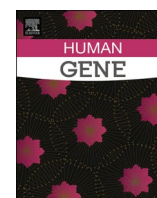




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

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



In silico transcriptional analysis of asymptomatic and severe COVID-19 patients reveals the susceptibility of severe patients to other comorbidities and non-viral pathological conditions

Poonam Sen, Harpreet Kaur^{*}

Pine Biotech, New Orleans, United States

ARTICLE INFO

Edited by: Serkan Yilmaz

Keywords:

SARS-CoV-2

2019-nCoV

COVID-19

Transcriptomics

Pathways

DGE (differentially expressed genes)

ARDS (acute respiratory distress syndrome)

ABSTRACT

COVID-19 is a severe respiratory disease caused by SARS-CoV-2, a novel human coronavirus. Patients infected with SARS-CoV-2 exhibit heterogeneous symptoms that pose pragmatic hurdles for implementing appropriate therapy and management of the COVID-19 patients and their post-COVID complications. Thus, understanding the impact of infection severity at the molecular level in the host is vital to understand the host response and accordingly its precise management. In the current study, we performed a comparative transcriptomics analysis of publicly available seven asymptomatic and eight severe COVID-19 patients. Exploratory data analysis employing Principal Component Analysis (PCA) showed the distinct clusters of asymptomatic and severe patients. Subsequently, the differential gene expression analysis using DESeq2 identified 1224 significantly upregulated genes ($\log_{2}FC \geq 1.5$, p-adjusted value < 0.05) and 268 significantly downregulated genes ($\log_{2}FC \leq -1.5$, p-adjusted value < 0.05) in severe samples in comparison to asymptomatic samples. Eventually, Gene Set Enrichment Analysis (GSEA) revealed the upregulation of anti-viral and anti-inflammatory pathways, secondary infections, Iron homeostasis, anemia, cardiac-related, etc.; while, downregulation of lipid metabolism, adaptive immune response, translation, recurrent respiratory infections, heme-biosynthetic pathways, etc. Conclusively, these findings provide insight into the enhanced susceptibility of severe COVID-19 patients to other health comorbidities including non-viral pathogenic infections, atherosclerosis, autoinflammatory diseases, anemia, male infertility, etc. owing to the activation of biological processes, pathways and molecular functions associated with them. We anticipate this study will facilitate the researchers in finding efficient therapeutic targets and eventually the clinicians in management of COVID-19 patients and post-COVID-19 effects in them.

1. Introduction

Since its first reported case at the end of 2019, an acute respiratory syndrome causing novel Coronavirus (2019-nCoV) outbreak in the human population has taken the world by storm. The 2019-nCoV was later officially named SARS-CoV-2 (Severe Acute Respiratory Syndrome related novel Coronavirus 2), and the disease caused by it COVID-19 (Coronavirus Disease 2019) (*Timeline: WHO's COVID-19 Response, 2022; Director-General, 2020; Gorbalenya et al., 2020*). The virus spread uncontrollably so much that in January 2020, World Health Organization declared COVID-19 as a “public health emergency of international concern” (PHEIC) and eventually as a pandemic in March 2020 (*Timeline: WHO's COVID-19 Response, 2022*). As of 1st April 2022, the total reported cases worldwide stand at 488,190,137 (Worldometers.info,

2022). The SARS-CoV-2 is an enveloped, positive single-stranded RNA virus that belongs to the Coronaviridae family, β -coronavirus genus, and is believed to have a zoonotic to human transmission (*Gorbalenya et al., 2020; Asselah et al., 2021; Zhou et al., 2020*). The trimeric spike (S) protein is key protein that play role in interaction with host; in which each monomer consists of two subunits, S1 and S2 (*Huang et al., 2020a*). The receptor-binding domain of the S1 subunit recognizes and binds to angiotensin-converting enzyme 2 (ACE2) receptor of the host, and the two-heptad repeat domain of S2 subunit mediates viral cell membrane fusion by forming a six-helical bundle (*Huang et al., 2020a; Letko et al., 2020*).

There are six other coronaviruses, i.e., 229E, OC43, NL63, HKU1, SARS-CoV, and MERS-CoV, which are already known to infect humans and cause respiratory and gastrointestinal problems (*Zaim et al., 2020*).

^{*} Corresponding author at: Curriculum Developer & Research Consultant, Pine Biotech, 1441 Canal St #229, New Orleans, LA 70112, United States.
E-mail address: hks04180@gmail.com (H. Kaur).

These human coronaviruses (HCoVs) are generally considered inconsequential except for our experience with SARS-CoV in 2003, MERS in 2012, and SARS-CoV-2 with the ongoing pandemic (Paules et al., 2020).

The mutations in the viral spike protein components, especially in its receptor-binding domain, have resulted in the generation of multiple variants, of which Delta variant (B.1.617.2) became a “variant of concern” (VOC) and posed a significant threat to human health (Guru-prasad, 2021; Dudas et al., 2021; Kumar et al., 2022). Our health sector has faced major challenges in tackling disease spread and providing management of symptoms in the patients (Fegert et al., 2020). Multiple drugs are introduced for symptomatic treatments, but none has been efficient to treat all symptoms caused due to the viral infection (Therapeutics and COVID-19: Living Guideline, 2022). Even a few drugs that were believed to be helpful in COVID-19 disease management were later found to cause other health concerns in the patients administered with these (Jean and Hsueh, 2020). The difficulty faced in devising standard therapeutic options is due to the high mutability rate of the virus, a complex interplay of virus-host interaction, and an individual's immune response to the infection (Ballow and Haga, 2021; Bakhshandeh et al., 2021; Varghese et al., 2020; Fung and Babik, 2021).

SARS-CoV-2 impacts individuals in peculiar ways (Ballow and Haga, 2021). Most infected subjects are asymptomatic or mildly symptomatic, but some develop severe symptoms (Ballow and Haga, 2021). Comorbidities such as diabetes mellitus, hypertension, cardiovascular disease (CVD), and advanced age further increase the risk of disease severity (Wu et al., 2020; Yang et al., 2020; Wallentin et al., 2020; Chen et al., 2021a). As in many asymptomatic or mild cases, diagnostic test reports false-negative results even in the presence of infection, and due to the shared spectrum of symptoms with other viral infections, it becomes difficult to discern COVID-19 from other viral infections (McIntosh, 2022). This makes disease management further complicated. Another primary concern associated with COVID-19 is high infectivity as it spreads by human contact and through air droplets and aerosols, making it difficult to control (van Doremalen et al., 2020; Rothan and Byrreddy, 2020). COVID-19 spread through fecal matter is speculated in some studies, though the presence of viral particles in the fecal samples of infected individuals is well documented and makes it an essential diagnostic tool (Brojna et al., 2021; Jones et al., 2020). The main clinical manifestations of SARS-CoV-2 in severe COVID-19 patients involve lower respiratory tract issues resulting in Acute Respiratory Distress Syndrome (ARDS) and hypoxia, fever, cytokine storm due to hyperactive immune system, brain fog, headache, cardiac arrest, and multi-organ damage and even death in severe cases (Wallentin et al., 2020; Huang et al., 2020b; Wang et al., 2020a; Sher, 2021; Scordo et al., 2021; Devaux et al., 2020). Most disease symptoms may persist for 10–15 days, with some may exist for a prolonged time (Klopfenstein et al., 2020; Faes et al., 2020). It is well known that even after the viral load declines significantly, many health issues persist in the COVID-19 recovered patients (Rogers et al., 2020; Lamprecht, 2020; Fernandez-de-Las-Penas et al., 2021). These post-COVID effects are observed mainly in hospitalized and severe patients and add to another layer of disease mismanagement (Carfi et al., 2020; Garrigues et al., 2020). So, the significant challenges of disease management include SARS-CoV-2's high infectivity rate, poor efficacy of available treatments, the complexity of symptoms, and less understanding of disease progression (Lee and Hsueh, 2020).

SARS-CoV-2, upon entry into the nasopharyngeal tract, interacts with the transmembrane serine protease 2 (TMPRSS2) and Angiotensin-Converting Enzyme 2 (ACE2) receptors present on the endothelial cells of the respiratory tract (Hoffmann et al., 2020). ACE2 receptors are also present in other organs, such as the gastrointestinal tract, lymph nodes, thymus, bone marrow, spleen, liver, kidney, skin, and brain. This might be the possible reason for the viral impact on these organs (Devaux et al., 2020; Lamers et al., 2020; Hamming et al., 2004; Li, 2013; Sungnak et al., 2020). As extensively studied, virus entry in these organs is mediated through the interaction of receptor-binding domain on spike

protein of virus and the ACE2 receptors present on host cells (Li, 2013; Magrone et al., 2020). Upon infection, the virus replicates inside the host cell using the host replication machinery. In response to all this, the host immune system fights to reduce the viral load by inhibiting the replication of viral RNA. The diverse symptoms results from the involvement of various biochemical pathways triggered by viral entry and replication, the host cellular response to control the spread of the infection (Meredith Wadman et al., 2020).

With the advancement in the RNA sequencing technology, one can view the transcriptomic landscape under a given condition and for a particular cell type. It is also instrumental in understanding the pathogenesis of a disease in the host (Su et al., 2020). Diverse scientific groups across the globe have developed numerous resources and tools to compile and analyze the data from host and pathogen (NCBI, 2022; COVID Data Tracker, 2022; Mathieu et al., 2021; Talwai et al., 2021; Wang et al., 2020b; Chen et al., 2021b; Patiyal et al., 2020; Ouyang et al., 2021; Holy et al., 2020; Zettler et al., 2021; Gong et al., 2020; Singh et al., 2020a; Faleiros et al., 2022; Riffe and Acosta, 2021; Mahdi et al., 2021; Park et al., 2022; Xu et al., 2021; Zheng et al., 2020; Mulholland et al., 2021; Khullar and Chandra, 2020; Benetti et al., 2021; Saffern, 2020; Vetrugno et al., 2021; Haider et al., 2021).

Due to the systemic effects of COVID-19 infection, it becomes more challenging to treat patients with complex symptoms. Thus, the current study is an attempt to understand the underlying changes taking place in the host at the molecular or transcriptional level with respect to severity of infection. Towards this, we extensively explored transcriptional profiles of severe and asymptomatic COVID-19 patients using various Insilco methods for the identification of key significant genes that are significantly differentially expressed between these two groups. Eventually, Gene Set Enrichment Analysis (GSEA) was performed on identified significantly differentially expressed genes to understand the biological pathways, processes and functions that are significantly disturbed between severe and asymptomatic COVID-19 patients.

2. Methods

2.1. Dataset and experimental design

In the current study, we obtained publicly available data (GSE178967) from the NCBI GEO (Gene Expression Omnibus). This dataset comprises RNA-Seq read counts and metadata information conducted on 108 SARS-CoV-2 subjects by the Stanford COVID-19 CTRU (Jagannathan et al., 2021). These COVID-19 subjects, confirmed by RT-PCR, were administered Peginterferon Lambda and placebo on day00. Peginterferon Lambda is a therapeutic drug for reducing the viral particles in COVID-19 patients (Feld et al., 2021). Whole blood samples for RNA extraction for high throughput sequencing were collected on day 00 (untreated) and day 05 (treated) from the day of drug administration. The available RNA sequencing data are the read counts aligned to transcripts or genes for 180 samples from day 00 and day 05 of 108 subjects. In the series matrix file (provided in GEO), the COVID-19 subjects are categorized as asymptomatic, moderately symptomatic, and severe (Jagannathan et al., 2021). We have also used the same categorization of subjects for our analysis. The series matrix file contains other clinically significant information such as age, gender, day from drug administration (Peginterferon Lambda and placebo), and viral shedding value. The details and data structure of the study are summarized in Table 1 and Supplementary Table S1, respectively. The summary of clinical information extracted from the GEO series matrix is provided in Supplementary Table S2.

2.2. Data preparation and normalization

2.2.1. Data pre-processing

The data contains sample IDs in row 1, transcript IDs (ENST ID) in column 1, gene symbols in column 2, and corresponding non-

Table 1

Detail of the study as derived from GEO (*Timeline: WHO's COVID-19 Response, 2022*).

GEO accession	Series GSE178967
Study title	Baseline signatures associated with clinical, virologic, and immunologic outcomes in patients with mild to moderate COVID-19
Organism	<i>Homo Sapiens</i>
Bio project	PRJNA741686
SRA	SRP325729
Platform	GPL24676
	Illumina NovaSeq 6000

normalized read counts in the matrix as integer values. The sample IDs belong to asymptomatic, moderately symptomatic, or severe subjects from day 0 or day 5 of peginterferon lambda and placebo administration. The RNA sequencing expression values of the dataset are non-normalized read counts (as mentioned in supplementary file information of the original dataset submitted in GEO) (Jagannathan et al., 2021). These read counts are the number of reads mapped and aligned to a particular transcript/gene region identified from the human reference genome. It is generally required to pre-process the read count data to get statistically significant results (Han et al., 2012a; Malley et al., 2016; Fan et al., 2021; Garca et al., 2014; Zelaya, 2019). We followed common pre-processing steps for both PCA and Differential Gene Expression (DGE) analysis, but the normalization steps were different based on the downstream analyses. The Principal Component Analysis is a dimensionality reduction unsupervised machine learning method that requires normalized data (Fan et al., 2021; Garca et al., 2014; Lever et al., 2017; Dinc et al., 2014). While DESeq2 is a DGE analysis tool that mandates data to be unprocessed read counts as integer values (Love et al., 2014). DESeq2 uses inbuilt methods to normalize for library size and hence does not require prior normalization (Love et al., 2014; Zhu et al., 2018; Vaja et al., 2021).

