scientific reports

OPEN

Transcriptomic analysis of differential expression between surviving and nonsurviving patients infected by the SARS-CoV-2 Delta variant

Vlasov Ivan[1,5✉], Usenko Tatiana[2,3,5], Panteleeva Alexandra[2,3], Nikolaev Mikhail[2,3], Izumchenko Artem[2,3], Panafidina Valeriia[2], Gavrilova Elena[2], Shlyk Irina[2], Miroshnikova Valentina[2,3], Shadrina Maria[1], Polushin Yurii[2], Pchelina Sofya[2,3,4] & Slominsky Petr[1]

For a more precise understanding of the course of the pathological process in patients with severe COVID-19, it is necessary to continue the search for factors that affect the course of the pathological process and the possibility of a favorable outcome in critically ill patients. Comparative RNA-seq analysis of the transcriptome of peripheral blood mononuclear cell (PBMCs) in patients with a severe clinical course of COVID-19 caused by the SARS-CoV-2 Delta strain revealed a number of differentially expressed genes that distinguish patients with different clinical outcomes (survivors vs. nonsurvivors) in the period of 30 days after admission to the hospital. Most of them are associated with the "negative regulation of viral process" and "negative regulation of immune response" clusters. Moreover, in surviving patients, there is increased expression of the key genes *C1QB*, *C1QA*, *ISG15*, *SERPING1*, *VSIG4*, *KLRD1*, *TRPM4*, and *HFE* incorporated in these clusters. Among these key genes, the *ISG15* gene, which links several clusters of gene ontology enrichments and encodes an interferon-induced ubiquitin-like protein, deserves special attention. Its product, ISG15, is known to be a primary substrate for SARS-CoV-2 protease PLpro, which plays a role in counteracting hosts' antiviral mechanisms.

In December of 2019, the first patients with a novel form of pneumonia caused by infection with SARS-CoV-2 virus were described in the Wuhan province of China. Previously, two other pneumonia-inducing coronaviruses—MERS-CoV and SARS-CoV—had been described. All three of the aforementioned coronaviruses cause a severe form of pneumonia (with an up to 36% lethality rate in the case of MERS-CoV); however, SARS-CoV-2 is the most transmissible[1]. Its high transmissibility and quick spread led to a global pandemic, resulting in more than 700 million cases and more than 7 million deaths worldwide[2].

RNA viruses, and SARS-CoV-2 in particular, are characterized by high genome instability, which is caused by errors in replication of viral RNA[3]. From the standpoint of SARS-CoV-2 infection pathophysiology, mutations in spike glycoprotein S are especially interesting, since this is the protein that allows SARS-CoV-2 to invade host cells via binding to angiotensin-converting enzyme 2 (ACE2)[4]. Overall, since 2019, more than 400 mutations have been identified in S glycoprotein[5–9]. However, only some of those noticeably affected virus properties. Such impactful mutations have caused several SARS-CoV-2 variants to emerge, causing several waves of pandemic, each wave dominated by its own virus variant.

[1]National Research Center «Kurchatov Institute», Moscow, Russian Federation. [2]Pavlov First Saint-Petersburg State Medical University, Saint-Petersburg, Russian Federation. [3]Petersburg Nuclear Physics Institute named by B.P. Konstantinov of National Research Center «Kurchatov Institute», Saint-Petersburg, Russian Federation. [4]Kurchatov Genome Center—PNPI, Saint-Petersburg, Russian Federation. [5]Vlasov Ivan and Usenko Tatiana contributed equally to this work. ✉email: invlasov@mail.ru

The first variant to emerge was the B1.1.7 variant, also known as the Alpha variant. It was initially identified in Great Britain in September of 2020 and later spread to more than 200 countries all over the world[10]. This variant has 23 mutations differentiating it from the original Wuhan variant, three of which—N501Y, D614G, and P681H—are considered to be most impactful in causing an increase in the transmissivity and severity of the virus. After that, several other impactful variants emerged, including Beta (B.1.351, identified in October 2020 in South Africa), Delta (B.1.617.2, identified in December 2020 in India), Gamma (B.1.1.28, identified in January 2021)[10] and Omicron (B.1.1.529, identified in November 2021 in South Africa). Currently, the Omicron variant and its subvariants dominate the landscape of SARS-CoV-2 genetic variability[11].

Mutations in variants of the virus (both in S glycoprotein and other genes) can affect the properties of the virus on both clinical and molecular levels. When compared to previously emerged variants, the Delta variant of SARS-CoV-2 carries more than ten novel mutations in spike protein[12,13], including two major mutations in receptor binding domain[13]. These mutations are posited to lead to increased infectivity, severity, and immune avoidance of this variant[12], as well as stronger affinity to ACE2[13].

Previously, we have shown that, in peripheral blood monocytes (PMBCs) of patients with severe COVID-19 caused by the Alpha variant of SARS-CoV-2, expression of genes related to low-density lipoprotein (LDL), the metabolic pathway was increased in patients with lethal outcomes[14]. As virus continues to evolve, it is necessary to re-examine previous findings and establish, if mechanisms that previously were found to be important in determination of the outcome still play the same role in COVID-19 cases, caused by the infection with new strains of virus. In this study, our goal was to replicate the design of our previous study while investigating a different SARS-CoV-2 variant. For that reason, we have investigated differential expression in primary blood monocytes between SARS-CoV-2 survivors and nonsurvivors, with both groups being confirmed to be infected by the Delta coronavirus variant. As in the previous paper, we have compared intensive care unit (ICU) patients with lethal and nonlethal outcomes at 30 days since hospitalization.

Materials and methods
Patients
From June 25, 2021, to August 12, 2021, 66 patients with severe pneumonia were admitted to the ICU of Pavlov First State Medical University of St. Petersburg. Patients were filtered based on the following inclusion criteria: Delta SARS-CoV-2 positive COVID-19 diagnosis as confirmed by RT-PCR, Russian ethnicity, age between 50 and 70 years, absence of chronic comorbidities (cancer, cerebrovascular diseases, heart failure, renal failure). Twenty patients met the aforementioned inclusion criteria. The situation of the patients was monitored for 30 days, and with ten patients dying and ten patients surviving in this period. Ten (five + five) patients from these groups were enrolled in the present study. The baseline demographic and clinical characteristics of the patients are summarized in Table 1. In all cases, the sample was collected at the moment of hospitalization to the ICU, the outcome was recorded 30 days after hospitalization, and RNA was extracted 7–21 h after sample collection.

The recruitment for this study was conducted from June 2021 until September 2021. The study was conducted in accordance with the World Medical Assembly Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. All blood samples were collected with the informed consent of the investigated patients. The study was approved by the Ethics Committee of Pavlov First State Medical University of St. Petersburg (Russia). All patients provided their written informed consent for participation in this study.

RNA isolation, library preparation, and sequencing
Peripheral blood mononuclear cells (PBMCs) were isolated from 8 mL of EDTA anticoagulated venous blood by a Ficoll-Paque gradient method (Ficoll-Paque PLUS, GE Healthcare, Chicago, IL, USA). After centrifugation, PBMCs were collected from the interface and washed twice with PBS (pH 7.4) to remove the platelet-rich plasma fraction. The PBMC cell pellets were aliquoted and resuspended in TRIzol Reagent (Thermo Fisher Scientific, Waltham, MA, USA) and immediately frozen at 80°C.

Survival Status	Age	RIN	Body Temperature at Hospitalization	Oxygen Saturation	Respiration Rate	Ground Glasses' Opacity Severity Based on CT-SKAN (quartile)	Day of Death/Discharge
Deceased	61	8.2	36.8	70	23	3	4
Survived	60	9.3	37.8	75	24	3	10
Survived	45	9.2	37.2	86	24	3	16
Deceased	48	9.6	36.9	90	24	4	9
Survived	63	8.6	37.5	90	16	4	34
Deceased	43	9.1	36.7	89	22	3	9
Deceased	73	9.3	36.7	89	19	3	12
Survived	43	7.6	38.6	90	17	2	17
Survived	58	9.1	37.1	94	17	2	10
Deceased	66	8.8	36.70	97	18	2	32
Deceased	45	9.7	37.3	97	16	2	26
Survived	59	8.1	37.1	75	30	2	18

Table 1. Clinical characteristics and RNA integrity number (RIN) of participating individuals.

Total RNA was extracted from PBMCs of SARS-COV-2 infected patients using TRIzolTM Reagent (Thermo Fisher Scientific, USA) following the manufacturer's protocol. All RNA samples were stored at 80ºC for subsequent analysis. RNA quality and quantity were analyzed using a NanoDrop spectrophotometer system (Thermo Fisher Scientific) and an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) with the RNA 6000 Nano Kit according to the manufacturer's instructions. All samples had an RNA integrity number (RIN) > 8. Then, total RNA was used to obtain polyA fractions using oligoT magnetic beads Dynabeads mRNA Purification Kit (Ambion, Austin, TX, USA) according to the manufacturer's protocol. Further, libraries for massive parallel sequencing were prepared from polyA RNA using a NEBNext Ultra II RNA Library Prep Kit (NEB, Ipswich, MA, USA) and NEBNext Multiplex Oligos for Illumina (Index Primers Set 1) according to the kit instructions. The concentration of the libraries was determined using Qubit dsDNA HS Assay Kit with a Qubit 2.0 Fluorometer (Thermo Fisher Scientific). The lengths of distribution fragments from the cDNA libraries were assessed on an Agilent 2100 Bioanalyzer (Agilent Technologies) using the Agilent High Sensitivity DNA Kit (Agilent Technologies). Unpaired (50-bp reads). Sequencing of these libraries was performed on the Illumina HiSeq1500 platform using the TruSeq SBS Kit V3 sequencing reagents (Illumina, San Diego, CA, USA). RNA-Seq data for each sample (raw and processed data) can be accessed in the Gene Expression Omnibus, accession no. GSE272392.

Data processing and analysis

RNA-seq data were analyzed to identify differentially expressed genes (DEGs) using three pipelines according to the block diagram in Fig. 1. Gene ontology (GO) analysis of identified DEGs was performed to reveal enriched pathways.

