequent shaking. Digestion was stopped using formic acid, and all the samples were vacuum-dried.

Mass Spectrometry and Database Search

The dried trypsin-digested peptides were reconstituted in a mixture of 20% acetonitrile (ACN) and 80% MQ containing 0.01% formic acid. The protein digests were analyzed using an Agilent 1290 Infinity II LC system coupled with an Agilent Advance Bio Q-TOF (6545XT). The column used for chromatography was an Agilent AdvanceBio Peptide Map (2.1 × 150 mm, 2.7 μ). Mobile phase A was Milli-Q water (MQ) (0.1% formic acid), and mobile phase B was ACN (0.1% formic acid). The peptides were separated by using a 90 min gradient flow at a flow rate of 0.4 mL/min. The MS and tandem MS (MS/MS) scans were obtained in the positive mode and stored in the centroid mode. The following MS data acquisition parameter V_{cap} was set at 3500 V, and the drying gas flow rate and the temperature were set at 12 L/min and 270 °C, respectively. The collision energy with a slope of 3.6 V/100 Da and an offset of 4.8 V was used for fragmentation. The precursor ion data were captured in a mass range of 200−1800 m/z, and product ion data were obtained in the range of 50−2900 m/z. Reference exclusion was given for 0.05 min after 1 spectrum. The raw data were analyzed using MaxQuant software (v1.6.2.10) and processed in MS Excel. The database analysis was performed against SARS-CoV-2 proteome and *Homo sapiens* proteome in the UniProt database (Proteome ID—UP000005640). The following search parameters were used for the database analysis: precursor mass tolerance: 10 ppm and fragment mass tolerance: 40 ppm, with cysteine carbamidomethylation as a fixed modification and methionine oxidation as variable modifications.

MS Data Repository

MS proteomics data obtained on nasopharyngeal swabs have been deposited to the ProteomeXchange Consortium via the PRIDE[20] partner repository under data set identifiers PXD021896 and 10.6019/PXD021896.

Gene Ontology and Pathway Analysis

All Uniprot IDs of host proteins found exclusively in COVID-19-positive samples were extracted and analyzed through DAVID Tool for conversion to Entrez IDs. These Entrez IDs were then used for the identification of Gene Ontology (GO) terms and pathways. The statistical significance of the genes enriched in positive samples was analyzed using the R package, clusterProfiler. To determine whether any terms annotate a specified list of genes at a frequency greater than that would be expected by chance, clusterProfiler calculates a p-value using
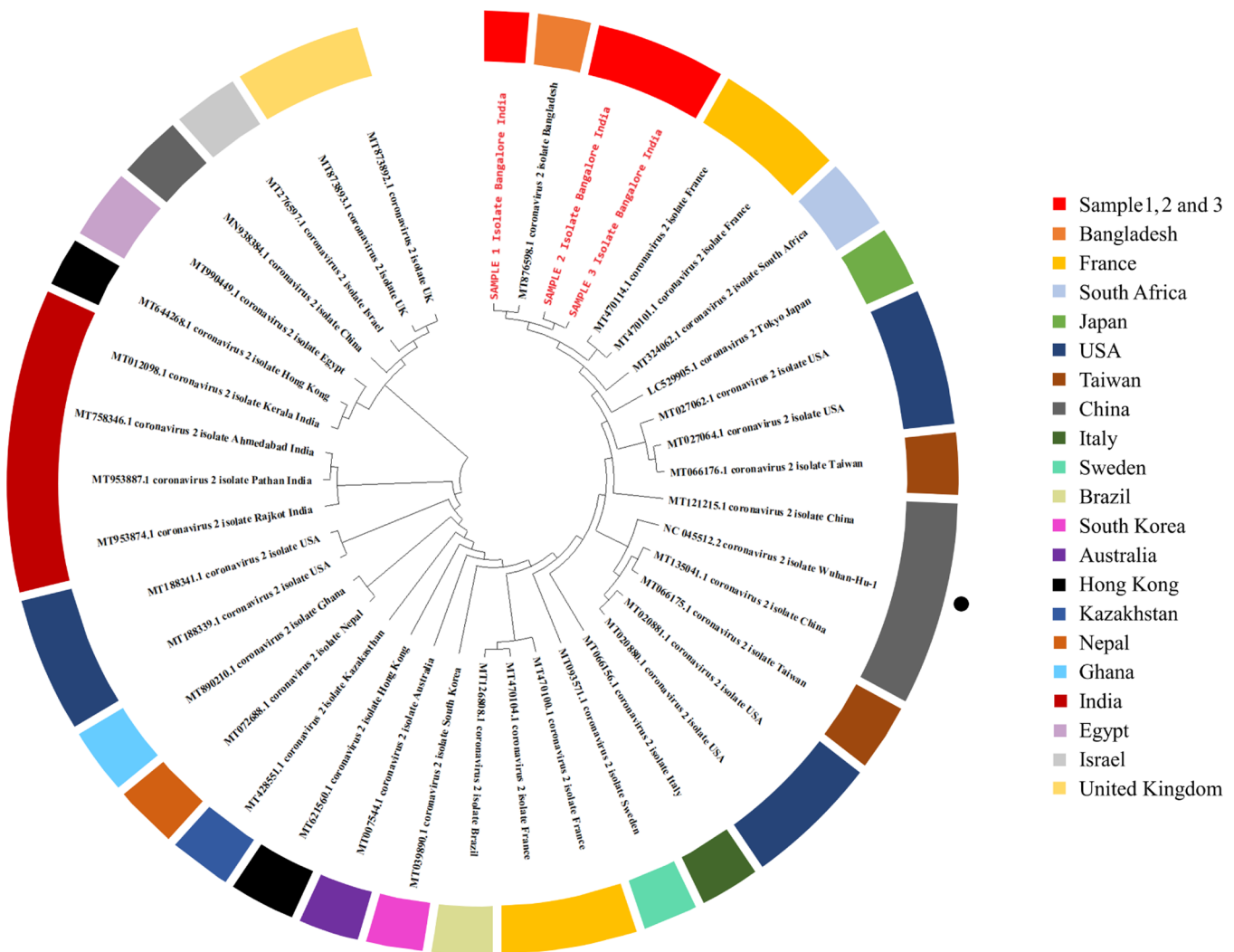
Figure 2. Phylogenetic analysis of SARS-CoV-2 isolates. Whole-genome phylogeny representing the relationship of Bangalore SARS-CoV-2 isolates based on the Maximum Likelihood method and Tamura model created using MEGA X. The phylogenetic analysis involved 40 SARS-CoV-2 sequences representing variants from 20 countries around the globe. The colors around the tree represent the country of origin for each isolate. Isolates from Bangalore are represented in red text showing close relation to the Bangladesh isolate. The black dot at the outer region of the circle marks the Wuhan-Hu-1 reference genome.

the hypergeometric distribution. Statistically enriched GO terms were then plotted and analyzed through the dot plot, category net plot (Cnetplot), and enrichment map (Emap).