The summary of workflow, including pre-processing and normalization, is depicted in Fig. 1. In pre-processing, we removed rows with NA, taken the average of duplicates genes using aggregate function in R, and filtered out the genes having zero or low expression. Studies suggest that low expression genes negatively impact the Differentially Expressed

Gene analysis (Sha et al., 2015). Thus, genes with low or zero variance filtered out using the `nzv` (non-zero variance) function of the “Caret package” available in R (Kuhn, 2008). Genes with zero variance across all samples are considered insignificant as these do not contribute to statistical significance and only increase time in the analysis (Malley et al., 2016; Han et al., 2012b). After removing genes with low expression values, we performed further pre-processing specific to PCA, and for DESeq2, we continued with the pre-processed and non-normalized data. Notably, we performed exploratory data analysis using PCA on normalized data while differential gene expression analysis using DESeq2 on raw read count values.

2.2.2. Data normalization

After the abovementioned pre-processing, subsequently, for PCA, we normalized the read counts by transforming them to log values and then performing center and scaling using the “Caret package” available in R (Kuhn, 2008). The data matrix that resulted from PCA normalization contains 180 samples with log-transformed read counts for 35,587 gene rows.

2.3. Analysis methodology

2.3.1. Exploratory data analysis using principal component analysis

We performed Principal Component Analysis (PCA) to identify the patterns in the dataset and variations between the samples in a group. PCA reduces the dimensions of a large dataset while retaining most of the variations. Hence, PCA assists in identifying sample clusters in a particular group and outliers (Jolliffe, 2002). We performed PCA on normalized data (comprising 180 samples with log-transformed read counts for 35,587 gene rows) using the `ggfortify` package in R. The first PCA includes all 180 samples of asymptomatic, moderately symptomatic, and severe subjects. Then we performed PCA for various groups as mentioned in Supplementary Table S3. One of these PCA, consists of asymptomatic and severe patients at Day 00 (untreated) which we believe will help us understand the host response mechanism in severe patients in comparison to asymptomatic. The total number of samples belonging to this group was 15, with seven asymptomatic and eight severe samples. With the help of scatterplots based on PCA components, we identified outliers, which were subsequently removed from the data

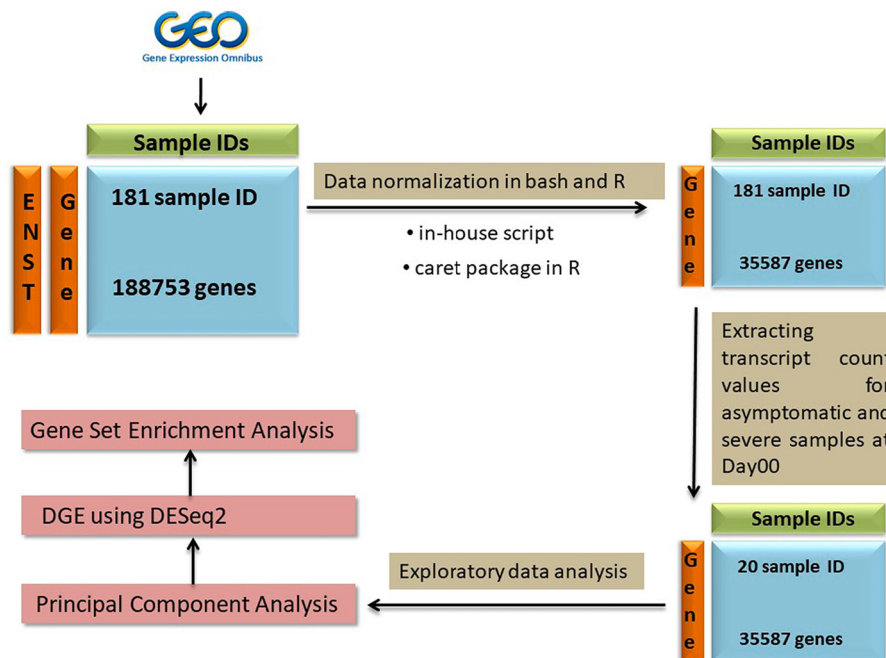


Fig. 1. The complete workflow of the study.

for the downstream PCA and DGE analysis on untreated (Day 00) group.

2.3.2. Differential gene expression analysis

After outliers removal using PCA, we performed differential gene expression analysis between severe and asymptomatic patients' samples using the DESeq2 package in R (Love et al., 2014). Notably, we considered only those genes as significantly expressed between groups with a p-adjusted value < 0.05 . This criterion of p-adjusted value is used in numerous studies (Love et al., 2014; Haynes, 2013; Armstrong, 2014; Kaur et al., 2020a; Menyhart et al., 2021; Jafari and Ansari-Pour, 2019; Streiner, 2015; Kaur et al., 2019; Dolaner et al., 2021; Kaur et al., 2020b). Further, we applied another filter, i.e., Log2 fold change (Log2FC) to identify significantly upregulated ($\text{Log2FC} \geq 1.5$) and downregulated ($\text{Log2FC} \leq -1.5$) genes in the severe patients in comparison to asymptomatic patients. Additionally, to understand patterns in gene expression between asymptomatic and severe patients, we constructed heatmaps using the heatmap function in R (Perry and heatmaps, 2021). Heatmap is a grid-like graphical representation of the expression of genes (in rows) in all the samples (in columns) taken into consideration (Khomtchouk et al., 2014).

2.3.3. Biological annotation

Subsequently, to understand the biological implication of significantly differentially expression genes obtained from DESeq2 analysis in severe patients, we performed gene enrichment analysis using the Enrichr (Chen et al., 2013; Kuleshov et al., 2016; Xie et al., 2021). We

queried the upregulated and downregulated gene sets independently in the Enrichr search engine (Chen et al., 2022). Enrichr gives various Gene Set Enrichment terms as output which can be analyzed for significance based on four ranking parameters, i.e., p-value, adjusted p-value, odds ratio, combined scores (Xie et al., 2021). Enrichr visualization bar graph shows top enrichment terms with significance depicted by the length and color of the bar. An enrichment term with a more extended bar and a lighter shade of red indicate higher significance than a term with a shorter bar and darker red color or grey color (Xie et al., 2021). A few of the top Gene Set Enrichment terms are, i.e., KEGG Human, WikiPathway, Gene Ontology (GO) terms, Jensen diseases, Human phenotype ontology, etc., based on p-value (< 0.05). To identify significant pathways involved in each enrichment term, we used the q-value (adjusted p-value) < 0.05 . Besides, we searched for the top significant and differentially expressed genes (from our analysis) in the literature to understand their already known role in COVID-19 pathogenesis.

3. Results

In the current study, we analyzed the transcriptomic profiles of asymptomatic and severe COVID-19 patients to compare the transcriptional changes and understand the biological implications of infection.

3.1. Data pre-processing

The data pre-processing able to remove genes without identifiers,

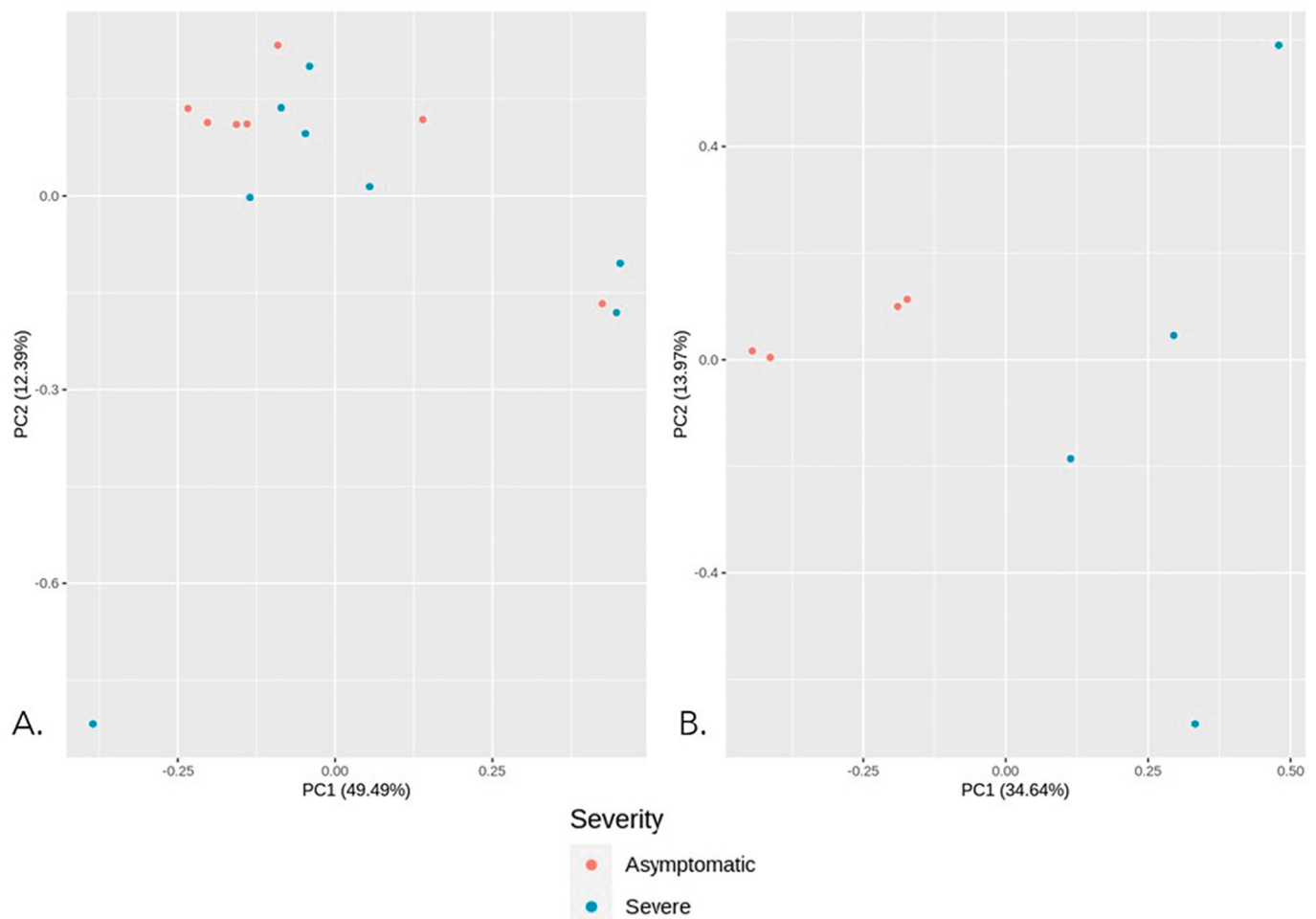


Fig. 2. Principal component analysis between untreated asymptomatic and severe groups. A. PCA between asymptomatic ($n = 7$, Day00) v/s severe samples ($n = 8$, Day00) B. PCA between asymptomatic ($n = 7$, Day00) v/s severe samples ($n = 8$, Day00) after outlier removal.

zero expression, and low variance in the data. Thus, the total number of transcripts reduced from 188,753 to 35,587 in the data. Subsequently, this dataset was used for exploratory and DGE analysis.

3.2. Exploratory data analysis

We analyzed each group's scatter plot and principal components to identify if any of the top Principal Components (PC) showed significant variations. The scatter plots for all Principal Component Analysis performed are provided in Fig. 2 and Supplementary Fig. 1 A-C. The scatter plot in Supplementary Fig. 1.A represents all three groups, i.e., untreated and treated asymptomatic, moderately symptomatic, and severe. However, three outliers can be observed at the bottom left of the plot; the clustering does not show any clear distinction between the three groups. PCA for remaining groups, i.e., severe male v/s female, severe below 45 years age v/s above 45 years age (Supplementary Fig. S1 B and C) also did not show any clear groups. The top principal components in these groups also did not show significant variations. Thus, we mainly shifted our focus to two critical groups, i.e., asymptomatic (untreated or day 00) and severe (untreated or day 00), since these conditions representing two contrasting viral infection conditions and interestingly, they also represent lesser within-group variation. The PCA between untreated asymptomatic ($n = 7$) and untreated severe samples ($n = 8$) represent nearly 61.8% variation in the data, wherein PC1 contributes 49.49%, and PC2 contributes ~12.39% variation (Fig. 2A). Using the clustering patterns in PCA, we identified seven samples as outliers. We removed those outlier samples and then performed PCA on the remaining eight samples (four severe and four asymptomatic), that represented nearly 48.61% variation in the data, where PC1 represents 34.64%, and PC2 represents 13.97% variation. Scatterplots based on the PC1 and PC2 of untreated asymptomatic and severe samples show clear distinction, and we got down to 4 samples in each group (Fig. 2B).

3.3. Differential gene expression analysis

Differential gene expression analysis between untreated severe and asymptomatic samples using DESeq2 identified 2837 genes as significantly differentially expressed (p adjusted value < 0.05). From these 2837 genes, 1224 genes were found to be significantly upregulated ($\text{Log2FC} \geq 1.5$, p -adjusted value < 0.05) and 268 genes as significantly downregulated ($\text{Log2FC} \leq -1.5$, p -adjusted value < 0.05) in severe samples in comparison to asymptomatic samples. The list of the total up- and downregulated genes is provided in Supplementary Tables S4 and S5, respectively. The volcano plot represents the pattern of differentially expressed genes (Fig. 3). Each dot in the plot represents a single gene with log2FC along the x-axis and $-\text{Log}_{10}(p\text{-value})$ along the y-axis. In the volcano plot, the genes depicted in black color are nonsignificant, while genes in blue and red color represent most significantly differentially expressed genes with $\text{padj} < 0.01$ and $\text{padj} < 0.05$, respectively. Further, heatmap (Fig. 4) represents the expression pattern of the top 50 genes (25 upregulated and 25 downregulated genes) in untreated severe COVID-19 samples in comparison to asymptomatic samples. The top 25 upregulated and down regulated genes are mentioned in Table 2. with respective gene description.