Pipeline 1

Fastq files for each sample were aligned to the GRCh38 genome (Gencode, release 37) using HISAT2 (v2.0.1b)[15] with default parameters. Quality control for each sample was performed using FastQC (version 0.11.9)[16] and RSeQC (version 4.0.0)[17]. Counting reads was done using featureCounts[18]. Differential expression analysis was done using DESeq function from DESeq2 with default settings[19] in R (version 4.0.3). The following design was used in the analysis: gene ~ age + group + RIN + time of tissue collection (ttr). Detected differential expression of genes was considered significant at p.value < 0.05 and fold change (FC) threshold > 1.5.
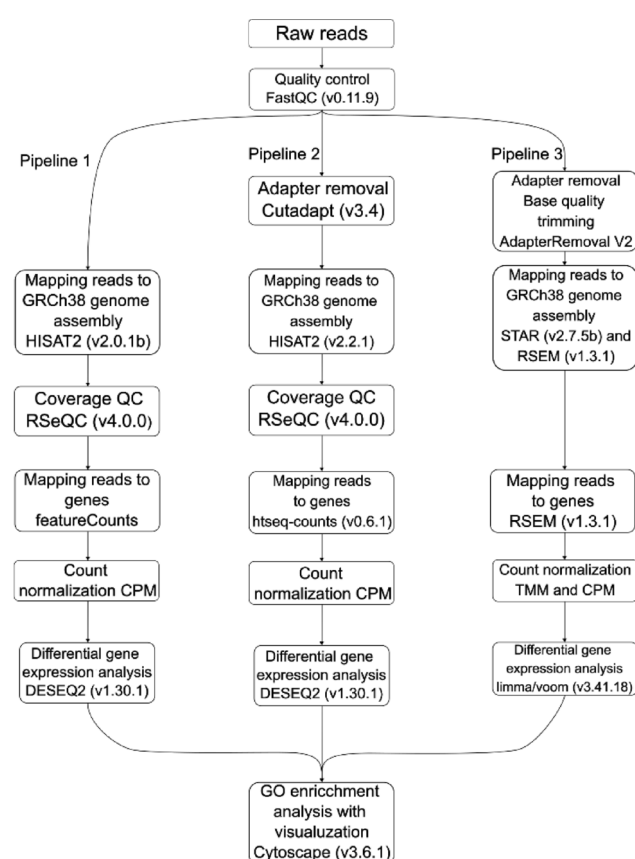
Fig. 1. Description of pipelines used during data processing.

Pipeline 2
Adapters were removed by Cutadapt (version 3.4). Fastq files for each sample were aligned to the GRCh38 genome (Gencode, release 37) using HISAT2 (version 2.2.1)[15] with default parameters to build an index of the reference genome and mapping reads to the genome. Quality control for each sample was performed by FastQC (version 0.11.9)[16] and RSeQC (version 4.0.0)[17]. Counts of the number of sequencing reads mapping to each gene after the alignment step were conducted using the htseq-count function from the HTSeq framework (version 0.6.1)[20]. Gene differential expression analysis of two groups (two biological replicates per condition) was performed using the DESeq2 package (version 1.30.1)[19] in R (version 4.0.3). The following design was used in the analysis: gene ~ age + group + RIN + ttr. Detected differential expression of genes was considered statistically significant at p.value < 0.05 and FC threshold > 1.5.

Pipeline 3
Ambiguous and low-quality bases were removed from the files obtained during the FastQ sequencing. The bases were removed with AdapterRemoval V2[21]. The trimmed files were aligned to the human reference genome GRCh38 and GRCh38.92 gene annotation using RSEM with the "rsem-prepare-reference" and "rsem-calculate-expression commands"[22], and the "-star" option was also used to generate STAR indices[23]. The resulting pseudocounts were normalized using the TMM algorithm implemented in the R "edgeR" package[24], as well as the "calcNormFactors" command and the CPM algorithm implemented in the R "limma" package, with the "voom" command[25]. To identify differential expressions, the normalized reads were processed using the commands "voom" (estimating the ratio of mean to variance, determining weights for observations), "lmFit" (creating a linear model describing observations), and "eBayes" (determining model parameters) from the R "limma" package[25]. In the linear model, the outcomes of the disease, age, time from the moment of tissue collection to isolation, and RNA integrity index were considered. DEGs were selected according to the following criteria: FC threshold > 1.5 and p-value moderated t test limma < 0.05.

Gene ontology enrichment analysis
Gene ontology (GO) enrichment analysis with biological processes terms was performed using Clue GO version 2.5.5[26] and Cluepedia version 1.5.5[27] for Cytoscape version 3.6.1 and was conducted for DEGs obtained using three pipelines. Significantly enriched terms were selected using the one-sided hypergeometric test with FDR correction ($p < 0.01$). Term groups were selected by ClueGO based on the number of common genes/term (> 40%). For enrichment, only terms of at least level 3 and no more than level 8 were considered, with which at least 3 DEGs were associated and with which the total number of DEGs associated was at least 4% of all genes associated with the term. All genes for which at least three reads were identified were used as a background for enrichment analysis.

Results
Differential gene expression analysis using DESeq2 and Limma/voom
We obtained PBMCs from each patient with severe COVID-19: five patients who survived (survivors) and five patients who died from infection in the ICU (nonsurvivors). Variation across biological replicates was low with Spearman correlation values within replicates for the two groups ranging from 0.86 to 0.89. When applying a threshold of FC > 1.5 and p.value.adjusted < 0.05. With FDR multiple testing corrections, no DEGs were identified. After applying a threshold of FC > 1.5 and p.value < 0.05, we identified a total of 429 DEGs (292 upregulated and 156 downregulated genes) using Pipeline 1, 345 DEGs (266 upregulated and 79 downregulated genes) using Pipeline 2, and 306 DEGs (191 upregulated and 115 downregulated genes) using Pipeline 3 in survivors compared with nonsurvivors among patients with severe Delta SARS-CoV-2 COVID-19 enrolled in this study. Lists of all DEGs of the three pipelines are available in Supplementary Table 1. Seventy-seven (60 upregulated and 17 downregulated) common DEGs were identified among all three pipelines (Supplementary Table 2). DEGs were considered to be common between pipelines if and only they were identified to be significantly differentially expressed in each pipeline.

Enrichment gene ontology analysis
To identify biological pathways in Delta SARS-COV-2 COVID-19 patients systematically, we performed GO enrichment analysis of significantly up- and downregulated genes for DEGs identified in each pipeline, as well as separately for the genes, common between all three pipelines. The functional modules were identified as mutually overlapping gene sets clustered together and named using GO hierarchical structure terms. A GO term enrichment was conducted for these genes. We considered "metabolic process" terms with $p < 0.05$ (Bonferroni corrected hypergeometric test) and all types of GO-term-to-gene connections. Significantly enriched terms are presented in Fig. 2 and Supplementary Table 3. Most significantly enriched terms were related to the processes of inflammation, innate immunity and viral replication. In Pipeline 1, such terms as GO:0070106 (interleukin-27-mediated signaling pathway), GO:0071639 (positive regulation of monocyte chemotactic protein-1 production), GO:0071605 (monocyte chemotactic protein-1 production, GO:0045071 (negative regulation of viral genome replication), and GO:0035456 (response to interferon-beta). In the case of Pipeline 2, terms with the lowest p.val were related to natural immunity and inflammation, such as GO:0002227 (innate immune response in mucosa), GO:0046597 (negative regulation of viral entry into host cell), GO:0006956 (complement activation), GO:0061844 (antimicrobial humoral immune response mediated by antimicrobial peptide), and GO:0045071 (negative regulation of viral genome replication). However, some unexpected terms were significantly enriched in Pipeline 2—muscle contraction (GO:000693) and oxidoreduction-driven active transmembrane transporter activity (GO:0015453). The latter term was also significantly enriched in Pipeline 3.
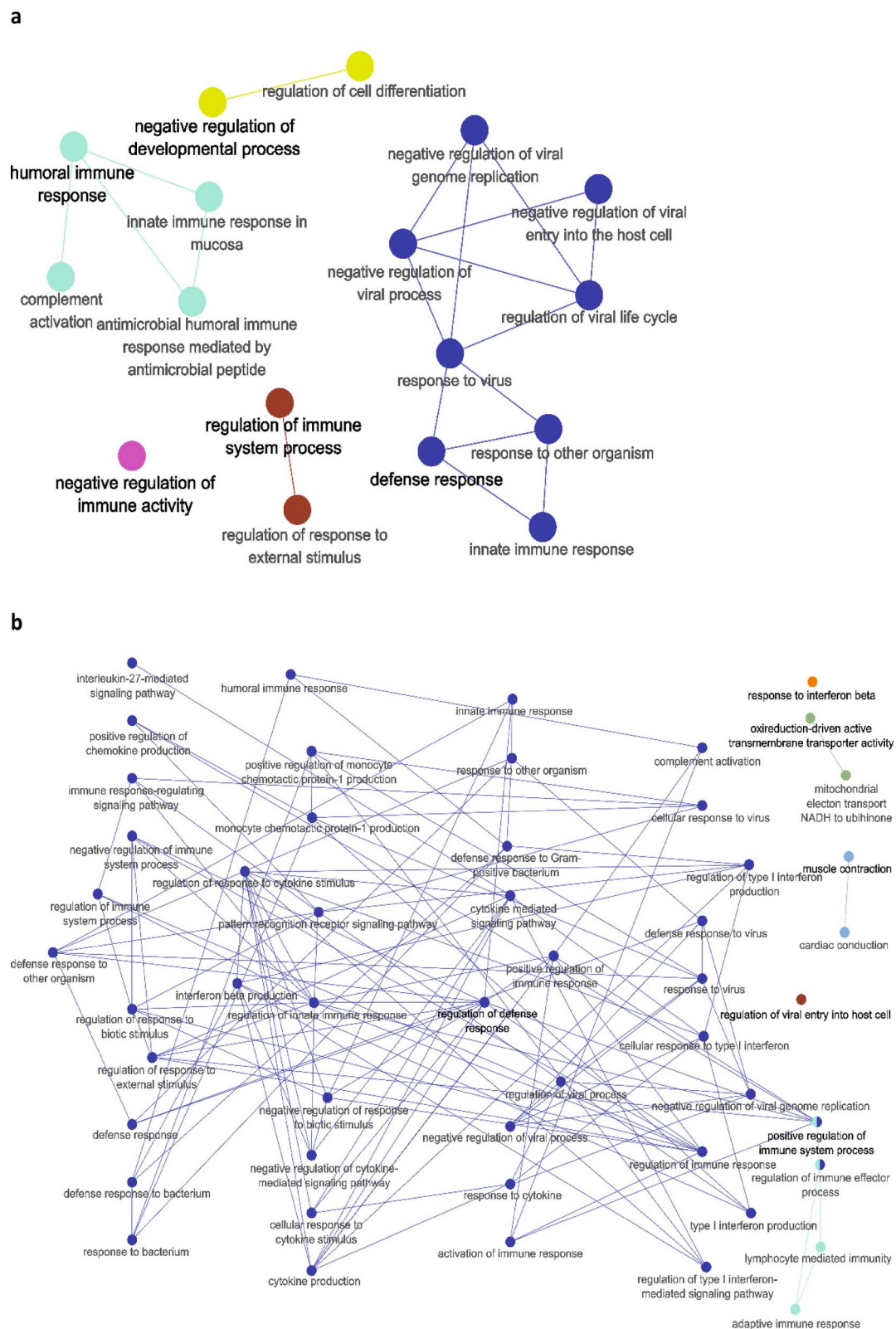
Fig. 2. GO term-term network analysis of the significantly enriched GO terms in differently expressed genes in survivors compared to nonsurvivors (multiple test correction by BH). (**A**) Pipeline 1; (**B**) Pipeline 2; (**C**) Pipeline 3.

