



# Analysis of transcriptomics data from COVID-19 patients: a pilot research

Dominik Hadzega<sup>1</sup> · Klaudia Babisova<sup>1</sup> · Michaela Hyblova<sup>1</sup> · Nikola Janostiakova<sup>2</sup> · Peter Sabaka<sup>3</sup> · Pavol Janega<sup>1</sup> · Gabriel Minarik<sup>1</sup>

Received: 30 August 2023 / Accepted: 3 January 2024 / Published online: 19 January 2024  
© The Author(s) 2024

## Abstract

During SARS-CoV-2 infection, the virus transforms the infected host cell into factories that produce new viral particles. As infection progresses, the infected cells undergo numerous changes in various pathways. One of these changes is the occurrence of a cytokine storm, which leads to severe symptoms. In this study, we examined the transcriptomic changes caused by COVID-19 by analyzing RNA-seq data obtained from COVID-19-positive patients as well as COVID-19-negative donors. RNA-seq data were collected for the purpose of identification of potential biomarkers associated with a different course of the disease. We analyzed the first datasets, consisting of 96 samples to validate our methods. The objective of this publication is to report the pilot results. To explore potential biomarkers related to disease severity, we conducted a differential expression analysis of human transcriptome, focusing on COVID-19 positivity and symptom severity. Given the large number of potential biomarkers we identified, we further performed pathway enrichment analysis with terms from Kyoto Encyclopedia of Genes and Genomics (KEGG) to obtain a more profound understanding of altered pathways. Our results indicate that pathways related to immune processes, response to infection, and multiple signaling pathways were affected. These findings align with several previous studies that also reported the influence of SARS-CoV-2 infection on these pathways.

**Keywords** RNA-seq · COVID-19 · SARS-CoV-2 · Gene enrichment analysis · Enriched pathways · Differentially expressed genes · Transcriptomics

## Introduction

SARS-CoV-2 virus first emerged in December 2019 when it infected a patient hospitalized with the disease now known as COVID-19 (Wu et al. 2020). It subsequently spread worldwide causing a global pandemic (Cucinotta and Vanelli 2020). For this pandemic event, much more powerful tools were available to study and manage the outbreak than ever before. Next-generation sequencing capabilities helped

scientists study the genetic code of the virus and its evolution over time (Oude Munnink et al. 2021).

One of these capabilities certainly represents the possibility to study how the human transcriptome in the tissue of an infected patient changes or how it differs in case of more severe symptoms. After infecting human tissue, SARS-CoV-2 alters normal host cell metabolism and signaling to adjust the intracellular environment for itself and its own replication. In general, this involves interfering with signaling pathways that regulate processes of DNA repair and replication, immune response, transcription, metabolism, cell cycle, and apoptosis. In particular, the pathways that are known to be altered include phosphoinositide 3-kinase (PI3K)/protein kinase B (AKT), Type I and III interferons, transforming growth factor- $\beta$  (TGF- $\beta$ ), Toll-like receptors (TLR), and nuclear factor kappa light chain enhancer (NF- $\kappa$ B) pathways. Also, the altering of the Ca<sup>2+</sup> signaling is believed to be part of the infection (Jamison et al. 2022). More severe phenotypes have been linked to excessive and uncontrolled immune response known as cytokine storms. It leads to the release of a large amount of pro-inflammatory

✉ Gabriel Minarik  
gabriel.minarik@medirex.sk

Dominik Hadzega  
Dominik.Hadzega@medirexgroupacademy.sk

<sup>1</sup> Medirex Group Academy, Nitra, Slovakia

<sup>2</sup> Comenius University in Bratislava, Medical Faculty,  
Institute of Medical Biology, Genetics and Clinical Genetics,  
Špitálska 24, Bratislava, Slovakia

<sup>3</sup> Department of Infectology and Geographical Medicine,  
Faculty of Medicine, Comenius University in Bratislava,  
Bratislava, Slovakia

cytokines into the blood-stream, followed by widespread inflammation and damage to tissues and organs (Henderson et al. 2020; Mehta et al. 2020; Wang et al. 2020).

This study is a pilot to our ongoing research on the impact of SARS-CoV-2 infection on the human transcriptome and bacterial transcript abundance (yet to be published) analyzed from RNA-sequencing of nasopharyngeal swabs. This study investigates the transcriptome profiles in COVID-19-positive and COVID-19-negative samples. We focused on differential gene expression and identification of pathways affected by the infection. Moreover, we compared patients infected with different SARS-CoV-2 variants.

## Materials and methods

### Study approval

Sample collection was performed as part of the clinical study approved by the Ethical Committee of Bratislava Self-Governing Region under the identifier 03228/2021/HF is-sued January 12, 2021. All patients completed questionnaires with relevant information regarding their health condition in relation to COVID-19 and signed informed consent.

### Samples

Nasopharyngeal swabs from patients suspected of having COVID-19 were obtained in two primary regimens. Firstly, patients with severe disease symptoms hospitalized at the partnering hospitals were enrolled in the study. Secondly, patients with mild or no disease symptoms were recruited in mobile testing facilities for SARS-CoV-2 by a company providing routine laboratory diagnostic services for the general population during the COVID-19 pandemic. Subgroups 1 to 4 were formed based on positivity and negativity of SARS-CoV-2 testing and the severity of COVID-19 symptoms according to the following scheme: group 1, positive with severe symptoms; group 2, positive with mild symptoms; group 3, no symptoms; group 4, negative. To classify cases, the WHO definition of symptoms was used (living guidance for clinical management of COVID-19, 2021 by World Health Organization). In total, 96 samples were included in the study (72 COVID-19-positive patients and 24 healthy sample donors). Nasopharyngeal swabs were collected from March 2021 to October 2022.

### Nucleic acid extraction

Nasopharyngeal swab specimens were collected from COVID-19 patients and individuals from control group and stored in the viRNAtrap collection medium (GeneSpector,

Czech Republic) at 4 °C. Total RNA was extracted using Sera-Xtracta Virus/Pathogen Kit (Cytiva, UK) in line with the manufacturer instructions. Four hundred microliters of the nasopharyngeal swab medium was used for the extraction with a final elution volume of 50 µl. RNA was quantified with the Qubit™ RNA High Sensitivity Assay Kit (Invitrogen). RNA isolates were stored at –80 °C.