3.3.1. Biological annotation - gene enrichment analysis

We queried significantly up and down regulated genes obtained from DGE analysis to the Enrichr search engine independently. The resulting bar plots represent the top enriched terms for upregulated genes (Figs. 5, Supplementary Fig. S2-S4) and downregulated genes (Figs. 6, Supplementary Fig. S5). We also extracted the complete results of all enriched terms for both upregulated (see Table S6-S21, Supplementary File 2) and downregulated gene sets (see Table S22-S36, Supplementary File 2) as tables. Besides, we also studied the significant terms and searched in the literature whether these are associated with COVID-19 pathogenesis previously. The key terms are briefly explained below:

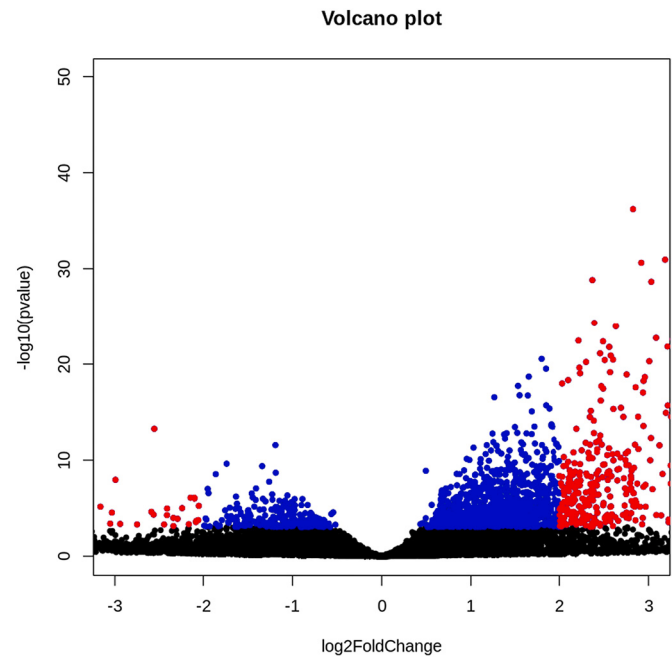


Fig. 3. Volcano plot based on p -value and log2FC . Each dot here represents a single gene. Black represents nonsignificant genes, blue and red represent genes differentially regulated at $\text{padj} < 0.01$ and $\text{padj} < 0.05$.

3.3.1.1. Gene set enrichment analysis of upregulated genes

3.3.1.1.1. Association with the viral infection and inflammatory response. Immune response terms that were found to be enriched for upregulated gene set include decreased interleukin-12b secretion MP:0008670; decreased B cell proliferation MP:0005093; abnormal interleukin level MP:0008751; impaired natural killer cell-mediated cytotoxicity MP:0005070; increased prostaglandin level MP:0009814; lymph node hyperplasia MP:0008102, Oncostatin M Signaling Pathway WP2374. Further, enriched terms found to associated with response to viral infection like Type II interferon signaling (I.F.N.G.) (WP619), IL-18 signaling pathway (WP4754), IL8 signaling (WP4754), Structural Pathway of Interleukin 1 (IL-1) (WP2637), IL-6 signaling pathway (WP364), decreased interferon-alpha secretion (MP:0008563), IL-4 signaling pathway (WP395), decreased interleukin-1 beta secretion (MP:0008658), abnormal T-helper 2 physiology (MP:0005466), abnormal macrophage physiology (MP:0002451), abnormal T-helper 1 physiology (MP:0005465), abnormal granulocyte physiology (MP:0002462), sepsis (MP:0005044). While the enriched terms related to anti-inflammatory and immune response are Activation of NLRP3 Inflammasome by SARS-CoV-2 (WP4876), abnormal inflammatory response (MP:0001845), IL-10 Anti-inflammatory Signaling Pathway WP4495.

3.3.1.1.2. Association with secondary infections. Interestingly, we found the enrichment of upregulated genes with terms that are associated with various infections other than COVID-19. These enriched terms are Influenza A, Epstein-Barr virus infection, Kaposi sarcoma-associated herpesvirus infection, *Staphylococcus aureus* infection, Measles, Human immunodeficiency virus 1 infection, Hepatitis C, increased susceptibility to bacterial infection (MP:0002412), Recurrent gram-negative bacterial infections (HP:0005420), increased susceptibility to fungal infection (MP:0005399), increased susceptibility to bacterial infection (MP:0002412), increased susceptibility to Picornaviridae infection (MP:0020937), Kaposi sarcoma-associated herpesvirus infection, increased susceptibility to Riboviria infection (MP:0020913), increased susceptibility to Herpesvirales infection (MP:0020916). Enrichment terms related to nutrients for upregulated gene set were Copper homeostasis WP3286, Vitamin B12 Disorders WP4271, and Zinc

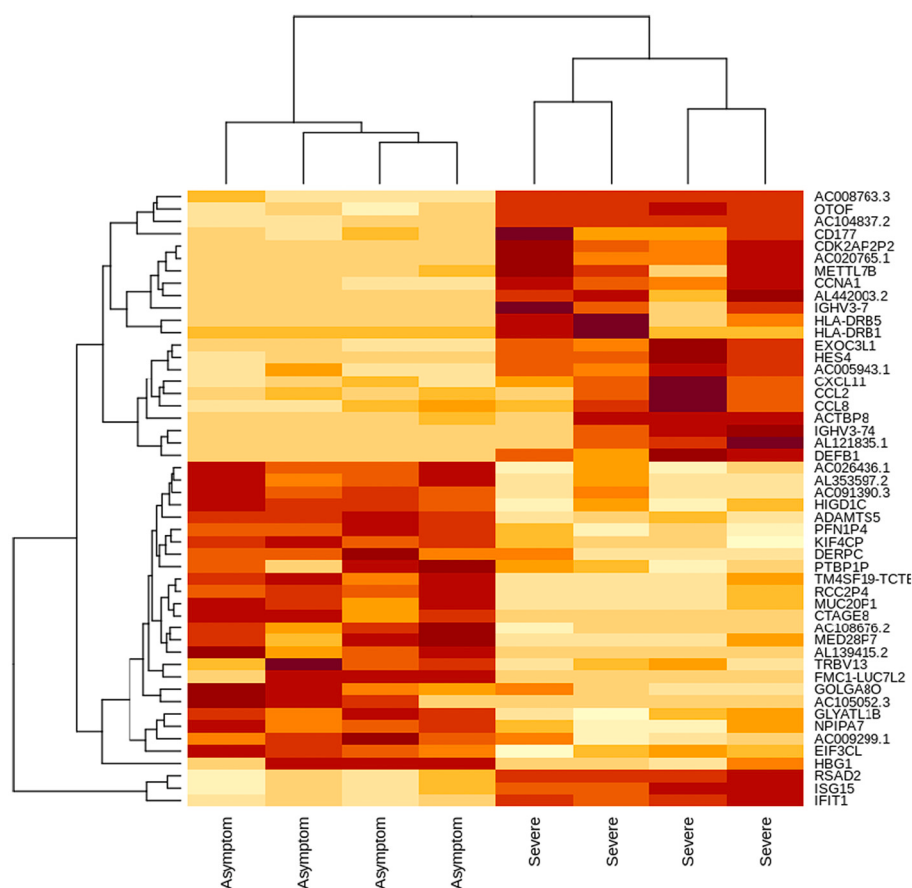


Fig. 4. Heatmap based on top 25 upregulated and downregulated genes from DESEQ2 of asymptomatic and severe samples. The color scale denotes the expression values in the heatmap. The red color's intensity represents upregulated genes, and the yellow color's intensity represents downregulated genes in the sample under consideration.

homeostasis WP3529. Iron homeostasis enrichment terms in upregulated gene sets are Ferroptosis WP4313, Folate Metabolism WP176, abnormal iron homeostasis MP:0005637, decreased spleen iron level MP:0008808, Abnormality of iron homeostasis (HP:0011031).

3.3.1.1.3. Association with organs other than the respiratory system. We also observed enriched pathways related to various organs, such as kidney-related glomerulonephritis MP:0002743; renal glomerular immunoglobulin deposits MP:0020519; and liver-related increased liver iron level MP:0008807. Heart-related Adrenergic signaling in cardiomyocytes (KEGG), myocarditis MP:0001856, Extracellular vesicles in the crosstalk of cardiac cells WP4300, ApoE, and miR-146 in inflammation and atherosclerosis WP3926, arrhythmogenic right ventricular dysplasia (Diseases) [implication of *JUP* gene in ARVD], Arrhythmogenic right ventricular cardiomyopathy (KEGG), cholesterol level (OMIM Diseases) [implication of *VNN1* gene], myocardial infarction (OMIM Diseases) [implication of *PSMA6* gene], cardiomyopathy, (OMIM Diseases) [implication of *MYBPC3* gene], Chronic obstructive pulmonary disease (HP:0006510), Abnormality of lateral ventricle (HP:0030047), Abnormality of the carotid arteries (HP:0005344); Arteriovenous malformation (HP:0100026); Arterial thrombosis (HP:0004420). Pathways enriched for the intestine are "Duodenal and small intestinal stenosis," "abnormal gut flora balance" MP:0010377, and those related to the skin were hypopigmented skin patches (HP:0001053), Urticaria (HP:0001025), Recurrent skin infections (HP:0001581), Hyper melanotic macule (HP:0001034), Recurrent bacterial skin infections (HP:0005406), Eczematoid dermatitis (HP:0000976) skin hemorrhage MP:0011514. Brain related neurological and behavioral pathways found were Inappropriate behavior (HP:0000719), Personality changes (HP:0000751), Diminished

motivation (HP:0000745), Dementia (HP:0000726), Memory impairment (HP:0002354), Restlessness (HP:0000711), Vertigo (HP:0002321), Neuroinflammation WP4919, Galanin receptor pathway WP4970, Meningitis (HP:0001287).

3.3.1.1.4. Association with other important pathways. Enrichment analysis of upregulated genes set shown the association with the male infertility WP4673, Abnormality of the preputium (HP:0100587), and Erectile abnormalities (HP:0100639). Further, we found the enrichment of upregulated genes in Ferritin, an inflammatory marker used in COVID-19 prognosis, Transcriptional cascade regulating adipogenesis WP4211, Fibrin Complement Receptor 3 Signaling Pathway WP4136. Other WikiPathway that are observed to be significantly upregulated in severe patients are IL1 and megakaryocytes in obesity (WP2865); Adipogenesis (WP236); Non-genomic actions of 1,25 dihydroxy vitamin D3 (WP4341); Vitamin D Receptor Pathway (WP2877); Myometrial relaxation and contraction pathways (WP289); Extracellular vesicles in the crosstalk of cardiac cells (WP4300). Pathways enriched related to blood cells are thrombocytopenia MP:0003179; abnormal myelopoiesis MP:0001601; impaired hematopoiesis MP:0001606; increased spleen weight MP:0004952. Descartes_Cell_Tissue 2021 shows Myeloid cells, Microglia, Antigen-presenting cells in the Thymus, Erythroblasts, Megakaryocytes in the Heart, Corneal and conjunctival epithelial cells in Eye, Vascular endothelial cells enrichment. Jensen diseases database indicates the association of upregulated genes with Arthritis, Peritonitis, Vasculitis, Periodontitis, Tularemia, Lupus Erythematosus, Boutonneuse fever, Hemochromatosis. The enriched GO cellular function(s) were azurophil granule (GO:0042582); ficolin-1-rich granule (GO:0101002); platelet alpha granule (GO:0031091). KEGG Human 2021 terms enriched in upregulated gene sets are NOD-like receptor signaling

Table 2

List of top 25 upregulated ($\text{Log2FC} \geq 1.5$, p-adjusted value <0.05) and down-regulated ($\text{Log2FC} \leq -1.5$, p-adjusted value <0.05) genes in severe COVID-19 subjects in comparison to asymptomatic subjects with their gene description.

TOP 25 UPREGULATED GENES		TOP 25 DOWNREGULATED GENES	
Gene name	Gene description	Gene name	Gene description
HLA-DRB1	Protein coding, Major Histocompatibility Complex, Class II, DR Beta 1 (Stelzer et al., 2016a; Stelzer et al., 2016b)	AC105052.3	sense overlapping (Lang et al., 2019)
HLA-DRB5	Protein coding, Major Histocompatibility Complex, Class II, DR Beta 5 (Stelzer et al., 2016a; Stelzer et al., 2016b)	FMC1-LUC7L2	Protein coding, FMC1-LUC7L2 Readthrough (Stelzer et al., 2016a; Stelzer et al., 2016b)
IGHV3-7	Protein coding, Immunoglobulin Heavy Variable 3-7 (Stelzer et al., 2016a; Stelzer et al., 2016b)	CTAGE8	Protein coding, CTAGE Family Member 8 (Stelzer et al., 2016a; Stelzer et al., 2016b)
ACTBP8	Pseudogene, ACTB Pseudogene 8 (Stelzer et al., 2016a; Stelzer et al., 2016b)	AL139415.2	Processed pseudogenes (Lang et al., 2019)
AC008763.3	Novel protein (Uhlen et al., 2015; Uhlen et al., 2010)	ADAMTS5	Protein coding, ADAM Metallopeptidase with Thrombospondin Type 1 Motif 5 (Stelzer et al., 2016a; Stelzer et al., 2016b)
CD177	Protein coding, CD177 Molecule (Stelzer et al., 2016a; Stelzer et al., 2016b)	MED28P7	Pseudogene, Mediator Complex Subunit 28 Pseudogene 7 (Stelzer et al., 2016a; Stelzer et al., 2016b)
AL442003.2	NA	MUC20P1	Pseudogene, Mucin 20, Cell Surface Associated Pseudogene 1 (Stelzer et al., 2016a; Stelzer et al., 2016b)
IGHV3-74	Protein coding, Immunoglobulin Heavy Variable 3-74 (Stelzer et al., 2016a; Stelzer et al., 2016b)	RCC2P4	Pseudogene, Regulator Of Chromosome Condensation 2 Pseudogene 4 (Stelzer et al., 2016a; Stelzer et al., 2016b)
ISG15	Protein coding, ISG15 Ubiquitin Like Modifier (Stelzer et al., 2016a; Stelzer et al., 2016b)	PFN1P4	Pseudogene, Profilin 1 Pseudogene 4 (Stelzer et al., 2016a; Stelzer et al., 2016b)
CCL2	Protein coding, C—C Motif Chemokine Ligand 2 (Stelzer et al., 2016a; Stelzer et al., 2016b)	AL353597.2	processed transcript, transcribed processed pseudogene (Lang et al., 2019)
CCL8	Protein coding, C—C Motif Chemokine Ligand 8 (Stelzer et al., 2016a; Stelzer et al., 2016b)	HBG1	Protein coding, Hemoglobin Subunit Gamma 1 [106, 107]
OTOF	Protein coding, Otoferlin (Stelzer et al., 2016a; Stelzer et al., 2016b)	GOLGA80	Protein coding, Golgin A8 Family Member O (Stelzer et al., 2016a; Stelzer et al., 2016b)
RSAD2	Protein coding, Radical S-Adenosyl Methionine Domain Containing 2 (Stelzer et al., 2016a; Stelzer et al., 2016b)	TM4SF19-TCTEX1D2	RNA Gene, TM4SF19-DYNLT2B Readthrough (NMD Candidate) (Stelzer