c
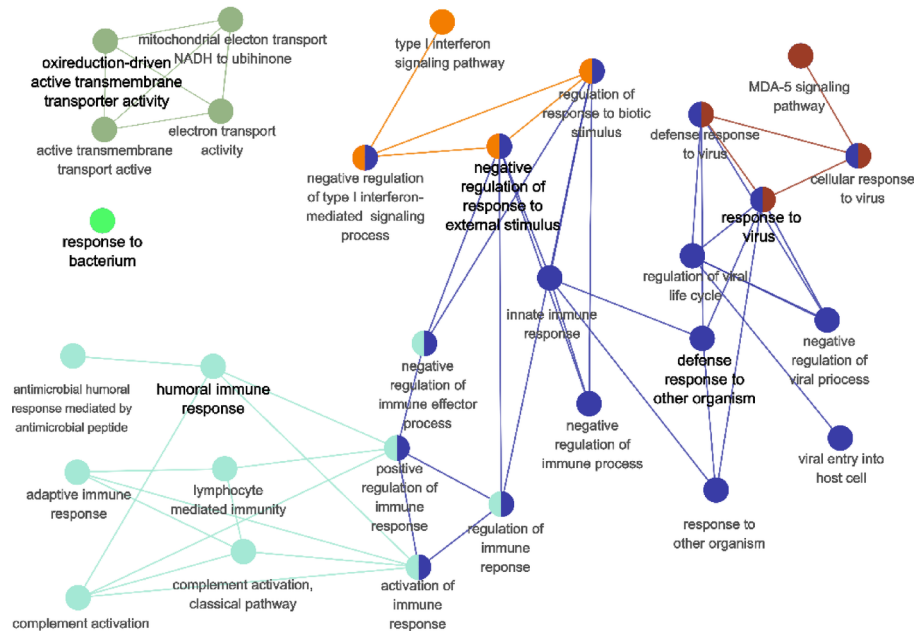
Figure 2. (continued)

In the case of Pipeline 3, most clusters also were related to the immune response—GO:0006958 (complement activation, classical pathway), GO:0039530 (MDA-5 signaling pathway), GO:0006956 (complement activation), GO:0006120 (mitochondrial electron transport, NADH to ubiquinone), and GO:0060339 (negative regulation of type I interferon-mediated signaling pathway).

Overlapping DEGs (77 total, 60 upregulated, 17 downregulated) in PBMCs of survivors compared with nonsurvivors between the three pipelines were identified and used for further analysis. The overlapping genes are presented in Supplementary Table 2. All but two (75 out of 77) the overlapping DEGs between the three pipelines had the same direction of fold change. Next, we conducted GO enrichment for 77 DEGs, common between three pipelines (Fig. 3). Fourteen significantly enriched terms were identified, 13 of which are related to aspects of immune response and viral processes (Fig. 3, Supplementary Table 4). A term–gene network has been constructed based on the relation of terms and genes from GO. Five groups were identified based on common DEGs among the terms.

Group one consists of six terms—GO:0050777 (negative regulation of immune response, the "main term" with the highest significance/lowest p value), GO:0002709 (regulation of T cell-mediated immunity), GO:0002697 (regulation of immune effector process), GO:0002698 (negative regulation of immune effector process), GO:0002449 (lymphocyte-mediated immunity), and GO:0002460 (adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains). Based on the main term and intersection of the terms, this group can be related to negative regulation of immune response. It is worth noting that every gene in this group is upregulated in survivors compared to nonsurvivors.

Group two consists of three terms—GO:0048525 (negative regulation of viral process), GO:0051607 (defense response to virus), and regulation of GO:1,903,900 (regulation of viral life cycle). The terms of this group are related to negative regulation of viral processes.

Group three consists of two terms—GO:0006956 (complement activation) and GO:0006959 (humoral immune response). It is worth noting that the term "complement activation" is the most significantly enriched of all the terms and has the highest percentage of related genes differentially expressed.

Both groups four and five consist of a single term—GO:0006936 (muscle contraction) and GO:0042742 (defense response to bacterium), respectively.

Based on this enrichment, gene *ISG15* is especially interesting, since it is related to three different GO term groups at the same time. Its expression in increased 2.8-fold in survivors compared to nonsurvivors. It is also notable as one of four genes related to five terms at the same time—GO:0050777 (negative regulation of immune response), GO:0048525 (negative regulation of viral process), GO:1,903,900 (regulation of viral life cycle), GO:0051607 (defense response to virus), and GO:0042742 (defense response to bacterium).

Eight genes are related to four and more significantly enriched terms—*C1QB*, *C1QA*, *ISG15*, *SERPING1*, *VSIG4*, *KLRD1*, *TRPM4*, and *HFE*. Expression of all these genes is increased in survivors compared to
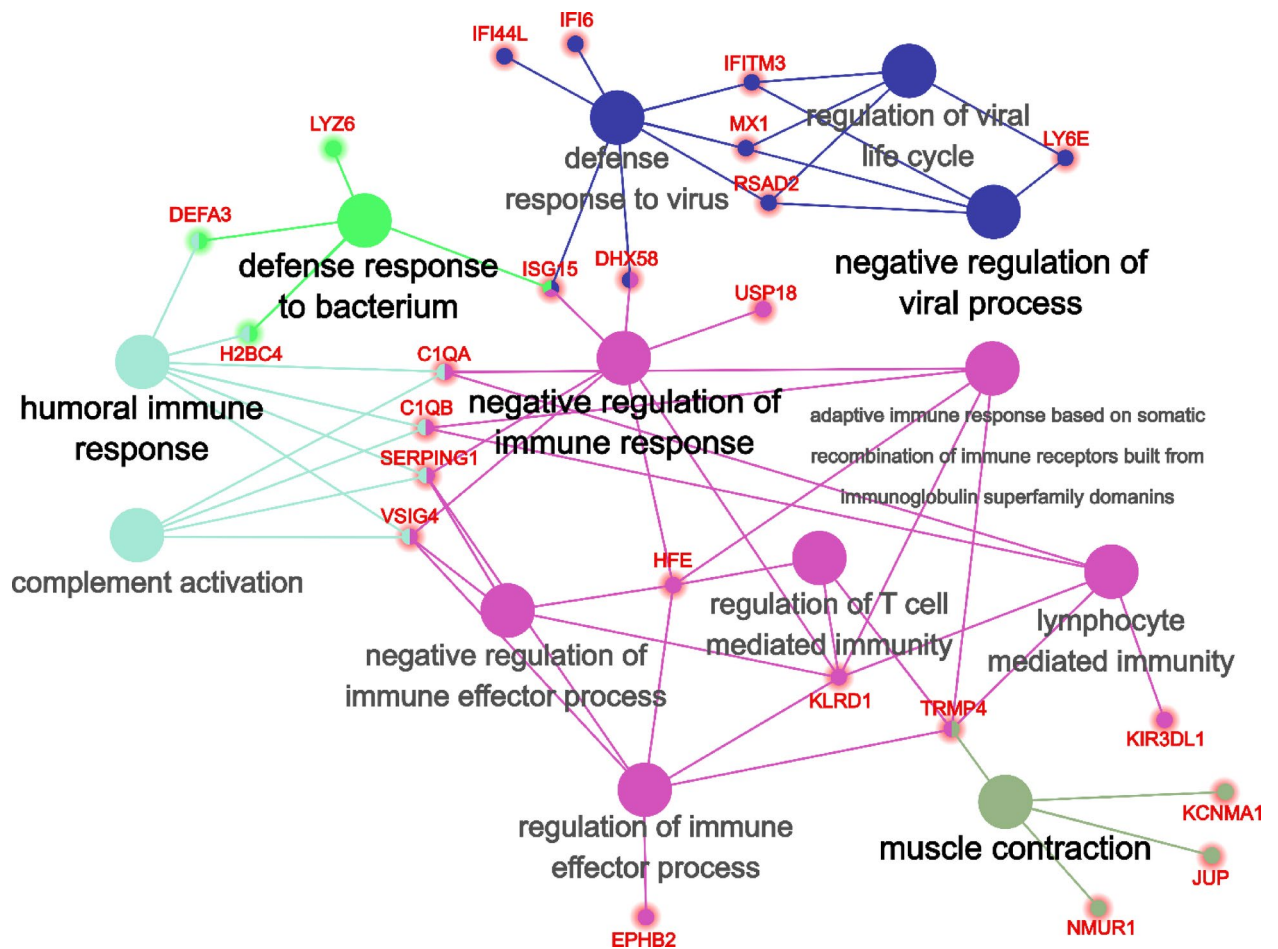
Fig. 3. GO gene–term network analysis of the top enriched GO terms in differentially expressed genes in survivors compared to nonsurvivors (multiple test correction by BH) in overlapping DEGs between all pipelines. Halos around genes indicate the direction of differential expression, with a red halo denoting increased expression in the survivor group as compared to the nonsurvivor group and a green halo denoting reduced expression in the survivor group as compared to the nonsurvivor group.

nonsurvivors. Genes of the first component of the complement system—*C1QB* and C1QA—were among the most overexpressed (a more than 4.5-fold increase). The same is the case for the gene encoding the immunoregulatory protein VSIG4. Average expression levels (in CPM) of five of these genes per group can be seen in Fig. 4.