RESULTS

Genome Sequence Reveals Emerging Mutations in SARS-CoV-2

Since the first SARS-CoV-2 genome sequence shared on 12th January 2020,[5] more than 390,000 genome sequences have been available at GISAID to date.[21] To correlate the sequence of prevailing SARS-CoV-2 in Bangalore, India, with those reported earlier, we carried out Illumina HiSeq X, NGS of SARS-CoV-2 (collected in June 2020). SARS-CoV-2 RNA from three nasopharyngeal swabs, tested positive by RT-PCR, was converted to cDNA and processed for NGS, as described in the Materials and Methods section. NGS analysis retrieved the complete genome sequence from all three samples. FastQ files generated from NGS were reference-mapped to the SARs-CoV-2 isolate, Wuhan-Hu-1 (accession no. NC 045512.2) with a 100% genome coverage. The alignment of these three

isolates with the reference genome showed the prevalence of SNPs in all three samples (Figure 1). Samples 1, 2, and 3 showed 11, 16, and 19 SNPs, respectively. A total of 27 variations were found in the isolates, out of which 4 are common to all and 11 are exclusive to samples 2 and 3. The four common mutations observed in all three isolates are c.241C > T, c.3037C > T, c.14408C > T and c.23403A > G. In total, we found nine mutations belonging to the category of most frequent mutations, out of which six (c.241C > T, c.3037C > T, c.14408C > T, c.23403A > G, c.25563G > T and c.28881G > A) are common in all the continents, c.26735C > T and c.28854C > T are specific to Asian isolates, and c.18877C > T is found in both American and Asian isolates.

The amino acid substitutions due to these point mutations are represented in the lower panel of Figure 1. Out of 27, 25 mutations are in the coding region, which generate 14 synonymous and 11 missense amino acid substitutions (Table S1). A total of 3 of these 11 missense mutations could be of high impact as they substitute charged to uncharged amino acid or vice versa and hence may impact the structure and consequently the function of the proteins.
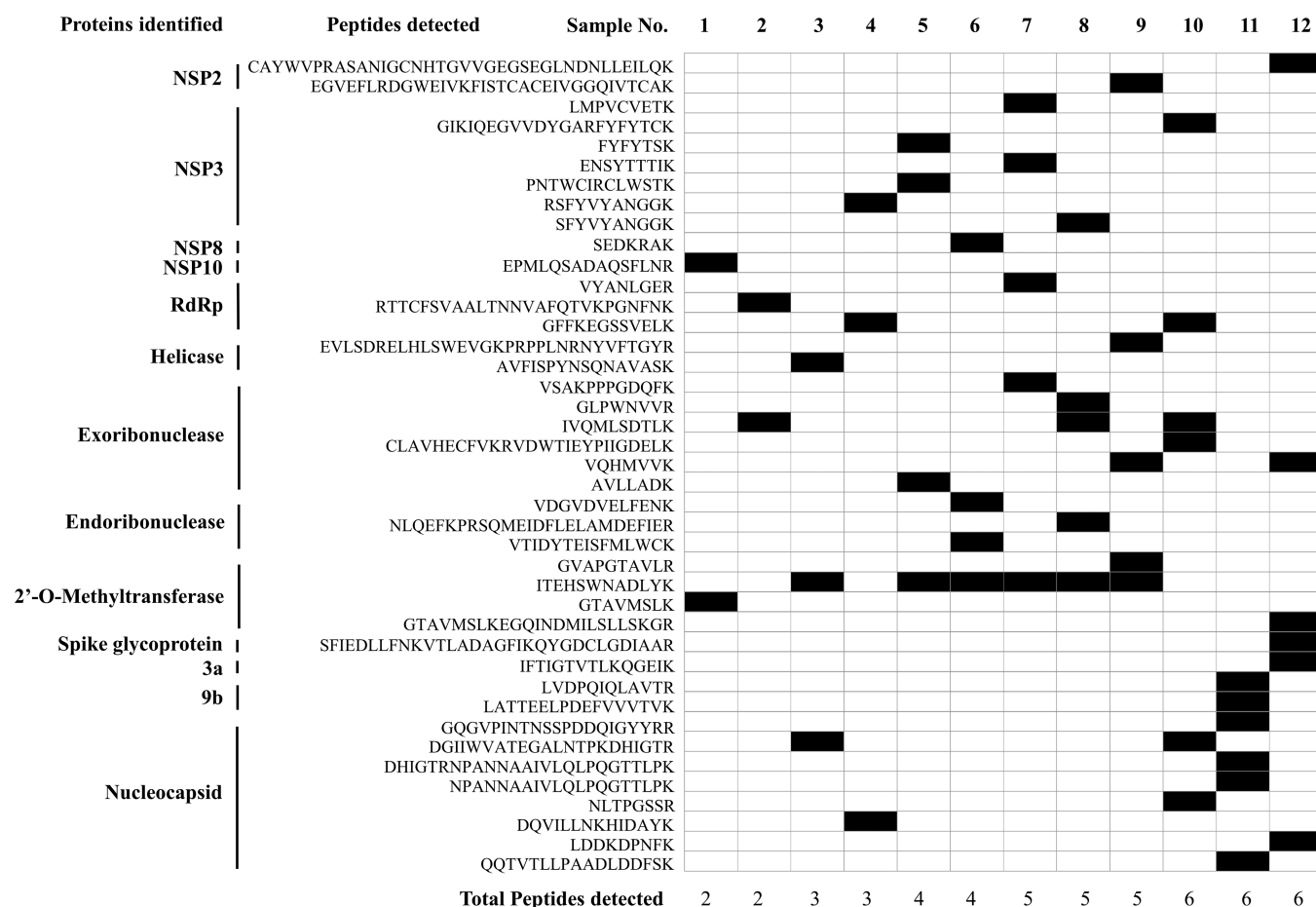
Figure 3. Clinical proteome of SARS-COV-2. The peptide map depicts SARS-CoV-2 peptides identified from the clinical nasopharyngeal swabs of 12 COVID-19 patients. Cells highlighted in black represent detected peptides in that sample. The sequence of the peptides along with the matched protein is listed on the left. Total peptides identified in the sample are indicated at the bottom. Sample numbers 1–12 are arranged in increasing order of total peptides detected.

These include p.D614G in spike glycoprotein, p.Q57H in ORF3a, and p.G204R in nucleocapsid.

Further, to see the evolutionary relationship of these isolates, we constructed a phylogenetic tree using MEGA X software. Phylogenetic analysis showed that samples 2 and 3 are more closely related and all three are in close relation to one of the Bangladesh isolates, which appears to have originated from France isolates. Distinct clade-wise assignment revealed that all the isolates belong to the G-derived clade (European origin), isolate 1 belongs to GR, and isolates 2 and 3 belong to GH (Table S2 and Figure S2).

