Article

Study of Progression of COVID-19 in Indian Population based on Transcriptomic Approach

Mrinmoy Kshattry[1], Sarangthem Dinamani Singh[1], G. Narahari Sastry[2,3], Pankaj Bharali[1,2]*

[1]Center for Infectious Diseases, Biological Science and Technology Division, CSIR-North East Institute of Science and Technology (CSIR-NEIST), Jorhat-785006, Assam, India
[2]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad-201002, India
[3]Advanced Computation and Data Sciences Division, CSIR-North East Institute of Science and Technology, Jorhat-785006, Assam, India
* Correspondence: Dr. Pankaj Bharali, Scientist, Center for Infectious Diseases, Biological Science and Technology Division, CSIR-North East Institute of Science and Technology (CSIR-NEIST), Jorhat-785006, Assam, India. Email: pankajbharali98@gmail.com / pbharali@neist.res.in

Abstract: The pandemic of COVID-19 ravaged most countries and made the healthcare system go for a toss. The impact of the disease is different in each patient and it progresses differently. Based on the severity, the COVID-19 infection is stratified into three main categories- mild, moderate, and severe. In this study, we performed a transcriptomic study of different stages and studied the progression of the disease. The study was based on an Indian population of 28 COVID-19 patients, which were classified into different groups. Our analysis has shown that as the disease progresses, the genes involved in the degranulation of the neutrophils and galactose metabolism increase. Furthermore, we identified the hub proteins in each stage. TB is one of the comorbidities of COVID-19 and a comparative study was done to identify the preserved module of genes in both. Enrichment analysis showed that the members of this module are significantly involved in translation and ribosome synthesis.

Availability: Data are available as supplementary information

Keywords: transcriptomics; COVID-19; tuberculosis; progression; non-coding RNA; hub proteins

1. Introduction

Since its outbreak in 2019, severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has caused irreparable loss to mankind. Belonging to the family betacoronavirus, the SARS-CoV-2 infection i.e., COVID-19 was declared a pandemic by WHO on March 11, 2020 [1]. Localized in the naso-oropharyngeal tract, the virus primarily targets the respiratory and vascular systems. Clinical symptoms of the symptomatic patients include fever, cough, nausea, shortness of breath, and diarrhea [2]. The infection can be classified as mild, moderate, and severe based on the symptoms. According to the National Institute of Health (https://www.covid19treatmentguidelines.nih.gov/), patients with mild infection show fever, cough, and muscle pain symptoms. As the infection progresses to the moderate stage, evidence of lower respiratory disease is noted and the saturation of oxygen level in blood is >= 94%. In the severe stage, a drop in oxygen saturation level in blood is observed below 94%, in the co-occurrence of respiratory rate > 30 breaths/min and lung filtrates > 50%.

With the advent of newer technologies and NGS technologies, more light has been shed on the mechanism of COVID-19 infection and its severity. Whole genome sequencing of the viral variants revealed that the infections caused by the Omicron variants were less severe than that of the Delta variants [3]. It is reported in India that over 17 states and 2 union territories the average seropositivity was 10.14% and there was high infection and seroprevalence of COVID-19, particularly in NE India during the second wave [4,5]. The

 doi:10.20944/preprints202212.0431.v1

Machine Learning based approach with the association of serological data identifies the infection status of the vaccine recipient [6]. In a study cohort of unvaccinated children and adolescents, COVID-19 infection (Omicron) is less associated with severe cases, as compared to that of the Delta variant [7]. Single-cell RNA sequencing of bronchoalveolar lavage revealed the association between immunotypes of Natural Killer cells (NK cells) and the severity of the disease [8]. In the NK cells, the ORF10 was found to be differentially expressed during the infection. In the NK cells of severe patients, ORF10 was highly overexpressed compared to moderate infection. In addition to it, bactericidal/permeability-increasing fold containing family A, member 1 (BPIFA1), a secretory protein that regulates interferon signaling in times of inflammation, was overexpressed in infected lung cells [9,10]. Host factors play an important role in determining the response of the body against infections. Chromosome 3 of the human contains the major genetic factors that determine the severity of the human. These regions of the chromosome regulate the expression of chemokine receptors such as CCR5 [11].

Transcriptomic analysis has helped to study the molecular signatures in individuals with COVID-19 infection. Transcriptomic profiles of blood taken from infected individuals showed that expression of the genes involved in the activation of granulocytes demarcated the severe patients from the mild ones [12]. Similarly, iterative clustering and guide-gene selection 2 (ICGS2) analysis of single-cell transcription of bronchoalveolar cells (BAL) identified a distinct pattern between severe and mild patients. BAL extracted from mild patients consisted of mainly macrophages. In contrast, BAL from severe patients was enriched in neutrophils and macrophages [13]. A study of a population of Dubai with different severity identified that the expression of angiotensin-converting enzyme-2 (ACE2) was highest in severe patients as compared to mild and moderate patients [14].

In this study, we perform a transcriptomic study of an Indian population to understand the progression of COVID-19 through stages of different severity. The FastQ files of the study are publicly available and were downloaded from NCBI. (GSE196822). A comparative study was performed against tuberculosis (TB), for which the data was downloaded from NCBI too. (GSE181143)

2. Methods

2.1. Transcriptomic analysis

The reads are downloaded from NCBI in FastQ format. Fastp was used in the pipeline for trimming the adapters [15]. Quality control of the trimmed reads was performed using FastQC. The modified reads were then aligned against the reference genome *Homo sapiens GRCh38* with the help of Hisat2 [16]. Information regarding the mapped and unmapped reads will be stored in BAM files. Using *samtools*, the aligned reads were extracted [17]. Quantification of the mapped reads for each gene was done using the tool *featureCounts*. The final step required the normalization and log transformation of the reads using the R package, *DESeq2* [18]. To identify differentially expressed genes that were significant in each stage, a two-fold cut-off was applied: $log2foldchange > 1$ and $padj < 0.05$.

2.2. Gene ontologies and pathways

The gene ontologies and the pathway were studied using the tool Enrichr [19,20,21]. The results which had an adjusted p-value of less than 0.1 were selected. The final results were arranged in descending order of their Rich factor (%). The rich factor is calculated as the ratio of the number of genes from the input list that were enriched in a function to the number of genes annotated in the pathway.

2.3. Construction of PPI network

For the analysis of the protein-protein interactions (PPI), the interactions were downloaded from the Human Interactome Project [22]. Additionally, the experimentally validated interactions from the Bioplex database were also included in the study [23,24,25].

 doi:10.20944/preprints202212.0431.v1

The network was constructed using the R package igraph and its degree was calculated [26]. The top 10 highly connected proteins (i.e., degree of the proteins) of a PPI network were considered hub proteins.