Discussion

Our study has identified a set of genes, the differential expression of which is associated with survival outcomes in severe COVID-19, particularly through their involvement in negative regulation of immune response, complement activation, and the defense response to the virus. To achieve this, we have conducted transcriptomic analysis of PBMCs derived from patients with severe SARS-CoV-2 Delta strain induced pneumonia and identified a number of genes differentially expressed between survivors and nonsurvivors in the period of 30 days after hospitalization in the ICU. To elucidate the role of identified DEGs in determination of survival outcome, we have conducted Gene Ontology Biological Processes (GO: BP) term enrichment. Most of the terms significantly enriched by DEGs were related to various immune and antiviral processes. A term–gene network has been constructed, based on the relation of terms and genes from GO. Significantly enriched terms can be split into five groups based on common genes—negative regulation of immune response (group 1), negative regulation of viral processes (group 2), humoral and complement-system activation (group 3), defense response to bacterium (group 4), and muscle contraction (group 5). Next, we will analyze potential role of significantly enriched gene ontology biological processes in COVID-19 and identify some key genes based on their position in the term–gene network and the current understanding of their role in the pathophysiology of COVID-19 and other diseases.

However, before this, it is worth mentioning that the small sample size impacts the generalizability of the study's findings, especially in what concerns differential expression of particular genes. Since Gene Ontology enrichments impose additional levels of statistical testing on the data, we believe that identification of biological processes as a whole in the outcome of COVID-19, and less so individual DEGs, is the main takeaway of the
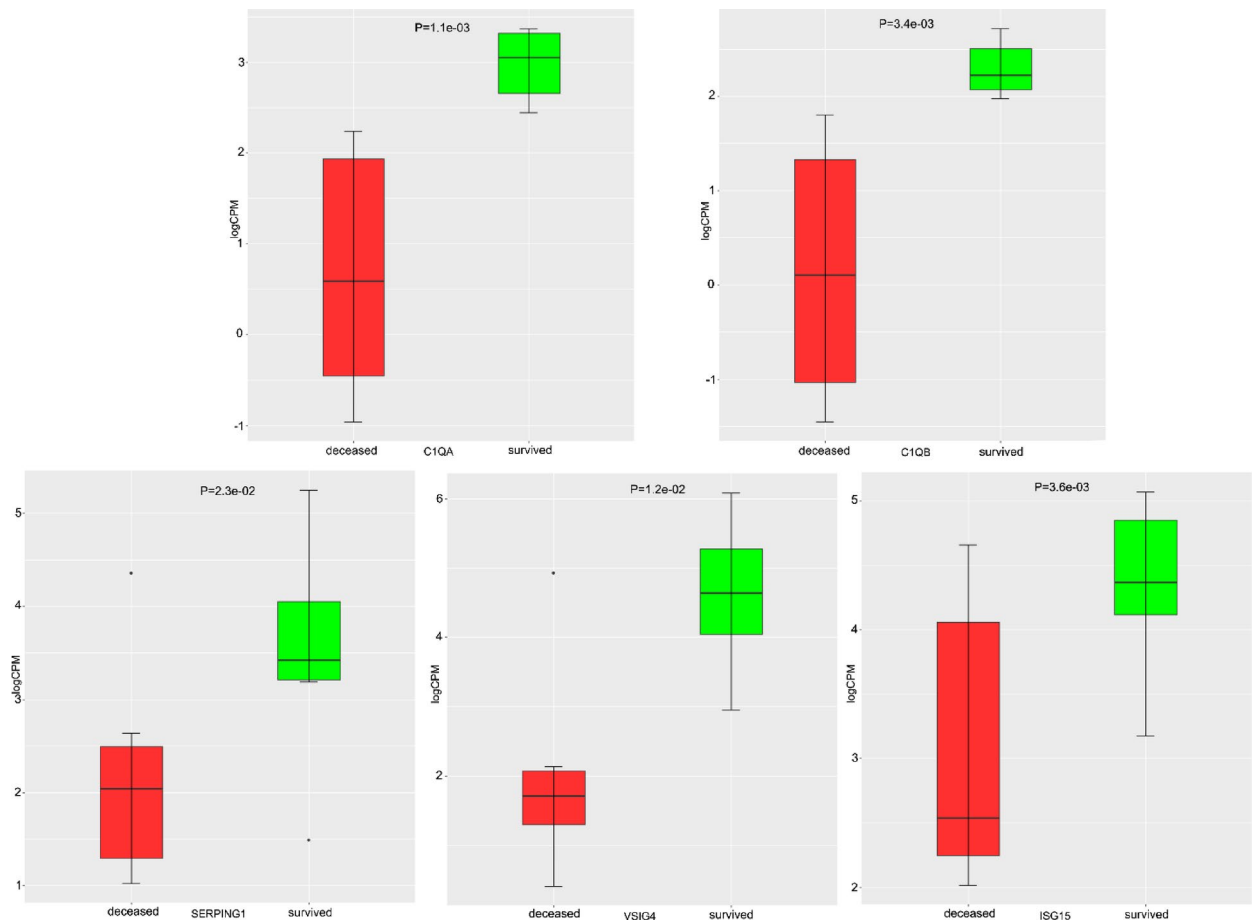
Fig. 4. Average (between three pipelines) expression in logarithm of counts per million (logCPM) per group. Red color denotes gene expression in the nonsurvivor group; green color denotes gene expression in the survivor group.

study. Differential expression of identified key genes needs to be further validated on larger independent cohorts in order to confirm their role as important determinant in COVID-19 outcome and potential therapeutic targets.

The dual role of immune regulation in COVID-19

It has been previously established that the role of immune response in COVID-19 is twofold—on one hand, strong and timely activation of the immune system is necessary to prevent viral replication and stop the infection, but, on the other hand, the immune system hyperactivation that develops in the severe cases of COVID-19 contributes to its severity and lethality[28]. Data obtained in our own study further contribute to this twofold picture. The significant enrichment of terms related to negative regulation of immune response and the fact that most of these genes are upregulated in survivors imply that the positive outcome of severe COVID-19 is partially based on downregulation of the immune system. However, significant enrichment of terms related to negative regulation of viral processes shows that strong antiviral immune response is still necessary for survival in severe COVID-19 cases. SARS-CoV-2's ability to evade the immune system and its ability to cause reduced interferon response are well documented[29]. The differential expression of genes related to these terms in the survivor group underlines the crucial role of effective immune response in counteracting COVID-19. The importance of interferon response is further stressed by group 2 terms being specifically related to three interferon induced DEGs—I*FI6*, *IFITM3*, and *IFI44L*—*a*ll of which are overexpressed in survivor group. It is further worth noting that only two DEGs are related to both groups 1 (negative regulation of immune response) and 2 (negative regulation of viral processes), with most being exclusively related to either group, suggesting that different sets of genes are involved in negative and positive components of immune system regulation. Taken together, these data points illustrate that the balance between immune activation and regulation appears critical for survival in severe COVID-19. While robust antiviral defense mechanisms are necessary, excessive immune responses, such as the dysregulated cytokine production observed in patients with severe and critically severe cases of COVID-19[30], can exacerbate the disease and cause additional harm.

Putting the terms "humoral response" (as a part of adaptive immunity) and "complement activation" (as a part of natural immunity) into a single group is most likely based on the involvement of the common genes, associated with both groups 1 and 3—*SERPING1*, *VSIG4*, *C1QB*, and *C1QA*. Before discussing these genes, however, it is worth examining the genes that are related to humoral response but not to complement activation.

Both of them, *DEFA3* and *H2BC4*, are also simultaneously related to the term "defense response to bacterium," and these genes are most likely the reason behind enrichment of the "humoral response" term. Enrichment of the term "defense response to bacterium" appears at first glance to be a result of an error. However, upon further examination, this enrichment could have important implications for its role in the outcome of COVID-19. Genes *LYZ*[31], *DEFA3*[32], and *H2BC4*[33] encode proteins with well-documented antimicrobial properties. Among them, *DEFA3* and *H2BC4* are related to both terms humoral immune response and defense response to bacterium, while *LYZ* is related exclusively to defense response to bacterium. All three genes of these are upregulated in the group of nonsurvivors compared to survivors. Most of the genes used in GO enrichment are the opposite—upregulated in survivors compared to nonsurvivors (60 out of 77)---which makes it highly unlikely that this happened by chance. This may imply that downregulation/absence of upregulation of the antimicrobial and humoral response may play an important role in a positive outcome of COVID-19.

