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 [1], Harpreet Kaur[1*]

1. Pine Biotech, New Orleans, U.S.A.

Emails of Authors

Poonam Sen: senpoonam1490@gmail.com

Harpreet Kaur: hks04180@gmail.com; harpreet@pine.bio

*** Correspondence**

Harpreet Kaur, PhD

Curriculum Developer & Research Consultant

Pine Biotech

1441 Canal St #229, New Orleans, LA 70112, United States

Email: hks04180@gmail.com; harpreet@pine.bio

ORCID ID: https://orcid.org/0000-0003-0421-8341

Homepage: https://www.linkedin.com/in/harpreet-kaur-phd-b06722134/

Abstract

COVID-19 is a severe respiratory disease caused by SARS-CoV-2, a novel human coronavirus. The host response to SARS-CoV-2 infection is not clearly understood. Patients infected with SARS-CoV-2 exhibit heterogeneous intensity of symptoms, i.e., asymptomatic, mild, and severe. Moreover, effects on organs also vary from person to person. These heterogeneous responses pose pragmatic hurdles for implementing appropriate therapy and management of COVID-19 patients. Post-COVID complications pose another major challenge in managing the health of these patients. Thus, understanding the impact of disease severity at the molecular level is vital to delineate the precise host response and management. In the current study, we performed a comprehensive transcriptomics analysis of publicly available seven asymptomatic and eight severe COVID-19 patients. Exploratory data analysis using Principal Component Analysis (PCA) showed the distinct clusters of asymptomatic and severe patients. Subsequently, the differential gene expression analysis using DESeq2 identified 1,224 significantly upregulated genes (logFC>= 1.5, p-adjusted value <0.05) and 268 significantly downregulated genes (logFC<= -1.5, p-adjusted value <0.05) in severe samples in comparison to asymptomatic samples. Eventually, Gene Set Enrichment Analysis (GSEA) of upregulated genes revealed significant enrichment of terms, i.e., anti-viral and anti-inflammatory pathways, secondary infections, Iron homeostasis, anemia, cardiac-related, etc. Gene set enrichment analysis of downregulated genes indicates lipid metabolism, adaptive immune response, translation, recurrent respiratory infections, heme-biosynthetic pathways, etc. In summary, severe COVID-19 patients are more susceptible to other health issues/concerns, non-viral pathogenic infections, atherosclerosis, autoinflammatory diseases, anemia, male infertility, etc. And eventually, these findings provide insight into the precise therapeutic management of severe COVID-19 patients and efficient disease management.

Keywords

SARS-CoV-2, 2019-nCoV, COVID-19, Transcriptomics, Pathways, DGE (Differentially Expressed Genes), ARDS (acute respiratory distress syndrome)

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 (**S**evere **A**cute **R**espiratory **S**yndrome related novel **Co**rona**v**irus **2**), and the disease caused by it COVID-19 (**Co**rona**vi**rus **D**isease 2019) [1-3]. 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 [1]. As of 1st April 2022, the total reported cases worldwide stand at 488,190,137 [4]. 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 [3, 5, 6]. The trimeric spike (S) protein that forms the virus's envelope plays an essential role in the virus-host cell interaction [7]. 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 [8]. 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 [9].

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 [10-12]. Our health sector has faced major challenges in tackling disease spread and providing management of symptoms in the patients [13]. Multiple drugs are introduced for symptomatic treatments, but none has been efficient to treat all symptoms caused due to the viral infection [14]. 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 [15]. 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 [16-19].

SARS-CoV-2 impacts individuals in peculiar ways [16]. Most infected subjects are asymptomatic or mildly symptomatic, but some develop severe symptoms [16]. Comorbidities such as diabetes mellitus, hypertension, cardiovascular disease (CVD), and advanced age further increase the risk of disease severity [20-23]. 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 [24]. 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 [25, 26]. 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 [27, 28]. 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 muti-organ damage and even death in severe cases [22, 29-33]. Most disease symptoms may persist for 10-15 days, with some may exist for a prolonged time [34, 35]. It is well known that even after the viral load declines significantly, many health issues persist in the COVID-19 recovered patients [36-38]. These post-COVID effects are observed mainly in hospitalized and severe patients and add to another layer of disease mismanagement [39, 40]. 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 [41].

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 [42]. 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 [33, 43-46]. 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 [45, 47]. 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 [48].

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 [49]. Diverse scientific groups across the globe have developed numerous resources and tools to compile and analyze the data from host and pathogen [50-73].

Due to the systemic effects of COVID-19 infection, it becomes more challenging to treat patients with complex symptoms. Hence, we believe studying the differential mechanism operating in asymptomatic and severe COVID-19 patients can help us manage disease manifestations in severe patients. Thus, we performed a comparative analysis of transcriptomics profiles of severe and asymptomatic COVID-19 patients using PCA and DESeq2. Exploratory analysis based on PCA of samples shows two clear, distinct clusters of severe and asymptomatic samples. The differential gene expression analysis revealed significantly altered transcriptomics patterns between these two groups. Subsequently, Gene Set Enrichment Analysis (GSEA) identified some of the key altered pathways and biological processes involved in severe patients compared to asymptomatic 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 [74]. 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 [75]. 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 [74]. 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.

Table 1: Detail of the study as derived from GEO [1]

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

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-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) [74]. 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 [76-80]. 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 [78, 79, 81, 82]. While DESeq2 is a DGE analysis tool that mandates data to be unprocessed read counts as integer values [83]. DESeq2 uses inbuilt methods to normalize for library size and hence does not require prior normalization [83-85].

The summary of workflow, including pre-processing and normalization, is depicted in Figure 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 [86]. Thus, genes with low or zero variance filtered out using the nzv (non-zero variance) function of the "Caret package" available in R [87]. 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 [77, 88]. 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 [87]. The data matrix that resulted from PCA normalization contains 180 samples with log-transformed read counts for 35,587 gene rows.

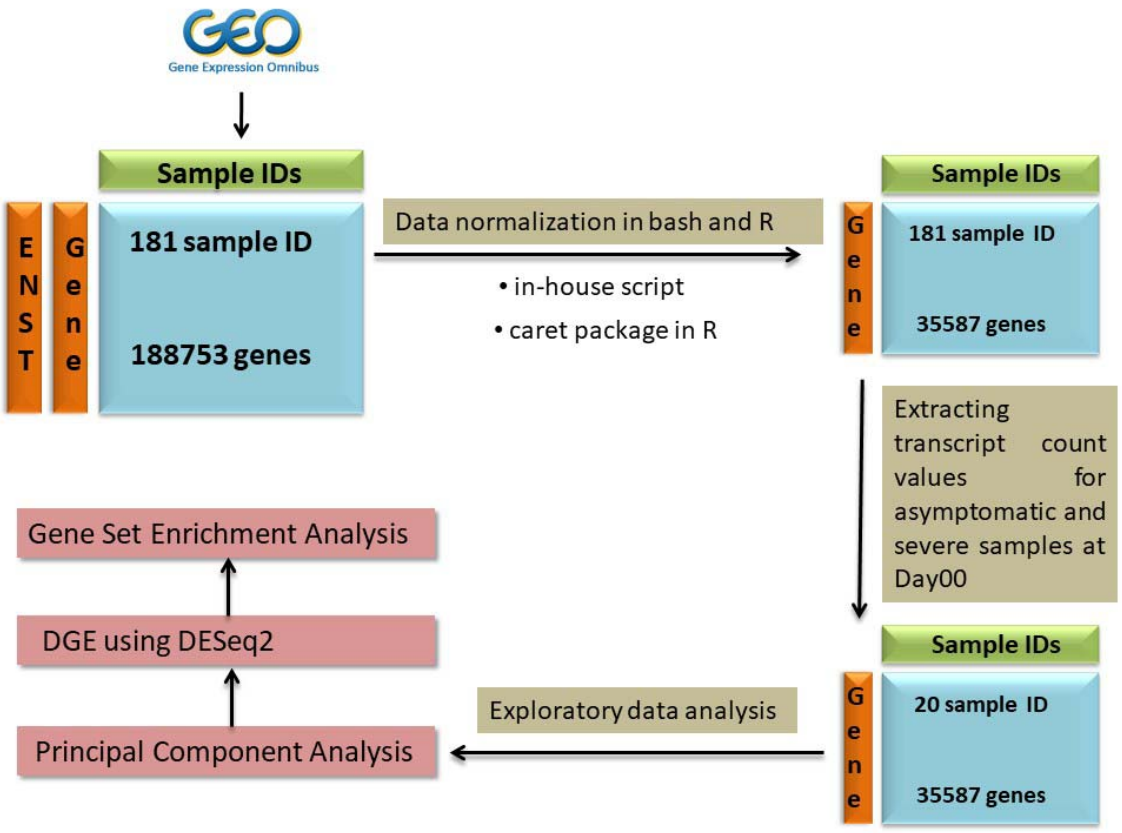


Figure 1. The Complete workflow of the study.

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 [89]. 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 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 [83]. 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 [83, 90-98]. Further, we applied another filter, i.e., Log2 fold change (Log2FC) to identify significantly upregulated (Log2FC >=1.5) and downregulated (Log2FC <=-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 [99]. Heatmap is a grid-like graphical representation of the expression of genes (in rows) in all the samples (in columns) taken into consideration [100].

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 [101-103]. We queried the upregulated and downregulated gene sets independently in the Enrichr search engine [104]. 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 [103].

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 [103]. 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. A publicly available RNA sequencing read count dataset was extracted, pre-processed, and normalized. We performed exploratory data analysis using PCA to understand variations between groups and to identify outliers. Subsequently, we performed differential gene expression analysis between these identified groups (Severe vs. Asymptomatic). Eventually, gene enrichment analysis was performed using the significantly differentially expressed gene sets to discern their biological involvement in viral immuno-pathogenesis.

3.1. Data Pre-processing

After the pre-processing, we are able to remove genes without identifiers, zero expression, and low variance in the data. Thus, the total number of genes 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 Figure 2 and Supplementary Figure 1 A-C. The scatter plot in Supplementary Figure 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 Figure 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 are 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 (Figure 2A). Using the clustering patterns in PCA, we identified seven samples as outliers. We removed these 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. So, 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 (Figure 2B). For a significant DGE analysis, the minimum number of samples in each group must be three, so we considered these four samples from both groups for further downstream analysis.