Table 2 (continued)

TOP 25 UPREGULATED GENES		TOP 25 DOWNREGULATED GENES	
Gene name	Gene description	Gene name	Gene description
METTL7B	Protein coding, Methyltransferase Like 7B (Stelzer et al., 2016a; Stelzer et al., 2016b)	DERPC	Protein coding, DERPC Proline and Glycine Rich Nuclear Protein (Stelzer et al., 2016a; Stelzer et al., 2016b)
HES4	Protein coding, Hes Family BHLH Transcription Factor 4 (Stelzer et al., 2016a; Stelzer et al., 2016b)	TRBV13	Protein coding, T Cell Receptor Beta Variable 13 [106, 107]
CDK2AP2P2	Pseudogene, PTGER4P2-CDK2AP2P2 Readthrough, Transcribed Pseudogene (Stelzer et al., 2016a; Stelzer et al., 2016b)	AC108676.2	Processed pseudogenes (Lang et al., 2019)
IFIT1	Protein coding, Interferon Induced Protein With Tetratricopeptide Repeats 1 (Stelzer et al., 2016a; Stelzer et al., 2016b)	AC009299.1	Processed pseudogenes (Lang et al., 2019)
AL121835.1	Processed pseudogenes (Lang et al., 2019)	HIGD1C	Protein coding, HIG1 Hypoxia Inducible Domain Family Member 1C (Stelzer et al., 2016a; Stelzer et al., 2016b)
CCNA1	Protein coding, Cyclin A1 (Stelzer et al., 2016a; Stelzer et al., 2016b)	PTBP1P	Pseudogene, Polypyrimidine Tract Binding Protein 1 Pseudogene (Stelzer et al., 2016a; Stelzer et al., 2016b)
DEFB1	Protein coding, Defensin Beta 1 (Stelzer et al., 2016a; Stelzer et al., 2016b)	EIF3CL	Protein coding, Eukaryotic Translation Initiation Factor 3 Subunit C Like (Stelzer et al., 2016a; Stelzer et al., 2016b)
AC104837.2	Processed pseudogenes (Lang et al., 2019)	KIF4CP	Pseudogene, Kinesin Family Member 4C, Pseudogene (Stelzer et al., 2016a; Stelzer et al., 2016b)
AC020765.1	Processed pseudogenes (Lang et al., 2019)	NPIPA7	Protein coding, Nuclear Pore Complex Interacting Protein Family Member A7 (Stelzer et al., 2016a; Stelzer et al., 2016b)
AC005943.1	nonsense mediated decay (Lang et al., 2019)	AC091390.3	unprocessed pseudogene (Lang et al., 2019)
EXOC3L1	Protein coding, Exocyst Complex Component 3 Like 1 (Stelzer et al., 2016a; Stelzer et al., 2016b)	GLYATL1B	Protein coding, Glycine-N-Acyltransferase Like 1B (Stelzer et al., 2016a; Stelzer et al., 2016b)
CXCL11	Protein coding, C-X-C Motif Chemokine Ligand 11 (Stelzer et al., 2016a; Stelzer et al., 2016b)	AC026436.1	Processed pseudogenes (Lang et al., 2019)

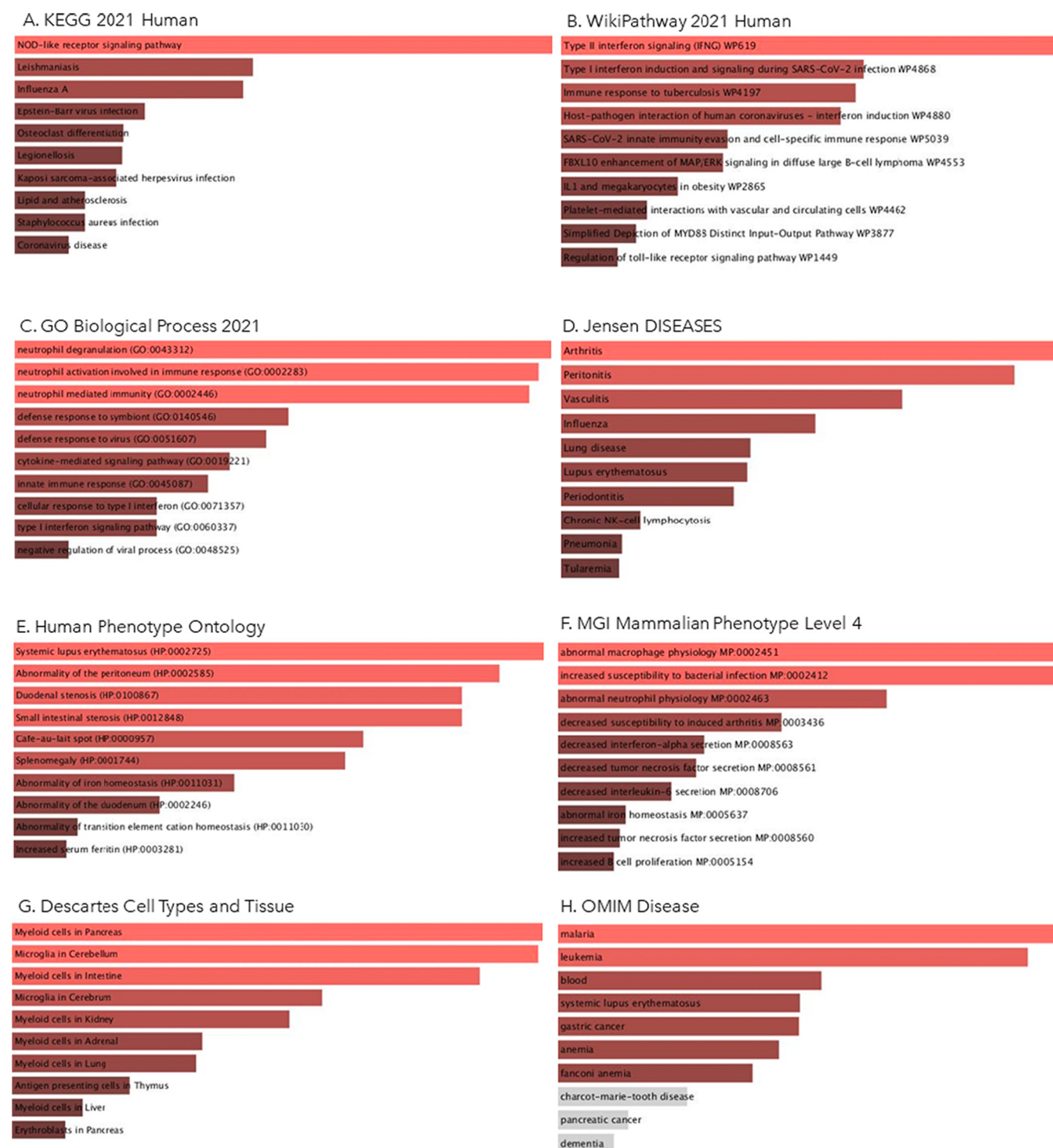


Fig. 5. Ontologies and pathways upregulated in DESeq2 analysis of severe and asymptomatic COVID-19 subjects using Enrichr database.

pathway; Osteoclast differentiation; Legionellosis; Lipid and atherosclerosis; *Staphylococcus aureus* infection; Measles; C-type lectin receptor signaling pathway; TNF signaling pathway; Rheumatoid arthritis; IL-17 signaling pathway.

3.3.1.2. Gene set enrichment analysis of downregulated genes. We observed that downregulated genes in severe patients are significantly associated with Hematopoietic cell lineage and Primary immunodeficiency. Besides, they were involved in lipid metabolism, adaptive

immune response, translation, recurrent respiratory infections, heme biosynthetic pathways, etc.

3.3.1.2.1. Association with metabolic pathways. Notably, some of the downregulated genes were found to be associated with metabolic pathways such as Arachidonic acid metabolism, Inositol phosphate metabolism, Histidine metabolism, Glycosylphosphatidylinositol (GPI)-anchor biosynthesis, Linoleic acid metabolism, beta-Alanine metabolism, Fructose, and mannose metabolism, Glycerophospholipid metabolism, Carbohydrate digestion, and absorption, through

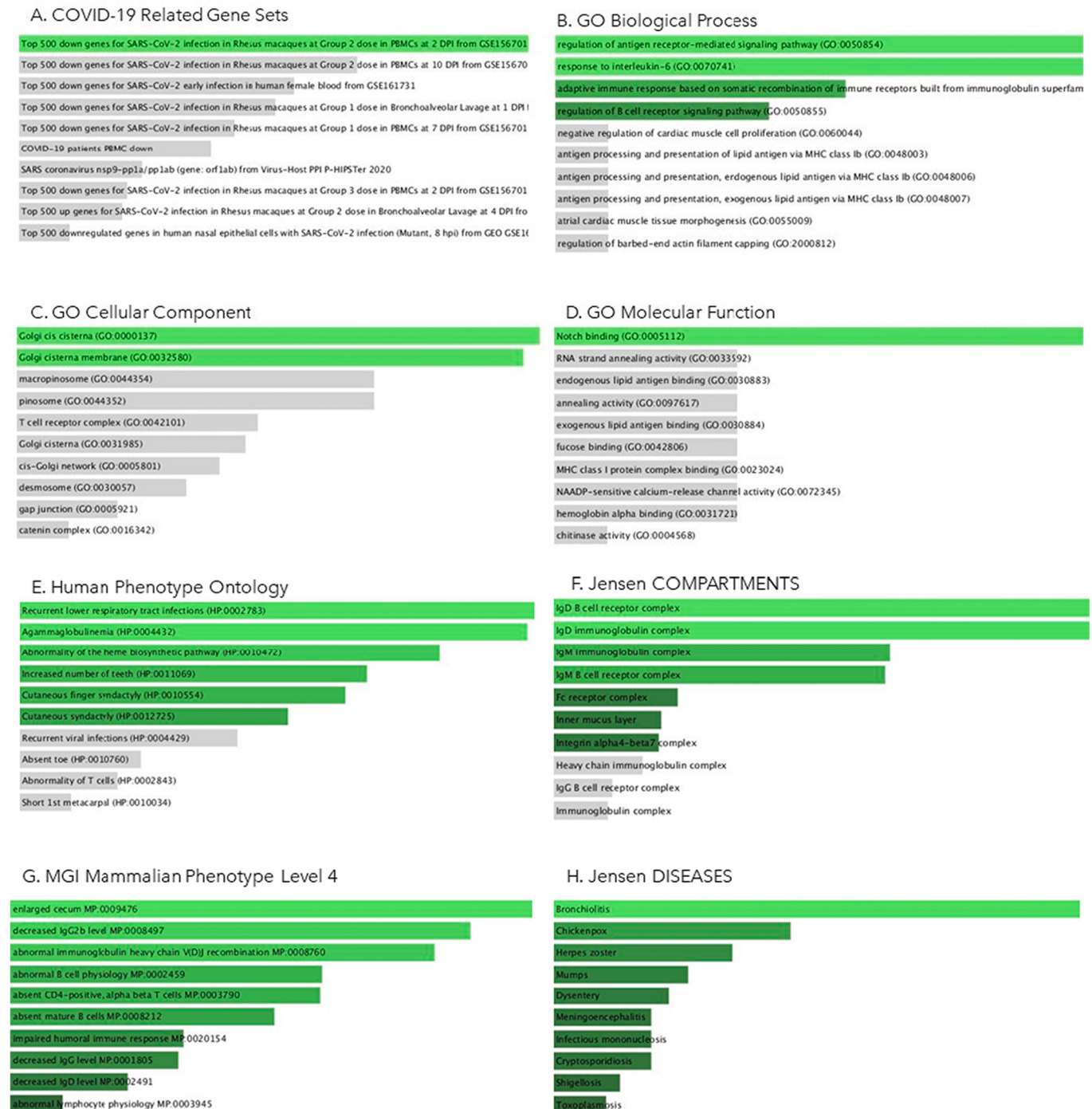


Fig. 6. Ontologies and pathways downregulated in DESeq2 analysis of severe and asymptomatic COVID-19 subjects using Enrichr database.

enrichment was not significant.

3.3.1.2.2. Association with adaptive immune response. Next, we observed downregulated genes are significantly enriched in GO biological processes that are associated with adaptive immune response, including regulation of antigen receptor-mediated signaling pathway (GO:0050854), response to interleukin-6 (GO:0070741), adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains (GO:0002460), regulation of B cell receptor signaling pathway (GO:0050855), regulation of antigen receptor-mediated signaling pathway (GO:0050854), response to interleukin-6 (GO:0070741), adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin

superfamily domains (GO:0002460), regulation of B cell receptor signaling pathway (GO:0050855).