The role of ISG15 as a potential key gene in determining COVID-19 outcomes

In the gene–term network (Fig. 3), gene *ISG15* occupies a key position, as it is the only gene that connects three different groups (1, 2, and 4) and is related to five different significantly enriched GO terms. *ISG15* encodes interferon-inducible ubiquitin-like protein[34], which binds covalently to various substrates originating from both host cell and virus in a process called "ISGylation." ISGylation plays an important role in inhibition of viral replication. The precise role and fate of ISGylated proteins are unknown; however, several possible mechanisms have been found. ISGylation can affect host protein localization, as in the case of filamin B[35], and it also could disrupt aggregation of viral protein complexes, as was observed in interaction of ISG15 with influenza B virus and human papilloma virus[36]. ISG15 can also disrupt the egress of virus from host cells, as has been shown using the example of interaction of ISG15 with human immunodeficiency virus 1 (HIV-1), in which release of HIV-1 virions, but not production of HIV-1 proteins, were impeded by overexpression of ISG15[37]. Besides interacting with various proteins through ISGylation, unconjugated ISG15 can act as a cytokine. Extracellular ISG15 acts as a ligand for integrin ITGAL[38], interacting with it and stimulating release of IFNγ and IL10 from NK cells through that interaction, with expression of IL10 being especially notable as an important factor in COVID-19 prognosis[30].

Previously, activation of *ISG15* expression was identified in expressional profiling of nasopharyngeal swabs of patients infected with COVID-19[39–41]. On the other hand, in vitro SARS-CoV-2 infection did not cause an increase in *ISG15* mRNA production, which could be explained by specific properties of the used cell culture[42]. The important role of *ISG15* can be also emphasized by the concentration and activity of ISG15 in a cell regulated by SARS-CoV-2. Papain-like proteinase PLpro plays an important role in this regulation. In SARS-CoV-2, PLpro is both necessary for normal processing of viral proteins and in regulation of posttranslational modification of host proteins, primarily ISG15, whereas in other coronaviruses it is primarily employed to prevent ubiquitination of cellular and viral proteins[43–45]. The main targets of PLpro are complexes of ISG15 with IRF3 (interferon responsive factor 3, regulator of type 1 interferon response) and MDA5 (a protein from the family of RIG-I-like receptors, stimulating the synthesis of interferon and proinflammatory cytokines when binding to viral RNA)[46,47]. This adaptation, allowing SARS-CoV-2 to target specifically ISG15 and ISGylated proteins, demonstrates the vital importance of ISG15 in antiviral immunity and suggests that increased expression and, consequently, activity of ISG15 play a crucial role in determining the outcomes of SARS-CoV-2 infection. Increased levels of *ISG15* mRNA in the survivor group may lead to increased presence of ISG15 protein and, in turn, allow some of it to avoid inactivation by SARS-CoV-2. Based on our data and previous research, we consider overexpression of *ISG15* to be a key event in determining a positive outcome of severe COVID-19 and believe it to be a valuable target for therapeutic intervention.

The complement system in COVID-19

Based on the raw number of related terms in the enrichment, eight key genes can be identified—*C1QB*, *C1QA*, *ISG15*, *SERPING1*, *VSIG4*, *KLRD1*, *TRPM4*, and *HFE*. Four of these genes are related to both groups 2 and 3 and to the terms "complement activation" and "negative regulation of immune response," as well as the aforementioned "humoral response"---*C1QA*, *C1QB*, *SERPING1*, and *VSIG4*. Going further, we would like to focus our attention on these four genes and on their role and the wider role of the complement system in COVID-19. Also, it is worth noting that all four of these genes are upregulated in survivors compared to nonsurvivors.

The complement system plays a role in antiviral infections by opsonizing viruses and virus-infected cells, directly neutralizing viruses outside of cells, lysing virus-infected cells, and inducing virus-specific immune and inflammatory responses[48]. The complement system can be activated through three possible pathways—classical, nonclassical, and lectin. Generally, it is considered that all the pathways of complement system activation lead to the same protein cascade and activation of the same effector mechanisms. However, based on our own and previously obtained data, it appears that the pathway of complement system activation could be important for determining outcomes of COVID-19. As previously mentioned, significantly increased expression of genes *C1QA* and *C1QB*, encoding both subunits of the first complement system component (related specifically to the classical pathway of complement activation), was observed in our comparison. Currently, it is not completely clear how interaction with SARS-CoV-2 activates the complement system. It has been shown previously that all of the major SARS-CoV-2 proteins (S, N, M, E) interact with the first component of the complement system, engaging the classic pathway of complement system activation[49]. However, additional data exist that contradict these finding and implies that SARS-CoV-2 proteins do not directly bind to C1Q, instead activating the complement through the lectin pathway[50]. Complement system proteins deposits are found in various tissues of COVID-19 patients, e.g., alveolar, kidney, and liver capillaries. However, most of those deposits consist of complement system proteins related to a nonclassical complement-system activation pathway[51]. A retrospective examination of concentration of several complement system components (C1Q, C3, C4, and C5)

in a cohort of 74 patients with COVID-19 has demonstrated that a high level of expression of nonclassical and lectin complement system activation pathway proteins are linked to negative outcomes of COVID-19, while no analogous connection was found in the case of classical pathway[52]. It has also been shown that in a cohort of 71 patients, levels of C1q in serum were decreased in severe cases of COVID-19 when compared with patients with mild COVID-19[53]. While not all changes in expression that can be observed in comparisons of mild and severe cases must necessarily also be present in comparisons of survival and nonsurvival outcomes, the assumption can be made that transcriptomic profiles of patients with mild cases of COVID-19 will be closer to those of surviving patients and, correspondingly, the transcriptomic profiles of patients with severe cases will be closer those of patients with terminal outcomes. Under this assumption, we can observe that our results are in accordance with previously obtained data, since we have observed an increase in the expression of both subunits of first complement component in the survivor group. Taken together with previously published data, increased expression of *C1QA* and *C1QB* in the survivor group implies that activation of the complement system via the classical pathway may contribute to a positive outcome of the disease. However, whether this relationship is casual in nature or is an association should be interpreted carefully, since it has also been shown that the concentration of C1q, unlike that of IL6 and IL10, is not an independent predictor of clinical outcomes, being instead dependent on the concentrations of other components of the "cytokine storm"[30].

The twofold role of the complement system in determining COVID-19 outcomes

Previously, the twofold role that immune system regulation plays in COVID-19 pathogenesis has been mentioned. The role of the complement system in SARS-CoV-2 pathogenesis also fits a twofold paradigm. On one hand, hyperactivation of the complement system has long been implicated in development of respiratory failure in both COVID-19[54] and other diseases[55], and targeting the complement system as an intervention appears a possible strategy for improving outcomes in severe COVID-19[56,57]. On the other hand, the complement system plays an indispensable role in combating infectious diseases[58], and inhibiting it as an intervention opens the door to a significant increase in the risk of treatment-induced infection-related side effects[56,57] due to disruption of its role in immune response. Several putative reasons for the twofold role of the complement system in severe SARS-CoV-2 pathogenesis can be posited. The first is timing, with early activation being beneficial and the later role of the complement becoming detrimental[48]; the second being the "Goldilocks zone"[58], with both too much activation and too little being detrimental; and the third one being that the particular pathway of activation determines whether complement activation will be detrimental or beneficial. Our own data could be considered as supporting the third hypothesis, seeing as we have identified that an increase in the expression of components of the complement system specifically involved in only the classic complement activation pathway is connected to survival outcomes. Previous research has also shown that outcomes of COVID-19 are associated with changes in concentration of components related to particular pathways of complement-system activation, but not just increase or decrease in activation of system overall[51–53]. Based on our data and previous research, increased expression of *C1QA* and *C1QB* could to be involved in pathogenesis and lethality of COVID-19, and could be considered a molecular marker of milder cases, and stimulation of their expression could be a possible option for therapeutic intervention. However, role of the complement system in SARS-CoV-2 is complex, and therapeutic interventions involving complement system are not without risk. Further research is required in order to better understand how complement system affects outcomes in SARS-CoV-2 infections.

VSIG4 and SERPING1—Potential key regulators of the complement system in COVID-19

Another key upregulated gene related to negative regulation of the immune system, humoral system, and complement activation is *VSIG4*, a gene encoding V-set and Ig domain-containing four proteins, also known as the "complement receptor of the immunoglobulin superfamily" (CRIg). This gene is expressed by macrophages and dendritic cells, participating in ingestion of complement-opsonized particles. VSIG4 also inhibits activation of a nonclassical complement pathway by binding to C3b[59,60], leading additional credence to the idea that the particular pathway of complement activation is one of the factors determining survival outcomes in COVID-19 patients, since its overexpression is linked to survival in our data. Besides participating in phagocytosis of complement-opsonized particles and complement system regulation, engagement of *VSIG4* with its ligands in macrophages also causes immunosuppressive and anti-inflammatory effects[61].

The role of *VSIG4* in COVID-19 pathogenesis is not fully understood. However, data points exist that point to its increased expression being both possibly detrimental and possibly beneficial. Reduced expression of VSIG4 leading to inhibition of anti-inflammatory activity was previously described in COPD (chronic obstructive pulmonary disease) as an important factor in its pathogenesis[62]. Since, based on our GO: BP enrichment, we consider negative regulation of the immune system being important in survival for severe COVID-19, its previously demonstrated immunosuppressive and anti-inflammatory activity could be beneficial. On the other hand, overexpression of this protein is considered an unfavorable factor in ARDS (acute respiratory distress syndrome)[63] and lung cancer[64,65], which could also mean its overexpression could be a detrimental factor in COVID-19, seeing how respiratory failure is one of the more dangerous symptoms in severe cases of COVID-19. Therefore, increased *VSIG4* can be both a detrimental and beneficial factor in severe COVID-19. Its overexpression in our data suggests that it could play a protective role in severe COVID-19 through modulation of the complement system response, inhibition of alternative pathway of complement activation and anti-inflammatory and immunosuppressive activity as a receptor. Based on its position in gene–term network and its role in several important processes in COVID-19 pathogenesis, we consider *VSIG4* to be one of the key genes and a potential biomarker and target for therapeutic intervention.