2.4. Preservation of modules of co-expression network of COVID-19 in TB

The gene co-expression network gives a cluster of genes whose expression is highly correlated. The significant DEGs were used as input to build a gene co-expression network for COVID-19 and TB using the R package WGCNA [27]. For the network, Pearson's correlations analysis of each gene pair was calculated to construct an adjacency matrix using the adjacency function of the WGCNA package. The adjacency matrix was, in turn, used to construct a scale-free co-expression network based on a soft-thresholding parameter βeta (β) to enrich strong correlations between gene pairs. For determining the scale-free nature of the network, a β was chosen such that the R2 of its corresponding β was greater than 0.85. The calculated adjacency matrix was used as an input for building a Topological Overlap Matrix (TOM) by using the function TOMsimilarity. Subsequently, the TOM was then used for performing hierarchical clustering using the flashClust function for module identification. Finally, the network modules for the test dataset (i.e., Covid-19) were identified using the, R package, dynamicTreeCut [28] with a minimum module size (i.e., number of genes in a module) being 30, and minimum sensitivity (deepSplit) = 2 for the gene dendrogram. The members of the modules were determined using the eigenvalue of the genes. Preservation analysis was performed to assess whether a module of a test network is preserved in the reference network or not. Modules were identified in COVID-19 and its preservation was studied in TB. The basic idea behind module preservation is to determine the preservation of genes within a module by comparing a test network (TB) with a reference network (COVID-19). The WGCNA function modulePreservation was executed to study the preservation analysis. Preservation of the module is conducted by calculating two parameters such as Zsummary and medianRank. A module with Zsummary greater than 8 is considered preserved between the two networks. However, this parameter is based on the module size. Hence, the value of medianRank is used to filter off further modules as it does not depend on the median size [29]. Modules with medianRank value lesser than 5 are considered preserved.

3. Results

3.1. Unique transcriptomic signatures at each stage

Transcriptomic analysis at each stage has shown unique gene signatures at each stage. The number of overexpressed genes was greater than that of under-expressed genes in each stage. Upon normalization, it was found that the number of under-expressed genes increased as the condition worsened whereas the relative number of overexpressed genes remained the same (Table 1). 275 genes were found to be significantly expressed in all the stages of the disease (Figure 1) Gene expression of 12% of the common genes gradually increased as the disease progressed. Enrichment analysis showed that these genes were significantly involved in inflammation, cytokine production, and the biosynthesis of an antibiotic such as neomycin and kanamycin (Figure 2). However, the genes whose expression decreased with the progression of infection were not enriched in any significant biological process. It was seen that the significant DEGs of Covid-19 were involved in biological processes such as innate immune response, gamma-aminobutyric acid signaling, and biological pathways such as diabetes mellitus and coagulation (Figure 3[a], [b]). The significant DEGs were used to construct a PPI network and the nodes with degrees greater than 21 were considered hubs. The genes that were identified as hubs were CREB3L1, KLRC1, KIFC3, SCN3B, TRAF1, CCDC33, CMTM5, C7orf34, NRSN1 and CEP70.

Table 1: The number of under-expressed and overexpressed genes in each stage of the disease.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 December 2022 doi:10.20944/preprints202212.0431.v1

Stage of Covid-19	Overexpressed genes (%)	Under expressed genes (%)	Total significant DEGs
Mild	282 (83.62 %)	42 (12.28 %)	342
Moderate	1803 (49.31 %)	887 (24.26 %)	3656
Severe	1388 (46.39 %)	841 (28.1 %)	2992

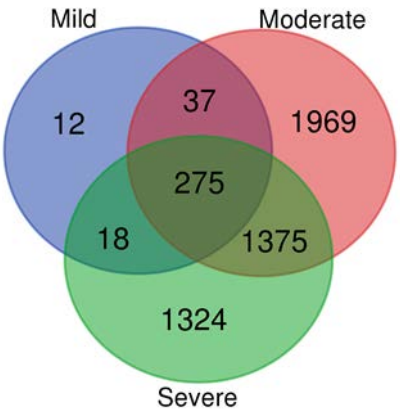


Figure 1: Venn diagram showing the union and intersection of significant DEGs from each stage.