### RT-qPCR

The presence of SARS-CoV-2 was determined by RT-qPCR using the COVID-19 Real-Time Multiplex RT-PCR Kit (Labsystems Diagnostics, Finland) and the ABI QuantStudio 6 RT-qPCR platform and Real-Time PCR System (Thermo Fisher, USA) utilizing the original manufacturers' protocols. Amplification cycles threshold of Ct value < 40 was needed to evaluate the sample as positive.

### RNA library preparation and sequencing

The metatranscriptomic libraries were prepared using KAPA RNA HyperPrep Kit with RiboErase (HMR) (Kapa Biosystems, South Africa) in line with the original protocol of the manufacturer. For quantity and quality control of prepared libraries, we used a Qubit 1X dsDNA High Sensitivity Assay Kit on Qubit 3.0 (Invitrogen) and Agilent High Sensitivity DNA Kit on Agilent 2100 Bioanalyzer (Agilent) instruments. Sequencing of pooled libraries was performed on NextSeq 500 and NextSeq 2000 (Illumina) platforms using 2 × 75 or 2 × 100 paired-end sequencing setup, respectively.

### Quality control and data preparation for analysis

First step of any analysis of RNA-seq data is quality control and this step was done by FastQC v0.11.9 (Andrews 2010). Reads were processed by Trimmomatic v0.39 (CROP:96 HEAD-CROP:10 LEADING:22 TRAILING:22 SLIDING-WINDOW:4:22 MINLEN:25 and our own set of adapter sequences were used in ILLUMINACLIP step) (Bolger et al. 2014). Parameters were chosen according to FastQC results.

### Reads mapping

After final affirmation of sufficient quality of reads by FastQC, reads were mapped to the human genome hg38 using the BWA-MEM algorithm v0.7.17 (Li 2013). Reads were mapped as paired set; otherwise, mapping parameters were set to default. The same procedure was applied on the SARS-CoV-2 genome. Mapping statistics were produced from a “.bam” outputs by Samtools flagstat (<https://github.com/bahlolab/bioinfotools/blob/master/SAMtools/flagstat.md>). (Accessed 23.05. 2023) n.d.).

## SARS-CoV-2 variant identification

For SARS-CoV-2 variant identification, we used the Galaxy pipeline, specifically “Mutation calling, viral genome reconstruction and lineage/clade assignment from SARS-CoV-2 sequencing data” (Maier and Batut 2023). For input, we used reads mapping on SARS-CoV-2 genome.

## Differentially expressed genes analysis

In search of human genes affected by the SARS-CoV-2 infection or the particular category of infection, we further analyzed the bam file with human-mapped reads. Gene expressions were quantified by FeatureCounts v.2.0.1 (Liao et al. 2014). Statistical comparison was performed by the R instance of DESeq2 v.1.38.3. (Love et al. 2014). Genes in the condition of adjusted  $p$ -value  $< 0.1$  were considered as significant hits. This is a default value according to DESeq2 manual (DESeq2 Manual (Accessed 23.05. 2023) n.d.).

## Identification of altered pathways

To find out which pathways were altered from the set of differentially expressed genes, we used the R instance of gProfiler2 v.0.2.1 (Kolberg et al. 2020). Genes found under  $p$ -value threshold 0.1 were used as a query for the analysis, ranked according to  $p$ -values and separated to down-regulated and up-regulated set. Databases of KEGG pathways, Wiki pathways, and Reactome were used in order to improve visibility; only the KEGG terms were selected for visualization. Set of genes which were part of differentially expressed gene analysis were set as custom background. Domain scope was set to “known.” For the visualization, barplot function in R was used.

## Results

In this study, we analyzed RNA-seq data from nasopharyngeal swabs. 96 samples were included in the study (72 COVID-19-positive patients and 24 negative donors). More detailed patient data are shown in Table 1.

**Table 1** Patient details. Here, we show specific numbers of patients in the respective subgroups as well as their sex and age statistics

Samples	Phenotype	Number of samples	Male/female	Age median	
<b>Negative</b>	-	24	12/12	33.5	
<b>Positive</b>	Asymptomatic	24	72	14/10	42
	Mild	25		14/11	46
	Severe	23		12/11	60
<b>Positive with variants</b>	Alpha	24	36	14/10	46
	Delta	8		4/4	48.5
	Omicron	4		0/4	43

## SARS-CoV-2 sequences in our samples

We assigned SARS-CoV-2 in our samples with specific clade and the WHO variant name. Twenty-four samples were assigned as Alpha variant (clade 20I), 8 samples as Delta (clade 21 J), 4 samples as Omicron (clade 21L), one sample was assigned to the 20C clade, and one was reported as recombinant.

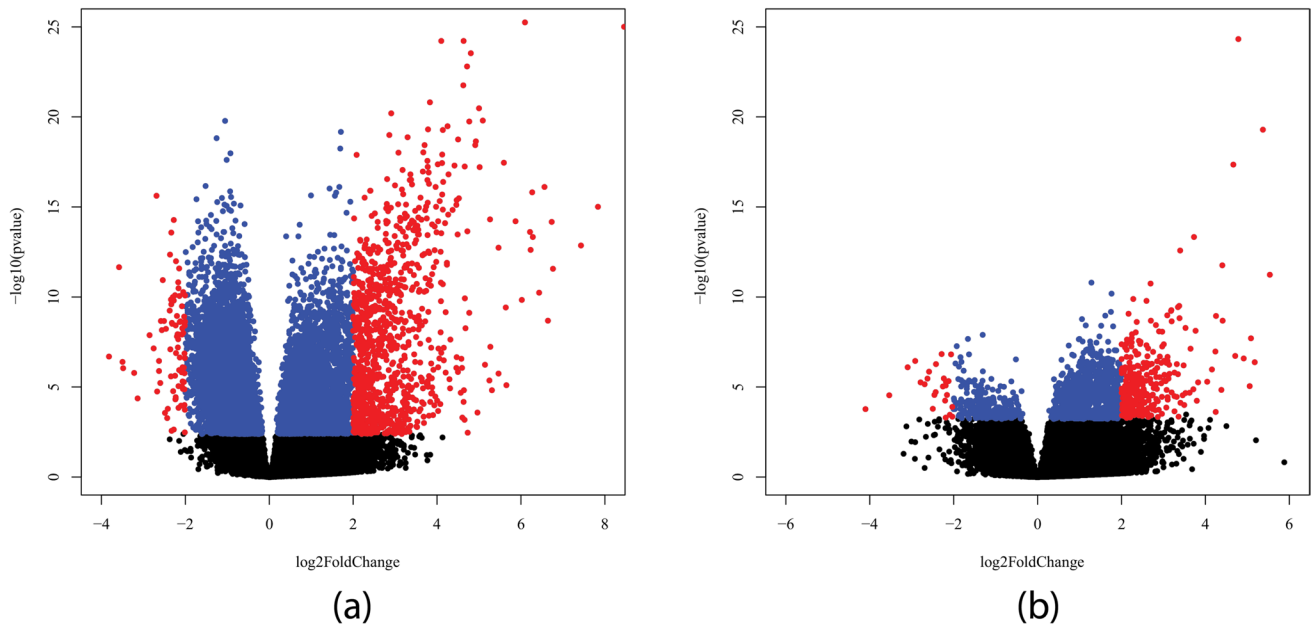
## Human transcriptome changes in COVID-19