Clinical Proteome of SARS CoV-2

Despite an ample collection of genome information, the information on the clinical proteome of SARS-CoV-2 is ill-explored. So far, only a handful of studies have reported the clinical proteome of SARS-CoV-2.[7,11,22−24] We carried out HRMS analysis of SARS-CoV-2 clinical proteome. Parts of the nasopharyngeal samples obtained from 12 COVID-19-positive patients with different clinical manifestations (Table S3) were analyzed using an Agilent Advance Bio Q-TOF (6545XT), as described in the Materials and Methods section. Briefly, proteins extracted (both intracellular and extracellular) from the nasopharyngeal swabs were reduced, alkylated, and trypsin-digested for proteome analysis. From MS/MS spectra recorded at false discovery rate (FDR) ≤1%, we were able to identify 41 unique peptides matching to 13 different viral proteins. The

maximum number of peptides was attributed to ORF1ab polyprotein. As shown in the peptide map in Figure 3, we detected eight peptides matching to nucleocapsid protein; seven peptides to NSP3; six peptides to exoribonuclease; four peptides to 2′-O-methyltransferase; three peptides each to RdRp (NSP12) and endoribonuclease; two peptides each to NSP2, helicase, and protein 9b; and one peptide each to NSP8, NSP10, spike glycoprotein, and protein 3a. Peptide ITEHSW-NADLYK (2′-O-methyltransferase) was detected in 50% of the samples (6/12). Sample 12 showed the maximum number of identified proteins (six attributed to six different peptides). The peptides for 2′-O-methyltransferase protein were detected in 2/3 of the samples (66.6%). All the RT-PCR-positive samples showed the presence of SARS-CoV-2 proteins. Our result suggests the potential of MS for highly sensitive and reliable diagnosis of SARS-CoV-2.

Host Responses to SARS CoV-2 Infection

We also looked for host protein dynamics upon SARS-CoV-2 infection by searching the MS data against the human proteome database (proteome ID UP000005640). For this, we analyzed nine samples of both COVID-19-positive and COVID-19-negative patients, with negative being those which tested negative by RT-PCR and which did not show any SARS-CoV-2 peptides. To characterize the pathways that get modulated by the viral infection, we compared the list of host proteins identified by liquid chromatography (LC)−MS/MS
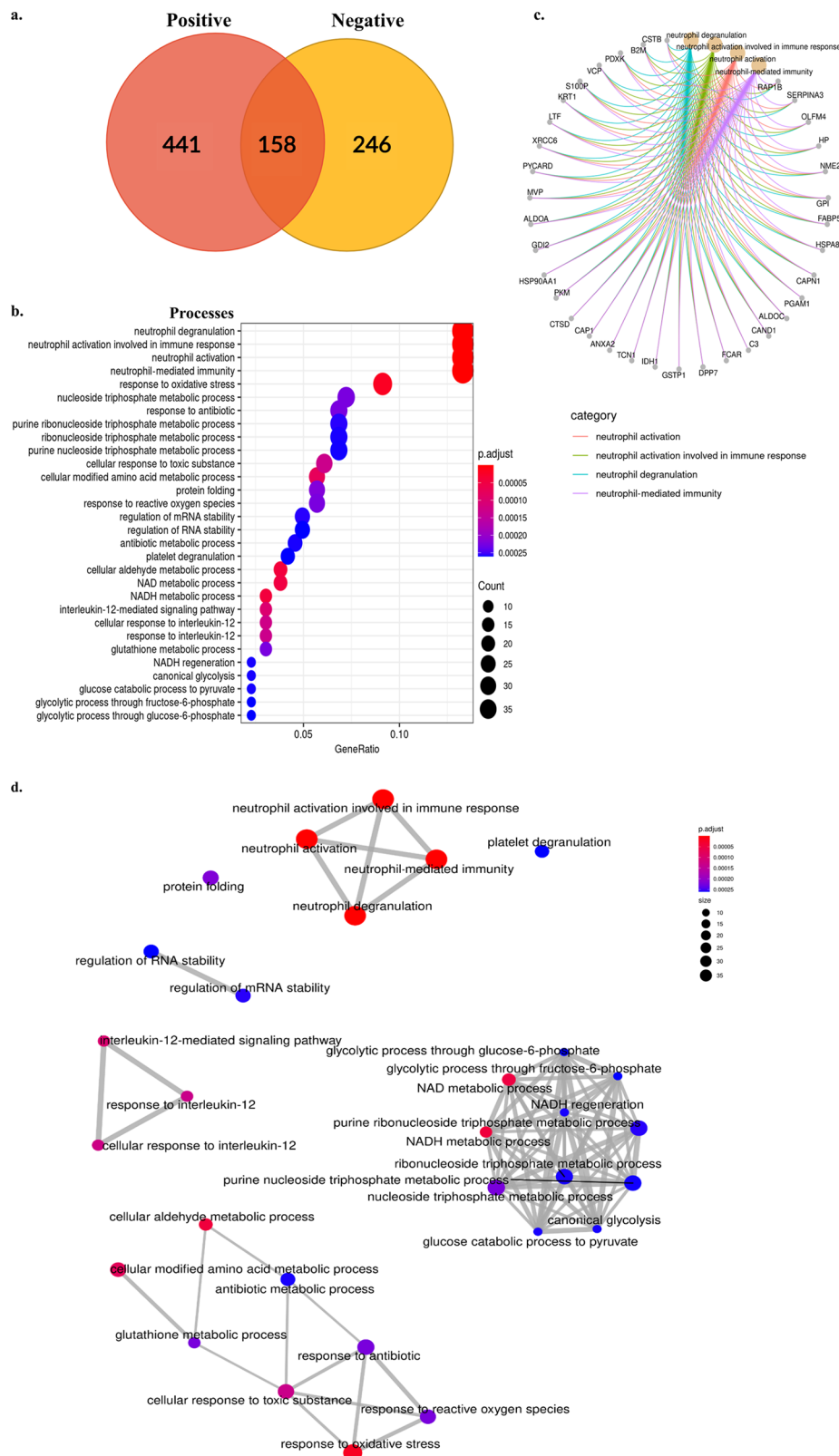
Figure 4. Clinical proteome and characterization of protein dynamics of the SARS-COV-2-infected cell. Proteomes of positive samples (nasopharyngeal swabs). (a) Venn diagram of the COVID-19-positive and COVID-19-negative host proteomes. (b) Pathway analysis of unique proteins identified in positive samples. The dot plot of top 30 pathways according to statistically enriched GO terms is plotted. The Y-axis of the plot represents pathways arranged in high-to-low order of gene counts. (c) Category net plot depicting linkages of genes and biological processes as a network for top four enriched pathways showing genes involved in them. (d) Enrichment map illustrates the cluster of functional modules by connecting overlapping gene sets of enriched terms into a network. Emap here represents the overall network and pathways of the unique proteome of COVID-19-positive clinical samples. Mainly, GO terms are organized in five networks, and functional nodules with genes involved in protein folding and platelet degranulation are predicted as an individual cluster.

in all positive and negative samples. Figure 4A represents the Venn diagram of host proteins. We identified 441 proteins to be uniquely present in positive samples, 246 exclusively in negative samples, and 158 found common to both groups.