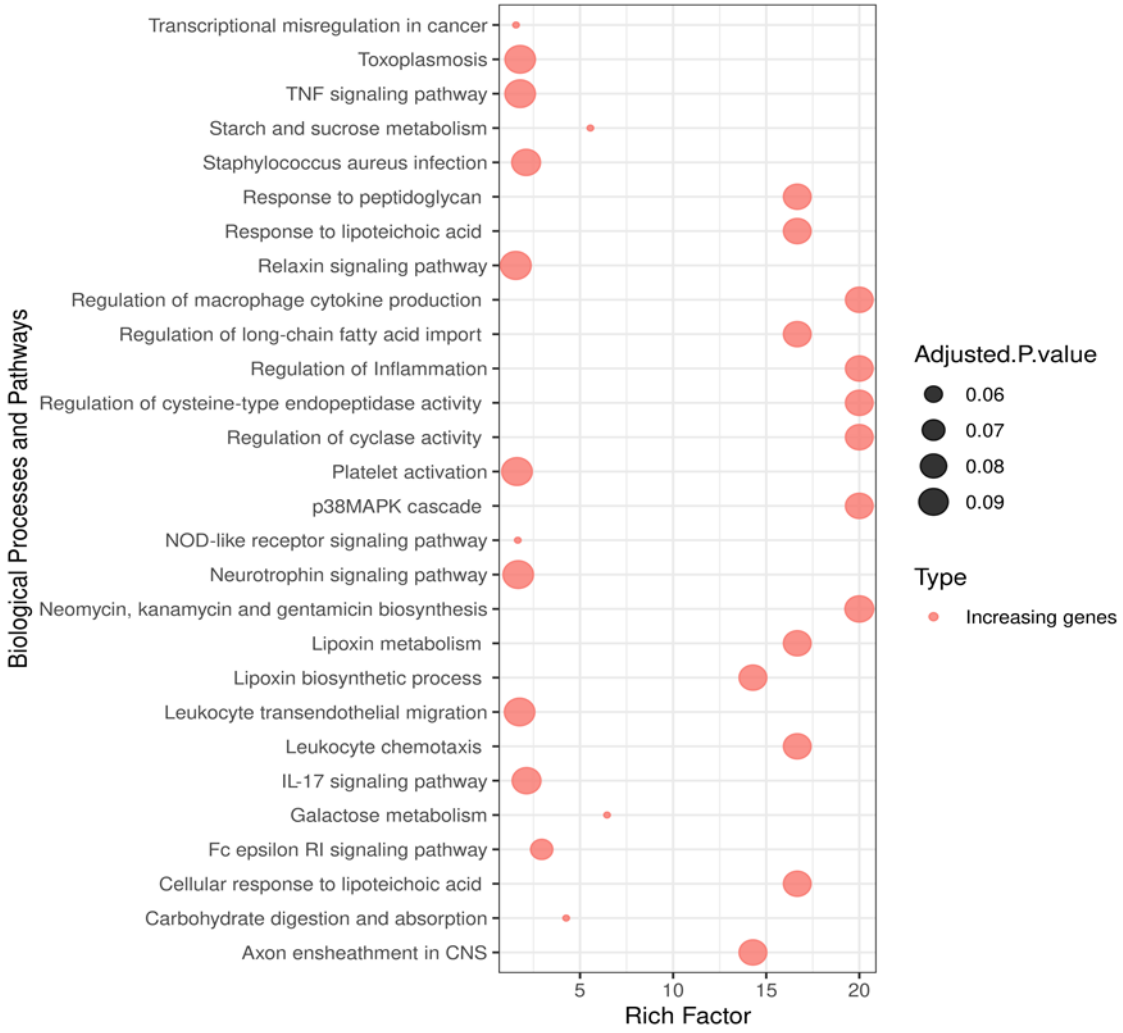


Figure 2: Ontology of the genes whose expression increases with the gradual progress of the disease.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 December 2022 doi:10.20944/preprints202212.0431.v1

3.2. Healthy vs Mild

The expression profile of the significant DEGs showed that the most significantly overexpressed gene was HIST1H1E. HIST1H1E is a gene whose expression is strongly associated with the risk of lung cancer [30]. Similarly, the most under-expressed gene was DZIP1, a component of the Hedgehog signaling pathway and is involved in mRNA regulation [31]. During the mild stage, the under-expressed genes were not found to be significantly enriched in any biological functions (Supplementary 1). On the other hand, the overexpressed genes were found to be involved in the metabolism of pantothenate and negatively regulated the production of interleukins such as IL-6 and IL-12 (Figure 4). The novel SARS-CoV-2 infection brings about hyperinflammation and is known to be caused by the markers such as IL-6 and IL-12 [32]. Pantothenate is a vitamin known to decrease inflammation and its excess metabolism may hinder the inflammation process [33]. Enrichment analysis has shown that the overexpressed genes are involved in pathways such as the degranulation of neutrophil and neutrophil immunity. Degree analysis of the PPI network showed that TLR2, ITGAM, MPO, FCGR3B, CD163, HIST1H1E, IL1R1, PYGL, HSPA1A, and HK3 acted as hub proteins in the network. These hubs play an important role in the inflammation process. TLR1, CD163, and FCGR3B are involved in the inflammatory process during Covid-19 [34,35] whereas HSPA1A and PYGL work as a countermeasure against the inflammation [36]. IL1R1 is a cytokine gene and along with ITGAM, is known to be strongly correlated with the severity of the infection [37,38]. The significant DEGs of the mild stage were insufficient enough to construct a significant PPI network and hence the hub genes were not calculated.

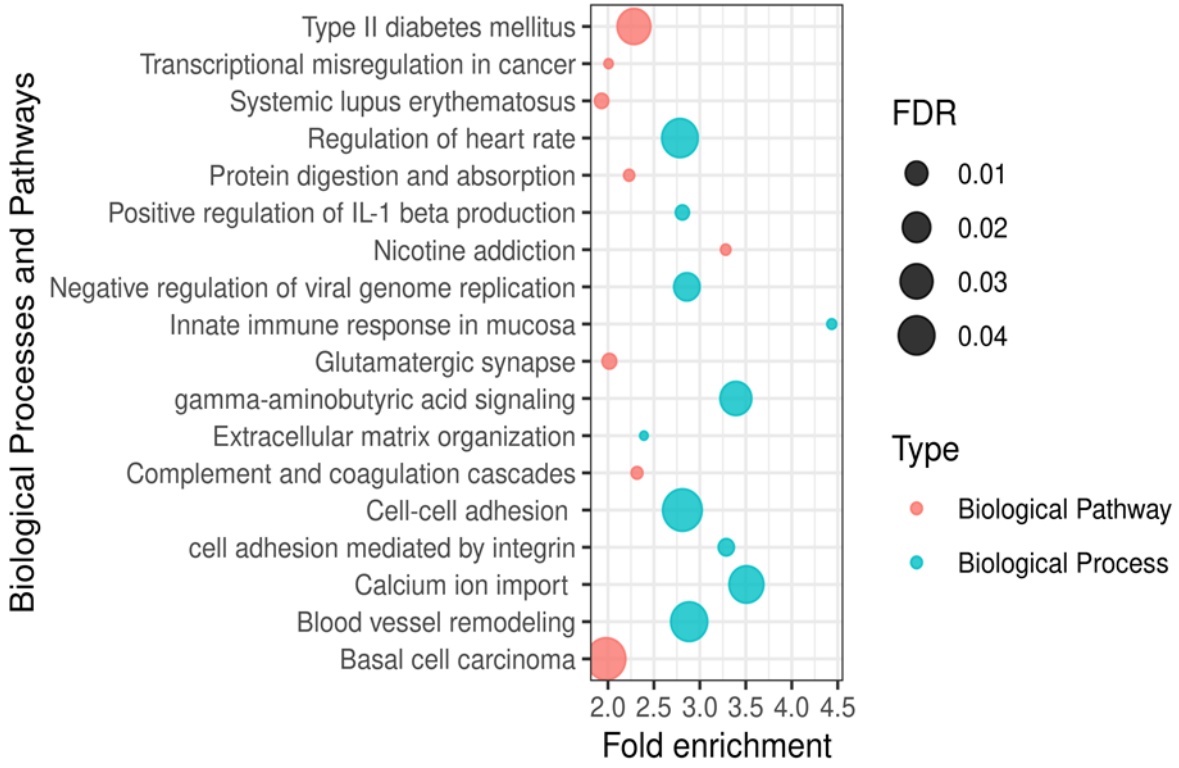


Figure 3[a]: Enrichment analysis of the significant genes of Covid-19.