3.3.1.2.3. Association with translation. Some of the downregulated genes, i.e., *EIF3CL*, *EIF4B*, *EIF5AL1*, *PASK*, were found to be involved (although not significantly enriched) in translation processes such as the formation of the translation preinitiation complex (GO:0001731), regulation of translational initiation (GO:0006446), positive regulation of translation (GO:0045727), regulation of translational elongation (GO:0006448), cytoplasmic translational initiation (GO:0002183), translation Factors WP107.

3.3.1.2.4. Association with recurrent respiratory diseases and abnormal Heme biosynthesis. Recurrent lower respiratory tract infections

(HP:0002783), Agammaglobulinemia (HP:0004432), Abnormality of the heme biosynthetic pathway (HP:0010472).

3.3.1.2.5. Other signaling pathways. Further, the TGF-beta signaling pathway, Notch signaling pathway, and Ferroptosis pathways were also associated with downregulated genes. Besides, downregulated genes are related to Cutaneous finger syndactyly (HP:0010554), Cutaneous syndactyly (HP:0012725), and Increased number of teeth (HP:0011069).

4. Discussion

The primary concern with highly transmissible COVID-19 disease is the lack of understanding of the disease-causing mechanisms, resulting in poor treatments and post COVID-19 complications (Cunningham et al., 2020). Clinical observations and scientific studies indicate that SARS-CoV-2 infection impacts not only respiratory organs but also other organs such as the brain, heart, kidney, gastrointestinal tract, etc. (Zaim et al., 2020; Faubel and Edelstein, 2016; Diao et al., 2021; Naicker et al., 2019; Chen et al., 2020a). The risk factors for COVID-19 severity include pre-existing comorbidities, particular age group of subjects, demographics, gender, etc. (Asselah et al., 2021; Zaim et al., 2020; Chen et al., 2021a; Akter et al., 2022; Chen et al., 2020b). The heterogeneous effects of the infection on various individuals pose a significant hurdle in the therapeutic management of COVID-19 patients. Thus, it is vital to delineate the molecular alterations occurring in different groups of patients based on the impact of infections. Towards this, the current study extensively explored the transcriptomics profiles of two contrasting groups of COVID-19 patients, i.e., severe, and asymptomatic to understand the underlying molecular changes that are occurring among them. The RNA sequencing data is derived from whole blood cells, a pool of immune cells, and significant biochemical products result of biochemical processes, making it a considerable tissue sample for transcriptomic profiling. Hence, we believe that the whole blood serves as a good source for understanding the immunopathology of COVID-19 subjects.

Firstly, Principal Component Analysis (PCA) results shown distinct asymptomatic and severe subjects clusters. Subsequently, differential gene expression analysis between these two groups identified 1224 upregulated ($\text{Log2FC} \geq 1.5$ and $p\text{-adjusted value} < 0.05$) and 268 downregulated ($\text{Log2FC} \leq -1.5$ and $p\text{-adjusted value} < 0.05$) genes in severe in comparison to asymptomatic COVID-19 subjects.

Further, gene enrichment analyses shown the enrichment in viral-infection associated pathways in general, as well as specific to COVID-19 infection. For instance, we observed the enrichment of upregulated genes in the type II interferon (IFNG) pathway. While type I IFNG is generally activated in viral response, studies have found suppression of Type I IFN in SARS-CoV infections (Totura et al., 2015; Lotfinejad et al., 2022). The enrichment studies observed an increased population of myeloid cells (in the pancreas, intestine, kidney, lung, liver) and microglia (in the brain), which form part of the innate immune response against the virus. These cell types have a known role in phagocytosis and anti-inflammation, biochemical pathways commonly observed in response to viral infection (Filgueira et al., 2021; Koushki et al., 2021). Antigen-presenting cells (APC.) in thymus enriched in severe patients also indicate an immune response to the virus. GO cellular components shown the enrichment of upregulated genes in increased ficolin and azurophil-rich granules secretion and literature indicated their association with the COVID-19 immune response (Charitos et al., 2021; Polycarpou et al., 2020; Singh et al., 2020b). Neutrophil count increases in COVID-19 infection (Akter et al., 2022; Reusch et al., 2021). Our enrichment analysis also found neutrophil activation and neutrophil mediated immunity. Neutrophil degranulation is enhanced in response to inflammatory reactions in the body (Rosa et al., 2021; Akgun et al., 2020; Meizlish et al., 2020). Previous studies have also reported increased neutrophil degranulation in response to COVID-19 infection in organisms other than humans (Rosa et al., 2021). One such study performed on the Rhesus macaque model shown increased neutrophil degranulation in young subjects compared to old subjects (Rosa et al.,

2021). As observed in severe patients, we propose that the upregulation of neutrophil degranulation occurs in response to the disease severity. Our subjects fall in the mean age of around 45 years, the old age group; there is need to compare neutrophil degranulation in COVID-19 response in an age-dependent study.

Further, our analysis observed “negative regulation of viral process” in the upregulated GO biological process, possibly explaining that the increased host immune response (anti-viral) reduces other viruses' multiplication. The possible reason for this is the activation of the anti-viral immune response that reduces the risks of other infections. Increased STING (stimulator of interferon genes), which mediates interferon expression, is a known prognosis of COVID-19 (Berthelot et al., 2020). Interestingly, we also observed that COVID-19 severe patients might have an increased risk of bacterial, fungal infection compared to asymptomatic patients. This observation of reduced viral infection but increased secondary infection aligns with previous studies on COVID-19 patients (Russell et al., 2021; De Bruyn et al., 2022). The immune response involved in viral and bacterial infection shares different immune components (Bourgoin et al., 2020; Tyagi and Bisen, 2019; Klaas and Crocker, 2012; Puryear et al., 2013; Bourgoin et al., 2019; Lester et al., 2019). A study reveals simultaneous expression of both IFN α and IFN γ inhibits the expression of biomarkers associated with viral and bacterial infection (Bourgoin et al., 2020). We believe that the complex interplay of viral and bacterial response factors and activation of viral response in the host inhibits the expression of host immune machinery to tackle bacterial infection and might be the probable reason for increased susceptibility to bacterial infections post COVID-19 infection. Previous research shows that NOD-like receptor signaling enhanced in response to SARS-CoV-1 infection results in disturbances in microbiota and increased secondary infection (Lotfinejad et al., 2022; Ray and Dittel, 2015). We also observed enrichment of the upregulated gene set in NOD-like receptor signaling, which might indicate gastrointestinal manifestations and increased susceptibility to bacterial infection in severe patients. This observation needs to be further confirmed by studying the host response expression induced by infections with various microorganisms.

Another significant observation in enriched terms for upregulated genes is the high coincidence of cardiac complications in COVID-19 patients, as evident in severe COVID-19 patients (Bakhshandeh et al., 2021; Huang et al., 2020b). Our study has observed coagulation dysfunction upregulated in severe patients, which could be the reason for the cardiac manifestation of COVID-19 infection. As mentioned previously, we have observed enrichment of APCs in the thymus, and there are studies linking the thymus' role in Arrhythmia (Long et al., 2020; Kellogg and Equils, 2021; Dai et al., 2018). We have also confirmed many “Disease-specific laboratory values” upregulated in severe patients. These are related to immunological response, inflammation response, and hypercoagulable state, increased aspartate aminotransferase (AST), and alanine aminotransferase (ALT), and increased interleukin 6 (IL-6), and decreased thrombocytes, reduced blood sodium.

We have also found that COVID 19 disease severity might impact fertility in the patients. It is known that ACE2 receptors are present in human male testicles (Reis et al., 2010), but the studies related to COVID-19's impact on testicular functionality are contradictory (Illiano et al., 2020; Zhao et al., 2003; Ding et al., 2004). Studies done to detect viral RNA in semen showed different results, with the majority indicating the absence of SARS-CoV-2 RNA (Pan et al., 2020; Song et al., 2020; Holtmann et al., 2020; Li et al., 2020; Khalili et al., 2020). So, we propose that if the viral particles are absent in the semen, the possibility of infertility in COVID-19 patients could be an inflammatory response to COVID-19 infection (Xu et al., 2006). The enrichment pathways related to downregulated gene sets also support the previous research findings, such as reduced notch binding signaling, absent mature B cells, CD4+, alpha, and beta T cells, etc. (Rosa et al., 2021; Bartleson et al., 2021).

As COVID-19 disease has a multifactorial response on the body, we

need more clinical features for prognosis, which can help us manage the diverse impact of COVID-19 on health. To address future novel virus disease management, we must not limit ourselves to real-time therapeutics. Instead, we must continuously build the concepts of generalized host response and disease progression on diverse tissues and subject groups.

5. Conclusion

Our unpreparedness with SARS-CoV-2 indicates the need for more stringent research to help us understand disease progression and devise strategies for other such outbreaks in the future. Our comparative study based on two contrasting COVID-19 infection conditions, i.e. severe and asymptomatic patients identified the alteration in key pathways and biological processes associated with various comorbidities. We observed upregulation of viral-specific immune response and inflammatory pathways. Besides, heightened organ-specific responses related to blood, heart, brain, intestine, and kidney enriched in severe subjects not limited to respiratory organs. Also, our study suggests that severe COVID-19 subjects become more prone to bacterial infections and less prone to other viral infections. Besides, we found the downregulation of lipid metabolism, adaptive immune response, translation, heme-biosynthetic pathways, etc. The major pathways highlighted in our study are associated with cardiac complications, autoinflammatory conditions, secondary infections, iron homeostasis and anemia, lipid metabolism, male infertility, etc. These altered pathways in severe patients might be indicative of post-COVID effects. We anticipate our study will facilitate researchers in finding better therapeutic targets and eventually the clinicians in managing COVID-19 patients and their post-COVID complications.

5.1. Limitation of the study

One of the major limitations associated with our study is a small number of samples in dataset. The lack of diversity in the data is another limitation of study since our study is mainly based on the data containing samples from the USA only. Furthermore, lacking the information on other health commodities of patients is another major challenge in deriving the precise biological implications. Despite small cohort, since we have sufficient numbers of samples to derive statistical significance of results, thus we believe if there is a large dataset available, similar strategy could be implemented to further validate these results and to confirm their clinical implication.

Authors contribution

Sen P performed the data analysis, produced the figures, and wrote the manuscript. Kaur H proposed the project, reviewed the data, provided assistance with analysis, drafted the manuscript, and supervised the whole project.

Funding

No funding.

Link for preprint

<https://www.biorxiv.org/content/10.1101/2022.04.16.488556v1>
(DOI: <https://doi.org/10.1101/2022.04.16.488556>).

Declaration of Competing Interest

The authors declare no financial and non-financial conflict of interest.

Acknowledgment

The authors are thankful to the OmicsLogic Research fellowship program, Pine Biotech, Inc., U.S.A.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.humgen.2022.201135>.

References

- Akgun, E., Tuzuner, M.B., Sahin, B., Kilercik, M., Kulah, C., Cakiroglu, H.N., Serteser, M., Unsal, I., Baykal, A.T., 2020. Proteins associated with neutrophil degranulation are upregulated in nasopharyngeal swabs from SARS-CoV-2 patients. *PLoS One* 15, e0240012.
- Akter, A., Ahmed, T., Tauheed, I., Akhtar, M., Rahman, S.I.A., Khaton, F., Ahmed, F., Ferdous, J., Afrad, M.H., Kawser, Z., Hossain, M., Khondaker, R., Hasnat, M.A., Sumon, M.A., Rashed, A., Ghosh, S., Calderwood, S.B., Charles, R.C., Ryan, E.T., Khatri, P., Maecker, H.T., Obermoser, G., Pulendran, B., Clemens, J.D., Banu, S., Shirin, T., LaRocque, R.C., Harris, J.B., Bhuiyan, T.R., Chowdhury, F., Qadri, F., 2022. Disease characteristics and serological responses in patients with differing severity of COVID-19 infection: A longitudinal cohort study in Dhaka, Bangladesh. *PLoS Negl. Trop. Dis.* 16, e0010102.
- Armstrong, R.A., 2014. When to use the Bonferroni correction. *Ophthalmic Physiol. Opt.* 34, 502–508.
- Asselah, T., Duranet, D., Pasmant, E., Lau, G., Schinazi, R.F., 2021. COVID-19: discovery, diagnostics and drug development. *J. Hepatol.* 74, 168–184.
- Bakhshandeh, B., Jahanafrooz, Z., Abbasi, A., Goli, M.B., Sadeghi, M., Mottaqi, M.S., Zamani, M., 2021. Mutations in SARS-CoV-2; Consequences in structure, function, and pathogenicity of the virus. *Microb. Pathog.* 154, 104831.
- Ballou, M., Haga, C.L., 2021. Why do some people develop serious COVID-19 disease after infection, while others only exhibit mild symptoms? *J Allergy Clin Immunol Pract* 9, 1442–1448.
- Bartleson, J.M., Radenkovic, D., Covarrubias, A.J., Furman, D., Winer, D.A., Verdin, E., 2021. SARS-CoV-2, COVID-19 and the ageing immune system. *Nat. Aging* 1, 769–782.
- Benetti, G., Cardobi, N., Cardano, G., Arena, C., Micheletto, C., Guariglia, S., Montemuzzi, S., Cavedon, C., 2021. CT-based radiomics as a tool to recognize COVID-19 positive patients. *Phys. Med.* 92, S46.
- Berthelot, J.-M., Lioté, F., Maugars, Y., Sibilia, J., 2020. Lymphocyte changes in severe COVID-19: delayed over-activation of STING? *Front. Immunol.* 11.
- Bourgoin, P., Soliveres, T., Ahzir, D., Arnoux, I., Meisel, C., Unterwaller, N., Morange, P.E., Michelet, P., Malergue, F., Markarian, T., 2019. Clinical research assessment by flow cytometry of biomarkers for infectious stratification in an emergency department. *Biomark. Med* 13, 1373–1386.
- Bourgoin, P., Biechele, G., Ait Belkacem, I., Morange, P.E., Malergue, F., 2020. Role of the interferons in CD64 and CD169 expressions in whole blood: relevance in the balance between viral- or bacterial-oriented immune responses. *Immun. Inflamm. Dis.* 8, 106–123.
- Brogna, B., Brogna, C., Petrillo, M., Conte, A.M., Benincasa, G., Montano, L., Piscopo, M., 2021. SARS-CoV-2 detection in fecal sample from a patient with typical findings of COVID-19 pneumonia on CT but negative to multiple SARS-CoV-2 RT-PCR tests on oropharyngeal and nasopharyngeal swab samples. *Medicina (Kaunas)* 57.
- Carfi, A., Bernabei, R., Landi, F., C.-P.-A.C.S.G. Gemelli Against, 2020. Persistent symptoms in patients after acute COVID-19. *JAMA* 324, 603–605.
- Charitos, P., Heijnen, I., Egli, A., Bassetti, S., Trendelenburg, M., Osthoff, M., 2021. Functional activity of the complement system in hospitalized COVID-19 patients: A prospective cohort study. *Front. Immunol.* 12, 765330.
- Chen, E., Kuleshov, M., Bailey, A., Xie, Z., Clarke, D.J.B., Evangelista, J.E., Wojciechowski, M., Kropiwnicki, E., Jagodnik, K., Jeon, M., Litz, S., Jones, M., Tan, C., Kou, Y., Clark, N., Rouillard, A., Fernandez, N., Duan, Q., Wang, Z., Koplev, S., Jenkins, S., Lachmann, A., McDermott, M., Monteiro, C., Gundersen, G., Ma'ayan, A., 2022. Enrichr Search Engine. The Ma'ayan Lab.
- Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., Ma'ayan, A., 2013. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.* 14, 128.
- Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., Xia, J., Yu, T., Zhang, X., Zhang, L., 2020b. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 395, 507–513.
- Chen, Q., Allot, A., Lu, Z., 2021b. LitCovid: an open database of COVID-19 literature. *Nucleic Acids Res.* 49, D1534–D1540.
- Chen, T., Wu, D., Chen, H., Yan, W., Yang, D., Chen, G., Ma, K., Xu, D., Yu, H., Wang, H., Wang, T., Guo, W., Chen, J., Ding, C., Zhang, X., Huang, J., Han, M., Li, S., Luo, X., Zhao, J., Ning, Q., 2020a. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. *BMJ* 368, m1091.
- Chen, Y., Klein, S.L., Garibaldi, B.T., Li, H., Wu, C., Osevala, N.M., Li, T., Margolick, J.B., Pawelec, G., Leng, S.X., 2021a. Aging in COVID-19: vulnerability, immunity and intervention. *Ageing Res. Rev.* 65, 101205.
- COVID Data Tracker, 2022. Centers for Disease Control and Prevention. US Department of Health and Human Services, CDC, Atlanta, GA.