We classified unique proteins from the proteome of positive samples to their GO terms and pathways to characterize host protein dynamics upon viral infection. The statistical significance of GO was further analyzed using the R package, clusterProfiler. In total, proteins unique to positive samples were classified into 244 GO terms. Figure 3B shows a dot plot for the top 30 GO terms according to their statistical significance. We found neutrophil-mediated immune responses including degranulation and the neutrophil activation pathway to be higher in positive samples (35 gene count). The abundance of proteins involved in these pathways was found to be enriched in positive samples. Additionally, we observed a large collection of proteins involved in the cellular response to oxidative stress and toxic substances and in metabolic pathways, for instance, the nucleoside/ribonucleoside triphosphate metabolic process, the NAD and NADH metabolic process, the amino acid metabolism, and the glycolytic process. Apart from this, we also observed an increased number of proteins involved in the RNA processing mechanism including regulation of RNA/mRNA stability, splicing, and localization to Cajal bodies. Among host immune responses, proteins involved in interleukin 12 and 7-mediated signaling pathways were also identified in the positive samples. Figure 3D shows an enriched map plot for the pathways identified for unique proteins present in positive samples. Our results demonstrate that SARS-CoV-2, like all viruses, manipulates the host in all aspects—biologically, molecularly, and cellularly for its survival.

We also analyzed unique proteins identified in negative samples that were not detected in positive nasopharyngeal swabs. The top 30 enriched GO terms are shown in the dot plot in Figure S3a. Although proteins involved in neutrophil-mediated immune responses were found in both positive and negative samples, their count in negative samples was lower (22) than in positive samples (35). Additionally, proteomic analysis of negative samples showed enrichment of some basic pathways in host epithelial cells involving epidermis development, cornification, and cellular oxidant detoxification which were not found enriched in COVID-19-positive samples. However, proteins that belong to the oxidative stress pathway were present in both the samples; their abundance is lower in negative samples (19 gene count) than in positive samples (24 gene count). Figure S3c demonstrates an enriched map plot for proteins exclusively identified in negative samples. Proteome analysis of SARS-CoV-2-infected cells and non-infected cells highlights viral−host interaction and the hijacking of host biological pathways by the virus for its survival and replication.

■ DISCUSSION

To strengthen the understanding of SARS-CoV-2 in terms of its origin, virulence, and pathogenesis, we have analyzed the virus genome, proteome, and host protein response in COVID-19 patients. Our genomic analysis based on these three samples from Bengaluru, which is an urban city in the southern part of India, confirms a high rate of mutation in Indian isolates. Earlier studies observed an average mutation rate of 7.23 mutations per sample.[25] This rate appears to vary between countries, while India (8.40), Kazakhstan (9.47), and Bangladesh (9.47) show a high mutation rate per sample as

compared to the world's average. While India's average mutation rate is estimated around 8.40, all three isolates in our analysis showed ≥11 mutations per sample. The observed lower mutation rate in previous studies could be because they were conducted during the early phase of the pandemic. In total, we found 27 SNPs, out of which 14 result in synonymous mutation, 11 are missense mutations, and 2 are extragenic in nature, observed in the 5′UTR region. 3 out of 11 missense mutations are charged to uncharged or vice versa amino acid substitution observed in N, S, and Orf3a proteins. A comparison of the predicted centroid secondary structure of wild-type (WT) and variant (mutant) 5′UTR showing positional entropy of each base showed that there is no change in entropy for 241C > T transition. However, entropy is significantly changed for 110C > T transition not only for that particular base but also for the stem it belongs to as compared to WT. This transition may impact viral RNA processing and expression through altered binding of predicted host RBPs—BRUNOL4, RBM38, and SRSF2.[26] So far, the virus has maintained its genome integrity by avoiding large-scale indels, and most of the mutations observed are in the form of SNPs (Figure 1). The silent mutations could cause a cumulative effect in the long term in terms of translational efficiency by changing codon usage. Missense mutations may or may not have a direct effect on the protein structure and function, but the emergence of new mutations with time may modulate its virulence. Also, mutations in the extragenic region may impact RNA folding, transcription, and replication ability.[27] Based on the marker mutations within the phylogenetic cluster, GISAID introduced a nomenclature system for major clades (Figure S2). Distinct clade-wise assignment reveals that all the isolates belong to the G-derived clade (European origin), with isolate 1 to GR and isolates 2 and 3 to GH specifically, highlighting the current prevalence of G-originated clades. Although there are six major clades currently, more clades can be classified in future with emerging and settling mutations.

To map the evolution and spread of SARS-CoV-2, a global whole-genome phylogenetic tree was created (Figure 2). We observed no direct correlation between isolates and geographic regions, although some isolates showed close relations with other isolates of the same country or neighboring countries. For instance, isolates from India mapped close to neighboring countries like Bangladesh, Nepal, and China (Hong Kong), indicating exchange of the virus between neighboring countries. A few of China isolates were found close to the Wuhan reference sequence in the phylogenetic tree, but most of the isolates showed a mosaic pattern of distribution (Figure 2).