According to our results, there were significant changes in the human transcriptome in COVID-19 patients. DESeq2 statistical test reported 16,365 genes with an adjusted  $p$ -value  $< 0.1$  compared to COVID-19 negative controls (Supplementary Table 1). Four thousand five hundred thirty-nine genes were reported with a  $p$ -value  $< 0.1$  when comparing mild and severe disease (Supplementary Table 2). Expression change profiles are shown as volcano plots in Fig. 1.

Our data also show variation by a specific SARS-CoV-2 variant (based on the WHO variant names). The distance in the transcriptional profile is visualized in Fig. 2a. A large difference was shown when comparing the Alpha (all 20I) and Delta (all 21 J) variants. There were 14,109 genes with  $p$ -value  $< 0.1$ , which is visualized in Fig. 2b. We speculate further on the reasons for this difference in the “Discussion” section. A comparison of patients with Delta and Omicron (albeit with a limited number of samples) brought 468 genes with a  $p$ -value  $< 0.1$ . However, it needs to be pointed out that all Omicron samples were from female patients.

## Pathways altered by the infection

Since no unequivocal information can be obtained from the large number of genes that meet the criteria for significant effect, we performed an enrichment analysis of the KEGG pathways. Comparison of COVID-19-positive and COVID-19-negative patients (Fig. 3) showed that most of the pathways comprising enriched genes with up-regulation are related to immune response to various diseases or to other immune-related pathways

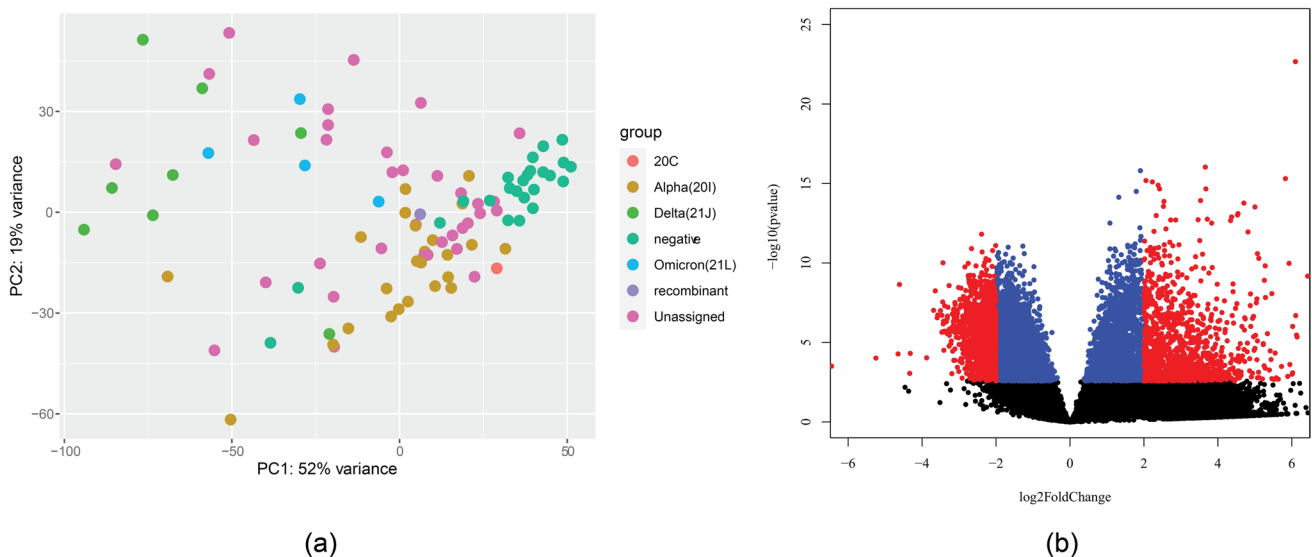


**Fig. 1** Volcano plots showing differential expression by plotting every gene as a dot positioned by  $\log_2$  of its fold change (axis  $x$ ) and  $-\log_{10}$  of its adjusted  $p$ -value (axis  $y$ ). The colors show whether or not the gene passes through threshold requirements—blue (adj.  $p$ -value  $< 0.01$ ) and red (adj.  $p$ -value  $< 0.01$  and  $\log_2$  fold change  $> 2$ ).

**a** Volcano plot visualizing differential expression—down- and up-regulated genes in the samples from COVID-19 patients. **b** Volcano plot visualizing differential expression between the samples from patients with severe symptoms (up-regulated right) compared to those with mild symptoms