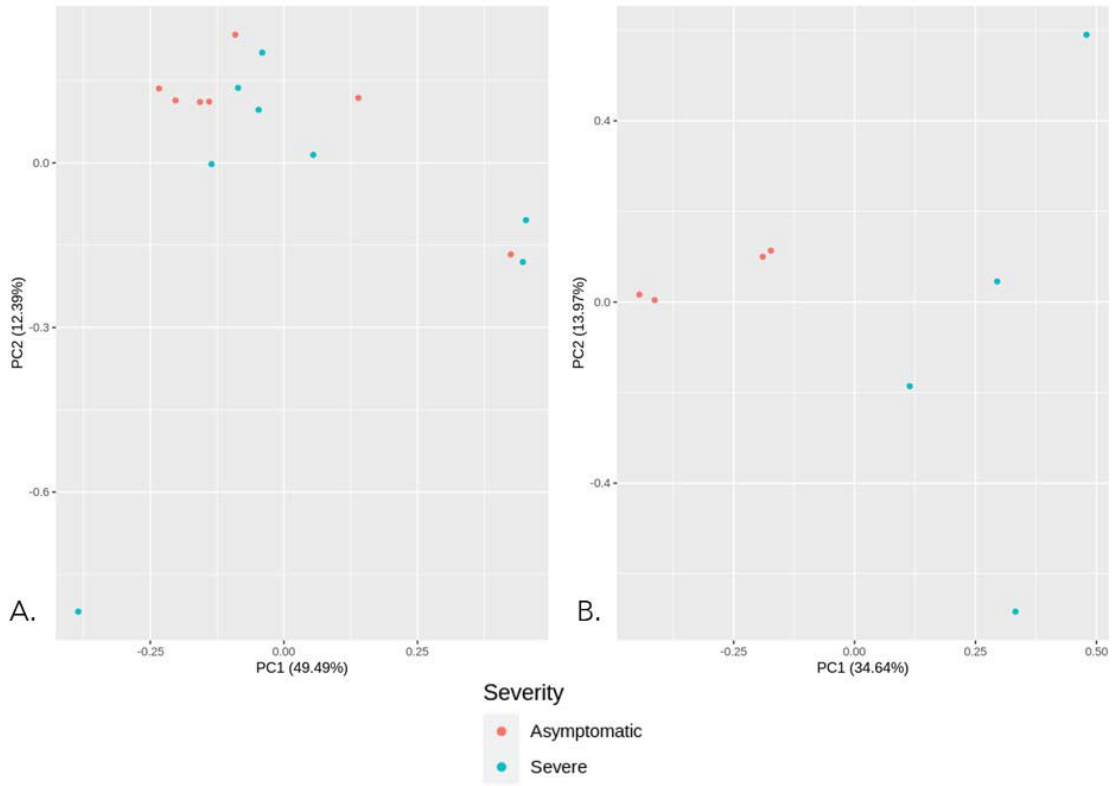


Figure 2. Principal Component Analysis between untreated asymptomatic and severe groups. A. PCA between asymptomatic (n=7, Day00) v/s severe samples (n=8, Day00) B. PCA between asymptomatic (n=7, Day00) v/s severe samples (n=8, Day00) after outlier removal.

3.3. Differential gene expression analysis

Differential gene expression analysis between untreated severe and asymptomatic samples using DESeq2 identified 2,837 genes as significantly differentially expressed (p adjusted value < 0.05). From these 2,837 genes, 1224 genes were found to be significantly upregulated (Log2FC>= 1.5, p-adjusted value <0.05) and 268 genes as significantly downregulated (Log2FC<= -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 (Figure 3). Each dot in the plot represents a single gene with log2FC along the x-axis and -Log10 (p-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 padj <0.01 and padj <0.05, respectively. Further, heatmap (Figure 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 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. The top 25 upregulated and down regulated genes ae mentioned in Table 2. with respective gene description.

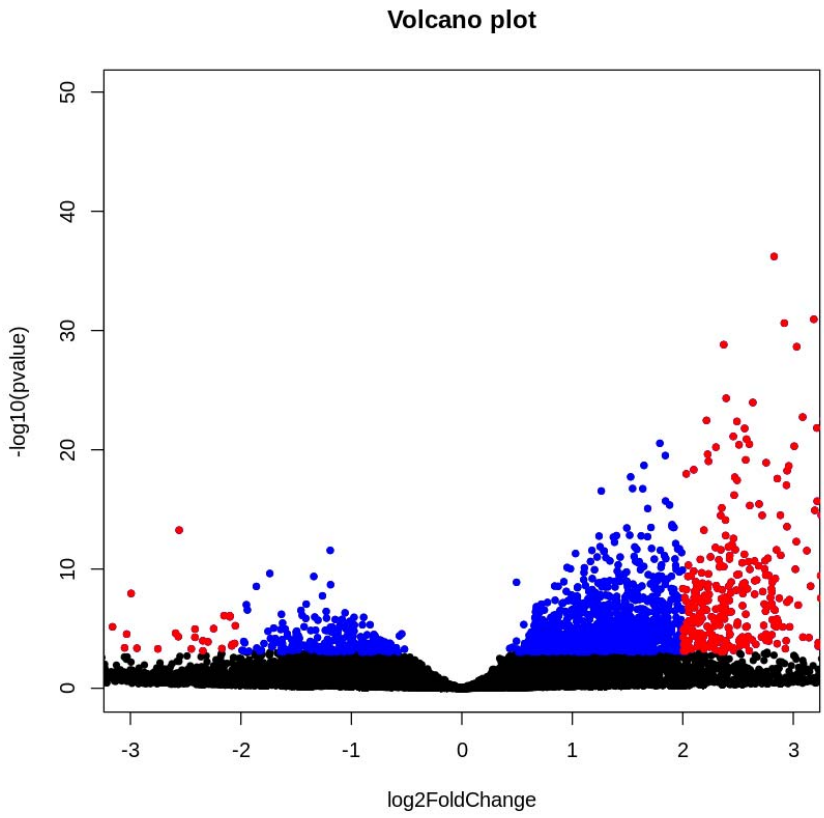


Figure 3. Volcano plot based on p-value and log2FC. Each dot here represents a single gene. Black represents nonsignificant genes, blue and red represent genes differentially regulated at padj <0.01 and padj <0.05.

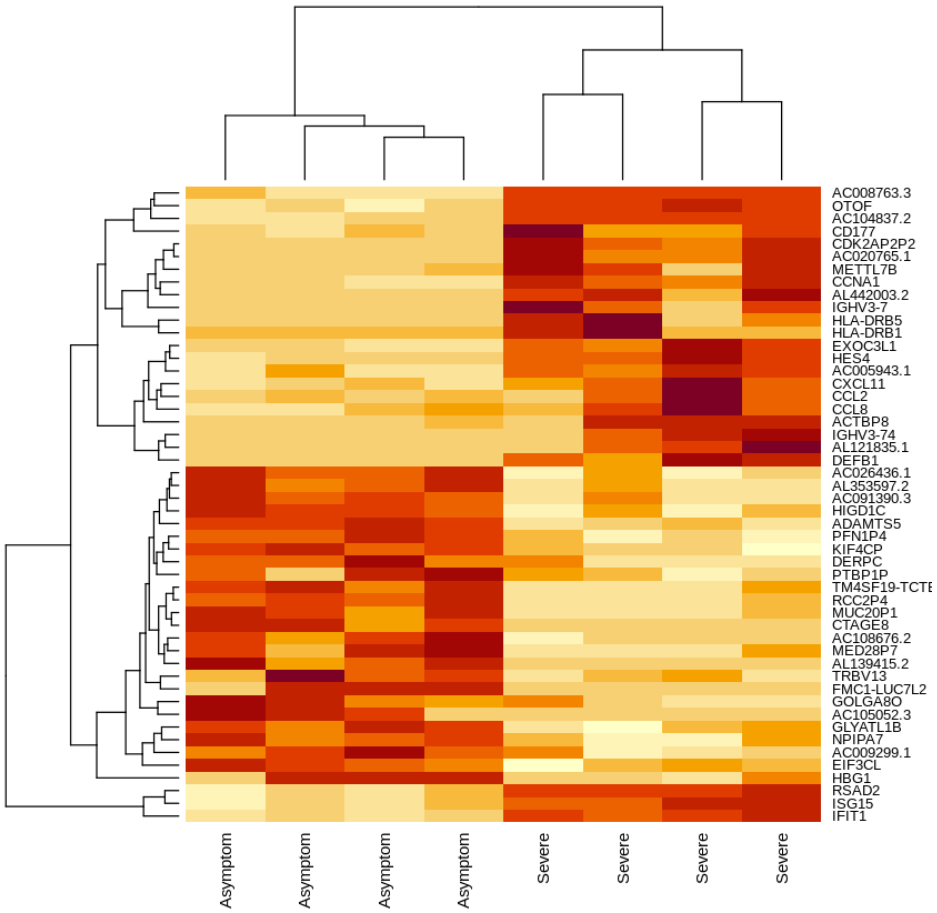


Figure 4. Heatmap based on top 25 upregulated and downregulated genes from DESEQ2 of asymptomatic and severe samples.

Table 2: List of top 25 upregulated (Log2FC>= 1.5, p-adjusted value <0.05) and downregulated (Log2FC<= -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 [105, 106]	***AC105052.3***	sense overlapping [107]
HLA-DRB5	Protein coding, Major Histocompatibility Complex,	***FMC1-LUC7L2***	Protein coding, FMC1-LUC7L2 Readthrough [105,

	Class II, DR Beta 5 [105, 106]		106]
IGHV3-7	Protein coding, Immunoglobulin Heavy Variable 3-7 [105, 106]	***CTAGE8***	Protein coding, CTAGE Family Member 8 [105, 106]
ACTBP8	Pseudogene, ACTB Pseudogene 8 [105, 106]	***AL139415.2***	Processed pseudogenes [107]
AC008763.3	Novel protein [108, 109]	***ADAMTS5***	Protein coding, ADAM Metallopeptidase with Thrombospondin Type 1 Motif 5 [105, 106]
CD177	Protein coding, CD177 Molecule [105, 106]	***MED28P7***	Pseudogene, Mediator Complex Subunit 28 Pseudogene 7 [105, 106]
AL442003.2	NA	***MUC20P1***	Pseudogene, Mucin 20, Cell Surface Associated Pseudogene 1 [105, 106]
IGHV3-74	Protein coding, Immunoglobulin Heavy Variable 3-74 [105, 106]	***RCC2P4***	Pseudogene, Regulator Of Chromosome Condensation 2 Pseudogene 4 [105, 106]
ISG15	Protein coding, ISG15 Ubiquitin Like Modifier [105, 106]	***PFN1P4***	Pseudogene, Profilin 1 Pseudogene 4 [105, 106]
CCL2	Protein coding, C-C Motif Chemokine Ligand 2 [105, 106]	***AL353597.2***	processed transcript, transcribed processed pseudogene [107]
CCL8	Protein coding, C-C Motif Chemokine Ligand 8 [105, 106]	***HBG1***	Protein coding, Hemoglobin Subunit Gamma 1 [105, 106]
OTOF	Protein coding, Otoferlin [105, 106]	***GOLGA8O***	Protein coding, Golgin A8 Family Member O [105, 106]