The final complement-system related differentially expressed key gene is *SERPING1*. *SERPING1* encodes C1-INH, a serine protease inhibitor (serpin), which plays a role in inhibition of the innate immune responses and complement system activation[66]. C1-INH is capable of inhibiting all three pathways of complement activation

and has shown the capability of limiting the complement and inflammation-related damage in multiple injury models[67,68]. C1-INH inhibits those pathways via forming covalent bonds with C1r and C1 s (classic pathway), MASP-1 and MASP-1 (lectin pathway), and C3b (alternative pathway)[66], and it is worth noting that, since this is a posttranslational mechanism, an observed increase of *C1QA* and *C1QB* expression on the transcriptional level may be not connected with increased activity of C1q. Being an interferon-induced gene, *SERPING1* could be posited to play a role of negative feedback mechanism, limiting immune system hyperactivation and, therefore, improving outcomes[66], which is especially important, considering the propensity of severe COVID-19 to cause immune-related damage to various tissues. *SERPING1* is implied to play a role in COVID-19 pathogenesis with its expression being shown to be increased compared to healthy individuals in reanalysis of three cases of from 2020[69]. It has also been shown that the T allele of variant rs78958998 is associated with increased risk of viremia in COVID-19 cases[70]. Overall, its beneficial role as a negative feedback mechanism against runaway immune processes fits very well with current understanding of the possible role of immune-mediated damage in severe COVID-19.

Currently, to our knowledge, no therapies for COVID-19 or other infectious diseases have been developed that specifically target *ISG15*, *SERPING1*, or *VSIG4* genes. As an IFN-induced gene, ISG15 can be upregulated by introduction of a wide variety of compounds, such as type I IFN, double-stranded RNA, lipopolysaccharide, or all-trans-retinoic acid[71]. However, the effect of such compounds is not specific to inducing *ISG15* expression, and they have a wide variety of effects on many different biological processes and induce a variety of different responses. IFN treatment appears to be the most promising among those, however evidence for efficacy and safety of IFN treatments in SARS-CoV-2 is mixed[72], leading to outcomes that are better than control in some and worse than control in other reports. Currently, treatment of hereditary angioedema with recombinant SERPING1 (Ruconest, Salix Pharmaceuticals) is approved in some jurisdictions[73], and it could be interesting to investigate it for efficacy in SARS-CoV-2 cases, although it is important to be mindful of how angioedema pathophysiology is different from that of COVID-19. As for *VSIG4*, anti-VSIG4 antibodies have been demonstrated to be able to induce immune response in vitro, in vivo, and ex vivo assays, but the relevance of these findings for SARS-CoV-2 remains to be shown, especially considering the fact that, in the original paper presenting those findings, the goal was to suppress activity of VSIG4, not activate it[74].

Determinants of outcome in the Delta vs. Alpha variant of COVID-19

Next, we would like to compare our data obtained in this study to the results of our own previous investigations of COVID-19. Previously, in a different paper, we have demonstrated that a lethal outcome in COVID-19 caused by SARS-CoV-2 Delta variant infection was associated with increased levels of pro-inflammatory cytokines[75]. However, in another paper, we have shown a downregulation of LDL particle-receptor pathway activity in ICU patients who survived severe COVID-19 infection[14], which was in agreement with results obtained in independent studies, including a meta-analysis of 22 published works[76–78]. It is worth considering why those findings have not been replicated in the present study. The methodology of the present study and the 2021 study are very similar, from patient selection to the analysis of differential expression. One of the significant differences is in the variant of SARS-CoV-2 causing the disease. The Alpha variant was the most widespread variant in St. Petersburg during the patient cohort collection of the 2021 study[14]. The Alpha variant is characterized by seven missense mutations (N501Y, A570D, D614G, P681H, T716I, S982 A, and D1118H) and three deletions (69/70 del and 144 del) in spike protein S. The Delta variant is characterized by a different set of mutations (T19R, G142D, Δ157–158, L452R, T478 K, D614G, P681R, and D950 N), and the identified differences could affect the interaction of S protein with ACE2 and immune-system proteins[10,79,80]. This could lead to both differences in virus infectivity and activation of different metabolic pathways in pathogenesis of COVID-19. The role of particular missense variants in interaction with cellular proteins is not clearly understood. For example, there have been no mutations identified in the gene encoding PLpro between Delta and Alpha variants of SARS-CoV-2[81,82]. However, no differences in the expression of *ISG15* between survivor and nonsurvivor groups have been identified in our data in the case of severe COVID-19 caused by infection by the Alpha variant of SARS-CoV-2. This suggests that the nature of our findings may be variant-specific. However, despite our best effort, several confounding variables are present, and could not be fully accounted for, as will be discussed further in the Study Limitations section. Therefore, interpretation of strain-specificity in our findings needs to be done cautiously. Nevertheless, the differences between the results of this and the previous study are evident. Significantly enriched biological processes in this study are almost exclusively related to the immune system, whereas there was a high number of metabolism-related significantly enriched processes in our previous study[14]. One possible explanation of such differences would be the role of evolving immune evasion of SARS-CoV-2. With SARS-CoV-2 constantly getting better at avoiding the immune system[83], the role of subtle differences of immune system regulation and activity in determining outcomes of COVID-19 should also increase.

Further study of patients infected by different variants of SARS-CoV-2 may elucidate the role of the variant-specific aspects of pathogenesis of COVID-19. We believe that continuous research of COVID-19 cases caused by different variants of constantly evolving SARS-CoV-2 is important. To accomplish this, our suggestion would be to attempt to conduct periodic studies of SARS-CoV-2 infection using various methods, such as RNA-seq, while keeping the design of each study as constant as possible. This would allow keeping as many of the confounding variables the same as possible and to investigate the variant-specific aspects of pathogenesis of COVID-19. In our work, taking together the present paper and a previous publication of a similarly designed investigation of COVID-19 infection caused by the Alpha variant[14], we have attempted to do exactly that. Fortunately, thanks to the spread of vaccinations, improving standard of care, and reduced load on health-care systems, number of lethal cases of COVID-19 has significantly gone down. However, this impairs our ability to adhere to the same study design in the future (comparison of survivors and nonsurvivors), which suggests that, going forward,

in a series of experiments on different strains and variants of SARS-CoV-2, we will need to employ different comparisons, such as, e.g., severe versus mild cases.

Conclusion

The transcriptomic analysis of PBMCs in patients with severe COVID-19 identified a number of genes differentially expressed between survivors versus nonsurvivors. Enrichment has been identified of several groups of significantly enriched GO biological-process terms related to various immune processes. Based on the significantly enriched terms and their related genes, several processes that may be important for surviving severe COVID-19 have been identified, such as upregulation of genes of the classical complement activation pathway, upregulation of genes related to negative regulation of the immune process, and upregulation of genes related to negative regulation of viral processes and downregulation of genes related to defense against microbial infection.

Several key genes have been identified based on this enrichment—*C1QB*, *C1QA*, *ISG15*, *SERPING1*, *VSIG4*, *KLRD1*, *TRPM4*, and *HFE*. Gene *ISG15*, increased expression of which can play a role in response to viral protease PLpro, is especially notable. The identification of key immune-modulating genes like *ISG15*, *SERPING1*, *C1QA*, *C1QB*, and *VSIG4* opens avenues for therapeutic approaches that aim to modulate the immune response and improve outcomes in severe COVID-19 cases via upregulation of expression of these genes. However, due to the small sample size used in the study and the use of p.values, not corrected for multiple testing, we believe that findings regarding particular differentially expressed genes should be considered to be preliminary and require confirmation on larger independent cohorts.

Previously, we have shown that downregulation of the LDL particle-receptor pathway activity plays a role in ICU patients that survive severe COVID-19 infection caused by the Alpha variant of SARS-CoV-2. On the other hand, most DEGs in comparison of survivors versus nonsurvivors of severe COVID-19 caused by the Delta variant are related to immune response. Such differences may be related to mutations in the structure of a viral protein, which lead to changes in pathogenesis of severe COVID-19.

Study limitations

1) The Delta variant of SARS-CoV-2 appears to be either extinct or almost extinct, with new studies of the Omicron variant and its subvariants being necessary to elucidate the differences between survivors and nonsurvivors in pathogenesis of severe COVID-19.
2) Despite our best efforts at replicating the design of the previous study, several changes have been made, such as in the number of included patients being changed from five versus three in a 2021 study to four versus five in the current study, and the number of technical replicates in 2021 of two being changed to one in the current study, which may impact the comparability of the two works.
3) Apart from changes in design, numerous changes to factors outside of our control have occurred since the previous study, with probably the most important being the ubiquity of vaccinated patients among ICU patients and changes in the ICU treatment protocols that are used. All of this makes identifying the true reasons behind the differences between the two studies almost impossible.
4) It is worth noting that the small sample size (five versus five) and use of p.values without multiple test corrections impacts the interpretability of our study. Due to the fact that enrichments pass through additional level of statistical tests, significantly enriched biological processes are probably the most generally applicable part of our results, whereas significant changes in the expression of particular genes (even the ones found in all three pipelines) need to be confirmed on independent cohorts before considering them for further examination or use.

Data availability

Raw FASTQ files and processed read counts are available from the Gene Expression Omnibus (GEO), accession no. GSE272392. In order to obtain additional data for this study, please contact the corresponding author at email provided on the title sheet.