3.3. Healthy vs Moderate

As observed in the previous stage, the overexpressed genes were most significantly involved in the degranulation of leukocytes. Along with it, the biological functions such as the generation of superoxide and p38 MAPK signaling were also enriched (Figure 5)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 December 2022 doi:10.20944/preprints202212.0431.v1

(Supplementary 2). Entry of virus particles through its receptor causes the accumulation of O_2^- in cells by bringing about the dysfunction of mitochondria [39]. MAPK signaling is responsible for the production of pro-inflammatory cytokines and is activated during the infection [40]. Similarly, these genes were enriched in biological pathways such as VEGF-signaling and renal cell carcinoma. VEGF (vascular endothelial growth factor) causes leakiness of the vascular tissues at the time of the infection and induces hypoxia. Additionally, VEGF causes inflammation of the lung cells. Inhibition of VEGF signaling has been found to drastically increase the oxygen support of the patients [41]. In the moderate stage, the genes involved in the targeting of the proteins to the ER membrane and the assembly of the ribosomal subunits were under-expressed. This shows that the formation and targeting of proteins are significantly disturbed at this stage. Non-structural proteins have been identified to interact with ribosomes at different states and interfere in the translation of the mRNAs [42]. Network analysis revealed that the hub genes in this stage i.e., RPS3A, RPL13, RPL3, RPS6, RPLP0, RPL35, RPL4, RPS8, RPL23A, and RPL19 were involved in the assembly of ribosomes and their proteins. All the hub genes in this stage were under-expressed.

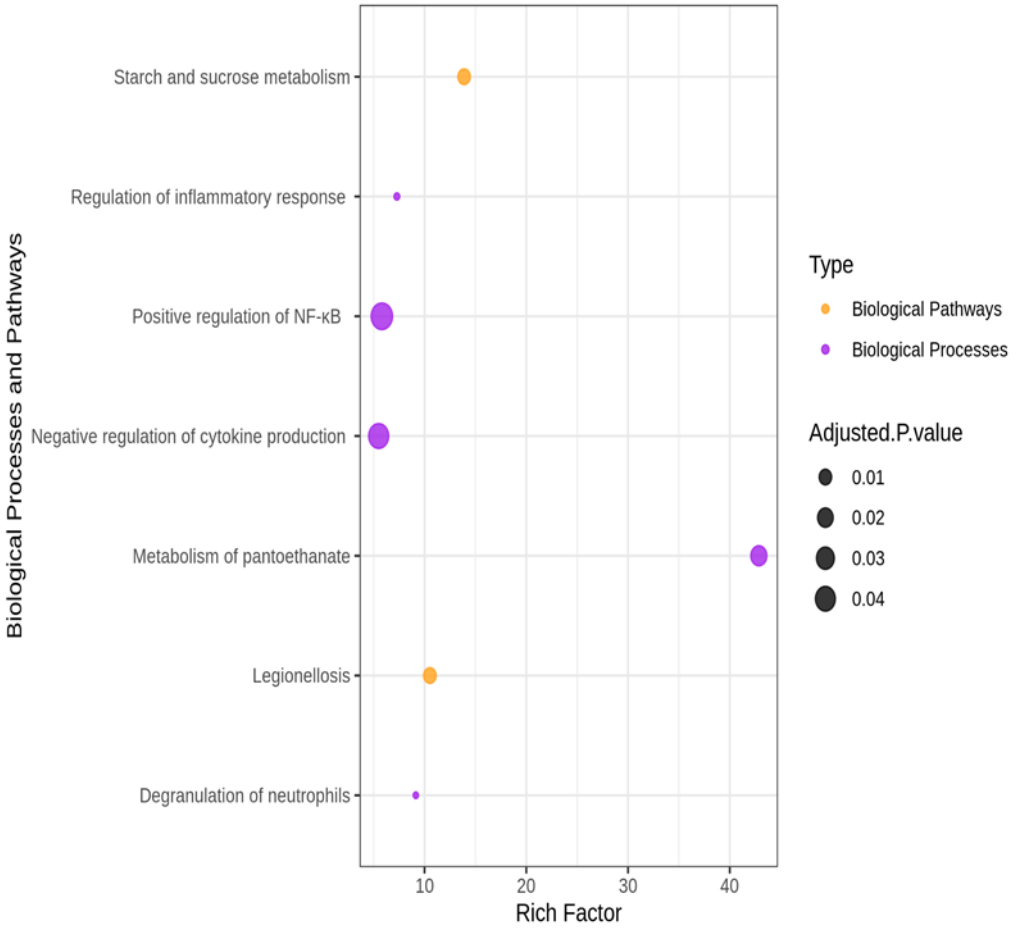


Figure 3[b]: Scatterplot showing the enrichment of overexpressed genes (mild stage) in various biological processes and pathways.