RSAD2	Protein coding, Radical S-Adenosyl Methionine Domain Containing 2 [105, 106]	***TM4SF19-TCTEX1D2***	RNA Gene, TM4SF19-DYNLT2B Readthrough (NMD Candidate) [105, 106]
METTL7B	Protein coding, Methyltransferase Like 7B [105, 106]	***DERPC***	Protein coding, DERPC Proline and Glycine Rich Nuclear Protein [105, 106]
HES4	Protein coding, Hes Family BHLH Transcription Factor 4 [105, 106]	***TRBV13***	Protein coding, T Cell Receptor Beta Variable 13 [105, 106]
CDK2AP2P2	Pseudogene, PTGER4P2-CDK2AP2P2 Readthrough, Transcribed Pseudogene [105, 106]	***AC108676.2***	Processed pseudogenes [107]
IFIT1	Protein coding, Interferon Induced Protein With Tetratricopeptide Repeats 1 [105, 106]	***AC009299.1***	Processed pseudogenes [107]
AL121835.1	Processed pseudogenes [107]	***HIGD1C***	Protein coding, HIG1 Hypoxia Inducible Domain Family Member 1C [105, 106]
CCNA1	Protein coding, Cyclin A1 [105, 106]	***PTBP1P***	Pseudogene, Polypyrimidine Tract Binding Protein 1 Pseudogene [105, 106]
DEFB1	Protein coding, Defensin Beta 1 [105, 106]	***EIF3CL***	Protein coding, Eukaryotic Translation Initiation Factor 3 Subunit C Like [105, 106]
AC104837.2	Processed pseudogenes [107]	***KIF4CP***	Pseudogene, Kinesin Family

			Member 4C, Pseudogene [105, 106]
AC020765.1	Processed pseudogenes [107]	***NPIPA7***	Protein coding, Nuclear Pore Complex Interacting Protein Family Member A7 [105, 106]
AC005943.1	nonsense mediated decay [107]	***AC091390.3***	unprocessed pseudogene [107]
EXOC3L1	Protein coding, Exocyst Complex Component 3 Like 1 [105, 106]	***GLYATL1B***	Protein coding, Glycine-N-Acyltransferase Like 1B [105, 106]
CXCL11	Protein coding, C-X-C Motif Chemokine Ligand 11 [105, 106]	***AC026436.1***	Processed pseudogenes [107]

3.2.5. Biological annotation - Gene Enrichment analysis

We queried all 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 (Figures 5, Supplementary Figure S2-S4) and downregulated genes (Figures 6, Supplementary Figure 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:

3.2.5.1. Gene Set Enrichment Analysis of upregulated Genes

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 Signalling 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.

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 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).

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).

Association with male infertility

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).

Association with other important pathways for understanding host response

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 Signalling 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 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.

3.2.5.2. 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.

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 enrichment was not significant.

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).

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

409 elongation (GO:0006448), cytoplasmic translational initiation (GO:0002183), translation
410 Factors WP107.

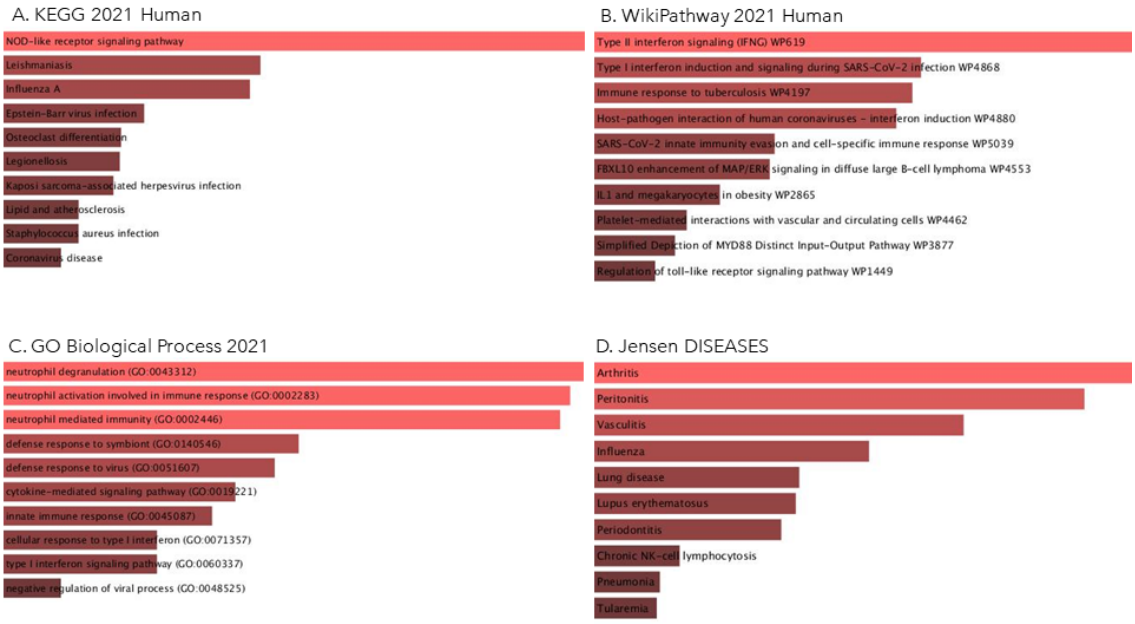
411 **Association with recurrent respiratory diseases and abnormal Heme biosynthesis**

412 Recurrent lower respiratory tract infections (HP:0002783), Agammaglobulinemia
413 (HP:0004432), Abnormality of the heme biosynthetic pathway (HP:0010472).