References

1. Samieefar, N. et al. Delta Variant: The New Challenge of COVID-19 Pandemic, an Overview of Epidemiological, Clinical, and Immune Characteristics., *Acta Biomed*, vol. 93, no. 1, p. e2022179, Mar. (2022). https://doi.org/10.23750/abm.v93i1.12210
2. WHO. WHO Coronavirus (COVID-19) Dashboard. https://data.who.int/dashboards/covid19/cases?n=c
3. Giovanetti, M. et al. Evolution patterns of SARS-CoV-2: snapshot on its genome variants. *Biochem. Biophys. Res. Commun.* **538**, 88–91. https://doi.org/10.1016/j.bbrc.2020.10.102 (Jan. 2021).
4. Nitin, P., Nandhakumar, R., Vidhya, B., Rajesh, S. & Sakunthala, A. COVID-19: invasion, pathogenesis and possible cure – A review. *J. Virol. Methods*. **300**, 114434. https://doi.org/10.1016/j.jviromet.2021.114434 (Feb. 2022).
5. Starr, T. N. et al. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell* **182** (5), 1295–1310e. https://doi.org/10.1016/j.cell.2020.08.012 (Sep. 2020).
6. Bindayna, K. M. & Crinion, S. Variant analysis of SARS-CoV-2 genomes in the middle East. *Microb. Pathog.* **153**, 104741. https://doi.org/10.1016/j.micpath.2021.104741 (Apr. 2021).
7. Gladkikh, A. et al. Characterization of a novel SARS-CoV-2 genetic variant with distinct Spike protein mutations. *Viruses* **13** (6). https://doi.org/10.3390/v13061029 (May 2021).
8. Sahin, E. et al. Oct., Genomic characterization of SARS-CoV-2 isolates from patients in Turkey reveals the presence of novel mutations in spike and nsp12 proteins., *J Med Virol*, vol. 93, no. 10, pp. 6016–6026, (2021). https://doi.org/10.1002/jmv.27188
9. Klink, G. V. et al. Spread of endemic SARS-CoV-2 lineages in Russia before April 2021. *PLoS One*. **17**, e0270717. https://doi.org/10.1371/journal.pone.0270717 (2022). no. 7.

10. Choi, J. Y. & Smith, D. M. SARS-CoV-2 Variants of Concern., *Yonsei Med J*, vol. 62, no. 11, pp. 961–968, Nov. (2021). https://doi.org/10.3349/ymj.2021.62.11.961

11. Dai, B. et al. Update on Omicron variant and its threat to vulnerable populations. *Public. Health Pract. (Oxf)*. **7**, 100494. https://doi.org/10.1016/j.puhip.2024.100494 (Jun. 2024).

12. Saberiyan, M. et al. SARS-CoV-2: phenotype, genotype, and characterization of different variants. *Cell. Mol. Biol. Lett.* **27** (1), 50. https://doi.org/10.1186/s11658-022-00352-6 (Jun. 2022).

13. Mohammadi, M., Shayestehpour, M. & Mirzaei, H. The impact of spike mutated variants of SARS-CoV2 [Alpha, Beta, Gamma, Delta, and Lambda] on the efficacy of subunit recombinant vaccines., *Braz J Infect Dis*, vol. 25, no. 4, p. 101606, Jul. (2021). https://doi.org/10.1016/j.bjid.2021.101606

14. Vlasov, I. et al. Dec., Transcriptomic Profiles Reveal Downregulation of Low-Density Lipoprotein Particle Receptor Pathway Activity in Patients Surviving Severe COVID-19., *Cells*, vol. 10, no. 12, (2021). https://doi.org/10.3390/cells10123495

15. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype., *Nat Biotechnol*, vol. 37, no. 8, pp. 907–915, Aug. (2019). https://doi.org/10.1038/s41587-019-0201-4

16. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

17. Wang, L., Wang, S. & Li, W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28** (16), 2184–2185. https://doi.org/10.1093/bioinformatics/bts356 (Aug. 2012).

18. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features., *Bioinformatics*, vol. 30, no. 7, pp. 923–30, Apr. (2014). https://doi.org/10.1093/bioinformatics/btt656

19. Love, M. I., Huber, W. & Anders, S. 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 (Dec. 2014).

20. Putri, G. H., Anders, S., Pyl, P. T., Pimanda, J. E. & Zanini, F. Analysing high-throughput sequencing data in Python with HTSeq 2.0. *Bioinformatics* **38** (10), 2943–2945. https://doi.org/10.1093/bioinformatics/btac166 (May 2022).

21. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes*. https://doi.org/10.1186/s13104-016-1900-2 (2016).

22. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **12** (1), 323. https://doi.org/10.1186/1471-2105-12-323 (Aug. 2011).

23. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29** (1), 15–21. https://doi.org/10.1093/bioinformatics/bts635 (Jan. 2013).

24. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data., *Bioinformatics*, vol. 26, no. 1, pp. 139–40, Jan. (2010). https://doi.org/10.1093/bioinformatics/btp616

25. Ritchie, M. E. et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43** (7), e47. https://doi.org/10.1093/nar/gkv007 (Apr. 2015).

26. Bindea, G. et al. ClueGO: A cytoscape plug-in to Decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* https://doi.org/10.1093/bioinformatics/btp101 (2009).

27. Bindea, G., Galon, J. & Mlecnik, B. CluePedia cytoscape plugin: pathway insights using integrated experimental and in Silico data. *Bioinformatics* https://doi.org/10.1093/bioinformatics/btt019 (2013).

28. Sapir, T. et al. COVID-19 and the immune response: A Multi-Phasic approach to the treatment of COVID-19. *Int. J. Mol. Sci.* **23** (15). https://doi.org/10.3390/ijms23158606 (Aug. 2022).

29. Mattoo, S. U. S., Kim, S. J., Ahn, D. G. & Myoung, J. Escape and Over-Activation of innate immune responses by SARS-CoV-2: two faces of a coin. *Viruses* **14** (3). https://doi.org/10.3390/v14030530 (Mar. 2022).

30. Tang, Y. et al. Aberrant cytokine expression in COVID-19 patients: associations between cytokines and disease severity. *Cytokine* **143**, 155523. https://doi.org/10.1016/j.cyto.2021.155523 (Jul. 2021).

31. Witholt, B., Heerikhuizen, H. V. & De Leij, L. How does lysozyme penetrate through the bacterial outer membrane? *Biochim. Biophys. Acta*. **443** (3), 534–544. https://doi.org/10.1016/0005-2736(76)90471-5 (Sep. 1976).

32. Ganz, T. et al. Oct., Defensins. Natural peptide antibiotics of human neutrophils., *J Clin Invest*, vol. 76, no. 4, pp. 1427–35, (1985). https://doi.org/10.1172/JCI112120

33. Tollin, M. et al. Antimicrobial peptides in the first line defence of human colon mucosa. *Peptides (N Y)*. **24** (4), 523–530. https://doi.org/10.1016/s0196-9781(03)00114-1 (Apr. 2003).

34. Loeb, K. R. & Haas, A. L. The interferon-inducible 15-kDa ubiquitin homolog conjugates to intracellular proteins., *J Biol Chem*, vol. 267, no. 11, pp. 7806–13, Apr. [Online]. (1992). Available: http://www.ncbi.nlm.nih.gov/pubmed/1373138

35. Jeon, Y. J. et al. Apr., ISG15 modification of filamin B negatively regulates the type I interferon-induced JNK signalling pathway., *EMBO Rep*, vol. 10, no. 4, pp. 374–80, (2009). https://doi.org/10.1038/embor.2009.23

36. Perng, Y. C. & Lenschow, D. J. ISG15 in antiviral immunity and beyond. *Nat. Rev. Microbiol.* **16** (7), 423–439. https://doi.org/10.1038/s41579-018-0020-5 (Jul. 2018).

37. Okumura, A., Lu, G., Pitha-Rowe, I. & Pitha, P. M. Innate antiviral response targets HIV-1 release by the induction of ubiquitin-like protein ISG15., *Proc Natl Acad Sci U S A*, vol. 103, no. 5, pp. 1440–5, Jan. (2006). https://doi.org/10.1073/pnas.0510518103

38. Swaim, C. D., Scott, A. F., Canadeo, L. A. & Huibregtse, J. M. Extracellular ISG15 Signals Cytokine Secretion through the LFA-1 Integrin Receptor., *Mol Cell*, vol. 68, no. 3, pp. 581–590.e5, Nov. (2017). https://doi.org/10.1016/j.molcel.2017.10.003

39. Lieberman, N. A. P. et al. In vivo antiviral host transcriptional response to SARS-CoV-2 by viral load, sex, and age. *PLoS Biol.* **18** (9), e3000849. https://doi.org/10.1371/journal.pbio.3000849 (Sep. 2020).

40. Vastrad, B., Vastrad, C. & Tengli, A. Bioinformatics analyses of significant genes, related pathways, and candidate diagnostic biomarkers and molecular targets in SARS-CoV-2/COVID-19. *Gene Rep.* **21**, 100956. https://doi.org/10.1016/j.genrep.2020.100956 (Dec. 2020).

41. Fang, L., Tang, T. & Hu, M. Identification of Differentially Expressed Genes in COVID-19 and Integrated Bioinformatics Analysis of Signaling Pathways., *Genet Res (Camb)*, vol. p. 2728757, 2021, (2021). https://doi.org/10.1155/2021/2728757

42. Islam, T. et al. Integrative transcriptomics analysis of lung epithelial cells and identification of repurposable drug candidates for COVID-19., *Eur J Pharmacol*, vol. 887, p. 173594, Nov. (2020). https://doi.org/10.1016/j.ejphar.2020.173594

43. Klemm, T. et al. Mechanism and Inhibition of the papain-like protease, PLpro, of SARS-CoV-2. *Embo J.* **39** (18), e106275. https://doi.org/10.15252/embj.2020106275 (2020).

44. Shin, D. et al. Nov., Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity., *Nature*, vol. 587, no. 7835, pp. 657–662, (2020). https://doi.org/10.1038/s41586-020-2601-5

45. Zhang, M. et al. ISGylation in innate antiviral immunity and pathogen defense responses: A review. *Front. Cell. Dev. Biol.* **9**, 788410. https://doi.org/10.3389/fcell.2021.788410 (2021).

46. Liu, G. et al. ISG15-dependent activation of the sensor MDA5 is antagonized by the SARS-CoV-2 papain-like protease to evade host innate immunity. *Nat. Microbiol.* **6** (4), 467–478. https://doi.org/10.1038/s41564-021-00884-1 (Apr. 2021).

47. Tecalco Cruz, A. C. Free ISG15 and protein isgylation emerging in SARS-CoV-2 infection. *Curr. Drug Targets*. **23** (7), 686–691. https://doi.org/10.2174/1389450123666220316094720 (2022).

48. Java, A. et al. The complement system in COVID-19: friend and foe? *JCI Insight*. **5**, 15, https://doi.org/10.1172/jci.insight.140711 (Aug. 2020).