 doi:10.20944/preprints202212.0431.v1

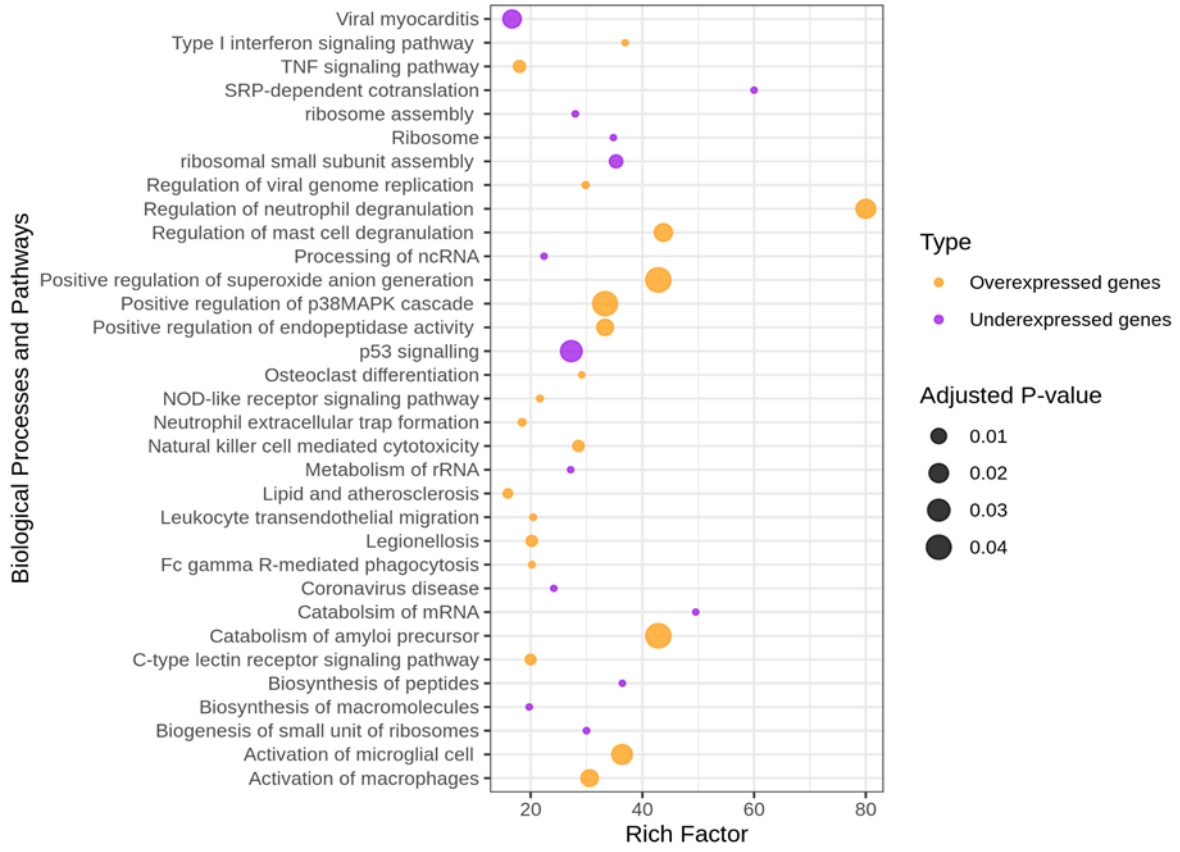


Figure 4: Scatterplot showing the enrichment of DEGs (moderate stage) in significant biological processes and pathways.

3.4. Healthy vs Severe

In the final stage of severity, the most significant biological processes were the degranulation of neutrophils, degranulation of platelets, and IL-1 β-production (Figure 6) (Supplementary 3). Blocking of IL-1 can significantly increase respiratory conditions in severe patients [43]. In terms of biological pathways, the most significant was PPAR signaling. Upon binding of the ligands to PPAR (Peroxisome proliferator-activated response), the ACE2 gets upregulated in the lung tissues of the infected patients [44]. The immunity in this stage is markedly compromised. Activation and proliferation of T and B cells are significantly under-expressed. The genes which are involved in the occurrence of type 1 diabetes mellitus (T1DM) have also been significantly under-expressed. T1DM is a common co-occurrence in Covid-19. The hub genes controlling the transcription machinery in this stage were RPS27A, CCNA2, RPL13A, RPL19, CCNB1, RPL3, MAPK3, RPS6, JUN, and RPL37A. Other than a synthesis of proteins, the hub genes are involved in the cell cycle, MAPK-signaling, and regulation of gene expression.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 December 2022 doi:10.20944/preprints202212.0431.v1

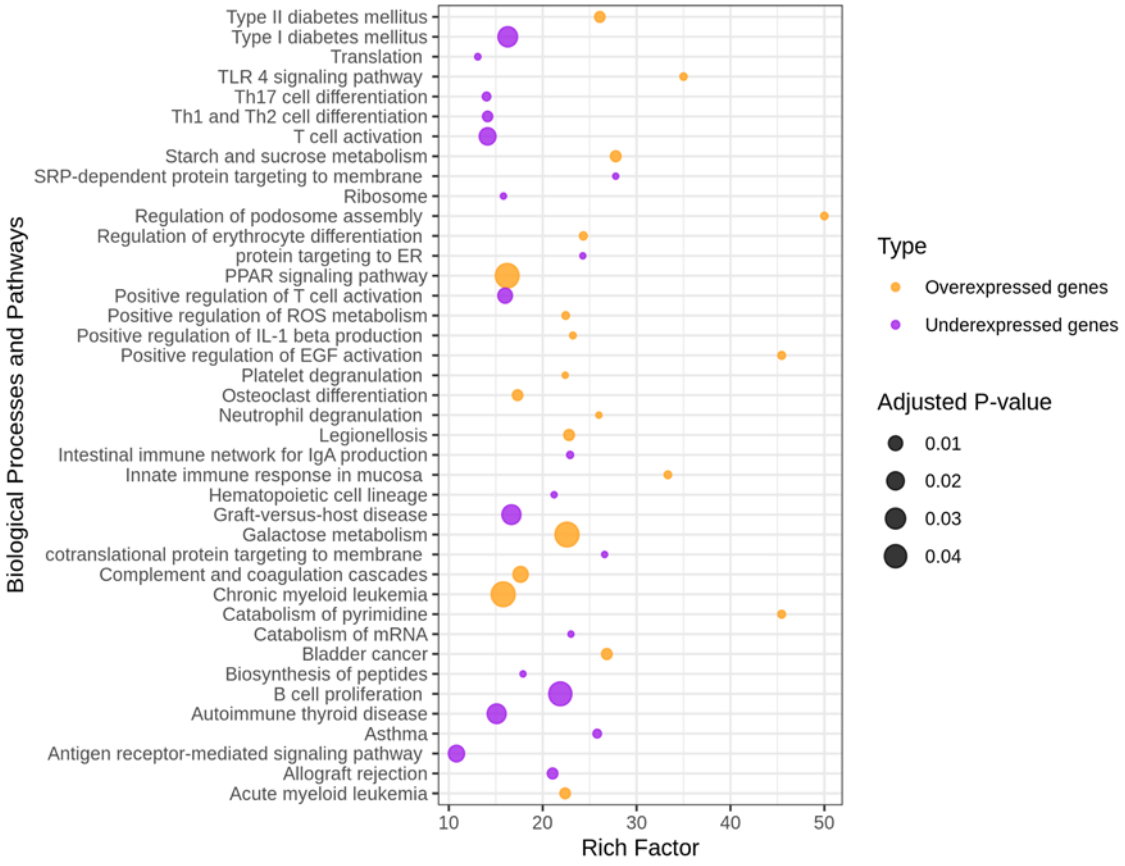
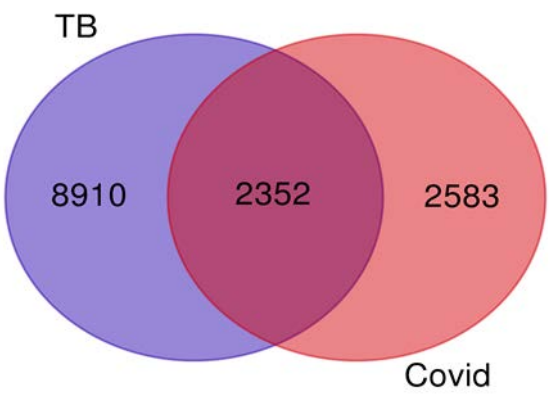


Figure 5: Scatterplot showing the enrichment of DEGs (severe stage) in significant biological processes and pathways.