Our HRMS data revealed important facets of COVID-19 disease biology. Previous studies also explored this area using viral cell lines, nasopharyngeal swabs, and the gargle solution.[9−11,28] So far, from the clinical swabs, only a few viral proteins have been identified. Reported studies mainly identified structural proteins of SARS-CoV-2. In this study, we analyzed 12 RT-PCR-positive nasopharyngeal swabs for the presence of SARS-CoV-2 peptides. In our sample size, we found 41 peptides matching to 13 different SARS-CoV-2 proteins. Most of the peptides matched to nucleocapsid protein and NSP3. A total of 29 peptides matched to ORF1ab polyprotein, 2 to protein 9b, 1 to protein 3a, and 9 to structural proteins S (1) and N (8). Detection of ORF9b with two peptides in sample 11 further confirms its expression in clinical
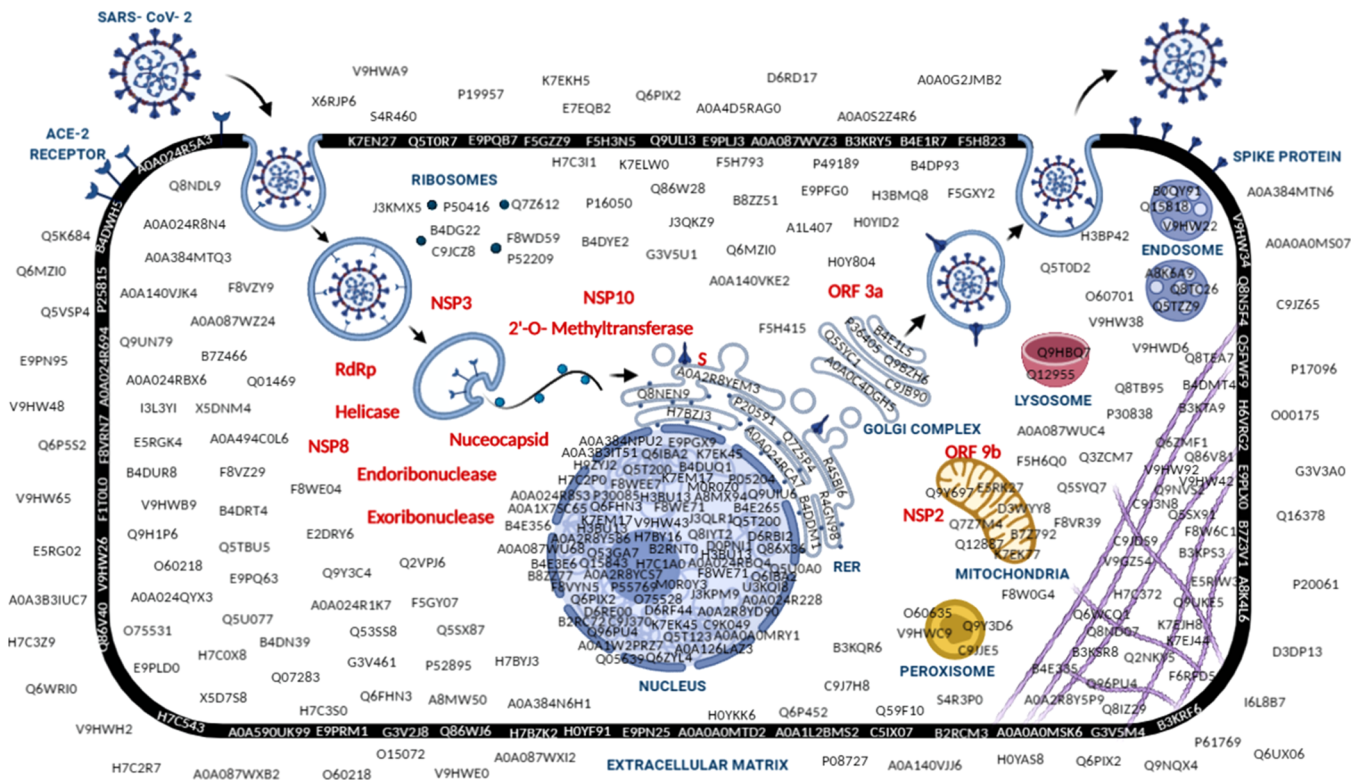
Figure 5. Overview of SARS-CoV-2 and host-cell proteomes. The figure depicts SARS-CoV-2 and host-epithelial cell proteomes identified in this study. Upon vesicular internalization, the coronavirus RNA undergoes a series of replications and translates into viral proteins. Viral proteins are indicated in red text, while host proteins identified in positive samples are shown in black. Viral structural proteins like spike glycoprotein undergo folding and trafficking through ER. The nucleocapsid form (RNA and N protein) individually assembled in the cytoplasm further joins structural proteins in the ER-Golgi intermediate compartment forming a complete virion particle. The assembled virion particles exit the infected cell by exocytosis and continue to transmit.

samples. Although previous studies predicted the function of ORF9b in suppressing Type I and III interferons,[29,30] no evidence of its expression was reported so far. This is the first study where ORF9b has been identified in clinical samples, which can further be explored to explain the severity of clinical cases and varying susceptibility of individuals. We observed that a number of peptides identified mostly correlate with *Ct* values in the RT-PCR test. The expression of ORF10 protein still remains elusive, although previous data implicate its role in ubiquitination; its presence in clinical swabs has not been shown in any study so far.

In comparison to RNA, proteins are more stable and better candidates for diagnosis. A greater number of peptides identified in a sample may indicate a greater viral load and proliferation, and samples showing maximum peptides could reflect severity of the disease. Through global proteomics, we identified temporal changes in the host proteome upon viral infection. In our sample size, we identified 441 proteins exclusively in positive samples. Pathway analysis of these proteins reveals alteration in basic host processes and heightened immune response (Figure 4), as reported previously, which showed that host cells try to combat the viral load by elevating immune responses, especially those mediated by macrophage, complement, and IL-6 signaling.[31,32] Here, we observed enrichment of neutrophil-mediated immune responses which are known to play a crucial role in airway infections and elicit antiviral response.[33] Not many studies are available which comment on the role of neutrophils in COVID-19, and their role is not very clear. Although they

are important for effective immune response, they can also be cytotoxic and lead to hyperinflammation through degranulation and lysis during severe pneumonia.[33−35] With increasing studies showing upregulated neutrophil genes and neutrophil-attracting chemokines in SARS-CoV-2, current literature studies also showed their role in damaging host inflammatory responses through involvement of neutrophil extracellular traps and an increased neutrophil-to-lymphocyte ratio observed in severe cases, indicating their association to viral pathology.[36−38] Hence, further studies would be useful in dissecting their role and to understand the mechanism involved in severe cases.

We also observed more proteins involved in the cellular response to IL-12 and IL-7, which has a role in the adaptive immune system, mainly the T-cell-mediated immune response. Several viruses, especially the causative agent of respiratory viral infections, alter the host redox balance for their survival and induce oxidative stress to facilitate their replication within the host.[39] Many recent studies have highlighted the role of oxidative stress in such viral infections including SARS-CoV and SARS-CoV-2.[40−43] As expected, we observed a heightened response to oxidative stress. An additional cluster of proteins identified in clinical swabs is mainly enriched in RNA processing (splicing, spliceosomal components), mRNA stability, localization to the Cajal body, and the RNA metabolism. This result relates to previous studies which also suggested splicing as a crucial pathway for SARS-CoV-2 survival.[28,44] Pathways involved in the metabolism, for instance, the carbon metabolism, RNA/DNA synthesis,

NAD/NADH synthesis, and the unsaturated fatty acid metabolism, are also among the enriched pathways observed in positive samples. Unsaturated fatty acids are components of phospholipids and involve in maintenance of membrane fluidity. Phospholipids along with sphingolipids mediate signal transduction and immune responses. Earlier studies reported phagocytosis and platelet degranulation-mediated alteration in the production of glycerophospholipid and the reduction of glycerophospholipid upon SARS-CoV-2 infection.[31,45] In addition, proteomic analysis of negative samples showed lower enrichment of proteins belonging to neutrophil-mediated immune response, which is in agreement with our conclusion that the neutrophil pathway is upregulated in the infected patients. Overall, through our study, we suggest alterations of host processes affecting cellular, metabolic, or biological functions. Furthermore, our global proteomics unravels cellular and molecular pathways for therapeutic interventions. Essentially, our result suggests that proteomics can not only offer timely and sensitive diagnosis but also reveal important facets of host−viral interactions.