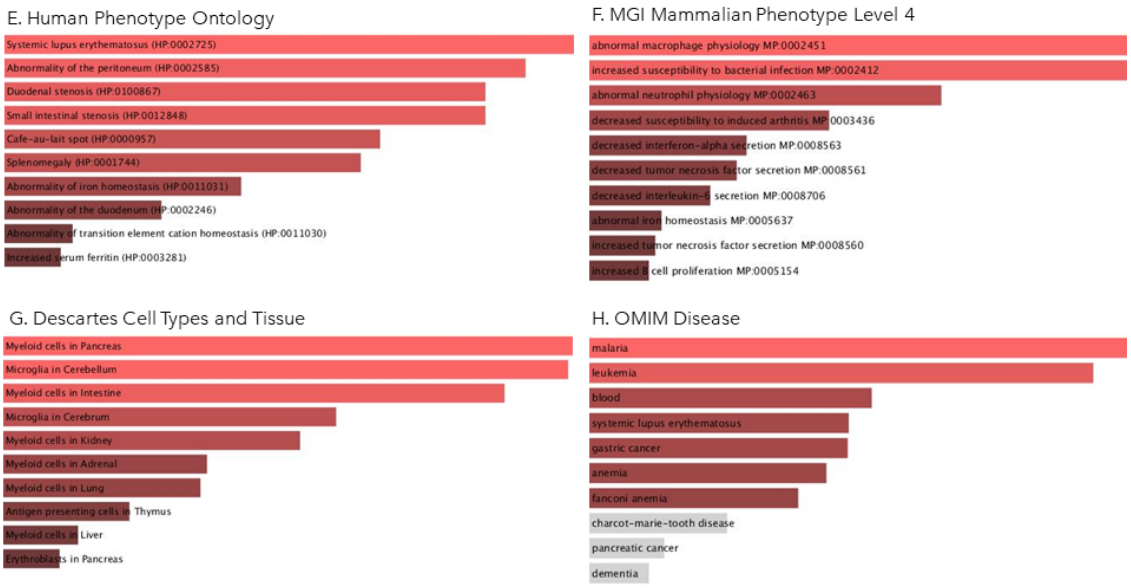
414 **Other signaling pathways**

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

419



420

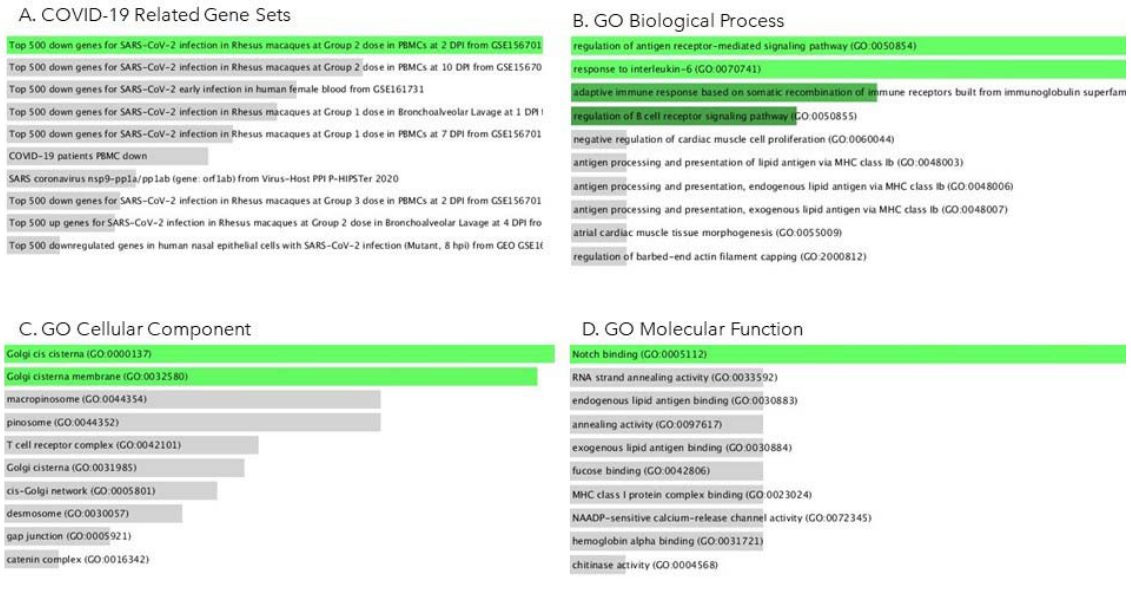


421

422 Figure 4: Ontologies and pathways upregulated in DESeq2 analysis of severe and asymptomatic

423 COVID-19 subjects using Enrichr database.

424



425

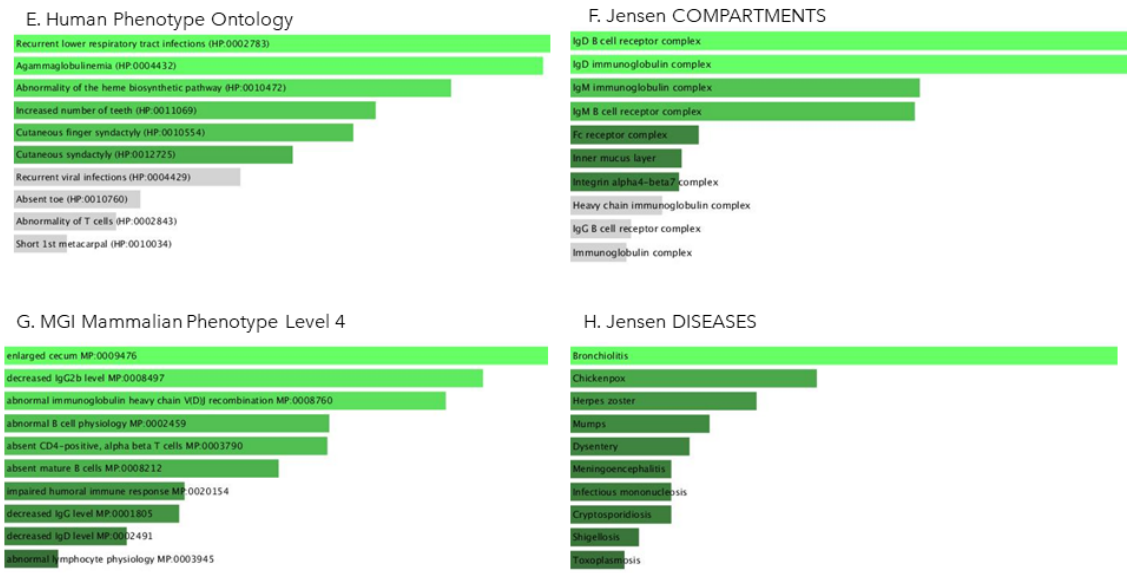


Figure 6: Ontologies and pathways downregulated in DESeq2 analysis of severe and asymptomatic COVID-19 subjects using Enrichr database.

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 [110]. 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. [8, 111-114]. The risk factors for COVID-19 severity include pre-existing comorbidities**,** particular age group of subjects, demographics, gender, etc. [5, 8, 23, 115, 116]. 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. This study extensively explored the transcriptomics profiles of two contrasting groups of COVID-19 patients, i.e., severe, and asymptomatic. 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. Identifying significant pathways involved in the patients might help manage the severity of the disease.

Exploratory analysis using Principal Component Analysis (PCA) shows distinct asymptomatic and severe subjects clusters. Subsequently, differential gene expression

analysis was performed between these two groups employing the DESeq2. We identified 2,837 genes as significantly differentially expressed between severe and asymptomatic COVID-19 subjects. To reduce false discovery rate (FDR) and increase statistical significance, we used a stringent filter (Bonferroni-padj value <0.05 and Log2fold Change (Log2FC)) and found 1,224 upregulated (Log2FC >= 1.5 and p-adjusted value <0.05) and 268 downregulated (Log2FC <= -1.5 and p-adjusted value <0.05) genes in severe in comparison to asymptomatic COVID-19 subjects. Further, to understand the alterations at the molecular and biological level, we queried differentially regulated genes in the Enrichr database.

Our study found pathways known to express in viral response, in general, and specific to COVID-19 infection. We observed the type II interferon (IFNG) pathway upregulated in severe subjects. While type I IFNG is generally activated in viral response, studies have found suppression of Type I IFN in SARS-CoV infections [117, 118]. 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 [119, 120]. Antigen-presenting cells (A.P.C.) in thymus enriched in severe patients also indicate an immune response to the virus. G.O. cellular components show increased ficolin and azurophil-rich granules secretion in severe subjects. These are also associated with the COVID-19 immune response [121-123]. Neutrophil count increases in COVID-19 infection [115, 124]. Our enrichment analysis also found neutrophil activation and neutrophil mediated immunity. Neutrophil degranulation is enhanced in response to inflammatory reactions in the body [125-127]. And previous studies have also reported increased neutrophil degranulation in response to COVID-19 infection in organisms other than humans [125]. One such study performed on the Rhesus macaque model shows increased neutrophil degranulation in young subjects compared to old subjects [125]. 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; we 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 G.O. 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 [128]. 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 [129, 130]. The immune response involved in viral and bacterial infection shares different immune components [131-136]. A study reveals simultaneous expression of both IFNα and IFNγ inhibits the expression of biomarkers associated with viral and bacterial infection [131]. 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 [118, 137]. We also observed NOD-like receptor signaling enrichment in the upregulated gene set, 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 [17, 29]. 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 Antigen-presenting cells in the thymus, and there are studies linking the thymus' role in Arrhythmia [138-140]. 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 [141], but the studies related to COVID-19's impact on testicular functionality are contradictory [142-144]. Studies done to detect viral RNA in semen showed different results, with the majority indicating the absence of SARS-CoV-2 RNA [145-149]. 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 [150]. 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. [125, 151].

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.

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 of 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 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 clinicians in managing COVID-19 patients and post-COVID complications, essentially, researchers in finding better therapeutic targets. However, analyzing more samples in both groups will help validate our findings.

Limitation of the Study

We face multiple challenges in the transcriptomic analysis of COVID-19 patients. There are limited host response data available. The lack of diversity of data (tissues, demographics, patient clinical characteristics) is also another limitation. Also, we need more data available from infection studies to establish a correlation between secondary infection and SARS-CoV-2. The study can be improved in the following ways. As observed in COVID-19 patients, the

immune response components also vary during infection, and hence many of these complications may also change with the course of disease progression.

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.

Acknowledgment

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

Funding

No funding

Conflict of Interest

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

Supplementary Data

Supplementary File 1: Supplementary Figures

Supplementary File 2: Supplementary Tables

References

[1] Timeline: WHO's COVID-19 response, World Health Organization.
[2] W. Director-General, WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020, World Health Organization, 2020.
[3] A.E. Gorbalenya, S.C. Baker, R.S. Baric, R.J. de Groot, C. Drosten, A.A. Gulyaeva, B.L. Haagmans, C. Lauber, A.M. Leontovich, B.W. Neuman, D. Penzar, S. Perlman, L.L.M. Poon, D.V. Samborskiy, I.A. Sidorov, I. Sola, J. Ziebuhr, V. Coronaviridae Study Group of the International Committee on Taxonomy of, The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2, Nature Microbiology, 5 (2020) 536-544.
[4] Worldometers.info, Worldometer: COVID-19 CORONAVIRUS PANDEMIC, Dover, Delaware, U.S.A.
[5] T. Asselah, D. Durantel, E. Pasmant, G. Lau, R.F. Schinazi, COVID-19: Discovery, diagnostics and drug development, J Hepatol, 74 (2021) 168-184.
[6] P. Zhou, X.L. Yang, X.G. Wang, B. Hu, L. Zhang, W. Zhang, H.R. Si, Y. Zhu, B. Li, C.L. Huang, H.D. Chen, J. Chen, Y. Luo, H. Guo, R.D. Jiang, M.Q. Liu, Y. Chen, X.R. Shen, X. Wang, X.S. Zheng, K. Zhao, Q.J. Chen, F. Deng, L.L. Liu, B. Yan, F.X. Zhan, Y.Y. Wang,