3.5. Comparison with comorbidities such as TB

Differential gene expression analysis has shown that a large number of genes were shared between TB and Covid-19 (Figure 6[a]). Hence, the WGCNA (Weighted Gene Co-expression Network Analysis) was performed by considering the coexpression network of Covid-19 as a test and TB as the control network. For the construction of a scale-free network, the β was chosen to be 10. (Supplementary 4) Module preservation analysis showed that some of the modules of the Covid-19 network were conserved in TB (Figure 6[b]). The magenta, purple and brown modules were found to be preserved in the co-expression network of TB (Supplementary 5). Members of the magenta module are involved in the translation and regulation of protein targeting whereas the members of the purple module were found to be significantly enriched in T-cell proliferation and differentiation. Similarly, the process was repeated by considering TB as a test and Covid-19 as a control (Supplementary 6).



 doi:10.20944/preprints202212.0431.v1

Figure 6[a]: Venn diagram showing the common differentially expressed genes between Covid-19 and TB.

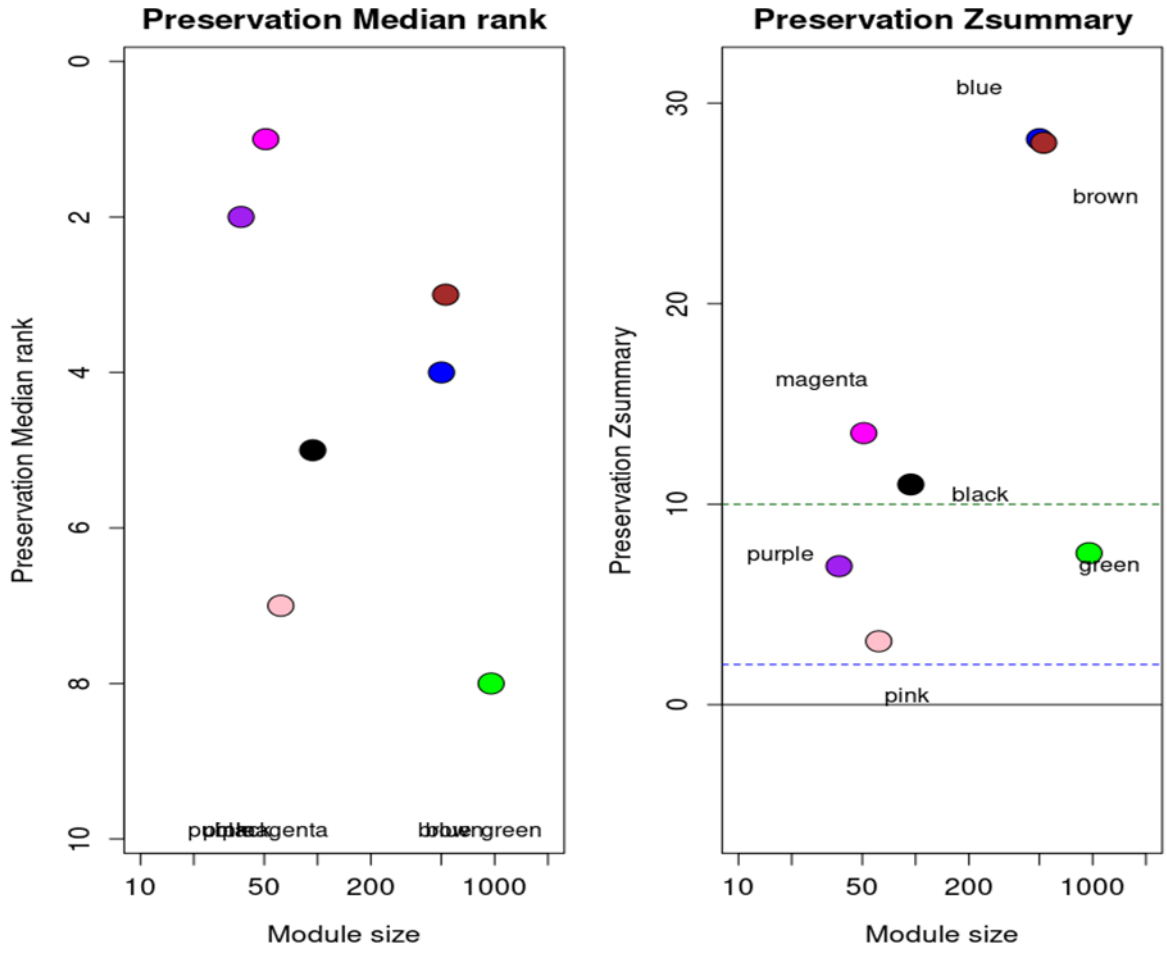


Figure 6[b]: Module preservation of Covid-19 modules in TB. Preservation analysis of the module is studied using the value *Median Rank* and *Zsummary*. A module of Zsummary value (z) with 2<z<10 is considered to be preserved in the test (TB) and the control (Covid-19) network. *Zsummary* value depends on the sample size and hence, median rank is a better metric to study the preservation of modules. Modules with medianRank closer to 0 are considered preserved. We identified three modules of Covid-19 which are preserved in TB i.e., magenta, purple, and brown.

4. Discussion

COVID-19 has wreaked havoc throughout the world and brought a toll on human lives. With wide-ranging symptoms, the disease condition worsens as the disease progresses. The stages of Covid-19 are classified based on the saturation level of oxygen in the blood. Next-generation sequencing has opened new avenues in the research of COVID-19 infection and can be used to study the progression of the disease with time. However, such a study has not been performed on an Indian population. Hence, the main objective of this study was to identify important genes and pathways unique to different stages of COVID-19 infection.

In this study, we have used an integrated system biology approach to identify the hub genes regulating the progression of this disease. We implemented the workflow on a publicly available transcriptomic dataset. It consists of 10 healthy samples, 9 mild samples, 10 moderate samples, and 9 severe samples. The significant genes were selected to study the biological pathways and processes that were enriched in each stage. Hub genes regulating the progression in each stage were identified and their involvement in COVID-19 infection was studied.

 doi:10.20944/preprints202212.0431.v1

Inflammation is a mechanism of the body to combat intruding pathogens. Unfortunately, uncontrolled inflammation can lead to multiple organ failures and eventually death [45]. Degranulation of mast cells and neutrophils is a common observation during COVID-19 infection. These cells release cytokines from their granules and in extreme condition causes the cytokine storm [46]. Other than inflammatory processes, the genes are significantly involved in the metabolism of starch and sucrose in all the stages. The level of sucrose in the body is strongly correlated with the expression of ACE2 in the lungs and increases the vulnerability to the virus [47]. During the infection, the DEGs have been found to be significantly enriched in IF-1, TNF, and NF-κβ signaling. With the progression, the number of significant genes involved in IF-1 signaling increases and in turn stimulates hyperinflammation [48]. A similar trend was observed in the case of TNF signaling. The hub genes responsible for the progression of each stage were involved in different biological processes. PYGL, HSPA1A and HK3. In the mild stage, the hub genes were involved in inflammation, capturing pathogens, removal of antigen-antibody complex, and metabolism of glycogen. As the disease progresses to moderate, the hub genes are involved in the assembly of ribosomes and protein synthesis.