49. Savitt, A. G. et al. SARS-CoV-2 exacerbates COVID-19 pathology through activation of the complement and Kinin systems. *Front. Immunol.* **12**, 767347. https://doi.org/10.3389/fimmu.2021.767347 (2021).

50. Ali, Y. M. et al. Lectin pathway mediates complement activation by SARS-CoV-2 proteins. *Front. Immunol.* **12**, 714511. https://doi.org/10.3389/fimmu.2021.714511 (Jul. 2021).
51. Macor, P. et al. Multi-organ complement deposition in COVID-19 patients. *MedRxiv Jan.* https://doi.org/10.1101/2021.01.07.21249116 (2021).
52. Defendi, F. et al. Complement alternative and Mannose-Binding lectin pathway activation is associated with COVID-19 mortality. *Front. Immunol.* **12**, 742446. https://doi.org/10.3389/fimmu.2021.742446 (2021).
53. Wu, Y. et al. Clinical characteristics and immune injury mechanisms in 71 patients with COVID-19. *mSphere* **5** (4). https://doi.org/10.1128/mSphere.00362-20 (Jul. 2020).
54. Holter, J. C. et al. Oct., Systemic complement activation is associated with respiratory failure in COVID-19 hospitalized patients., *Proc Natl Acad Sci U S A*, vol. 117, no. 40, pp. 25018–25025, (2020). https://doi.org/10.1073/pnas.2010540117
55. Zilow, G., Sturm, J. A., Rother, U. & Kirschfink, M. Complement activation and the prognostic value of C3a in patients at risk of adult respiratory distress syndrome. *Clin. Exp. Immunol.* **79** (2), 151–157. https://doi.org/10.1111/j.1365-2249.1990.tb05171.x (Feb. 1990).
56. Annane, D. et al. Eculizumab as an emergency treatment for adult patients with severe COVID-19 in the intensive care unit: A proof-of-concept study. *EClinicalMedicine* **28**, 100590. https://doi.org/10.1016/j.eclinm.2020.100590 (Nov. 2020).
57. Zelek, W. M. et al. Nov., Complement Inhibition with the C5 Blocker LFG316 in Severe COVID-19., *Am J Respir Crit Care Med*, vol. 202, no. 9, pp. 1304–1308, (2020). https://doi.org/10.1164/rccm.202007-2778LE
58. Afzali, B., Noris, M., Lambrecht, B. N. & Kemper, C. The state of complement in COVID-19. *Nat. Rev. Immunol.* **22** (2), 77–84. https://doi.org/10.1038/s41577-021-00665-1 (Feb. 2022).
59. He, J. Q., Wiesmann, C. & van Campagne, M. A role of macrophage complement receptor CRIg in immune clearance and inflammation. *Mol. Immunol.* **45** (16), 4041–4047. https://doi.org/10.1016/j.molimm.2008.07.011 (Oct. 2008).
60. van Lookeren Campagne, M. & Verschoor, A. Pathogen clearance and immune adherence 'revisited': Immuno-regulatory roles for CRIg. *Semin Immunol.* **37**, 4–11. https://doi.org/10.1016/j.smim.2018.02.007 (Jun. 2018).
61. Small, A. G., Al-Baghdadi, M., Quach, A., Hii, C. & Ferrante, A. Complement receptor Immunoglobulin: a control point in infection and immunity, inflammation and cancer. *Swiss Med. Wkly.* **146**, w14301. https://doi.org/10.4414/smw.2016.14301 (2016).
62. Liu, Y., Liu, H., Li, C., Ma, C. & Ge, W. Proteome profiling of lung tissues in chronic obstructive pulmonary disease (COPD): platelet and macrophage dysfunction contribute to the pathogenesis of COPD. *Int. J. Chron. Obstruct Pulmon Dis.* **15**, 973–980. https://doi.org/10.2147/COPD.S246845 (2020).
63. Zhao, C. et al. Identification of an alveolar Macrophage-Related core gene set in acute respiratory distress syndrome. *J. Inflamm. Res.* **14**, 2353–2361. https://doi.org/10.2147/JIR.S306136 (2021).
64. Liao, Y. et al. Jul., VSIG4 expression on macrophages facilitates lung cancer development., *Lab Invest*, vol. 94, no. 7, pp. 706–15, (2014). https://doi.org/10.1038/labinvest.2014.73
65. Guo, T., Ma, H. & Zhou, Y. Bioinformatics analysis of microarray data to identify the candidate biomarkers of lung adenocarcinoma. *PeerJ* **7**, e7313. https://doi.org/10.7717/peerj.7313 (2019).
66. Hausburg, M. A. et al. C1 esterase inhibitor-mediated immunosuppression in COVID-19: Friend or foe? *Clinical immunology communications*, vol. 2, pp. 83–90, Dec. (2022). https://doi.org/10.1016/j.clicom.2022.05.001
67. Davis, A. E., Mejia, P. & Lu, F. Biological activities of C1 inhibitor., *Mol Immunol*, vol. 45, no. 16, pp. 4057–63, Oct. (2008). https://doi.org/10.1016/j.molimm.2008.06.028
68. Jiang, H., Wagner, E., Zhang, H. & Frank, M. M. Complement 1 inhibitor is a regulator of the alternative complement pathway., *J Exp Med*, vol. 194, no. 11, pp. 1609–16, Dec. (2001). https://doi.org/10.1084/jem.194.11.1609
69. Hausburg, M. A., Banton, K. L., Roshon, M. & Bar-Or, D. Clinically distinct COVID-19 cases share notably similar immune response progression: A follow-up analysis. *Heliyon* **7** (1), e05877. https://doi.org/10.1016/j.heliyon.2020.e05877 (Jan. 2021).
70. Roy-Vallejo, E. et al. Occurrence of SARS-CoV-2 viremia is associated with genetic variants of genes related to COVID-19 pathogenesis. *Front. Med. (Lausanne).* **10**, 1215246. https://doi.org/10.3389/fmed.2023.1215246 (2023).
71. Mustachio, L. M. et al. Evidence for the ISG15-Specific deubiquitinase USP18 as an antineoplastic target. *Cancer Res.* **78** (3), 587–592. https://doi.org/10.1158/0008-5472.CAN-17-1752 (Feb. 2018).
72. Jhuti, D. et al. Interferon Treatments for SARS-CoV-2: Challenges and Opportunities., *Infect Dis Ther*, vol. 11, no. 3, pp. 953–972, Jun. (2022). https://doi.org/10.1007/s40121-022-00633-9
73. Bouton, M. C. et al. The under-appreciated world of the Serpin family of Serine proteinase inhibitors. *EMBO Mol. Med.* **15** (6), e17144. https://doi.org/10.15252/emmm.202217144 (Jun. 2023).
74. Sazinsky, S. et al. Antibodies targeting human or mouse VSIG4 repolarize Tumor-Associated macrophages providing the potential of potent and specific clinical Anti-Tumor response induced across multiple Cancer types. *Int. J. Mol. Sci.* **25** (11). https://doi.org/10.3390/ijms25116160 (Jun. 2024).
75. Usenko, T. et al. Fraction of plasma exomeres and low-density lipoprotein cholesterol as a predictor of fatal outcome of COVID-19. *PLoS One.* **18, no.** https://doi.org/10.1371/journal.pone.0278083 (2023). 2 February, Feb.
76. Wei, X. et al. Hypolipidemia is associated with the severity of COVID-19. *J. Clin. Lipidol.* **14** (3), 297–304. https://doi.org/10.1016/j.jacl.2020.04.008 (May 2020).
77. Hu, X., Chen, D., Wu, L., He, G. & Ye, W. Declined serum high density lipoprotein cholesterol is associated with the severity of COVID-19 infection, *Clinica Chimica Acta*, vol. 510, pp. 105–110, Nov. (2020). https://doi.org/10.1016/j.cca.2020.07.015
78. Zinellu, A. et al. Cholesterol and COVID-19 Severity, and Mortality: A Systematic Review and Meta-Analysis With Meta-Regression., *Front Public Health*, vol. 9, p. 705916, Aug. (2021). https://doi.org/10.3389/fpubh.2021.705916
79. Hemmer, C. J., Löbermann, M. & Reisinger, E. C. [COVID-19: epidemiology and mutations: an update]. *Radiologe* **61** (10), 880–887. https://doi.org/10.1007/s00117-021-00909-0 (Oct. 2021).
80. Tao, K. et al. The biological and clinical significance of emerging SARS-CoV-2 variants. *Nat. Rev. Genet.* **22** (12), 757–773. https://doi.org/10.1038/s41576-021-00408-x (Dec. 2021).
81. Arya, R. et al. Structural insights into SARS-CoV-2 proteins. *J. Mol. Biol.* **433** (2), 166725. https://doi.org/10.1016/j.jmb.2020.11.024 (Jan. 2021).
82. Kumar, A. et al. Mar., SARS-CoV-2-specific virulence factors in COVID-19., *J Med Virol*, vol. 93, no. 3, pp. 1343–1350, (2021). https://doi.org/10.1002/jmv.26615
83. Rubio-Casillas, A., Redwan, E. M. & Uversky, V. N. SARS-CoV-2: A master of immune evasion. *Biomedicines* **10** (6). https://doi.org/10.3390/biomedicines10061339 (Jun. 2022).

Author contributions

I.V. and P.S. have written the main manuscript textI.V., T.U. and A.P. have written the pipeline code and analyzed the data to produce the resultsI.V., M.S., T.U., S.P. and P.S. actively participated in editing of the final manuscriptM.N., Y.P., E.G., I.A.and V.M. procured and curated the patient data.V.P. verified the results and verified the virus strainI.A., I.S. and Y.P. diagnosed and recruited the patients and procured the patient samplesP.S. and S.P. have planned the study designI.V. have drawn the figures and constructed the tablesS.P. have procured fundingM.S., S.P. and P.S. have administrated the project.

Declarations

Conflict of interest
The authors declare that they have no conflict of interest.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-00280-3.

Correspondence and requests for materials should be addressed to V.I.

Reprints and permissions information is available at www.nature.com/reprints.

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