584 G.F. Xiao, Z.L. Shi, A pneumonia outbreak associated with a new coronavirus of probable
585 bat origin, Nature, 579 (2020) 270-273.
586 [7] J. Zhang, T. Xiao, Y. Cai, B. Chen, Structure of SARS-CoV-2 spike protein, Current
587 Opinion in Virology, 50 (2021) 173-182.
588 [8] S. Zaim, J.H. Chong, V. Sankaranarayanan, A. Harky, COVID-19 and Multiorgan
589 Response, Current Problems in Cardiology, 45 (2020) 100618.
590 [9] C.I. Paules, H.D. Marston, A.S. Fauci, Coronavirus Infections—More Than Just the
591 Common Cold, JAMA, 323 (2020) 707-708.
592 [10] L. Guruprasad, Human SARS CoV-2 spike protein mutations, Proteins: Structure,
593 Function, and Bioinformatics, 89 (2021) 569-576.
594 [11] G. Dudas, S.L. Hong, B.I. Potter, S. Calvignac-Spencer, F.S. Niatou-Singa, T.B.
595 Tombolomako, T. Fuh-Neba, U. Vickos, M. Ulrich, F.H. Leendertz, K. Khan, C. Huber, A.
596 Watts, I. Olendraite, J. Snijder, K.N. Wijnant, A. Bonvin, P. Martres, S. Behillil, A. Ayouba,
597 M.F. Maidadi, D.M. Djomsi, C. Godwe, C. Butel, A. Simaitis, M. Gabrielaite, M. Katenaite, R.
598 Norvilas, L. Raugaite, G.W. Koyaweda, J.K. Kandou, R. Jonikas, I. Nasvytiene, Z.
599 Zemeckiene, D. Gecys, K. Tamusauskaite, M. Norkiene, E. Vasiliunaite, D. Ziogiene, A.
600 Timinskas, M. Sukys, M. Sarauskas, G. Alzbutas, A.A. Aziza, E.K. Lusamaki, J.M. Cigolo,
601 F.M. Mawete, E.L. Lofiko, P.M. Kingebeni, J.M. Tamfum, M.R.D. Belizaire, R.G. Essomba,
602 M.C.O. Assoumou, A.B. Mboringong, A.B. Dieng, D. Juozapaite, S. Hosch, J. Obama, M.O.
603 Ayekaba, D. Naumovas, A. Pautienius, C.D. Rafai, A. Vitkauskiene, R. Ugenskiene, A.
604 Gedvilaite, D. Cereskevicius, V. Lesauskaite, L. Zemaitis, L. Griskevicius, G. Baele,
605 Emergence and spread of SARS-CoV-2 lineage B.1.620 with variant of concern-like
606 mutations and deletions, Nat Commun, 12 (2021) 5769.
607 [12] S. Kumar, T.S. Thambiraja, K. Karuppanan, G. Subramaniam, Omicron and Delta
608 variant of SARS-CoV-2: A comparative computational study of spike protein, Journal of
609 Medical Virology, 94 (2022) 1641-1649.
610 [13] J.M. Fegert, B. Vitiello, P.L. Plener, V. Clemens, Challenges and burden of the
611 Coronavirus 2019 (COVID-19) pandemic for child and adolescent mental health: a narrative
612 review to highlight clinical and research needs in the acute phase and the long return to
613 normality, Child Adolesc Psychiatry Ment Health, 14 (2020) 20.
614 [14] Therapeutics and COVID-19: living guideline, WHO-2019-nCoV-therapeutics-2022.2,
615 World Health Organization, 2022.
616 [15] S.-S. Jean, P.-R. Hsueh, Old and re-purposed drugs for the treatment of COVID-19,
617 Expert Review of Anti-infective Therapy, 18 (2020) 843-847.
618 [16] M. Ballow, C.L. Haga, Why Do Some People Develop Serious COVID-19 Disease After
619 Infection, While Others Only Exhibit Mild Symptoms?, J Allergy Clin Immunol Pract, 9 (2021)
620 1442-1448.
621 [17] B. Bakhshandeh, Z. Jahanafrooz, A. Abbasi, M.B. Goli, M. Sadeghi, M.S. Mottaqi, M.
622 Zamani, Mutations in SARS-CoV-2; Consequences in structure, function, and pathogenicity
623 of the virus, Microb Pathog, 154 (2021) 104831.
624 [18] P.M. Varghese, A.G. Tsolaki, H. Yasmin, A. Shastri, J. Ferluga, M. Vatish, T. Madan, U.
625 Kishore, Host-pathogen interaction in COVID-19: Pathogenesis, potential therapeutics and
626 vaccination strategies, Immunobiology, 225 (2020) 152008.
627 [19] M. Fung, J.M. Babik, COVID-19 in Immunocompromised Hosts: What We Know So Far,
628 Clin Infect Dis, 72 (2021) 340-350.
629 [20] C. Wu, X. Chen, Y. Cai, J. Xia, X. Zhou, S. Xu, H. Huang, L. Zhang, X. Zhou, C. Du, Y.
630 Zhang, J. Song, S. Wang, Y. Chao, Z. Yang, J. Xu, X. Zhou, D. Chen, W. Xiong, L. Xu, F.
631 Zhou, J. Jiang, C. Bai, J. Zheng, Y. Song, Risk Factors Associated With Acute Respiratory
632 Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in
633 Wuhan, China, JAMA Intern Med, 180 (2020) 934-943.
634 [21] J. Yang, Y. Zheng, X. Gou, K. Pu, Z. Chen, Q. Guo, R. Ji, H. Wang, Y. Wang, Y. Zhou,
635 Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a
636 systematic review and meta-analysis, Int J Infect Dis, 94 (2020) 91-95.
637 [22] L. Wallentin, J. Lindback, N. Eriksson, Z. Hijazi, J.W. Eikelboom, M.D. Ezekowitz, C.B.
638 Granger, R.D. Lopes, S. Yusuf, J. Oldgren, A. Siegbahn, 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 (2020) 4037-4046.
[23] Y. Chen, S.L. Klein, B.T. Garibaldi, H. Li, C. Wu, N.M. Osevala, T. Li, J.B. Margolick, G. Pawelec, S.X. Leng, Aging in COVID-19: Vulnerability, immunity and intervention, Ageing Res Rev, 65 (2021) 101205.
[24] K. McIntosh, COVID-19: Epidemiology, virology, and prevention, 2022.
[25] N. van Doremalen, T. Bushmaker, D.H. Morris, M.G. Holbrook, A. Gamble, B.N. Williamson, A. Tamin, J.L. Harcourt, N.J. Thornburg, S.I. Gerber, J.O. Lloyd-Smith, E. de Wit, V.J. Munster, Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1, N Engl J Med, 382 (2020) 1564-1567.
[26] H.A. Rothan, S.N. Byrareddy, The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak, J Autoimmun, 109 (2020) 102433.
[27] B. Brogna, C. Brogna, M. Petrillo, A.M. Conte, G. Benincasa, L. Montano, M. Piscopo, 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 (2021).
[28] D.L. Jones, M.Q. Baluja, D.W. Graham, A. Corbishley, J.E. McDonald, S.K. Malham, L.S. Hillary, T.R. Connor, W.H. Gaze, I.B. Moura, M.H. Wilcox, K. Farkas, 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 (2020) 141364.
[29] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet, 395 (2020) 497-506.
[30] J. Wang, M. Jiang, X. Chen, L.J. Montaner, 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 (2020) 17-41.
[31] L. Sher, Post-COVID syndrome and suicide risk, QJM, 114 (2021) 95-98.
[32] K.A. Scordo, M.M. Richmond, N. Munro, Post-COVID-19 Syndrome: Theoretical Basis, Identification, and Management, AACN Adv Crit Care, 32 (2021) 188-194.
[33] C.A. Devaux, J.M. Rolain, D. Raoult, ACE2 receptor polymorphism: Susceptibility to SARS-CoV-2, hypertension, multi-organ failure, and COVID-19 disease outcome, J Microbiol Immunol Infect, 53 (2020) 425-435.
[34] T. Klopfenstein, N.J. Kadiane-Oussou, L. Toko, P.Y. Royer, Q. Lepiller, V. Gendrin, S. Zayet, Features of anosmia in COVID-19, Med Mal Infect, 50 (2020) 436-439.
[35] C. Faes, S. Abrams, D. Van Beckhoven, G. Meyfroidt, E. Vlieghe, N. Hens, B.C.G.o.C.-H. Surveillance, Time between Symptom Onset, Hospitalisation and Recovery or Death: Statistical Analysis of Belgian COVID-19 Patients, International Journal of Environmental Research and Public Health, 17 (2020) 7560.
[36] J.P. Rogers, E. Chesney, D. Oliver, T.A. Pollak, P. McGuire, P. Fusar-Poli, M.S. Zandi, G. Lewis, A.S. David, 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 (2020) 611-627.
[37] B. Lamprecht, [Is there a post-COVID syndrome?], Pneumologe (Berl), 17 (2020) 398-405.
[38] C. Fernandez-de-Las-Penas, D. Palacios-Cena, V. Gomez-Mayordomo, M.L. Cuadrado, L.L. Florencio, Defining Post-COVID Symptoms (Post-Acute COVID, Long COVID, Persistent Post-COVID): An Integrative Classification, Int J Environ Res Public Health, 18 (2021).
[39] A. Carfi, R. Bernabei, F. Landi, C.-P.-A.C.S.G. Gemelli Against, Persistent Symptoms in Patients After Acute COVID-19, JAMA, 324 (2020) 603-605.
[40] E. Garrigues, P. Janvier, Y. Kherabi, A. Le Bot, A. Hamon, H. Gouze, L. Doucet, S. Berkani, E. Oliosi, E. Mallart, F. Corre, V. Zarrouk, J.D. Moyer, A. Galy, V. Honsel, B. Fantin, Y. Nguyen, Post-discharge persistent symptoms and health-related quality of life after hospitalization for COVID-19, J Infect, 81 (2020) e4-e6.

694 [41] P.-I. Lee, P.-R. Hsueh, Emerging threats from zoonotic coronaviruses-from SARS and
695 MERS to 2019-nCoV, Journal of Microbiology, Immunology and Infection, 53 (2020) 365-
696 367.
697 [42] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Kruger, T. Herrler, S. Erichsen, T.S.
698 Schiergens, G. Herrler, N.H. Wu, A. Nitsche, M.A. Muller, C. Drosten, S. Pohlmann, SARS-
699 CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven
700 Protease Inhibitor, Cell, 181 (2020) 271-280 e278.
701 [43] M.M. Lamers, J. Beumer, J. van der Vaart, K. Knoops, J. Puschhof, T.I. Breugem,
702 R.B.G. Ravelli, J. Paul van Schayck, A.Z. Mykytyn, H.Q. Duimel, E. van Donselaar, S.
703 Riesebosch, H.J.H. Kuijpers, D. Schipper, W.J. van de Wetering, M. de Graaf, M.
704 Koopmans, E. Cuppen, P.J. Peters, B.L. Haagmans, H. Clevers, SARS-CoV-2 productively
705 infects human gut enterocytes, Science, 369 (2020) 50-54.
706 [44] I. Hamming, W. Timens, M.L. Bulthuis, A.T. Lely, G. Navis, H. van Goor, Tissue
707 distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in
708 understanding SARS pathogenesis, J Pathol, 203 (2004) 631-637.
709 [45] F. Li, Receptor recognition and cross-species infections of SARS coronavirus, Antiviral
710 Res, 100 (2013) 246-254.
711 [46] W. Sungnak, N. Huang, C. Becavin, M. Berg, R. Queen, M. Litvinukova, C. Talavera-
712 Lopez, H. Maatz, D. Reichart, F. Sampaziotis, K.B. Worlock, M. Yoshida, J.L. Barnes,
713 H.C.A.L.B. Network, SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells
714 together with innate immune genes, Nat Med, 26 (2020) 681-687.
715 [47] T. Magrone, M. Magrone, E. Jirillo, Focus on Receptors for Coronaviruses with Special
716 Reference to Angiotensin- Converting Enzyme 2 as a Potential Drug Target - A Perspective,
717 Endocr Metab Immune Disord Drug Targets, 20 (2020) 807-811.
718 [48] J.C.-F. Meredith Wadman, Jocelyn Kaiser, Catherine Matacic, How does coronavirus
719 kill? Clinicians trace a ferocious rampage through the body, from brain to toes. [Jun;2020];
720 2020, Science, 2020.
721 [49] Y. Su, D. Chen, D. Yuan, C. Lausted, J. Choi, C.L. Dai, V. Voillet, V.R. Duvvuri, K.
722 Scherler, P. Troisch, P. Baloni, G. Qin, B. Smith, S.A. Kornilov, C. Rostomily, A. Xu, J. Li, S.
723 Dong, A. Rothchild, J. Zhou, K. Murray, R. Edmark, S. Hong, J.E. Heath, J. Earls, R. Zhang,
724 J. Xie, S. Li, R. Roper, L. Jones, Y. Zhou, L. Rowen, R. Liu, S. Mackay, D.S. O'Mahony,
725 C.R. Dale, J.A. Wallick, H.A. Algren, M.A. Zager, I.S.-S.C.B. Unit, W. Wei, N.D. Price, S.
726 Huang, N. Subramanian, K. Wang, A.T. Magis, J.J. Hadlock, L. Hood, A. Aderem, J.A.
727 Bluestone, L.L. Lanier, P.D. Greenberg, R. Gottardo, M.M. Davis, J.D. Goldman, J.R. Heath,
728 Multi-Omics Resolves a Sharp Disease-State Shift between Mild and Moderate COVID-19,
729 Cell, 183 (2020) 1479-1495 e1420.
730 [50] NCBI, NCBI Visual Data Dashboard.
731 [51] COVID Data Tracker, Centers for Disease Control and Prevention, Atlanta, GA: US
732 Department of Health and Human Services, CDC.
733 [52] E. Mathieu, H. Ritchie, E. Ortiz-Ospina, M. Roser, J. Hasell, C. Appel, C. Giattino, L.
734 Rodes-Guirao, Author Correction: A global database of COVID-19 vaccinations, Nat Hum
735 Behav, 5 (2021) 956-959.
736 [53] A. Talwai, V. Wing, Y. Itzkovich, A. Galaznik, A. Chatterjee, R. Jain, J. LaBonte, PIN83
737 The COVID-19 Research Database: Building One of the Largest PRO Bono Real-World
738 DATA Repositories, Value in Health, 24 (2021) S121.
739 [54] L. Lu Wang, K. Lo, Y. Chandrasekhar, R. Reas, J. Yang, D. Eide, K. Funk, R. Kinney, Z.
740 Liu, W. Merrill, P. Mooney, D. Murdick, D. Rishi, J. Sheehan, Z. Shen, B. Stilson, A.D. Wade,
741 K. Wang, C. Wilhelm, B. Xie, D. Raymond, D.S. Weld, O. Etzioni, S. Kohlmeier, CORD-19:
742 The Covid-19 Open Research Dataset, ArXiv, (2020).
743 [55] Q. Chen, A. Allot, Z. Lu, LitCovid: an open database of COVID-19 literature, Nucleic
744 Acids Res, 49 (2021) D1534-D1540.
745 [56] S. Patiyal, D. Kaur, H. Kaur, N. Sharma, A. Dhall, S. Sahai, P. Agrawal, L. Maryam, C.
746 Arora, G.P.S. Raghava, A Web-Based Platform on Coronavirus Disease-19 to Maintain
747 Predicted Diagnostic, Drug, and Vaccine Candidates, Monoclon Antib Immunodiagn
748 Immunother, 39 (2020) 204-216.