5. Conclusion

In summary, we performed transcriptomic analysis for different stages of COVID-19 infection and studied the enrichment of the significant genes. In the mild stage, the significant DEGs were enriched in the Hedgehog signaling and negatively regulated the production of interleukins such as IL-6 and IL-12. Subsequently, the significant DEGs are enriched in the generation of superoxide and production of IL-1 β production during the moderate and severe stages respectively. PPI networks were constructed for each stage and the hub proteins were identified.

Supplementary Materials: The supplementary materials are uploaded as Supp_1, Supp_2, Supp_3, Supp_4, Supp_5, and Supp_6.

Author Contributions: Methodology, M.K.; S.D.S. and P.B.; software, M.K.; S.D.S. and P.B.; validation, M.K.; S.D.S. and P.B.; investigation, M.K.; S.D.S. and P.B.; formal analysis, M.K.; S.D.S. and P.B.; resources, M.K.; S.D.S. and P.B.; writing - original draft preparation, M.K.; S.D.S. and P.B.; writing – review and editing, P.B. and G.N.S.; visualization, M.K.; S.D.S. and P.B.; supervision, P.B. and G.N.S.; project administration, P.B. and G.N.S.; funding acquisition, P.B. and G.N.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by grants [GPP-0393] from the Foundation for Innovative New Diagnostics, Switzerland.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data availability Statement: Not applicable.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References:

1. Ghosh, P., Ghosh, R., & Chakraborty, B. COVID-19 in India: statewide analysis and prediction. *JMIR public health and surveillance*. 2020, *6*, e20341.
2. Cascella, M., Rajnik, M., Aleem, A., Dulebohn, S. C., & Di Napoli, R. Features, evaluation, and treatment of coronavirus (COVID-19). *Statpearls [internet]*.2022.
3. Wrenn, J. O., Pakala, S. B., Vestal, G., Shilts, M. H., Brown, H. M., Bowen, S. M., ... & Das, S. R. COVID-19 severity from Omicron and Delta SARS-CoV-2 variants. *Influenza and Other Respiratory Viruses*.2022, 16, 832-836.
4. Naushin, S., Sardana, V., Ujjainiya, R., Bhatheja, N., Kutum, R., Bhaskar, A. K., ... & Sengupta, S. Insights from a Pan India seroepidemiological survey (phenome-India cohort) for SARS-CoV2. *Elife*.2021, *10*, e66537.

5. Wahengbam, R., Bharali, P., Manna, P., Phukan, T., Singh, M. G., Gogoi, G., ... & Sastry, G. N. Seroepidemiological and genomic investigation of SARS-CoV-2 spread in North East region of India. *Indian Journal of Medical Microbiology*.2022.
6. Singh, P., Ujjainiya, R., Prakash, S., Naushin, S., Sardana, V., Bhatheja, N., ... & Dash, D. A machine learning-based approach to determine infection status in recipients of BBV152 (Covaxin) whole-virion inactivated SARS-CoV-2 vaccine for serological surveys. *Computers in biology and medicine*. 2022, *146*, 105419.
7. Butt, A. A., Dargham, S. R., Loka, S., Shaik, R. M., Chemaitelly, H., Tang, P., ... & Abu-Raddad, L. J. (2022). Coronavirus disease 2019 disease severity in children infected with the omicron variant. *Clinical Infectious Diseases*.2022, *75*, 361-367.
8. Maucourant, C., Filipovic, I., Ponzetta, A., Aleman, S., Cornillet, M., Hertwig, L., ... & Karolinska COVID-19 Study Group. Natural killer cell immunotypes related to COVID-19 disease severity. *Science immunology*.2020, 50, eabd6832.
9. Liu, T., Jia, P., Fang, B., & Zhao, Z. Differential expression of viral transcripts from single-cell RNA sequencing of moderate and severe COVID-19 patients and its implications for case severity. *Frontiers in microbiology*.2020, *11*, 603509.
10. Britto, C. J., Niu, N., Khanal, S., Huleihel, L., Herazo-Maya, J. D., Thompson, A., ... & Cohn, L. E. BPIFA1 regulates lung neutrophil recruitment and interferon signaling during acute inflammation. *American Journal of Physiology-Lung Cellular and Molecular Physiology*.2019, *316*, L321-L333.
11. Zeberg, H. The major genetic risk factor for severe COVID-19 is associated with protection against HIV. *Proceedings of the National Academy of Sciences*.2022, *119*, e2116435119.
12. Aschenbrenner, A. C., Mouktaroudi, M., Krämer, B., Oestreich, M., Antonakos, N., Nuesch-Germano, M., ... & Ulas, T. Disease severity-specific neutrophil signatures in blood transcriptomes stratify COVID-19 patients. *Genome medicine*.2021, *13*, 1-25.
13. Shaath, H., Vishnubalaji, R., Elkord, E., & Alajez, N. M. (2020). Single-cell transcriptome analysis highlights a role for neutrophils and inflammatory macrophages in the pathogenesis of severe COVID-19. *Cells*.2020, *9*, 2374.
14. Jain, R., Ramaswamy, S., Harilal, D., Uddin, M., Loney, T., Nowotny, N., ... & Abou Tayoun, A. Host transcriptomic profiling of COVID-19 patients with mild, moderate, and severe clinical outcomes. *Computational and structural biotechnology journal*.2021, *19*, 153-160.
15. Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*.2018, *34*, i884-i890.
16. Kim, D., Langmead, B., & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nature methods*.2015, *12*, 357-360.
17. Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., ... & Li, H. Twelve years of SAMtools and BCFtools. *Gigascience*.2021, *10*, giab008.
18. Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*.2014, *15*, 1-21.
19. Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., ... & Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC bioinformatics*.2013, *14*, 1-14.
20. Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., ... & Ma'ayan, A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic acids research*.2016, *44*, W90-W97.
21. Xie, Z., Bailey, A., Kuleshov, M. V., Clarke, D. J., Evangelista, J. E., Jenkins, S. L., ... & Ma'ayan, A. Gene set knowledge discovery with enrichr. *Current protocols*.2021, *1*, e90.
22. Luck, K., Kim, D. K., Lambourne, L., Spirohn, K., Begg, B. E., Bian, W., ... & Calderwood, M. A.A reference map of the human binary protein interactome. *Nature*.2020, *580*, 402-408.
23. Huttlin, E. L., Ting, L., Bruckner, R. J., Gebreab, F., Gygi, M. P., Szpyt, J., ... & Gygi, S. P. The BioPlex network: a systematic exploration of the human interactome. *Cell*.2015, *162*, 425-440.
24. Huttlin, E. L., Bruckner, R. J., Paulo, J. A., Cannon, J. R., Ting, L., Baltier, K., ... & Harper, J. W. Architecture of the human interactome defines protein communities and disease networks. *Nature*.2017, *545*, 505-509.
25. Huttlin, E. L., Bruckner, R. J., Navarrete-Perea, J., Cannon, J. R., Baltier, K., Gebreab, F., ... & Gygi, S. P. Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. *Cell*.2021, *184*, 3022-3040.
26. Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, complex systems*.2006, *1695*, 1-9.
27. Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics*.2008, *9*, 1-13.
28. Langfelder, P., Zhang, B., & Horvath, S. (2008). Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics*.2008, *24*, 719-720.
29. Bakhtiarizadeh, M. R., Hosseinpour, B., Shahhoseini, M., Korte, A., & Gifani, P. Weighted gene co-expression network analysis of endometriosis and identification of functional modules associated with its main hallmarks. *Frontiers in genetics*.2018, *9*, 453.
30. Jin, G., Zhu, M., Yin, R., Shen, W., Xu, L., Hu, Z., ... & Shen, H. Low-frequency coding variants in BAT2, HIST1H1E, and BPIFB1 are associated with lung cancer risk in Chinese populations. *Cancer Research*.2015, *75*, 4598-4598.
31. ShigunovShigunov, P., Sotelo-Silveira, J., Stimamiglio, M. A., Kuligovski, C., Irigoín, F., Badano, J. L., ... & Dallagiovanna, B. (2014). Ribonomic analysis of human DZIP1 reveals its involvement in ribonucleoprotein complexes and stress granules. *BMC molecular biology*.2014, *15*, 1-14.