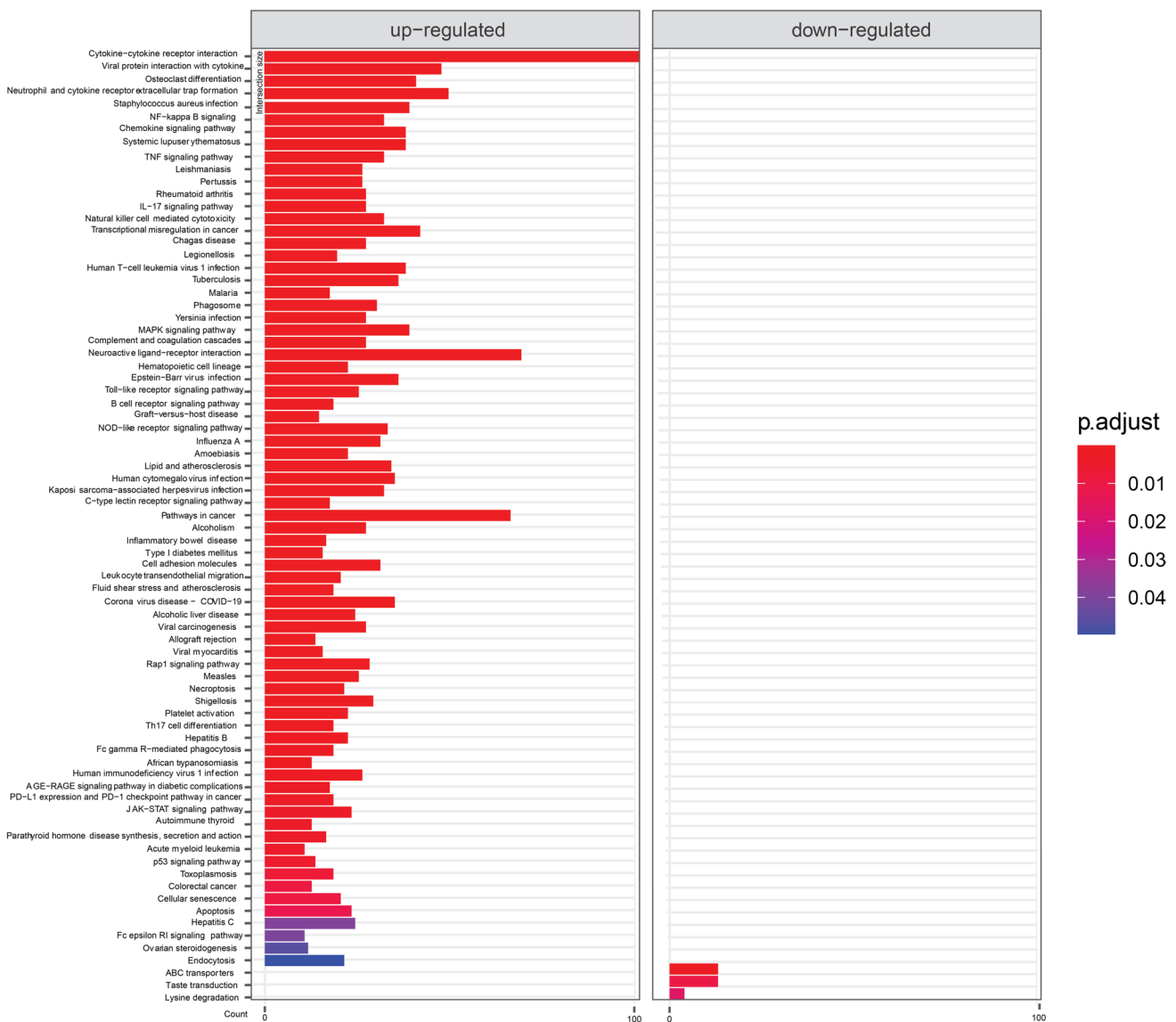
(cytokine-cytokine receptor interaction, neutrophil extracellular trap formation, viral protein interaction with cytokine, natural killer cell, B cell receptor, NOD-like receptor, Fc gamma R – mediated phagocytosis, Th17 cell differentiation). Other pathways with enriched genes

comprised the signaling pathways mostly related to immune system (NF-kappa B signaling pathway, TNF signaling pathway, chemokine signaling pathway, IL-17 signaling pathway, MAPK signaling pathway, C-type lectin receptor signaling pathway, Rap1 signaling pathway, JAK-STAT



**Fig. 2 a** Principal component analysis (PCA) analysis plot of gene expression in different SARS-CoV-2 strains. **b** Volcano plot visualizing differential expression between samples from patients with the

Alpha variant (20I) and those with the Delta variant (21 J). Colors show if gene passes through threshold requirements—blue (adj.  $p$ -value  $< 0.01$ ) and red (adj.  $p$ -value  $< 0.01$  and  $\log_2$  fold change  $> 2$ )



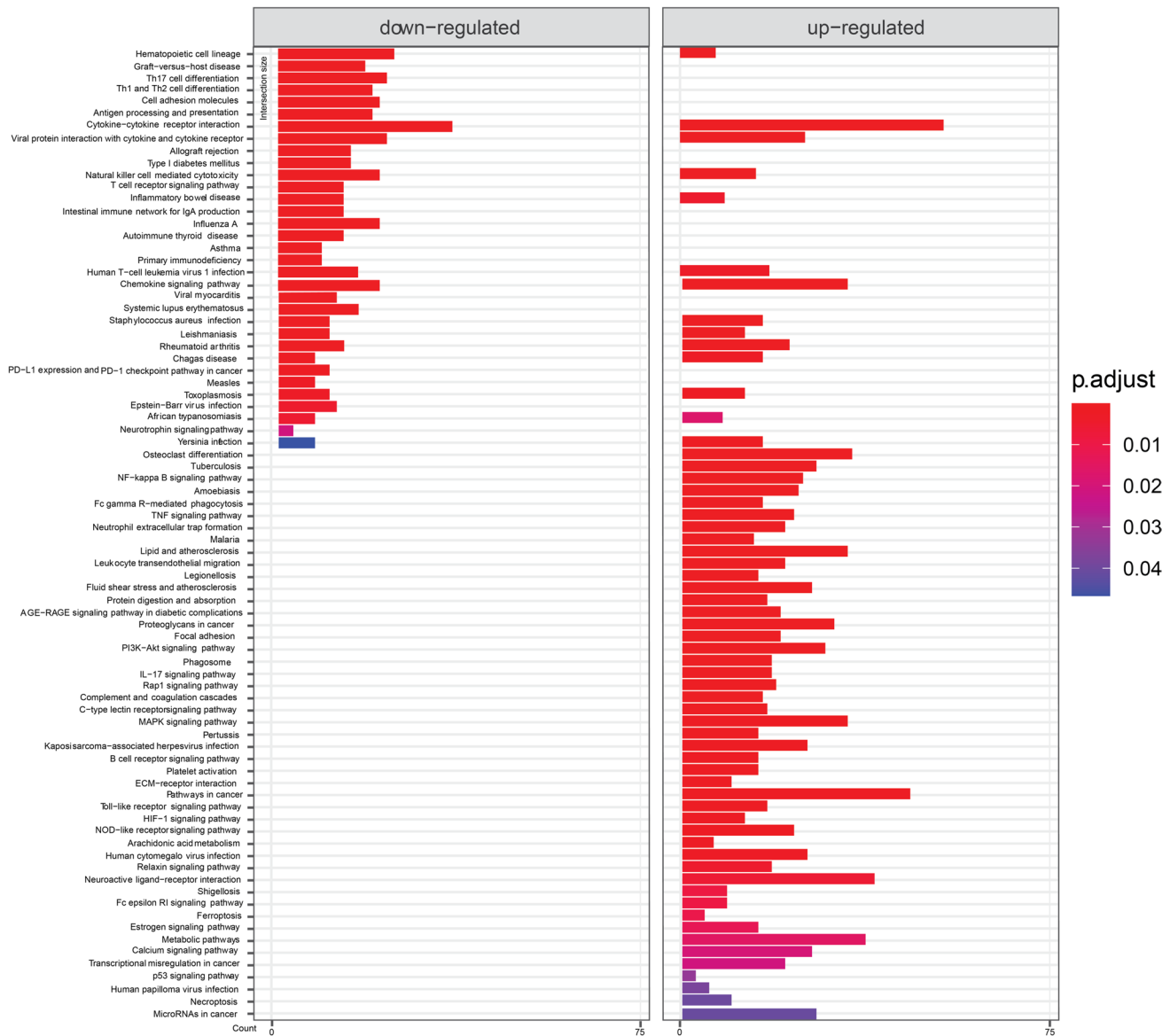
**Fig. 3** KEGG terms mapping on differentially expressed genes for COVID-19 positive samples. Only the results that meet requirements of the default threshold of adjusted  $p$ -value  $< 0.05$  are visualized. Length of the bars is showing gene counts for the term it belongs to