- Cunningham, A.C., Goh, H.P., Koh, D., 2020. Treatment of COVID-19: old tricks for new challenges. *Crit. Care* 24, 91.
- Dai, X., Zhang, D., Wang, C., Wu, Z., Liang, C., 2018. The pivotal role of Thymus in atherosclerosis mediated by immune and inflammatory response. *Int. J. Med. Sci.* 15, 1555–1563.
- De Bruyn, A., Verellen, S., Bruckers, L., Geebelen, L., Callebaut, I., De Pauw, I., Stessel, B., Dubois, J., 2022. Secondary infection in COVID-19 critically ill patients: a retrospective single-center evaluation. *BMC Infect. Dis.* 22, 207.
- Devaux, C.A., Rolain, J.M., Raoult, D., 2020. ACE2 receptor polymorphism: susceptibility to SARS-CoV-2, hypertension, multi-organ failure, and COVID-19 disease outcome. *J. Microbiol. Immunol. Infect.* 53, 425–435.
- Diao, B., Wang, C., Wang, R., Feng, Z., Zhang, J., Yang, H., Tan, Y., Wang, H., Wang, C., Liu, L., Liu, Y., Liu, Y., Wang, G., Yuan, Z., Hou, X., Ren, L., Wu, Y., Chen, Y., 2021. Human kidney is a target for novel severe acute respiratory syndrome coronavirus 2 infection. *Nat. Commun.* 12, 2506.
- Dinc, I., Sigdel, M., Dinc, S., Sigdel, M.S., Pusey, M.L., Aygun, R.S., 2014. Evaluation of normalization and PCA on the performance of classifiers for protein crystallization images. *Proc IEEE Southeastcon* 2014.
- Ding, Y., He, L., Zhang, Q., Huang, Z., Che, X., Hou, J., Wang, H., Shen, H., Qiu, L., Li, Z., Geng, J., Cai, J., Han, H., Li, X., Kang, W., Weng, D., Liang, P., Jiang, S., 2004. Organ distribution of severe acute respiratory syndrome (SARS) associated coronavirus (SARS-CoV) in SARS patients: implications for pathogenesis and virus transmission pathways. *J. Pathol.* 203, 622–630.
- Director-General, W., 2020. WHO Director-General's Opening Remarks at the Media Briefing on COVID-19 - 11 March 2020. World Health Organization.
- Dolaner, S., Kaur, H., Brodsky, E., Panov, J., Mazumder, M., 2021. Identification of lncRNAs as therapeutic targets in chronic lymphocytic leukemia. *Columbia Undergraduate Sci. J. (CUSJ)* 15 (2021).
- van Doremalen, N., Bushmaker, T., Morris, D.H., Holbrook, M.G., Gamble, A., Williamson, B.N., Tamin, A., Harcourt, J.L., Thornburg, N.J., Gerber, S.I., Lloyd-Smith, J.O., de Wit, E., Munster, V.J., 2020. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *N. Engl. J. Med.* 382, 1564–1567.
- Dudas, G., Hong, S.L., Potter, B.I., Calvignac-Spencer, S., Niatou-Singa, F.S., Tombolomako, T.B., Fuh-Neba, T., Vickos, U., Ulrich, M., Leendertz, F.H., Khan, K., Huber, C., Watts, A., Olendraitte, I., Snijder, J., Wijnant, K.N., Bonvin, A., Martres, P., Behillil, S., Ayoub, A., Maidadi, M.F., Djoms, D.M., Godwe, C., Butel, C., Simaitis, A., Gabriellaite, M., Katenaite, M., Norvilas, R., Raugaite, L., Koyaweda, G. W., Kandou, J.K., Jonikas, R., Nasvytiene, I., Zemeckiene, Z., Gecys, D., Tamauskaite, K., Norkiene, M., Vasilunaite, E., Ziogiene, D., Timinskas, A., Sukys, M., Sarauskas, M., Alzbutas, G., Aziza, A.A., Lusakaki, E.K., Cigolo, J.M., Mawete, F.M., Lofiko, E.L., Kingebeji, P.M., Tamfum, J.M., Belzair, M.R.D., Essomba, R.G., Assoumou, M.C.O., Mborongong, A.B., Dieng, A.B., Juozapaitis, D., Hosch, S., Obama, J., Ayekaba, M.O., Naumovas, D., Pautienius, A., Rafai, C.D., Vitkauskienė, A., Ugenskiene, R., Gedvilaitė, A., Cereskevicius, D., Lesauskaitė, V., Zemaitis, L., Griskevicius, L., Baele, G., 2021. Emergence and spread of SARS-CoV-2 lineage B.1.620 with variant of concern-like mutations and deletions. *Nat. Commun.* 12, 5769.
- Faes, C., Abrams, S., Van Beckhoven, D., Meyfroidt, G., Vlieghe, E., Hens, N., B.C.G.o.C.-H. Surveillance, 2020. Time between symptom onset, hospitalisation and recovery or death: statistical analysis of Belgian COVID-19 patients. *Int. J. Environ. Res. Public Health* 17, 7560.
- Faleiros, D.E., van den Bos, W., Botto, L., Scarano, F., 2022. TU Delft COVID-app: A tool to democratize CFD simulations for SARS-CoV-2 infection risk analysis. *Sci. Total Environ.* 826, 154143.
- Fan, C., Chen, M., Wang, X., Wang, J., Huang, B., 2021. A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. *Front. Energy Res.* 9.
- Faubel, S., Edelstein, C.L., 2016. Mechanisms and mediators of lung injury after acute kidney injury. *Nat. Rev. Nephrol.* 12, 48–60.
- Fegert, J.M., Vitiello, B., Plener, P.L., Clemens, V., 2020. Challenges and burden of the Coronavirus 2019 (COVID-19) pandemic for child and adolescent mental health: a narrative review to highlight clinical and research needs in the acute phase and the long return to normality. *Child Adolesc. Psychiatry Ment. Health* 14, 20.
- Feld, J.J., Kandel, C., Biondi, M.J., Kozak, R.A., Zahoor, M.A., Lemieux, C., Borgia, S.M., Boggild, A.K., Powis, J., McCready, J., Tan, D.H.S., Chan, T., Coburn, B., Kumar, D., Humar, A., Chan, A., O'Neil, B., Noureldin, S., Booth, J., Hong, R., Smookler, D., Aleyadeh, W., Patel, A., Barber, B., Casey, J., Hiebert, R., Mistry, H., Choong, I., Hislop, C., Santer, D.M., Lorne Tyrrell, D., Glenn, J.S., Gehring, A.J., Janssen, H.L.A., Hansen, B.E., 2021. Peginterferon lambda for the treatment of outpatients with COVID-19: a phase 2, placebo-controlled randomised trial. *Lancet Respir. Med.* 9, 498–510.
- Fernandez-de-Las-Penas, C., Palacios-Cena, D., Gomez-Mayordomo, V., Cuadrado, M.L., Florencio, L.L., 2021. Defining post-COVID symptoms (post-acute COVID, Long COVID, persistent post-COVID): an integrative classification. *Int. J. Environ. Res. Public Health* 18.
- Filgueira, L., Larionov, A., Lannes, N., 2021. The influence of virus infection on microglia and accelerated brain aging. *Cells* 10.
- Fung, M., Babik, J.M., 2021. COVID-19 in immunocompromised hosts: what we know so far. *Clin. Infect. Dis.* 72, 340–350.
- Garca, S., Luengo, J., Herrera, F., 2014. Data Preprocessing in Data Mining. Springer Publishing Company, Incorporated.
- Garrigues, E., Janvier, P., Kherabi, Y., Le Bot, A., Hamon, A., Gouze, H., Doucet, L., Berkani, S., Oliosi, E., Mallart, E., Corre, F., Zarrouk, V., Moyer, J.D., Galy, A., Honsel, V., Fantin, B., Nguyen, Y., 2020. Post-discharge persistent symptoms and health-related quality of life after hospitalization for COVID-19. *J. Inf. Secur.* 81, e4–e6.
- Gong, J., Ou, J., Qiu, X., Jie, Y., Chen, Y., Yuan, L., Cao, J., Tan, M., Xu, W., Zheng, F., Shi, Y., Hu, B., 2020. A tool for early prediction of severe coronavirus disease 2019 (COVID-19): a multicenter study using the risk nomogram in Wuhan and Guangdong, China. *Clin. Infect. Dis.* 71, 833–840.
- Gorbalenya, A.E., Baker, S.C., Baric, R.S., de Groot, R.J., Drosten, C., Gulyaeva, A.A., Haagmans, B.L., Lauber, C., Leontovich, A.M., Neuman, B.W., Penzar, D., Perlman, S., Poon, L.L.M., Samborskiy, D.V., Sidorov, I.A., Sola, I., Ziebuhr, J., V. Coronaviridae Study Group of the International Committee on Taxonomy of, 2020. The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* 5, 536–544.
- Guruprasad, L., 2021. Human SARS CoV-2 spike protein mutations, proteins: structure. *Funct. Bioinform.* 89, 569–576.
- Haider, R., Shamsi, T.S., Khan, N.A., 2021. Machine learning based decipherment of cell population Data: a promising hospital front-door screening tool for COVID-19. *Am. J. Clin. Pathol.* 156, S101–S102.
- Hamming, I., Timens, W., Bulthuis, M.L., Lely, A.T., Navis, G., van Goor, H., 2004. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J. Pathol.* 203, 631–637.
- Han, J., Kamber, M., Pei, J., 2012a. 3 - Data preprocessing. In: Han, J., Kamber, M., Pei, J. (Eds.), *Data Mining (Third Edition)*. Morgan Kaufmann, Boston, pp. 83–124.
- Han, J., Kamber, M., Pei, J., 2012b. *Data Mining: Concepts and Techniques*. Morgan Kaufmann, Boston.
- Haynes, W., 2013. Bonferroni correction. In: Dubitzky, W., Wolkenhauer, O., Cho, K.-H., Yokota, H. (Eds.), *Encyclopedia of Systems Biology*. Springer, New York, New York, NY, p. 154.
- Hoffmann, M., Kleine-Weber, H., Schroeder, S., Kruger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A., Muller, M.A., Drosten, C., Pohlmann, S., 2020. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181, 271–280 e278.
- Holtmann, N., Edimiris, P., Andree, M., Doehmen, C., Baston-Buest, D., Adams, O., Krussel, J.S., Bielfeld, A.P., 2020. Assessment of SARS-CoV-2 in human semen: a cohort study. *Fertil. Steril.* 114, 233–238.
- Holy, C., Shah, S., Elangovanraaj, N., Krishnan, D., Gupta, S., Trivedi, P., Devulapally, M., Mohapatra, A., Johnston, S., Coplan, P., 2020. PIN117 identification of patients with COVID-19 infection prior to the new COVID-19 diagnostic code - a premier database analysis. *Value Health* 23, S563.
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo, L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin, Q., Wang, J., Cao, B., 2020b. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395, 497–506.
- Huang, Y., Yang, C., Xu, X.F., Xu, W., Liu, S.W., 2020a. Structural and functional properties of SARS-CoV-2 spike protein: potential antiviral drug development for COVID-19. *Acta Pharmacol. Sin.* 41, 1141–1149.
- Illiano, E., Trama, F., Costantini, E., 2020. Could COVID-19 have an impact on male fertility? *Andrologia* 52, e13654.
- Jafari, M., Ansari-Pour, N., 2019. Why, when and how to adjust your P values? *Cell J.* 20, 604–607.
- Jagannathan, P., Hu, Z., van der Ploeg, K., Martinez Mori, D.A., 2021. Baseline Signatures Associated with Clinical, Virologic, and Immunologic Outcomes in Patients with Mild to Moderate COVID-19.
- Jean, S.-S., Hsueh, P.-R., 2020. Old and re-purposed drugs for the treatment of COVID-19. *Expert Rev. Anti-Infect. Ther.* 18, 843–847.
- Jolliffe, I.T., 2002. *Principal Component Analysis for Special Types of Data*. Springer.
- Jones, D.L., Baluja, M.Q., Graham, D.W., Corbushley, A., McDonald, J.E., Malham, S.K., Hillary, L.S., Connor, T.R., Gaze, W.H., Moura, I.B., Wilcox, M.H., Farkas, K., 2020. Shedding of SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the environment-based spread of COVID-19. *Sci. Total Environ.* 749, 141364.
- Kaur, H., Bhalla, S., Raghava, G.P.S., 2019. Classification of early and late stage liver hepatocellular carcinoma patients from their genomics and epigenomics profiles. *PLoS One* 14, e0221476.
- Kaur, H., Dhall, A., Kumar, R., Raghava, G.P.S., 2020a. Identification of platform-independent diagnostic biomarker panel for hepatocellular carcinoma using large-scale transcriptomics Data. *Front. Genet.* 10.
- Kaur, H., Bhalla, S., Garg, D., Mehta, N., Raghava, G.P.S., 2020b. Analysis and prediction of cholangiocarcinoma from transcriptomic profile of patients. *J. Hepatol.* 73.
- Kellogg, C., Equils, O., 2021. The role of the thymus in COVID-19 disease severity: implications for antibody treatment and immunization. *Hum. Vaccin Immunother* 17, 638–643.
- Khalili, M.A., Leisegang, K., Majzoub, A., Finelli, R., Panner Selvam, M.K., Henkel, R., Mojan, M., Agarwal, A., 2020. Male fertility and the COVID-19 pandemic: systematic review of the literature. *World J. Mens Health* 38, 506–520.
- Khomtchouk, B.B., Van Booven, D.J., Wahlestedt, C., 2014. HeatmapGenerator: high performance RNAseq and microarray visualization software suite to examine differential gene expression levels using an R and C++ hybrid computational pipeline. *Source Code Biol. Med.* 9, 30.
- Khullar, G., Chandra, M., 2020. Virtual dermatopathology: a potential educational tool during COVID-19 pandemic. *Dermatol. Ther.* 33, e13755.
- Klaas, M., Crocker, P.R., 2012. Sialoadhesin in recognition of self and non-self. *Semin. Immunopathol.* 34, 353–364.
- Klopfenstein, T., Kadiane-Oussou, N.J., Toko, L., Royer, P.Y., Lepiller, Q., Gendrin, V., Zayet, S., 2020. Features of anaemia in COVID-19. *Med. Mal. Infect.* 50, 436–439.
- Koushki, K., Salemi, M., Miri, S.M., Arjeini, Y., Keshavarz, M., Ghaemi, A., 2021. Role of myeloid-derived suppressor cells in viral respiratory infections; hints for discovering therapeutic targets for COVID-19. *Biomed. Pharmacother.* 144, 112346.