[57] S. Ouyang, Y. Wang, K. Zhou, J. Xia, LitCovid-AGAC: cellular and molecular level annotation data set based on COVID-19, Genomics Inform, 19 (2021) e23.
[58] C. Holy, S. Shah, N. Elangovanraaj, D. Krishnan, S. Gupta, P. Trivedi, M. Devulapally, A. Mohapatra, S. Johnston, P. Coplan, PIN117 Identification of Patients with COVID-19 Infection Prior to the New COVID-19 Diagnostic Code - a Premier Database Analysis, Value in Health, 23 (2020) S563.
[59] M. Zettler, A. Gajra, B. Feinberg, PRS30 COVID-19 RAPID Antigen Test False Positives and False Negatives Reported to the FDA Manufacturer and User Facility Device Experience Database, Value in Health, 24 (2021) S218-S218.
[60] J. Gong, J. Ou, X. Qiu, Y. Jie, Y. Chen, L. Yuan, J. Cao, M. Tan, W. Xu, F. Zheng, Y. Shi, B. Hu, 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 (2020) 833-840.
[61] A.K. Singh, G. Mishra, A. Maurya, G.T. Kulkarni, R. Awasthi, Biofabrication: An interesting tool to create in vitro model for COVID-19 drug targets, Med Hypotheses, 144 (2020) 110059.
[62] D.E. Faleiros, W. van den Bos, L. Botto, F. Scarano, TU Delft COVID-app: A tool to democratize CFD simulations for SARS-CoV-2 infection risk analysis, Sci Total Environ, 826 (2022) 154143.
[63] T. Riffe, E. Acosta, t.C.-D. team, Data Resource Profile: COVerAGE-DB: a global demographic database of COVID-19 cases and deaths, International Journal of Epidemiology, 50 (2021) 390-390f.
[64] A. Mahdi, P. Błaszczyk, P. Dłotko, D. Salvi, T.-S. Chan, J. Harvey, D. Gurnari, Y. Wu, A. Farhat, N. Hellmer, A. Zarebski, B. Hogan, L. Tarassenko, OxCOVID19 Database, a multimodal data repository for better understanding the global impact of COVID-19, Scientific Reports, 11 (2021).
[65] M.S. Park, H. Jo, H. Lee, S.Y. Jung, H.J. Hwang, Machine Learning-Based COVID-19 Patients Triage Algorithm Using Patient-Generated Health Data from Nationwide Multicenter Database, Infect Dis Ther, (2022).
[66] Z. Xu, Y. Li, Q. Lei, L. Huang, D.-y. Lai, S.-j. Guo, H.-w. Jiang, H. Hou, Y.-x. Zheng, X.-n. Wang, J. Wu, M.-l. Ma, B. Zhang, H. Chen, C. Yu, J.-b. Xue, H.-n. Zhang, H. Qi, S. Yu, M. Lin, Y. Zhang, X. Lin, Z. Yao, H. Sheng, Z. Sun, F. Wang, X. Fan, S.-C. Tao, COVID-ONE-hi: The One-stop Database for COVID-19 Specific Humoral Immunity and Clinical Parameters, Genomics, Proteomics & Bioinformatics, (2021).
[67] Q. Zheng, F.K. Jones, S.V. Leavitt, L. Ung, A.B. Labrique, D.H. Peters, E.C. Lee, A.S. Azman, B. Adhikari, B. Wahl, C. Sarnowski, D.A. Antiporta, D.J. Erchick, J. Perez-Saez, J. Ssekasanvu, K.H. Lee, L. White, N.S. Kostandova, N.P. Menezes, N.W. Albaugh, N. Gupta, S.S. Jiwani, S.T. Hegde, S. Srivastava, T. Aung, Y. Zhang, G. Norton, A. Kalra, A. Khaitan, D. Shah, J. Kaur, K. Kasi, L. Patel, L.S. Dhingra, M. Agarwal, S. Garg, U. Goel, V.J.S. Gill, E. Khan, A. Patwari, P. Khaloo, D. Joshi, E. Blagg, E. Pence, H.K. Nelson, J. Fan, L. Forbes, M. Schlussel, S. Etyemez, S. Song, U. Mohan, Y. Sun, S. Jang, N. Frumento, A. Sivakumar, A.-M. Hartner, V. Karandikar, Z. Yan, E.R. Beiter, J. Song, L. Wedlund, M. Singer, R.S. Rahman, Z.M. Virk, A. Abar, B. Tiu, T.M. Adamson, K. Paudel, H. Yao, Y. Wang, E.R. Li-Rodenborn, I. Ozdemir, M.-G. McLean, S. Rattigan, B.A. Borgert, C.A. Moreno, N.E. Quigley, C. Li, N. Kaur, C. Gimbrone, S.E. Scales, J.C. Zuniga-Moya, P.A. Babigumira, C.C. Igwe, H.E. Wang, L. Hsieh, S. Misra, K. Bruton, D. Byng, M. Miranda-Schaeubinger, M.N. Uddin, J.R. Ticehurst, E. Laney, A. Bhadauria, V. Gupta, M.C. Selles, A. Kartik, A. Singh, D. Garg, J. Saini, J. Kaur, M. Kaur, L. Denis, I. Ekong, 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, S. Lee, N. Petersen, F. Zielonka, A. Chuang, C. Saussier, D.A. Dutra, E.K. Conlan, L.L.Y. Specter, M. Alhariri, R. Karahan, T. Yen, Y. Bouchene, A. Sultanov, N. Singh, H.-C. Collaboration, HIT-COVID, a

global database tracking public health interventions to COVID-19, Scientific Data, 7 (2020) 286.
[68] R.H. Mulholland, E. Vasileiou, C.R. Simpson, C. Robertson, L.D. Ritchie, U. Agrawal, M. Woolhouse, J.L. Murray, H.R. Stagg, A.B. Docherty, C. McCowan, R. Wood, S.J. Stock, A. Sheikh, Cohort Profile: Early Pandemic Evaluation and Enhanced Surveillance of COVID-19 (EAVE II) Database, Int J Epidemiol, 50 (2021) 1064-1074.
[69] G. Khullar, M. Chandra, Virtual dermatopathology: A potential educational tool during COVID-19 pandemic, Dermatol Ther, 33 (2020) e13755.
[70] G. Benetti, N. Cardobi, G. Cardano, C. Arena, C. Micheletto, S. Guariglia, S. Montemezzi, C. Cavedon, CT-based radiomics as a tool to recognize COVID-19 positive patients, Phys Med, 92 (2021) S46-S46.
[71] M. Saffern, Cancer therapy tool informs COVID-19 vaccines, Nat Rev Immunol, 20 (2020) 352.
[72] G. Vetrugno, F. Foti, M. Di Pumpo, M. Cicconi, F. D'Ambrosio, D.I. La Milia, R. Pastorino, S. Boccia, G. Damiani, P. Laurenti, Machine learning and COVID-19: a tool for healthcare setting choice by primary care physicians, Eur J Public Health, 31 (2021) ckab164.720.
[73] R. Haider, T.S. Shamsi, N.A. Khan, Machine learning based decipherment of Cell Population Data: a promising hospital front-door screening tool for COVID-19, Am J Clin Pathol, 156 (2021) S101-S102.
[74] Baseline signatures associated with clinical, virologic, and immunologic outcomes in patients with mild to moderate COVID-19, 2021.
[75] J.J. Feld, C. Kandel, M.J. Biondi, R.A. Kozak, M.A. Zahoor, C. Lemieux, S.M. Borgia, A.K. Boggild, J. Powis, J. McCready, D.H.S. Tan, T. Chan, B. Coburn, D. Kumar, A. Humar, A. Chan, B. O'Neil, S. Noureldin, J. Booth, R. Hong, D. Smookler, W. Aleyadeh, A. Patel, B. Barber, J. Casey, R. Hiebert, H. Mistry, I. Choong, C. Hislop, D.M. Santer, D. Lorne Tyrrell, J.S. Glenn, A.J. Gehring, H.L.A. Janssen, B.E. Hansen, Peginterferon lambda for the treatment of outpatients with COVID-19: a phase 2, placebo-controlled randomised trial, Lancet Respir Med, 9 (2021) 498-510.
[76] J. Han, M. Kamber, J. Pei, 3 - Data Preprocessing, in: J. Han, M. Kamber, J. Pei (Eds.) Data Mining (Third Edition), Morgan Kaufmann, Boston, 2012, pp. 83-124.
[77] B. Malley, D. Ramazzotti, J.T.-y. Wu, Data Pre-processing, in: M.I.T.C. Data (Ed.) Secondary Analysis of Electronic Health Records, Springer International Publishing, Cham, 2016, pp. 115-141.
[78] C. Fan, M. Chen, X. Wang, J. Wang, B. Huang, A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data, Frontiers in Energy Research, 9 (2021).
[79] S. Garca, J. Luengo, F. Herrera, Data Preprocessing in Data Mining, Springer Publishing Company, Incorporated2014.
[80] C.V.G. Zelaya, Towards Explaining the Effects of Data Preprocessing on Machine Learning, 2019 IEEE 35th International Conference on Data Engineering (ICDE), 2019, pp. 2086-2090.
[81] J. Lever, M. Krzywinski, N. Altman, Principal component analysis, Nature Methods, 14 (2017) 641-642.
[82] I. Dinc, M. Sigdel, S. Dinc, M.S. Sigdel, M.L. Pusey, R.S. Aygun, Evaluation of Normalization and PCA on the Performance of Classifiers for Protein Crystallization Images, Proc IEEE Southeastcon, 2014 (2014).
[83] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biology, 15 (2014) 550.
[84] A. Zhu, J.G. Ibrahim, M.I. Love, Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences, Bioinformatics, 35 (2018) 2084-2092.
[85] R. Vaja, H. Kaur, M. Mazumder, E. Brodsky, In silico Analysis of Transcriptomic Profiling and Affected Biological pathways in Multiple Sclerosis, bioRxiv, (2021) 2021.2008.2015.456398.