Altogether, our proteo-genomic study revealed genome-wide SNPs, SARS-CoV-2 proteome, and the dynamics of the host proteome in clinical samples. In combination with etiological and patient severity details, multi-omics studies can not only predict the progression of SARS-CoV-2 but also help in identifying drug targets to offer counter treatment for this as well as future pandemics.

CONCLUSIONS

In this proteo-genomic study, we analyzed the clinical proteome of SARS-CoV-2 and the variations accumulated in the genome since its identification. The clinical landscape of SARS-CoV-2 and the host proteome highlighted correlation between the viral proteins and host responses (Figure 5). Through our proteomics study, we confirmed the expression of various (13 in this study) viral proteins in the host cell. Pathway analysis of the host proteome indicated enrichment of proteins majorly involved in immune response, the metabolism, and RNA processing. We identified several COVID-19 peptides within a 90 min MS-acquisition window, which are unique to SARS-CoV-2, confirming SARS-CoV-2 infection. Although more studies can further deepen the understanding of viral biopathology, this study offers both proteo-genomic analysis of SARS-CoV-2, confirming a high rate of mutation in Indian isolates, and the expression of the viral protein (Orf9b) in the host cell, which suppresses the host innate immune response. Enrichment of proteins involved in neutrophil-mediated immune response pointed toward the cross talk between the host and pathogen. Our study highlighted the potential of MS as a specific and sensitive diagnostic tool and laid down the foundation for future studies. Further studies combined with patient severity details can help in predicting the prognosis of viral infection.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00808.

List of SNPs identified in the sequenced genome; clade-wise arrangement of the sequenced genome; clinical characteristics of the samples; biology of SARS-CoV-2; GISAID nomenclature system for major clades; and proteome analysis of uninfected cells from SARS-CoV-2 negative samples (PDF), Variant summary and annotation (XLS), List of SARS-CoV-2 peptides identified in clinical swabs (XLSX), List of host proteins identified in positive samples (XLSX) List of host proteins identified in negative samples (XLSX) (ZIP)

AUTHOR INFORMATION

Corresponding Author

Utpal Tatu − Department of Biochemistry, Indian Institute of Science, Bangalore 560012, India; orcid.org/0000-0003-0404-8872; Phone: +91 (080) 22932823; Email: tatu@iisc.ac.in; Fax: 23600814

Authors

Sheetal Tushir − Department of Biochemistry, Indian Institute of Science, Bangalore 560012, India; orcid.org/0000-0002-8286-4448

Sathisha Kamanna − Department of Biochemistry, Indian Institute of Science, Bangalore 560012, India

Sujith S. Nath − Department of Biochemistry, Indian Institute of Science, Bangalore 560012, India

Aishwarya Bhat − Department of Biochemistry, Indian Institute of Science, Bangalore 560012, India

Steffimol Rose − Department of Biochemistry, Indian Institute of Science, Bangalore 560012, India

Advait R. Aithal − Department of Biochemistry, Indian Institute of Science, Bangalore 560012, India

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jproteome.0c00808

Author Contributions

ST and SK contributed equally. ST analyzed genomic and proteomic data. SK performed the MS work. SK and ST analyzed the MS data. SN and ST performed MS sample preparation. AB contributed to phylogenetic and genomic analysis. SR and ST performed pathway analysis for proteomic data. AA coordinated the clinical COVID-19 samples. UT conceived the idea of this study. UT and ST wrote the manuscript.

Notes

The authors declare no competing financial interest.
The MS and proteomics data obtained on clinical nasopharyngeal swabs have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with data set identifiers PXD021896 and 10.6019/PXD021896.

ABBREVIATIONS

SARS-CoV-2, Severe acute respiratory syndrome coronavirus 2; COVID-19, Coronavirus disease 2019; RT-PCR, Reverse transcription-polymerase chain reaction; GISAID, Global

Initiative on Sharing All Influenza Data; HRMS, High-resolution mass spectrometry; NGS, Next-generation sequencing; SNPs, Single-nucleotide polymorphisms; GO, Gene Ontology

■ REFERENCES

(1) Drosten, C.; Günther, S.; Preiser, W.; van der Werf, S.; Brodt, H.-R.; Becker, S.; Rabenau, H.; Panning, M.; Kolesnikova, L.; Fouchier, R. A. M.; Berger, A.; Burguière, A.-M.; Cinatl, J.; Eickmann, M.; Escriou, N.; Grywna, K.; Kramme, S.; Manuguerra, J.-C.; Müller, S.; Rickerts, V.; Stürmer, M.; Vieth, S.; Klenk, H.-D.; Osterhaus, A. D. M. E.; Schmitz, H.; Doerr, H. W. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* **2003**, *348*, 1967−1976.

(2) Ksiazek, T. G.; Erdman, D.; Goldsmith, C. S.; Zaki, S. R.; Peret, T.; Emery, S.; Tong, S.; Urbani, C.; Comer, J. A.; Lim, W.; Rollin, P. E.; Dowell, S. F.; Ling, A.-E.; Humphrey, C. D.; Shieh, W.-J.; Guarner, J.; Paddock, C. D.; Rota, P.; Fields, B.; DeRisi, J.; Yang, J.-Y.; Cox, N.; Hughes, J. M.; LeDuc, J. W.; Bellini, W. J.; Anderson, L. J. A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.* **2003**, *348*, 1953−1966.

(3) Zaki, A. M.; van Boheemen, S.; Bestebroer, T. M.; Osterhaus, A. D. M. E.; Fouchier, R. A. M. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* **2012**, *367*, 1814−1820.

(4) World Health Organization. *W. H. O. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)*. https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf, 2020.

(5) Wu, F.; Zhao, S.; Yu, B.; Chen, Y.-M.; Wang, W.; Song, Z.-G.; Hu, Y.; Tao, Z.-W.; Tian, J.-H.; Pei, Y.-Y.; Yuan, M.-L.; Zhang, Y.-L.; Dai, F.-H.; Liu, Y.; Wang, Q.-M.; Zheng, J.-J.; Xu, L.; Holmes, E. C.; Zhang, Y.-Z. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, *579*, 265−269.

(6) Snijder, E. J.; van der Meer, Y.; Zevenhoven-Dobbe, J.; Onderwater, J. J. M.; van der Meulen, J.; Koerten, H. K.; Mommaas, A. M. Ultrastructure and origin of membrane vesicles associated with the severe acute respiratory syndrome coronavirus replication complex. *J. Virol.* **2006**, *80*, 5927−5940.