signaling pathway, p53 signaling pathway). Other KEGG pathways included osteoclast differentiation, cell adhesion molecules, alcoholism, and alcoholic liver disease. For the down-regulated genes, 3 pathways were enriched: ABC transporters, taste transfer, and lysine degradation.

KEGG pathway analysis of mild symptoms compared to severe symptoms is shown in Fig. 4. Again, infectious disease- and immunity-related concepts dominated. Cytokine interactions-related terms were significant for both down- and up-regulated genes. From signaling pathways, there are NF-Kappa B, neutrophin, TNF, PI3K-Akt, IL-17, Rap 1, MAPK, Toll-like receptor, C-type lectin, B cell receptor, HIF-1, NOD-like receptor, calcium, and estrogen-signaling pathways.

## Discussion

This article presents a pilot study of the human nasopharyngeal transcriptome and metatranscriptome of patients with COVID-19. By sequencing samples and subsequent data analysis, we studied differences in gene expression between 72 COVID-19-positive and 24 COVID-19-negative patients, between patients with severe and mild symptoms, and between those who were infected with the Alpha and Delta SARS-CoV-2 variants. We observed substantial changes on transcriptomic levels between the above groups. The most suitable way to present the results appears to be pathway analysis visualization, which demonstrates the involvement of numerous mostly immunity-related pathways, signaling



**Fig. 4** KEGG terms mapping on genes differentially expressed in samples from patients with severe symptoms compared to those with mild symptoms. Only the results that meet requirements of the default

threshold of adjusted  $p$ -value  $< 0.05$  are visualized. Length of the bars is showing gene counts for the term it belongs to

pathways directly or indirectly connected to immunity, and specific disease terms (Figs. 3 and 4). Results point to cytokine, chemokine activity and inflammation. Observation of KEGG terms referring to various diseases (not only COVID-19) might be caused by similarities in the effect on tissue between infections or other disease events.

For COVID-19-linked genes, the most of these terms were pathways with many of the concerned genes up-regulated in samples from COVID-19-positive patients. It seems signaling pathways of the innate immunity were most affected. Genes of three important modulators of innate immunity—Nod-like (NLR), Toll-like receptor (TLR), and C-type lectin receptor (CLR) signaling pathways—were

up-regulated. Multiple of other seemingly up-regulated signaling pathways (judging by up-regulation of their genes) are part of processes connected to NLR and TLR signaling. Those are NF- $\kappa$ B signaling, MAPK signaling, and JAK-STAT signaling (Hata et al. 2002; Kawasaki and Kawai 2014). AGE-RAGE signaling genes might be involved through activating pathways resulting in NF- $\kappa$ B activity (Tobon-Velasco et al. 2014). Counting mentioned pathways, many up-regulated processes point to elevated cytokine or chemokine activity and promotion of inflammatory processes (Kawasaki and Kawai 2014; Rawlings et al. 2004). Apart from mentioned, it is the case for TNF signaling, chemokine signaling, IL-17 of Fc epsilon RI-mediated

signaling (Aggarwal 2003; Hata et al. 2002; Zhu et al. 2004). From remaining significant signaling pathways, p53 signaling is involved in different processes like cell cycle control and apoptosis (Hernández Borrero and El-Deiry 2021). Rap1 signaling have various roles like cell proliferation and adhesion; however, it might contribute through regulating MAPK (Zhang et al. 2017; York et al. 1998). Downregulated pathways identified were only ABC transporters, taste transduction, and lysine degradation.

Severe patients differed from those with mild symptoms with numerous altered pathways as well. In fact, same pathways as those making difference between COVID-19-positive and COVID-19-negative patients seem to be crucial (Fig. 4). However, here some genes of cytokine and chemokine pathways were downregulated in severe patients. Th1, Th2, and Th17 cells differentiation genes were downregulated in severe patients, while Th17 cell differentiation actually seemed supported by up-regulated genes in COVID-19 patients compared to negative donors. PI3K-Akt signaling pathway genes were up-regulated. Other unique results compared to COVID-19 patients in general were estrogen and calcium signaling pathways. It seems severity of COVID-19 is followed by even more inflammatory processes.