32. Atal, S., & Fatima, Z. (2020). IL-6 inhibitors in the treatment of serious COVID-19: a promising therapy? *Pharmaceutical medicine*.2020, *34*, 223-231.
33. Mikkelsen, K., & Apostolopoulos, V. Vitamin B1, B2, B3, B5, and B6 and the immune system. In *Nutrition and immunity*.2019, 115-125.
34. Zheng, M., Karki, R., Williams, E. P., Yang, D., Fitzpatrick, E., Vogel, P., ... & Kanneganti, T. D. TLR2 senses the SARS-CoV-2 envelope protein to produce inflammatory cytokines. *Nature immunology*.2021, *22*, 829-838.
35. Nassir, N., Tambi, R., Bankapur, A., Al Heialy, S., Karuvantevida, N., Khansaheb, H. H., ... & Uddin, M. Single-cell transcriptome identifies FCGR3B upregulated subtype of alveolar macrophages in patients with critical COVID-19. *Iscience*.2021, *24*, 103030.
36. Krause, M., Gerchman, F., & Friedman, R. (2020). Coronavirus infection (SARS-CoV-2) in obesity and diabetes comorbidities: is heat shock response determinant for the disease complications? *Diabetology & Metabolic Syndrome*.2020, *12*, 1-9.
37. Wang, R. Genetic variation of interleukin-1 receptor type 1 is associated with severity of COVID-19 disease. *Journal of Infection*.2022, *84*, e19-e21.
38. Chua, R. L., Lukassen, S., Trump, S., Hennig, B. P., Wendisch, D., Pott, F., ... & Eils, R. COVID-19 severity correlates with airway epithelium–immune cell interactions identified by single-cell analysis. *Nature biotechnology*.2020, *38*, 970-979.
39. Silvagno, F., Vernone, A., & Pescarmona, G. P. (2020). The role of glutathione in protecting against the severe inflammatory response triggered by COVID-19. *Antioxidants*.2020, *9*, 624.
40. Goel, S., Saheb Sharif-Askari, F., Saheb Sharif Askari, N., Madkhana, B., Alwaa, A. M., Mahboub, B., ... & Halwani, R. SARS-CoV-2 Switches 'on'MAPK and NFκB Signaling via the Reduction of Nuclear DUSP1 and DUSP5 Expression. *Frontiers in Pharmacology*.2021, *12*, 631879.
41. Pang, J., Xu, F., Aondio, G., Li, Y., Fumagalli, A., Lu, M., ... & Cao, Y. (2021). Efficacy and tolerability of bevacizumab in patients with severe Covid-19. *Nature communications*.2021, *12*, 1-10.
42. Wei, J., & Hui, A. (2022). Review of Ribosome Interactions with SARS-CoV-2 and COVID-19 mRNA Vaccine. *Life*.2022, *12*, 57.
43. Van De Veerdonk, F. L., & Netea, M. G. Blocking IL-1 to prevent respiratory failure in COVID-19. *Critical Care*.2020, *24*, 1-6.
44. Fantacuzzi, M., Amoroso, R., & Ammazzalorso, A. (2022). PPAR Ligands Induce Antiviral Effects Targeting Perturbed Lipid Metabolism during SARS-CoV-2, HCV, and HCMV Infection. *Biology*.2022, *11*, 114.
45. Zhang, B., Zhou, X., Qiu, Y., Song, Y., Feng, F., Feng, J., ... & Wang, J. (2020). Clinical characteristics of 82 cases of death from COVID-19. *PloS one*.2020, *15*, e0235458.
46. Hu, B., Huang, S., & Yin, L. The cytokine storm and COVID-19. *Journal of medical virology*.2021, *93*, 250-256.
47. Levitt, D., Delery, E., Simon, L., & Molina, P. High-Fat High-Sucrose Diet Increases ACE2 Receptor Expression in Lung and Pancreatic Islets in SIV-Infected Rhesus Macaques: Implications for Increased Risk for SARS-CoV-2 Infection. *The FASEB Journal*.2021, *35*.
48. Lee, J. S., & Shin, E. C. (2020). The type I interferon response in COVID-19: implications for treatment. *Nature Reviews Immunology*.2020, *20*, 585-586.