[86] Y. Sha, J.H. Phan, M.D. Wang, Effect of low-expression gene filtering on detection of differentially expressed genes in RNA-seq data, Annu Int Conf IEEE Eng Med Biol Soc, 2015 (2015) 6461-6464.
[87] M. Kuhn, Building Predictive Models in R Using the caret Package, Journal of Statistical Software, 28 (2008) 1 - 26.
[88] J. Han, M. Kamber, J. Pei, Data Mining: Concepts and Techniques, Morgan Kaufmann, Boston, 2012.
[89] I.T. Jolliffe, Principal component analysis for special types of data, Springer2002.
[90] W. Haynes, Bonferroni Correction, in: W. Dubitzky, O. Wolkenhauer, K.-H. Cho, H. Yokota (Eds.) Encyclopedia of Systems Biology, Springer New York, New York, NY, 2013, pp. 154-154.
[91] R.A. Armstrong, When to use the Bonferroni correction, Ophthalmic and Physiological Optics, 34 (2014) 502-508.
[92] H. Kaur, A. Dhall, R. Kumar, G.P.S. Raghava, Identification of Platform-Independent Diagnostic Biomarker Panel for Hepatocellular Carcinoma Using Large-Scale Transcriptomics Data, Frontiers in Genetics, 10 (2020).
[93] O. Menyhart, B. Weltz, B. Győrffy, MultipleTesting.com: A tool for life science researchers for multiple hypothesis testing correction, PLOS ONE, 16 (2021) e0245824.
[94] M. Jafari, N. Ansari-Pour, Why, When and How to Adjust Your P Values?, Cell J, 20 (2019) 604-607.
[95] D.L. Streiner, Best (but oft-forgotten) practices: the multiple problems of multiplicity—whether and how to correct for many statistical tests, The American Journal of Clinical Nutrition, 102 (2015) 721-728.
[96] H. Kaur, S. Bhalla, G.P.S. Raghava, Classification of early and late stage liver hepatocellular carcinoma patients from their genomics and epigenomics profiles, PLOS ONE, 14 (2019) e0221476.
[97] S. Dolaner, H. Kaur, E. Brodsky, J. Panov, M. Mazumder, Identification of LncRNAs as Therapeutic Targets in Chronic Lymphocytic Leukemia, Columbia Undergraduate Science Journal (CUSJ), Vol. 15 (2021) (2021).
[98] H. Kaur, S. Bhalla, D. Garg, N. Mehta, G.P.S. Raghava, analysis and prediction of cholangiocarcinoma from transcriptomic profile of patients, Journal of Hepatology, vol. 73 (2020).
[99] M. Perry, heatmaps (2021). Flexible Heatmaps for Functional Genomics and Sequence Features, Bioconductor.
[100] B.B. Khomtchouk, D.J. Van Booven, C. Wahlestedt, 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 (2014) 30.
[101] E.Y. Chen, C.M. Tan, Y. Kou, Q. Duan, Z. Wang, G.V. Meirelles, N.R. Clark, A. Ma'ayan, Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool, BMC Bioinformatics, 14 (2013) 128.
[102] M.V. Kuleshov, M.R. Jones, A.D. Rouillard, N.F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S.L. Jenkins, K.M. Jagodnik, A. Lachmann, M.G. McDermott, C.D. Monteiro, G.W. Gundersen, A. Ma'ayan, Enrichr: a comprehensive gene set enrichment analysis web server 2016 update, Nucleic Acids Res, 44 (2016) W90-97.
[103] Z. Xie, A. Bailey, M.V. Kuleshov, D.J.B. Clarke, J.E. Evangelista, S.L. Jenkins, A. Lachmann, M.L. Wojciechowicz, E. Kropiwnicki, K.M. Jagodnik, M. Jeon, A. Ma'ayan, Gene Set Knowledge Discovery with Enrichr, Curr Protoc, 1 (2021) e90.
[104] E. Chen, M. Kuleshov, A. Bailey, Z. Xie, D.J.B. Clarke, J.E. Evangelista, M. Wojciechowicz, E. Kropiwnicki, K. Jagodnik, M. Jeon, S. Litz, M. Jones, C. Tan, Y. Kou, N. Clark, A. Rouillard, N. Fernandez, Q. Duan, Z. Wang, S. Koplev, S. Jenkins, A. Lachmann, M. McDermott, C. Monteiro, G. Gundersen, A. Ma'ayan, Enrichr search engine, the Ma'ayan Lab.
[105] G. Stelzer, N. Rosen, I. Plaschkes, S. Zimmerman, M. Twik, S. Fishilevich, T.I. Stein, R. Nudel, I. Lieder, Y. Mazor, S. Kaplan, D. Dahary, D. Warshawsky, Y. Guan-Golan, A.

912 Kohn, N. Rappaport, M. Safran, D. Lancet, The GeneCards Suite: From Gene Data Mining
913 to Disease Genome Sequence Analyses, Curr Protoc Bioinformatics, 54 (2016) 1 30 31-31
914 30 33.
915 [106] G. Stelzer, N. Rosen, I. Plaschkes, S. Zimmerman, M. Twik, S. Fishilevich, T.I. Stein,
916 R. Nudel, I. Lieder, Y. Mazor, S. Kaplan, D. Dahary, D. Warshawsky, Y. Guan-Golan, A.
917 Kohn, N. Rappaport, M. Safran, D. Lancet, GeneCards®: The Human Gene Database,
918 Springer, Singapore, 2016.
919 [107] B. Lang, A. Armaos, G.G. Tartaglia, RNAct: Protein-RNA interaction predictions for
920 model organisms with supporting experimental data, Nucleic Acids Res, 47 (2019) D601-
921 D606.
922 [108] M. Uhlen, L. Fagerberg, B.M. Hallstrom, C. Lindskog, P. Oksvold, A. Mardinoglu, A.
923 Sivertsson, C. Kampf, E. Sjostedt, A. Asplund, I. Olsson, K. Edlund, E. Lundberg, S. Navani,
924 C.A. Szigyarto, J. Odeberg, D. Djureinovic, J.O. Takanen, S. Hober, T. Alm, P.H. Edqvist, H.
925 Berling, H. Tegel, J. Mulder, J. Rockberg, P. Nilsson, J.M. Schwenk, M. Hamsten, K. von
926 Feilitzen, M. Forsberg, L. Persson, F. Johansson, M. Zwahlen, G. von Heijne, J. Nielsen, F.
927 Ponten, Proteomics. Tissue-based map of the human proteome, Science, 347 (2015)
928 1260419.
929 [109] M. Uhlen, P. Oksvold, L. Fagerberg, E. Lundberg, K. Jonasson, M. Forsberg, M.
930 Zwahlen, C. Kampf, K. Wester, S. Hober, H. Wernerus, L. Bjorling, F. Ponten, Towards a
931 knowledge-based Human Protein Atlas, Nat Biotechnol, 28 (2010) 1248-1250.
932 [110] A.C. Cunningham, H.P. Goh, D. Koh, Treatment of COVID-19: old tricks for new
933 challenges, Critical Care, 24 (2020) 91.
934 [111] S. Faubel, C.L. Edelstein, Mechanisms and mediators of lung injury after acute kidney
935 injury, Nat Rev Nephrol, 12 (2016) 48-60.
936 [112] B. Diao, C. Wang, R. Wang, Z. Feng, J. Zhang, H. Yang, Y. Tan, H. Wang, C. Wang,
937 L. Liu, Y. Liu, Y. Liu, G. Wang, Z. Yuan, X. Hou, L. Ren, Y. Wu, Y. Chen, Human kidney is a
938 target for novel severe acute respiratory syndrome coronavirus 2 infection, Nat Commun, 12
939 (2021) 2506.
940 [113] S. Naicker, C.W. Yang, S.J. Hwang, B.C. Liu, J.H. Chen, V. Jha, The Novel
941 Coronavirus 2019 epidemic and kidneys, Kidney Int, 97 (2020) 824-828.
942 [114] T. Chen, D. Wu, H. Chen, W. Yan, D. Yang, G. Chen, K. Ma, D. Xu, H. Yu, H. Wang,
943 T. Wang, W. Guo, J. Chen, C. Ding, X. Zhang, J. Huang, M. Han, S. Li, X. Luo, J. Zhao, Q.
944 Ning, Clinical characteristics of 113 deceased patients with coronavirus disease 2019:
945 retrospective study, BMJ, 368 (2020) m1091.
946 [115] A. Akter, T. Ahmed, I. Tauheed, M. Akhtar, S.I.A. Rahman, F. Khaton, F. Ahmmed, J.
947 Ferdous, M.H. Afrad, Z. Kawser, M. Hossain, R. Khondaker, M.A. Hasnat, M.A. Sumon, A.
948 Rashed, S. Ghosh, S.B. Calderwood, R.C. Charles, E.T. Ryan, P. Khatri, H.T. Maecker, G.
949 Obermoser, B. Pulendran, J.D. Clemens, S. Banu, T. Shirin, R.C. LaRocque, J.B. Harris,
950 T.R. Bhuiyan, F. Chowdhury, F. Qadri, Disease characteristics and serological responses in
951 patients with differing severity of COVID-19 infection: A longitudinal cohort study in Dhaka,
952 Bangladesh, PLoS Negl Trop Dis, 16 (2022) e0010102.
953 [116] N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei, J.
954 Xia, T. Yu, X. Zhang, L. Zhang, Epidemiological and clinical characteristics of 99 cases of
955 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study, Lancet, 395 (2020)
956 507-513.
957 [117] A.L. Totura, A. Whitmore, S. Agnihothram, A. Schafer, M.G. Katze, M.T. Heise, R.S.
958 Baric, Toll-Like Receptor 3 Signaling via TRIF Contributes to a Protective Innate Immune
959 Response to Severe Acute Respiratory Syndrome Coronavirus Infection, mBio, 6 (2015)
960 e00638-00615.
961 [118] P. Lotfinejad, Z. Asadzadeh, S. Najjary, M.H. Somi, K. Hajiasgharzadeh, A.
962 Mokhtarzadeh, A. Derakhshani, E. Roshani, B. Baradaran, COVID-19 Infection: Concise
963 Review Based on the Immunological Perspective, Immunol Invest, 51 (2022) 246-265.
964 [119] L. Filgueira, A. Larionov, N. Lannes, The Influence of Virus Infection on Microglia and
965 Accelerated Brain Aging, Cells, 10 (2021).