(7) Hussein, H. A.; Hassan, R. Y. A.; Chino, M.; Febbraio, F. Point-of-Care Diagnostics of COVID-19: From Current Work to Future Perspectives. *Sensors* **2020**, *20*, 4289.

(8) Durrant, J. D.; Kochanek, S. E.; Casalino, L.; Ieong, P. U.; Dommer, A. C.; Amaro, R. E. Mesoscale All-Atom Influenza Virus Simulations Suggest New Substrate Binding Mechanism. *ACS Cent. Sci.* **2020**, *6*, 189−196.

(9) Gouveia, D.; Miotello, G.; Gallais, F.; Gaillard, J.-C.; Debroas, S.; Bellanger, L.; Lavigne, J.-P.; Sotto, A.; Grenga, L.; Pible, O.; Armengaud, J. Proteotyping SARS-CoV-2 Virus from Nasopharyngeal Swabs: A Proof-of-Concept Focused on a 3 Min Mass Spectrometry Window. *J. Proteome Res.* **2020**, *19*, 4407−4416.

(10) Ihling, C.; Tänzler, D.; Hagemann, S.; Kehlen, A.; Hüttelmaier, S.; Arlt, C.; Sinz, A. Mass Spectrometric Identification of SARS-CoV-2 Proteins from Gargle Solution Samples of COVID-19 Patients. *J. Proteome Res.* **2020**, *19*, 4389−4392.

(11) Nikolaev, E. N.; Indeykina, M. I.; Brzhozovskiy, A. G.; Bugrova, A. E.; Kononikhin, A. S.; Starodubtseva, N. L.; Petrotchenko, E. V.; Kovalev, G. I.; Borchers, C. H.; Sukhikh, G. T. 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.* **2020**, *19*, 4393−4397.

(12) Leggett, R. M.; Ramirez-Gonzalez, R. H.; Clavijo, B. J.; Waite, D.; Davey, R. P. Sequencing quality assessment tools to enable data-driven informatics for high throughput genomics. *Front. Genet.* **2013**, *4*, 288.

(13) Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **2011**, *17*, 10.

(14) *Severe Acute Respiratory Syndrome Coronavirus 2 Isolate Wuhan-Hu-1, Complete Genome*. https://www.ncbi.nlm.nih.gov/nuccore/NC_045512.2?report=genbank, 2020.

(15) Li, H.; Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **2010**, *26*, 589−595.

(16) Brouard, J.-S.; Schenkel, F.; Marete, A.; Bissonnette, N. The GATK joint genotyping workflow is appropriate for calling variants in RNA-seq experiments. *J. Anim. Sci. Biotechnol.* **2019**, *10*, 44.

(17) Cingolani, P.; Platts, A.; Wang, L. L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S. J.; Lu, X.; Ruden, D. M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **2012**, *6*, 80−92.

(18) Tamura, K.; Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **1993**, *10*, 512−526.

(19) Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K.; MEGA, X. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547−1549.

(20) Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; Pérez, E.; Uszkoreit, J.; Pfeuffer, J.; Sachsenberg, T.; Yılmaz, Ş.; Tiwary, S.; Cox, J.; Audain, E.; Walzer, M.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaíno, J. A. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **2019**, *47*, D442−d450.

(21) GISAID. https://www.gisaid.org/ (accessed January 20, 2020).

(22) Zecha, J.; Lee, C.-Y.; Bayer, F. P.; Meng, C.; Grass, V.; Zerweck, J.; Schnatbaum, K.; Michler, T.; Pichlmair, A.; Ludwig, C.; Kuster, B. Data, Reagents, Assays and Merits of Proteomics for SARS-CoV-2 Research and Testing. *Mol. Cell. Proteomics* **2020**, *19*, 1503−1522.

(23) Barh, D.; Tiwari, S.; Weener, M. E.; Azevedo, V.; Góes-Neto, A.; Gromiha, M. M.; Ghosh, P. Multi-omics-based identification of SARS-CoV-2 infection biology and candidate drugs against COVID-19. *Comput. Biol. Med.* **2020**, *126*, 104051.

(24) Su, Y.; Chen, D.; Yuan, D.; Lausted, C.; Choi, J.; Dai, C. L.; Voillet, V.; Duvvuri, V. R.; Scherler, K.; Troisch, P.; Baloni, P.; Qin, G.; Smith, B.; Kornilov, S. A.; Rostomily, C.; Xu, A.; Li, J.; Dong, S.; Rothchild, A.; Zhou, J.; Murray, K.; Edmark, R.; Hong, S.; Heath, J. E.; Earls, J.; Zhang, R.; Xie, J.; Li, S.; Roper, R.; Jones, L.; Zhou, Y.; Rowen, L.; Liu, R.; Mackay, S.; O'Mahony, D. S.; Dale, C. R.; Wallick, J. A.; Algren, H. A.; Zager, M. A.; Wei, W.; Price, N. D.; Huang, S.; Subramanian, N.; Wang, K.; Magis, A. T.; Hadlock, J. J.; Hood, L.; Aderem, A.; Bluestone, J. A.; Lanier, L. L.; Greenberg, P. D.; Gottardo, R.; Davis, M. M.; Goldman, J. D.; Heath, J. R. Multi-Omics Resolves a Sharp Disease-State Shift between Mild and Moderate COVID-19. *Cell* **2020**, *183*, 1479−1495.e20.

(25) Mercatelli, D.; Giorgi, F. M. Geographic and Genomic Distribution of SARS-CoV-2 Mutations. *Front. Microbiol.* **2020**, *11*, 1800.

(26) Mukherjee, M.; Goswami, S. Global cataloguing of variations in untranslated regions of viral genome and prediction of key host RNA binding protein-microRNA interactions modulating genome stability in SARS-CoV-2. *PloS One* **2020**, *15*, No. e0237559.

(27) Kim, D.; Lee, J.-Y.; Yang, J.-S.; Kim, J. W.; Kim, V. N.; Chang, H. The Architecture of SARS-CoV-2 Transcriptome. *Cell* **2020**, *181*, 914−921.e10.

(28) Grenga, L.; Gallais, F.; Pible, O.; Gaillard, J.-C.; Gouveia, D.; Batina, H.; Bazaline, N.; Ruat, S.; Culotta, K.; Miotello, G.; Debroas, S.; Roncato, M.-A.; Steinmetz, G.; Foissard, C.; Desplan, A.; Alpha-Bazin, B.; Almunia, C.; Gas, F.; Bellanger, L.; Armengaud, J. Shotgun proteomics analysis of SARS-CoV-2-infected cells and how it can optimize whole viral particle antigen production for vaccines. *Emerg. Microb. Infect.* **2020**, *9*, 1712−1721.