There are multiple processes that have been reported in previous studies to be involved in changes caused by infections in general and are regarded as key pathways altered during the SARS-CoV-2 infection. We identified most of those key pathways in pathway enrichment using KEGG, Reactome, and Wiki Pathway databases. KEGG pathway enrichment detected terms related to TLR receptors. In fact, TLR receptors are known to be the key component in reacting to the SARS-CoV-2 infection. For example, TLR3 hyperactivation can lead to a cytokine storm and the subsequent severe COVID-19 (Jamison et al. 2022). Other components, such as interleukins (IL-2, IL-6, IL-7, IL-10) and TNF- $\alpha$  are connected to severe COVID-19 symptoms, where cytokine storm (a sudden increase in cytokines) occurs instead of a healthy immune response (Fricke-Galindo and Falfán-Valencia 2021). TNF signaling pathway was detected as a significant KEGG pathway in both comparisons of disease positivity and symptom severity. Using the REAC pathway analysis of differentially expressed genes in COVID-19-positive cases, we detected interleukin IL-2, IL-7, and IL-10 as well as IL-1, IL-4, IL-13, and IL-36. KEGG pathway enrichment detected IL-17 signaling. Next, NF- $\kappa$ B related terms were enriched for both gene sets (genes affected by COVID-19 severity or positivity) not only using the KEGG database, but also Wiki Pathway and Reactome. It has been reported that severe COVID-19 is characterized by an inflammatory profile dominated by NF- $\kappa$ B activity (Huang et al. 2020). Next, terms related with TGF- $\beta$  and JAK-STAT signaling were enriched. JAK-STAT signaling was detected from differentially expressed

genes in COVID-19-positive patients. TGF-beta terms were recognized by Wiki Pathway enrichment in both positivity and severity-affected gene sets. In fact, the cytokine TGF- $\beta$  functions as an activator or suppressor of the Janus kinases (JAKs), and these processes are believed to be affected by the studied viral infection (Tan et al. 2020). Another pathway reported to be altered by SARS-CoV-2 infection, the PI3K-Akt signaling pathway (Al-Qahtani et al. 2022; Li et al. 2021), was also detected in our study (for both severity and positivity-affected gene sets).

A recent study of a similar transcriptome analysis of COVID-19-positive patients, comparing positive patients and negative controls using gene enrichment analysis (albeit with a different software: edgeR and Metascape), has been published by Rhoades et al. (2021). The authors mapped differentially expressed genes to GO terms (while we worked with the KEGG pathways, shown in Fig. 3). However, they too reported terms associated with innate and adaptive host defense pathways, which correlates with our results and is not surprising. Similarly, they found terms related to hematopoiesis (with the CD53 and IKZF1 genes reported, which were included in our results too), inflammatory response, B-cell activation, and leukocyte chemotaxis/mobility (with the CCL2 and CCR1 genes reported). They reported interferon-stimulated genes FITM3, and ISG15, which were also identified in our set of potentially affected genes. Other reported genes included type I interferon signaling genes IRF7 and STAT1 (our results failed to detect STAT1), UN-ATP-1 signaling (with reported genes BCL3 and MYD88, also featured in our results). We did not observe neuronal death pathways, nor the SNCA gene mentioned in the paper by Rhoades et al. (however, MDK, another listed gene, can be also found in our results). We did not observe many downregulated genes-related pathways (in the infected samples), whereas Rhoades et al. reported GO terms associated with tissue homeostasis and cellular organization. In the study conducted by Bass et al., up-regulation of innate immune response through several inflammatory genes and NF- $\kappa$ B was reported. This corroborates our observations, although samples collected from other organs and not from nasopharyngeal swabs were combined. IL-17 signaling was found to be increased in moderate and severe patients. We found it to be altered in patients with more severe symptoms and COVID-19 in general (Bass et al. 2021). Similar studies were conducted on blood samples (Daamen et al. 2022; Jackson et al. 2022; Thompson et al. 2023). Another study focused also on the comparison between different infections, but the authors also report results on nasopharyngeal swabs transcriptomics in COVID-19. Their pathway analysis showed prominent activation of genes related to interferon (IFN) signaling and IFN-stimulating genes (ISGs), but inhibition of interleukin-6 (IL-6) and IL-8 signaling genes (including IRAK1 and MAP2K7) (Ng et al. 2021).

A comparison of the results between the Alpha, Delta, and Omicron groups allows for certain degree of speculation. We observed significant differences between the Alpha and Delta groups, but we cannot say what exactly caused the difference in gene expression. In terms of patient background, we know that a larger number of Alpha patients experienced mild symptoms, but there were also patients with severe symptoms. Delta patients, on the other hand, experienced severe symptoms or no symptoms at all. There were 15 male and 8 female Alpha patients in our sample. Both younger and older patients were represented in each group. Some of the patients with severe symptoms were treated with antibiotics. Comparison of Delta and Omicron patients (albeit with a limited sample size) resulted in 468 genes with  $p$ -values  $< 0.1$ . KDM5D and DDX3Y had the highest scores, but these genes are not known to be related to immune response processes. However, they are related to spermatogenesis (Akimoto et al. 2008; Navarro-Costa et al. 2010). All Omicron samples happened to be (randomly) from women, so it is impossible to conclude what portion of the results were influenced by this factor in particular. We also did not elucidate the influence of the other significantly and highly expressed genes, CRYBG3 and PHACTR4. However, the moderately expressed CXCL8 is known to be elevated in viral infections, producing protein known as interleukin-8 (Baggiolini and Clark-Lewis 1992).

The study has several potential drawbacks. Patients with severe symptoms had a higher median age (68 years) compared to patients with mild symptoms (37 years), asymptomatic donors (42 years), or COVID-19-negative individuals from the control group (37 years). This is because the majority of patients hospitalized with severe symptoms were elderly, and therefore, the samples were the most accessible in this regard. Overall, 54.3% of the samples were from female patients. Only in the group of severely symptomatic patients, there was some imbalance (69.5% of samples from male patients). Our study has some typical disadvantages of standard bulk sequencing procedure. We did not separate cell types from each other, so signal from RNA-seq is mix from various cells. Also, there are limitations associated with the sampling procedure (such as different personal). For the purpose of minimalization of various confounding effects, multiple procedures at the level of data analysis were performed (quality control and preprocessing of data, data normalization). Also number of samples should be sufficient to answer significant part of biological variability unassociated with studied topic. There are multiple similar studies, which means that many of our results provide value as a validation study. However, many of these studies present results from different types of tissue or from a blood. The samples in our study originated from nasopharyngeal tissue, and there are not so many studies available with this

particular characteristic. As a sample, nasopharyngeal tissue has some drawbacks; however, it allows direct analysis of initial infection site instead of late infection sites in the lower respiratory tract.

## Conclusions

In this study, we investigated nasopharyngeal tissue transcriptome changes in COVID-19 disease. We identified pathways altered by the infection as such or its severity. Most of them are probably components of immune response or alterations caused directly by the virus. Many enriched pathway terms related to different diseases suggest similarity between these different pathological events. Then, there were cytokine-related terms, immune cell-related terms, and signaling pathways—most of them corresponding with the results of previous publications covering the studied topic.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12223-024-01130-x>.