[120] K. Koushki, M. Salemi, S.M. Miri, Y. Arjeini, M. Keshavarz, A. Ghaemi, Role of myeloid-derived suppressor cells in viral respiratory infections; Hints for discovering therapeutic targets for COVID-19, Biomed Pharmacother, 144 (2021) 112346.
[121] P. Charitos, I. Heijnen, A. Egli, S. Bassetti, M. Trendelenburg, M. Osthoff, Functional Activity of the Complement System in Hospitalized COVID-19 Patients: A Prospective Cohort Study, Front Immunol, 12 (2021) 765330.
[122] A. Polycarpou, M. Howard, C.A. Farrar, R. Greenlaw, G. Fanelli, R. Wallis, L.S. Klavinskis, S. Sacks, Rationale for targeting complement in COVID-19, EMBO Mol Med, 12 (2020) e12642.
[123] A. Singh, N. Sood, V. Narang, A. Goyal, Morphology of COVID-19-affected cells in peripheral blood film, BMJ Case Rep, 13 (2020).
[124] N. Reusch, E. De Domenico, L. Bonaguro, J. Schulte-Schrepping, K. Bassler, J.L. Schultze, A.C. Aschenbrenner, Neutrophils in COVID-19, Front Immunol, 12 (2021) 652470.
[125] B.A. Rosa, M. Ahmed, D.K. Singh, J.A. Choreno-Parra, J. Cole, L.A. Jimenez-Alvarez, T.S. Rodriguez-Reyna, B. Singh, O. Gonzalez, R. Carrion, Jr., L.S. Schlesinger, J. Martin, J. Zuniga, M. Mitreva, D. Kaushal, S.A. Khader, IFN signaling and neutrophil degranulation transcriptional signatures are induced during SARS-CoV-2 infection, Commun Biol, 4 (2021) 290.
[126] E. Akgun, M.B. Tuzuner, B. Sahin, M. Kilercik, C. Kulah, H.N. Cakiroglu, M. Serteser, I. Unsal, A.T. Baykal, Proteins associated with neutrophil degranulation are upregulated in nasopharyngeal swabs from SARS-CoV-2 patients, PLoS One, 15 (2020) e0240012.
[127] M.L. Meizlish, A.B. Pine, J.D. Bishai, G. Goshua, E.R. Nadelmann, M. Simonov, C.H. Chang, H. Zhang, M. Shallow, P. Bahel, K. Owusu, Y. Yamamoto, T. Arora, D.S. Atri, A. Patel, R. Gbyli, J. Kwan, C.H. Won, C. Dela Cruz, C. Price, J. Koff, B.A. King, H.M. Rinder, F.P. Wilson, J. Hwa, S. Halene, W. Damsky, D. van Dijk, A.I. Lee, H. Chun, A neutrophil activation signature predicts critical illness and mortality in COVID-19, medRxiv, (2020).
[128] J.-M. Berthelot, F. Lioté, Y. Maugars, J. Sibilia, Lymphocyte Changes in Severe COVID-19: Delayed Over-Activation of STING?, Frontiers in Immunology, 11 (2020).
[129] C.D. Russell, C.J. Fairfield, T.M. Drake, L. Turtle, R.A. Seaton, D.G. Wootton, L. Sigfrid, E.M. Harrison, A.B. Docherty, T.I. de Silva, C. Egan, R. Pius, H.E. Hardwick, L. Merson, M. Girvan, J. Dunning, J.S. Nguyen-Van-Tam, P.J.M. Openshaw, J.K. Baillie, M.G. Semple, A. Ho, I.C. investigators, 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 (2021) e354-e365.
[130] A. De Bruyn, S. Verellen, L. Bruckers, L. Geebelen, I. Callebaut, I. De Pauw, B. Stessel, J. Dubois, Secondary infection in COVID-19 critically ill patients: a retrospective single-center evaluation, BMC Infect Dis, 22 (2022) 207.
[131] P. Bourgoin, G. Biechele, I. Ait Belkacem, P.E. Morange, F. Malergue, 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 (2020) 106-123.
[132] R. Tyagi, P.S. Bisen, Immune response activation and immunomodulation, BoD–Books on Demand2019.
[133] M. Klaas, P.R. Crocker, Sialoadhesin in recognition of self and non-self, Semin Immunopathol, 34 (2012) 353-364.
[134] W.B. Puryear, H. Akiyama, S.D. Geer, N.P. Ramirez, X. Yu, B.M. Reinhard, S. Gummuluru, Interferon-inducible mechanism of dendritic cell-mediated HIV-1 dissemination is dependent on Siglec-1/CD169, PLoS pathogens, 9 (2013) e1003291.
[135] P. Bourgoin, T. Soliveres, D. Ahriz, I. Arnoux, C. Meisel, N. Unterwalder, P.E. Morange, P. Michelet, F. Malergue, T. Markarian, Clinical research assessment by flow cytometry of biomarkers for infectious stratification in an Emergency Department, Biomark Med, 13 (2019) 1373-1386.

[136] P.J. Lester, K.H. Buick, J.W. Baty, A. Felden, J. Haywood, Different bacterial and viral pathogens trigger distinct immune responses in a globally invasive ant, Sci Rep, 9 (2019) 5780.
[137] A. Ray, B.N. Dittel, Interrelatedness between dysbiosis in the gut microbiota due to immunodeficiency and disease penetrance of colitis, Immunology, 146 (2015) 359-368.
[138] B. Long, W.J. Brady, A. Koyfman, M. Gottlieb, Cardiovascular complications in COVID-19, Am J Emerg Med, 38 (2020) 1504-1507.
[139] C. Kellogg, O. Equils, The role of the thymus in COVID-19 disease severity: implications for antibody treatment and immunization, Hum Vaccin Immunother, 17 (2021) 638-643.
[140] X. Dai, D. Zhang, C. Wang, Z. Wu, C. Liang, The Pivotal Role of Thymus in Atherosclerosis Mediated by Immune and Inflammatory Response, International Journal of Medical Sciences, 15 (2018) 1555-1563.
[141] A.B. Reis, F.C. Araujo, V.M. Pereira, A.M. Dos Reis, R.A. Santos, F.M. Reis, Angiotensin (1-7) and its receptor Mas are expressed in the human testis: implications for male infertility, J Mol Histol, 41 (2010) 75-80.
[142] E. Illiano, F. Trama, E. Costantini, Could COVID-19 have an impact on male fertility?, Andrologia, 52 (2020) e13654.
[143] J.M. Zhao, G.D. Zhou, Y.L. Sun, S.S. Wang, J.F. Yang, E.H. Meng, D. Pan, W.S. Li, X.S. Zhou, Y.D. Wang, J.Y. Lu, N. Li, D.W. Wang, B.C. Zhou, T.H. Zhang, [Clinical pathology and pathogenesis of severe acute respiratory syndrome], Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi, 17 (2003) 217-221.
[144] Y. Ding, L. He, Q. Zhang, Z. Huang, X. Che, J. Hou, H. Wang, H. Shen, L. Qiu, Z. Li, J. Geng, J. Cai, H. Han, X. Li, W. Kang, D. Weng, P. Liang, S. Jiang, 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 (2004) 622-630.
[145] F. Pan, X. Xiao, J. Guo, Y. Song, H. Li, D.P. Patel, A.M. Spivak, J.P. Alukal, X. Zhang, C. Xiong, P.S. Li, J.M. Hotaling, No evidence of severe acute respiratory syndrome-coronavirus 2 in semen of males recovering from coronavirus disease 2019, Fertil Steril, 113 (2020) 1135-1139.
[146] C. Song, Y. Wang, W. Li, B. Hu, G. Chen, P. Xia, W. Wang, C. Li, F. Diao, Z. Hu, X. Yang, B. Yao, Y. Liu, Absence of 2019 novel coronavirus in semen and testes of COVID-19 patientsdagger, Biol Reprod, 103 (2020) 4-6.
[147] N. Holtmann, P. Edimiris, M. Andree, C. Doehmen, D. Baston-Buest, O. Adams, J.S. Kruessel, A.P. Bielfeld, Assessment of SARS-CoV-2 in human semen-a cohort study, Fertil Steril, 114 (2020) 233-238.
[148] D. Li, M. Jin, P. Bao, W. Zhao, S. Zhang, Clinical Characteristics and Results of Semen Tests Among Men With Coronavirus Disease 2019, JAMA Netw Open, 3 (2020) e208292.
[149] M.A. Khalili, K. Leisegang, A. Majzoub, R. Finelli, M.K. Panner Selvam, R. Henkel, M. Mojgan, A. Agarwal, Male Fertility and the COVID-19 Pandemic: Systematic Review of the Literature, World J Mens Health, 38 (2020) 506-520.
[150] J. Xu, L. Qi, X. Chi, J. Yang, X. Wei, E. Gong, S. Peh, J. Gu, Orchitis: a complication of severe acute respiratory syndrome (SARS), Biol Reprod, 74 (2006) 410-416.
[151] J.M. Bartleson, D. Radenkovic, A.J. Covarrubias, D. Furman, D.A. Winer, E. Verdin, SARS-CoV-2, COVID-19 and the Ageing Immune System, Nat Aging, 1 (2021) 769-782.