- Kuhn, M., 2008. Building predictive models in R using the caret package. *J. Stat. Softw.* 28, 1–26.
- Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., McDermott, M.G., Monteiro, C.D., Gundersen, G.W., Ma'ayan, A., 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 44, W90–W97.
- Kumar, S., Thambiraja, T.S., Karuppanan, K., Subramanian, G., 2022. Omicron and Delta variant of SARS-CoV-2: A comparative computational study of spike protein. *J. Med. Virol.* 94, 1641–1649.
- Lamers, M.M., Beumer, J., van der Vaart, J., Knoop, K., Puschhof, J., Breugem, T.I., Ravelli, R.B.G., Paul van Schayck, J., Mykityn, A.Z., Duimel, H.Q., van Donselaar, E., Riesebosch, S., Kuijpers, H.J.H., Schipper, D., van de Wetering, W.J., de Graaf, M., Koopmans, M., Cuppen, E., Peters, P.J., Haagmans, B.L., Clevers, H., 2020. SARS-CoV-2 productively infects human gut enterocytes. *Science* 369, 50–54.
- Lamprecht, B., 2020. Is there a post-COVID syndrome? *Pneumologie (Berl)* 17, 398–405.
- Lang, B., Armaso, A., Tartaglia, G.G., 2019. RNAi: protein-RNA interaction predictions for model organisms with supporting experimental data. *Nucleic Acids Res.* 47, D601–D606.
- Lee, P.-I., Hsueh, P.-R., 2020. Emerging threats from zoonotic coronaviruses—from SARS and MERS to 2019-nCoV. *J. Microbiol. Immunol. Infect.* 53, 365–367.
- Lester, P.J., Buick, K.H., Baty, J.W., Felden, A., Haywood, J., 2019. Different bacterial and viral pathogens trigger distinct immune responses in a globally invasive ant. *Sci. Rep.* 9, 5780.
- Letko, M., Marzi, A., Munster, V., 2020. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat. Microbiol.* 5, 562–569.
- Lever, J., Krzywinski, M., Altman, N., 2017. Principal component analysis. *Nat. Methods* 14, 641–642.
- Li, D., Jin, M., Bao, P., Zhao, W., Zhang, S., 2020. Clinical characteristics and results of semen tests among men with coronavirus disease 2019. *JAMA Netw. Open* 3, e208292.
- Li, F., 2013. Receptor recognition and cross-species infections of SARS coronavirus. *Antivir. Res.* 100, 246–254.
- Long, B., Brady, W.J., Koyfman, A., Gottlieb, M., 2020. Cardiovascular complications in COVID-19. *Am. J. Emerg. Med.* 38, 1504–1507.
- Lotfinejad, P., Asadzadeh, Z., Najjary, S., Somi, M.H., Hajiasgharzadeh, K., Mokhtarzadeh, A., Derakhshani, A., Roshani, E., Baradaran, B., 2022. COVID-19 infection: concise review based on the immunological perspective. *Immunol. Investig.* 51, 246–265.
- Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550.
- Magrone, T., Magrone, M., Jirillo, E., 2020. Focus on receptors for coronaviruses with special reference to angiotensin-converting enzyme 2 as a potential drug target - A perspective. *Endocr Metab Immune Disord Drug Targets* 20, 807–811.
- Mahdi, A., Błaszczyk, P., Dlotko, P., Salvi, D., Chan, T.-S., Harvey, J., Gurnari, D., Wu, Y., Farhat, A., Hellmer, N., Zarebski, A., Hogan, B., Tarasenko, L., 2021. OxCOVID19 database, a multimodal data repository for better understanding the global impact of COVID-19. *Sci. Rep.* 11.
- Malley, B., Ramazzotti, D., Wu, J.T.-Y., 2016. Data Pre-processing. In: *Data, M.I.T.C. (Ed.), Secondary Analysis of Electronic Health Records*. Springer International Publishing, Cham, pp. 115–141.
- Mathieu, E., Ritchie, H., Ortiz-Ospina, E., Roser, M., Hasell, J., Appel, C., Giattino, C., Rodes-Guirao, L., 2021. Author correction: A global database of COVID-19 vaccinations. *Nat. Hum. Behav.* 5, 956–959.
- McIntosh, K., 2022. COVID-19: Epidemiology, Virology, and Prevention.
- Meizlish, M.L., Pine, A.B., Bishai, J.D., Goshua, G., Nadelmann, E.R., Simonov, M., Chang, C.H., Zhang, H., Shallow, M., Bahel, P., Owusu, K., Yamamoto, Y., Arora, T., Atri, D.S., Patel, A., Gbyli, R., Kwan, J., Won, C.H., Dela Cruz, C., Price, C., Koff, J., King, B.A., Rinder, H.M., Wilson, F.P., Hwa, J., Halene, S., Damsky, W., van Dijk, D., Lee, A.L., Chun, H., 2020. A neutrophil activation signature predicts critical illness and mortality in COVID-19. *Blood Adv.* <https://doi.org/10.1182/bloodadvances.2020003568>.
- Menyhart, O., Weltz, B., Györfy, B., 2021. MultipleTesting.com: A tool for life science researchers for multiple hypothesis testing correction. *PLoS One* 16, e0245824.
- Meredith Wadman, J.C.-F., Kaiser, Jocelyn, Maticic, Catherine, Jun 2020. How does coronavirus kill? Clinicians trace a ferocious rampage through the body, from brain to toes. *Science* 2020.
- Mulholland, R.H., Vasileiou, E., Simpson, C.R., Robertson, C., Ritchie, L.D., Agrawal, U., Woolhouse, M., Murray, J.L., Stag, H.R., Docherty, A.B., McCowan, C., Wood, R., Stock, S.J., Sheikh, A., 2021. Cohort profile: early pandemic evaluation and enhanced surveillance of COVID-19 (EAVE II) database. *Int. J. Epidemiol.* 50, 1064–1074.
- Naicker, S., Yang, C.W., Hwang, S.J., Liu, B.C., Chen, J.H., Jha, V., Coronavirus, The Novel, 2019. Epidemic and kidneys. *Kidney Int.* 97 (2020), 824–828.
- NCBI, 2022. NCBI Visual Data Dashboard.
- Ouyang, S., Wang, Y., Zhou, K., Xia, J., 2021. LitCovid-AGAC: cellular and molecular level annotation data set based on COVID-19. *Genomics Inform.* 19, e23.
- Pan, F., Xiao, X., Guo, J., Song, Y., Li, H., Patel, D.P., Spivak, A.M., Alukal, J.P., Zhang, X., Xiong, C., Li, P.S., Hotaling, J.M., 2020. No evidence of severe acute respiratory syndrome-coronavirus 2 in semen of males recovering from coronavirus disease 2019. *Fertil. Steril.* 113, 1135–1139.
- Park, M.S., Jo, H., Lee, H., Jung, S.Y., Hwang, H.J., 2022. Machine learning-based COVID-19 patients triage algorithm using patient-generated health data from nationwide multicenter database. *Infect. Dis. Ther.* 11, 787–805. <https://doi.org/10.1007/s40121-022-00600-4>.
- Patiyal, S., Kaur, D., Kaur, H., Sharma, N., Dhali, A., Sahai, S., Agrawal, P., Maryam, L., Arora, C., Raghava, G.P.S., 2020. A web-based platform on coronavirus disease-19 to maintain predicted diagnostic, drug, and vaccine candidates. *Monoclon. Antib. Immunodiagn. Immunother.* 39, 204–216.
- Paules, C.I., Marston, H.D., Fauci, A.S., 2020. Coronavirus infections—more than just the common cold. *JAMA* 323, 707–708.
- Perry, M., heatmaps, 2021. Flexible Heatmaps for functional genomics and sequence features. *Bioconductor. R package version 1.16.0*.
- Polycarpou, A., Howard, M., Farrar, C.A., Greenlaw, R., Fanelli, G., Wallis, R., Klavinskis, L.S., Sacks, S., 2020. Rationale for targeting complement in COVID-19. *EMBO Mol. Med.* 12, e12642.
- Puryear, W.B., Akiyama, H., Geer, S.D., Ramirez, N.P., Yu, X., Reinhard, B.M., Gummuru, S., 2013. Interferon-inducible mechanism of dendritic cell-mediated HIV-1 dissemination is dependent on Siglec-1/CD169. *PLoS Pathog.* 9, e1003291.
- Ray, A., Dittel, B.N., 2015. Interrelatedness between dysbiosis in the gut microbiota due to immunodeficiency and disease penetrance of colitis. *Immunology* 146, 359–368.
- Reis, A.B., Araujo, F.C., Pereira, V.M., Dos Reis, A.M., Santos, R.A., Reis, F.M., 2010. Angiotensin (1-7) and its receptor Mas are expressed in the human testis: implications for male infertility. *J. Mol. Histol.* 41, 75–80.
- Reusch, N., De Domenico, E., Bonaguro, L., Schulte-Schrepping, J., Bassler, K., Schultze, J.L., Aschenbrenner, A.C., 2021. Neutrophils in COVID-19. *Front. Immunol.* 12, 652470.
- Riffe, T., Acosta, E., T.C.-D. team, 2021. Data resource profile: COVERAGE-DB: a global demographic database of COVID-19 cases and deaths. *Int. J. Epidemiol.* 50, 390–390f.
- Rogers, J.P., Chesney, E., Oliver, D., Pollak, T.A., McGuire, P., Fusar-Poli, P., Zandi, M.S., Lewis, G., David, A.S., 2020. Psychiatric and neuropsychiatric presentations associated with severe coronavirus infections: a systematic review and meta-analysis with comparison to the COVID-19 pandemic. *Lancet. Psychiatry* 7, 611–627.
- Rosa, B.A., Ahmed, M., Singh, D.K., Chorenno-Parra, J.A., Cole, J., Jimenez-Alvarez, L.A., Rodriguez-Reyna, T.S., Singh, B., Gonzalez, O., Carrion Jr., R., Schlesinger, L.S., Martin, J., Zuniga, J., Mitreva, M., Kaushal, D., Khader, S.A., 2021. IFN signaling and neutrophil degranulation transcriptional signatures are induced during SARS-CoV-2 infection. *Commun. Biol.* 4, 290.
- Rothan, H.A., Byrareddy, S.N., 2020. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J. Autoimmun.* 109, 102433.
- Russell, C.D., Fairfield, C.J., Drake, T.M., Turtle, L., Seaton, R.A., Wootton, D.G., Sigfrid, J., Harrison, E.M., Docherty, A.B., de Silva, T.I., Egan, C., Pius, R., Hardwick, H.E., Merson, L., Girvan, M., Dunning, J., Nguyen-Van-Tam, J.S., Openshaw, P.J.M., Baillie, J.K., Semple, M.G., Ho, A., I.C. Investigators, 2021. Co-infections, secondary infections, and antimicrobial use in patients hospitalised with COVID-19 during the first pandemic wave from the ISARIC WHO CCP-UK study: a multicentre, prospective cohort study. *Lancet Microbe* 2, e354–e365.
- Saffern, M., 2020. Cancer therapy tool informs COVID-19 vaccines. *Nat. Rev. Immunol.* 20, 352.
- Scordo, K.A., Richmond, M.M., Munro, N., 2021. Post-COVID-19 Syndrome: theoretical basis, identification, and management. *AACN Adv. Crit. Care* 32, 188–194.
- Sha, Y., Phan, J.H., Wang, M.D., 2015. Effect of low-expression gene filtering on detection of differentially expressed genes in RNA-seq data. *Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2015, 6461–6464.
- Sher, L., 2021. Post-COVID syndrome and suicide risk. *QJM* 114, 95–98.
- Singh, A., Sood, N., Narang, V., Goyal, A., 2020b. Morphology of COVID-19-affected cells in peripheral blood film. *BMJ Case Rep.* 13.
- Singh, A.K., Mishra, G., Maurya, A., Kulkarni, G.T., Awasthi, R., 2020a. Biofabrication: an interesting tool to create in vitro model for COVID-19 drug targets. *Med. Hypotheses* 144, 110059.
- Song, C., Wang, Y., Li, W., Hu, B., Chen, G., Xia, P., Wang, W., Li, C., Diao, F., Hu, Z., Yang, X., Yao, B., Liu, Y., 2020. Absence of 2019 novel coronavirus in semen and testes of COVID-19 patients. *Reprod. Biol. Med.* 103, 4–6.
- Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Stein, T.I., Nudel, R., Lieder, I., Mazor, Y., Kaplan, S., Dahary, D., Warshawsky, D., Guan-Golan, Y., Kohn, A., Rappaport, N., Safran, M., Lancet, D., 2016a. The GeneCards Suite: from gene data mining to disease genome sequence analyses. *Curr. Protoc. Bioinformatics* 54, 1 30 31–31 30 33.
- Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Stein, T.I., Nudel, R., Lieder, I., Mazor, Y., Kaplan, S., Dahary, D., Warshawsky, D., Guan-Golan, Y., Kohn, A., Rappaport, N., Safran, M., Lancet, D., 2016b. GeneCards®: The Human Gene Database. Springer, Singapore.
- Streiner, D.L., 2015. Best (but oft-forgotten) practices: the multiple problems of multiplicity—whether and how to correct for many statistical tests. *Am. J. Clin. Nutr.* 102, 721–728.
- 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., Unit, I.S.-S.C.B., 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., 2020. Multi-omics resolves a sharp disease-state shift between mild and moderate COVID-19. *Cell* 183, 1479–1495 e1420.
- Sungnak, W., Huang, N., Becavin, C., Berg, M., Queen, R., Litvinukova, M., Talavera-Lopez, C., Maatz, H., Reichart, D., Sampaziotis, F., Worlock, K.B., Yoshida, M., Barnes, J.L., H.C.A.L.B. Network, 2020. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat. Med.* 26, 681–687.