**Author contribution** Conceived of or designed study: GM, MH, DH. Performed research: DH, KB, NJ. Analyzed data: DH. Wrote the paper: DH, MH, GM, KB. Data curation and resources by: PS, PJ. Project administration and funding acquisition: PJ, GM. All authors read and approved the final manuscript.

**Funding** This research was funded by OP Integrated Infrastructure for the project: Serious diseases of civilization and COVID-19, ITMS: 313011AVH7, co-financed by the European Regional Development Fund.

**Data availability** Data are available on database ENA (European Nucleotide Archive), under project PRJEB62682. <https://www.ebi.ac.uk/ena/browser/view/PRJEB62682>.

## Declarations

**Ethics approval** The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Bratislava self-governing region (12.1.2021).

**Informed consent** Informed consent was obtained from all subjects involved in the study.

**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will



need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aggarwal BB (2003) Signalling pathways of the TNF superfamily: a double-edged sword. *Nat Rev Immunol* 3:745–756. <https://doi.org/10.1038/nri1184>
- Akimoto C, Kitagawa H, Matsumoto T, Kato S (2008) Spermatogenesis-specific association of SMCY and MSH5. *Genes Cells* 13:623–633. <https://doi.org/10.1111/j.1365-2443.2008.01193.x>
- Al-Qahtani AA, Pantazi I, Alhamlan FS, Alotheid H, Matou-Nasri S, Sourvinos G, Vergadi E, Tsatsanis C (2022) SARS-CoV-2 modulates inflammatory responses of alveolar epithelial type II cells via PI3K/AKT pathway. *Front Immunol* 13. <https://doi.org/10.3389/fimmu.2022.1020624>
- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. Accessed 23 May 2023
- Baggiolini M, Clark-Lewis I (1992) Interleukin-8, a chemotactic and inflammatory cytokine. *FEBS Lett* 307:97–101. [https://doi.org/10.1016/0014-5793\(92\)80909-Z](https://doi.org/10.1016/0014-5793(92)80909-Z)
- Bass A, Liu Y, Dakshanamurthy S (2021) Single-cell and bulk RNASeq profiling of COVID-19 patients reveal immune and inflammatory mechanisms of infection-induced organ damage. *Viruses* 13:2418. <https://doi.org/10.3390/v13122418>
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Cucinotta D, Vanelli M (2020) WHO Declares COVID-19 a Pandemic. *Acta Biomed* 91:157–160
- Daamen AR, Bachali P, Bonham CA, Somerville L, Sturek JM, Grammer AC, Kadl A, Lipsky PE (2022) COVID-19 patients exhibit unique transcriptional signatures indicative of disease severity. *Front Immunol* 13. <https://doi.org/10.3389/fimmu.2022.989556>
- DESeq2 manual (Accessed 23.05. 2023) (n.d.) Available Online: <https://bioconductor.org/packages/devel/bioc/manuals/DESeq2/man/DESeq2.pdf>
- Fricke-Galindo I, Falfán-Valencia R (2021) Genetics insight for COVID-19 susceptibility and severity: a review. *Front Immunol* 12:622176. <https://doi.org/10.3389/fimmu.2021.622176>
- Hata K, Andoh A, Shimada M, Fujino S, Bamba S, Araki Y, Okuno T, Fujiyama Y, Bamba T (2002) IL-17 stimulates inflammatory responses via NF-kappaB and MAP kinase pathways in human colonic myofibroblasts. *Am J Physiol Gastrointest Liver Physiol* 282(6):G1035–G1044. <https://doi.org/10.1152/ajpgi.00494.2001>
- Henderson LA, Canna SW, Schulert GS, Volpi S, Lee PY, Kernan KF, Caricchio R, Mahmud S, Hazen MM, Halyabar O, Hoyt KJ, Han J, Grom AA, Gattorno M, Ravelli A, Benedetti F, Behrens EM, Cron RQ, Nigrovic PA (2020) On the alert for cytokine storm: immunopathology in COVID-19. *Arthritis Rheumatol* 72:1059–1063. <https://doi.org/10.1002/art.41285>
- Hernández Borrero LJ, El-Deiry WS (2021) Tumor suppressor p53: biology, signaling pathways, and therapeutic targeting. *Biochim Biophys Acta Rev Cancer* 1876:188556. <https://doi.org/10.1016/j.bbcan.2021.188556>. <https://github.com/bahlolab/bioinfotools/blob/master/SAMtools/flagstat.md> (Accessed 23.05. 2023). (n.d.)
- Huang J, Hume AJ, Abo KM, Werder RB, Villacorta-Martin C, Alysandratos KD, Beermlann ML, Simone-Roach C, Lindstrom-Vautrin J, Olejnik J, Suder EL, Bullitt E, Hinds A, Sharma A, Bosmann M, Wang R, Hawkins F, Burks EJ, Saeed M, Wilson AA, Mühlberger E, Kotton DN (2020) SARS-CoV-2 infection of pluripotent stem cell-derived human lung alveolar type 2 cells elicits a rapid epithelial-intrinsic inflammatory response. *Cell Stem Cell* 27:962–973. e7. <https://doi.org/10.1016/j.stem.2020.09.013>
- Jackson H, Rivero CI, Broderick C, Habgood-Coote D, D'Souza G, Nichols S, Vito O, Gómez-Rial J, Rivero-Velasco C, Rodríguez-Núñez N, Barbeito-Castifeiras G, Pérez-Freixo H, Barreiro-de Acosta M, Cunnington AJ, Herberg JA, Wright VJ, Gómez-Carballa A, Salas A, Levin M, Martinon-Torres F, Cristina CS (2022) Characterisation of the blood RNA host response underpinning severity in COVID-19 patients. *Sci Rep* 12:12216. <https://doi.org/10.1038/s41598-022-15547-2>
- Jamison DA, Anand Narayanan S, Trovão NS, Guarnieri JW, Topper MJ, Moraes-Vieira PM, Zaksas V, Singh KK, Wurtele ES, Beheshti A (2022) A comprehensive SARS-CoV-2 and COVID-19 review, Part 1: Intracellular overdrive for SARS-CoV-2 infection. *Eur J Hum Genet* 30:889–898. <https://doi.org/10.1038/s41431-022-01108-8>
- Kawasaki T, Kawai T (2014) Toll-like Receptor Signaling Pathways. *Front Immunol* 5:461. <https://doi.org/10.3389/fimmu.2014.00461>
- Kolberg L, Raudvere U, Kuzmin I, Vilo J, Peterson H (2020) gprofiler2 -- an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. *F1000Res* 9:709. <https://doi.org/10.12688/f1000research.24956.2>
- Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. [arXiv:1303.3997v2](https://arxiv.org/abs/1303.3997v2) [q-bio.GN]
- Li F, Li J, Wang PH, Yang N, Huang J, Ou J, Xu T, Zhao X, Liu T, Huang X, Wang Q, Li M, Yang L, Lin Y, Cai Y, Chen H, Zhang Q (2021) SARS-CoV-2 spike promotes inflammation and apoptosis through autophagy by ROS-suppressed PI3K/AKT/mTOR signaling. *Biochim Biophys Acta Mol Basis Dis* 1867:166260. <https://doi.org/10.1016/j.bbadis.2021.166260>
- Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30:923–930. <https://doi.org/10.1093/bioinformatics/btt656>
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>
- Maier W, Batut B (2023) Mutation calling, viral genome reconstruction and lineage/clade assignment from SARS-CoV-2 sequencing data. *Galaxy Training at training.galaxyproject.org*. <https://training.galaxyproject.org/training-material/topics/variant-analysis/tutorials/sars-cov-2-variant-discovery/tutorial.html>
- Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ (2020) COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet* 395:1033–1034. [https://doi.org/10.1016/S0140-6736\(20\)30628-0](https://doi.org/10.1016/S0140-6736(20)30628-0)
- Navarro-Costa P, Plancha CE, Gonçalves J (2010) Genetic dissection of the AZF regions of the human Y chromosome: thriller or filler for male (in)fertility? *J Biomed Biotechnol* 2010:1–18. <https://doi.org/10.1155/2010/936569>
- Ng DL, Granados AC, Santos YA, Servellita V, Goldgof GM, Meydan C, Sotomayor-Gonzalez A, Levine AG, Balcerak J, Han LM, Akagi N, Truong K, Neumann NM, Nguyen DN, Bapat SP, Cheng J, Martin CS, Federman S, Foox J, Gliwa AS, Reyes K, Chao-Yang P, Guevara H, Wadford D, Miller S, Mason C, Chiu CY (2021) A diagnostic host response biosignature for COVID-19 from RNA profiling of nasal swabs and blood. *Sci Adv* 7:eabe5984. <https://doi.org/10.1126/sciadv.abe5984>
- Oude Munnink BB, Worp N, Nieuwenhuijse DF, Sikkema RS, Haagmans B, Fouchier RAM, Koopmans M (2021) The next phase of SARS-CoV-2 surveillance: real-time molecular epidemiology. *Nat Med* 27:1518–1524. <https://doi.org/10.1038/s41591-021-01472-w>
- Rawlings JS, Rosler KM, Harrison DA (2004) The JAK/STAT signaling pathway. *J Cell Sci* 117:1281–1283. <https://doi.org/10.1242/jcs.00963>