(29) Jiang, H.-w.; Zhang, H.-n.; Meng, Q.-f.; Xie, J.; Li, Y.; Chen, H.; Zheng, Y.-x.; Wang, X.-n.; Qi, H.; Zhang, J.; Wang, P.-H.; Han, Z.-G.; Tao, S.-c. SARS-CoV-2 Orf9b suppresses type I interferon responses by targeting TOM70. *Cell. Mol. Immunol.* **2020**, *17*, 998−1000.

(30) Han, L.; Zhuang, M.-W.; Zheng, Y.; Zhang, J.; Nan, M.-L.; Wang, P.-H.; Gao, C. SARS-CoV-2 ORF9b Antagonizes Type I and

III Interferons by Targeting Multiple Components of RIG-I/MDA-5-MAVS, TLR3-TRIF, and cGAS-STING Signaling Pathways. **2020**, bioRxiv, 2020.08.16.252973.

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

(32) Messner, C. B.; Demichev, V.; Wendisch, D.; Michalick, L.; White, M.; Freiwald, A.; Textoris-Taube, K.; Vernardis, S. I.; Egger, A.-S.; Kreidl, M.; Ludwig, D.; Kilian, C.; Agostini, F.; Zelezniak, A.; Thibeault, C.; Pfeiffer, M.; Hippenstiel, S.; Hocke, A.; von Kalle, C.; Campbell, A.; Hayward, C.; Porteous, D. J.; Marioni, R. E.; Langenberg, C.; Lilley, K. S.; Kuebler, W. M.; Mülleder, M.; Drosten, C.; Suttorp, N.; Witzenrath, M.; Kurth, F.; Sander, L. E.; Ralser, M. Ultra-High-Throughput Clinical Proteomics Reveals Classifiers of COVID-19 Infection. *Cell Syst.* **2020**, *11*, 11−24.e4.

(33) Camp, J. V.; Jonsson, C. B. A Role for Neutrophils in Viral Respiratory Disease. *Front. Immunol.* **2017**, *8*, 550.

(34) Haick, A. K.; Rzepka, J. P.; Brandon, E.; Balemba, O. B.; Miura, T. A. Neutrophils are needed for an effective immune response against pulmonary rat coronavirus infection, but also contribute to pathology. *J. Gen. Virol.* **2014**, *95*, 578−590.

(35) Zhu, B.; Zhang, R.; Li, C.; Jiang, L.; Xiang, M.; Ye, Z.; Kita, H.; Melnick, A. M.; Dent, A. L.; Sun, J. BCL6 modulates tissue neutrophil survival and exacerbates pulmonary inflammation following influenza virus infection. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 11888−11893.

(36) Veras, F. P.; Pontelli, M.; Silva, C.; Toller-Kawahisa, J.; de Lima, M.; Nascimento, D.; Schneider, A.; Caetite, D.; Rosales, R.; Colon, D.; Martins, R.; Castro, I.; Almeida, G.; Lopes, M. I.; Benatti, M.; Bonjorno, L.; Giannini, M.; Luppino-Assad, R.; Almeida, S.; Vilar, F.; Santana, R.; Bollela, V.; Martins, M.; Miranda, C.; Borges, M.; Pazin-Filho, A.; Cunha, L.; Zamboni, D.; Dal-Pizzol, F.; Leiria, L.; Siyuan, L.; Batah, S.; Fabro, A.; Mauad, T.; Dolhnikoff, M.; Duarte-Neto, A.; Saldiva, P.; Cunha, T.; Alves-Filho, J. C.; Arruda, E.; Louzada-Junior, P.; Oliveira, R.; Cunha, F. SARS-CoV-2 triggered neutrophil extracellular traps (NETs) mediate COVID-19 pathology. *J. Exp. Med.* **2020**, *217*, No. e20201129.

(37) Barnes, B. J.; Adrover, J. M.; Baxter-Stoltzfus, A.; Borczuk, A.; Cools-Lartigue, J.; Crawford, J. M.; Daßler-Plenker, J.; Guerci, P.; Huynh, C.; Knight, J. S.; Loda, M.; Looney, M. R.; McAllister, F.; Rayes, R.; Renaud, S.; Rousseau, S.; Salvatore, S.; Schwartz, R. E.; Spicer, J. D.; Yost, C. C.; Weber, A.; Zuo, Y.; Egeblad, M. Targeting potential drivers of COVID-19: Neutrophil extracellular traps. *J. Exp. Med.* **2020**, *217*, No. e20200652.

(38) Liu, J.; Liu, Y.; Xiang, P.; Pu, L.; Xiong, H.; Li, C.; Zhang, M.; Tan, J.; Xu, Y.; Song, R.; Song, M.; Wang, L.; Zhang, W.; Han, B.; Yang, L.; Wang, X.; Zhou, G.; Zhang, T.; Li, B.; Wang, Y.; Chen, Z.; Wang, X. Neutrophil-to-lymphocyte ratio predicts critical illness patients with 2019 coronavirus disease in the early stage. *J. Transl. Med.* **2020**, *18*, 206.

(39) Gjyshi, O.; Bottero, V.; Veettil, M. V.; Dutta, S.; Singh, V. V.; Chikoti, L.; Chandran, B. Kaposi's sarcoma-associated herpesvirus induces Nrf2 during de novo infection of endothelial cells to create a microenvironment conducive to infection. *PLoS Pathog.* **2014**, *10*, No. e1004460.

(40) Cecchini, R.; Cecchini, A. L. SARS-CoV-2 infection pathogenesis is related to oxidative stress as a response to aggression. *Med. Hypotheses* **2020**, *143*, 110102.

(41) Laforge, M.; Elbim, C.; Frère, C.; Hémadi, M.; Massaad, C.; Nuss, P.; Benoliel, J.-J.; Becker, C. Tissue damage from neutrophil-induced oxidative stress in COVID-19. *Nat. Rev. Immunol.* **2020**, *20*, 515−516.

(42) Schönrich, G.; Raftery, M. J.; Samstag, Y. Devilishly radical NETwork in COVID-19: Oxidative stress, neutrophil extracellular traps (NETs), and T cell suppression. *Adv. Biol. Regul.* **2020**, *77*, 100741.

(43) Delgado-Roche, L.; Mesta, F. Oxidative Stress as Key Player in Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) Infection. *Arch. Med. Res.* **2020**, *51*, 384−387.

(44) Bojkova, D.; Klann, K.; Koch, B.; Widera, M.; Krause, D.; Ciesek, S.; Cinatl, J.; Münch, C. Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* **2020**, *583*, 469−472.

(45) Rouzer, C. A.; Ivanova, P. T.; Byrne, M. O.; Brown, H. A.; Marnett, L. J. Lipid Profiling Reveals Glycerophospholipid Remodeling in Zymosan-Stimulated Macrophages†. *Biochemistry* **2007**, *46*, 6026−6042.