- Talwai, A., Wing, V., Itzkovich, Y., Galaznik, A., Chatterjee, A., Jain, R., LaBonte, J., 2021. PIN83 the COVID-19 research database: building one of the largest PRO bono real-world DATA repositories. *Value Health* 24, S121.
- Therapeutics and COVID-19: Lliving Gguideline, 2022. WHO-2019-nCoV-Therapeutics-2022.2. World Health Organization.
- Timeline: WHO's COVID-19 Response, 2022. World Health Organization.
- Totura, A.L., Whitmore, A., Agnihothram, S., Schafer, A., Katze, M.G., Heise, M.T., Baric, R.S., 2015. Toll-like receptor 3 signaling via TRIF contributes to a protective innate immune response to severe acute respiratory syndrome coronavirus infection. *mBio* 6 e00638–00615.
- Tyagi, R., Bisen, P.S., 2019. Immune response activation and immunomodulation. BoD–Books on Demand. <https://doi.org/10.5772/intechopen.73708>. <https://www.intechopen.com/books/6963>.
- Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., Wester, K., Hober, S., Werner, H., Bjorling, L., Ponten, F., 2010. Towards a knowledge-based human protein atlas. *Nat. Biotechnol.* 28, 1248–1250.
- Uhlen, M., Fagerberg, L., Hallstrom, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szegedy, C.A., Odeberg, J., Djureinovic, D., Takanen, J.O., Hober, S., Alm, T., Edqvist, P.H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J.M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J., Ponten, F., 2015. Proteomics. Tissue-based map of the human proteome. *Science* 347, 1260419.
- Vaja, R., Kaur, H., Mazumder, M., Brodsky, E., 2021. In silico analysis of transcriptomic profiling and affected biological pathways in multiple sclerosis. *Immunogenet Open Access* 7, 167. <https://doi.org/10.35248/IGOA.22.7.167>.
- Varghese, P.M., Tsolaki, A.G., Yasmin, H., Shastri, A., Ferluga, J., Vathis, M., Madan, T., Kishore, U., 2020. Host-pathogen interaction in COVID-19: pathogenesis, potential therapeutics and vaccination strategies. *Immunobiology* 225, 152008.
- Vetruigno, G., Foti, F., Di Pumpo, M., Cicconi, M., D'Ambrosio, F., La Milia, D.I., Pastorino, R., Boccia, S., Damiani, G., Laurenti, P., 2021. Machine learning and COVID-19: a tool for healthcare setting choice by primary care physicians. *Eur. J. Pub. Health* 31 ckab164.720.
- Wallentin, L., Lindback, J., Eriksson, N., Hijazi, Z., Eikelboom, J.W., Ezekowitz, M.D., Granger, C.B., Lopes, R.D., Yusuf, S., Oldgren, J., Siegbahn, A., 2020. Angiotensin-converting enzyme 2 (ACE2) levels in relation to risk factors for COVID-19 in two large cohorts of patients with atrial fibrillation. *Eur. Heart J.* 41, 4037–4046.
- Wang, J., Jiang, M., Chen, X., Montaner, L.J., 2020a. Cytokine storm and leukocyte changes in mild versus severe SARS-CoV-2 infection: review of 3939 COVID-19 patients in China and emerging pathogenesis and therapy concepts. *J. Leukoc. Biol.* 108, 17–41.
- Wang, L. Lu, Lo, K., Chandrasekhar, Y., Reas, R., Yang, J., Eide, D., Funk, K., Kinney, R., Liu, Z., Merrill, W., Mooney, P., Murdick, D., Rishi, D., Sheehan, J., Shen, Z., Stilson, B., Wade, A.D., Wang, K., Wilhelm, C., Xie, B., Raymond, D., Weld, D.S., Etzioni, O., Kohlmeier, S., 2020b. CORD-19: the Covid-19 open research dataset. *ArXiv*. <https://doi.org/10.48550/arXiv.2004.10706>.
- Worldometers.info, 2022. Worldometer: COVID-19 CORONAVIRUS PANDEMIC, Dover, Delaware, U.S.A.
- Wu, C., Chen, X., Cai, Y., Xia, J., Zhou, X., Xu, S., Huang, H., Zhang, L., Zhou, X., Du, C., Zhang, Y., Song, J., Wang, S., Chao, Y., Yang, Z., Xu, J., Zhou, X., Chen, D., Xiong, W., Xu, L., Zhou, F., Jiang, J., Bai, C., Zheng, J., Song, Y., 2020. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 Pneumonia in Wuhan, China. *JAMA Intern. Med.* 180, 934–943.
- Xie, Z., Bailey, A., Kuleshov, M.V., Clarke, D.J.B., Evangelista, J.E., Jenkins, S.L., Lachmann, A., Wojciechowski, M.L., Kropiwnicki, E., Jagodnik, K.M., Jeon, M., Ma'ayan, A., 2021. Gene set knowledge discovery with Enrichr. *Curr. Protoc.* 1, e90.
- Xu, J., Qi, L., Chi, X., Yang, J., Wei, X., Gong, E., Peh, S., Gu, J., 2006. Orchitis: a complication of severe acute respiratory syndrome (SARS). *Biol. Reprod.* 74, 410–416.
- Xu, Z., Li, Y., Lei, Q., Huang, L., Lai, D.-Y., Guo, S.-J., Jiang, H.-W., Hou, H., Zheng, Y.-X., Wang, X.-N., Wu, J., Ma, M.-L., Zhang, B., Chen, H., Yu, C., Xue, J.-B., Zhang, H.-N., Qi, H., Yu, S., Lin, M., Zhang, Y., Lin, X., Yao, Z., Sheng, H., Sun, Z., Wang, F., Fan, X., Tao, S.-C., 2021. COVID-ONE-hi: the one-stop database for COVID-19 specific humoral immunity and clinical parameters, Genomics, Proteomics & Bioinformatics.
- Yang, J., Zheng, Y., Gou, X., Pu, K., Chen, Z., Guo, Q., Ji, R., Wang, H., Wang, Y., Zhou, Y., 2020. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. *Int. J. Infect. Dis.* 94, 91–95.
- Zaim, S., Chong, J.H., Sankaranarayanan, V., Harky, A., 2020. COVID-19 and multiorgan response. *Curr. Probl. Cardiol.* 45, 100618.
- Zelaya, C.V.G., 2019. Towards Explaining the Effects of Data Preprocessing on Machine Learning, 2019 IEEE 35th International Conference on Data Engineering (ICDE), pp. 2086–2090.
- Zettler, M., Gajra, A., Feinberg, B., 2021. PRS30 COVID-19 RAPID antigen test false positives and false negatives reported to the FDA manufacturer and user facility device experience database. *Value Health* 24, S218.
- Zhao, J.M., Zhou, G.D., Sun, Y.L., Wang, S.S., Yang, J.F., Meng, E.H., Pan, D., Li, W.S., Zhou, X.S., Wang, Y.D., Lu, J.Y., Li, N., Wang, D.W., Zhou, B.C., Zhang, T.H., 2003. Clinical pathology and pathogenesis of severe acute respiratory syndrome. *Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi* 17, 217–221.
- Zheng, Q., Jones, F.K., Leavitt, S.V., Ung, L., Labrique, A.B., Peters, D.H., Lee, E.C., Azman, A.S., Adhikari, B., Wahl, B., Sarnowski, C., Antiporta, D.A., Erchick, D.J., Perez-Saez, J., Ssekasanvu, J., Lee, K.H., White, L., Kostandova, N.S., Menezes, N.P., Albaugh, N.W., Gupta, N., Jiawani, S.S., Hegde, S.T., Srivastava, S., Aung, T., Zhang, Y., Norton, G., Kalra, A., Khaitan, A., Shah, D., Kaur, J., Kasi, K., Patel, L., Dhingra, L.S., Agarwal, M., Garg, S., Goel, U., Gill, V.J.S., Khan, E., Patwari, A., Khaloo, P., Joshi, D., Blagg, E., Pence, E., Nelson, H.K., Fan, J., Forbes, L., Schluskel, M., Etyemez, S., Song, S., Mohan, U., Sun, Y., Jang, S., Frumentio, N., Sivakumar, A., Hartner, A.-M., Karandikar, V., Yan, Z., Beiter, E.R., Song, J., Wedlund, L., Singer, M., Rahman, R.S., Virk, Z.M., Abar, A., Tiu, B., Adamson, T.M., Paudel, K., Yao, H., Wang, Y., Li-Rodenborn, E.R., Ozdemir, I., McLean, M.-G., Rattigan, S., Borgert, B.A., Moreno, C.A., Quigley, N.E., Li, C., Kaur, N., Gimbrone, C., Scales, S.E., Zuniga-Moya, J.C., Babigumira, P.A., Igwe, C.C., Wang, H.E., Hsieh, L., Misra, S., Bruton, K., Byng, D., Miranda-Schaeubinger, M., Uddin, M.N., Ticehurst, J.R., Laney, E., Bhadauria, A., Gupta, V., Selles, M.C., Kartik, A., Singh, A., Garg, D., Saini, J., Kaur, J., Kaur, M., Denis, L., Ekong, I., Renyuanouyang, T., Fred, A.S.-H., Liu, M.R., Petersen, P., Agbadi, I., Segawa, V., Scott, Y., Shen, J., Okeeffe, Z., Brennan, M., Singh, A., Saini, M., Ndukwe, A., Vanajan, J., Minder, E., Lemaitre, L., Pi, M., Saini, M., Cabrera-Aguas, H.E., Zannat, A., Deng, N.-L.H., Nguyen, P.J., Hinson, L., Buysse, S., Kaur, C., Zhang, C., Saini, D.Y., Shu, H., Alemi, P., Shivshanker, R.B., Singh, T.B., McKay, X. Wang, Lee, S., Petersen, N., Zielonka, F., Chuang, A., Saussier, C., Dutra, D.A., Conlan, E.K., Specter, L.L.Y., Alhariri, M., Karahan, R., Yen, T., Bouchene, Y., Sultanov, A., Singh, N., H.-C. Collaboration, 2020. HIT-COVID, a global database tracking public health interventions to COVID-19. *Scientific Data* vol. 7, 286.
- Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L., Chen, H.D., Chen, J., Luo, Y., Guo, H., Jiang, R.D., Liu, M.Q., Chen, Y., Shen, X.R., Wang, X., Zheng, X.S., Zhao, K., Chen, Q.J., Deng, F., Liu, L.L., Yan, B., Zhan, F.X., Wang, Y.Y., Xiao, G.F., Shi, Z.L., 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579, 270–273.
- Zhu, A., Ibrahim, J.G., Love, M.I., 2018. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics* 35, 2084–2092.