- Rhoades NS, Pinski AN, Monsibais AN, Jankeel A, Doratt BM, Cinco IR, Ibraim I, Messaoudi I (2021) Acute SARS-CoV-2 infection is associated with an increased abundance of bacterial pathogens, including *Pseudomonas aeruginosa* in the nose. *Cell Rep* 36:109637. <https://doi.org/10.1016/j.celrep.2021.109637>
- Tan L, Wang Q, Zhang D, Ding J, Huang Q, Tang YQ, Wang Q, Miao H (2020) Correction: Lymphopenia predicts disease severity of COVID-19: a descriptive and predictive study. *Signal Transduct Target Ther* 5:61. <https://doi.org/10.1038/s41392-020-0159-1>
- Thompson RC, Simons NW, Wilkins L, Cheng E, Del Valle DM, Hoffman GE, Cervia C, Fennessy B, Mouskas K, Francoeur NJ, Johnson JS, Lepow L, Le Berichel J, Chang C, Beckmann AG, Wang Y, Nie K, Zaki N, Tuballes K, Sebra R, Merad M, Gnjatich S, Tuballes K, Barcessat V, Cedillo MA, Yuan D, Huckins L, Roussos P, Marron T, Glicksberg BS, Nadkarni G, Heath JR, Boyman O, Kim-Schulze S, Schadt EE, Charney AW, Beckmann ND (2023) Molecular states during acute COVID-19 reveal distinct etiologies of long-term sequelae. *Nat Med* 29:236–246. <https://doi.org/10.1038/s41591-022-02107-4>
- Tobon-Velasco J, Cuevas E, Torres-Ramos M (2014) Receptor for AGEs (RAGE) as MEDIATOR of NF- $\kappa$ B pathway activation in neuroinflammation and oxidative stress. *CNS Neurol Disord Drug Targets* 13:1615–1626. <https://doi.org/10.2174/1871527313666140806144831>
- Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, Wang B, Xiang H, Cheng Z, Xiong Y, Zhao Y, Li Y, Wang X, Peng Z (2020) Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* 323:1061–1069. <https://doi.org/10.1001/jama.2020.1585>
- Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, Yuan ML, Zhang YL, Dai FH, Liu Y, Wang QM, Zheng JJ, Xu L, Holmes EC, Zhang YZ (2020) A new coronavirus associated with human respiratory disease in China. *Nature* 579:265–269. <https://doi.org/10.1038/s41586-020-2008-3>
- York RD, Yao H, Dillon T, Ellig CL, Eckert SP, McCleskey EW, Stork PJS (1998) Rap1 mediates sustained MAP kinase activation induced by nerve growth factor. *Nature* 392:622–626. <https://doi.org/10.1038/33451>
- Zhang YL, Wang RC, Cheng K, Ring BZ, Su L (2017) Roles of Rap1 signaling in tumor cell migration and invasion. *Cancer Biol Med* 14:90–99. <https://doi.org/10.20892/j.issn.2095-3941.2016.0086>
- Zhu M, Liu Y, Koonpaew S, Granillo O, Zhang W (2004) Positive and negative regulation of Fc $\epsilon$ RI-mediated signaling by the adaptor protein LAB/NTAL. *J Exp Med* 200:991–1000. <https://doi.org/10.1084/jem.20041223>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.