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

Computers in Biology and Medicine 148 (2022) 105895

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Single cell gene expression profiling of nasal ciliated cells reveals distinctive biological processes related to epigenetic mechanisms in patients with severe COVID-19☆

Luis Diambra [a], Andres M. Alonso [b], Silvia Sookoian [c,d,*], Carlos J. Pirola [c,e,*]

[a] Centro Regional de Estudios Genómicos, Universidad Nacional de La Plata, CONICET, 1900 La Plata, Buenos Aires, Argentina
[b] Laboratorio de Parasitología Molecular, Instituto Tecnológico Chascomús (INTECH) CONICET-UNSAM, Chascomús, Argentina
[c] Universidad de Buenos Aires, Facultad de Medicina, Instituto de Investigaciones Médicas Alfredo Lanari, Ciudad Autónoma de Buenos Aires, Argentina
[d] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)– Universidad de Buenos Aires, Facultad de Medicina, Instituto de Investigaciones Médicas (IDIM), Departamento de Hepatología Clínica y Molecular, Ciudad Autónoma de Buenos Aires, Argentina
[e] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)– Universidad de Buenos Aires, Facultad de Medicina, Instituto de Investigaciones Médicas (IDIM), Departamento de Genética y Biología Molecular de Enfermedades Complejas, Ciudad Autónoma de Buenos Aires, Argentina

ARTICLE INFO

Keywords:
Single-cell transcriptomics
COVID-19
Histones
Methylation
SARS-CoV-2
Mitochondria

ABSTRACT

Objective: To explore the molecular processes associated with cellular regulatory programs in patients with COVID-19, including gene activation or repression mediated by epigenetic mechanisms. We hypothesized that a comprehensive gene expression profiling of nasopharyngeal epithelial cells might expand our understanding of the pathogenic mechanisms of severe COVID-19.
Methods: We used single-cell RNA sequencing (scRNAseq) profiling of ciliated cells (n = 12,725) from healthy controls (SARS-CoV-2 negative n = 13) and patients with mild/moderate (n = 13) and severe (n = 14) COVID-19. ScRNAseq data at the patient level were used to perform gene set and pathway enrichment analyses. We prioritized candidate miRNA-target interactions and epigenetic mechanisms.
Results: We found that mild/moderate COVID-19 compared to healthy controls had upregulation of gene expression signatures associated with mitochondrial function, misfolded proteins, and membrane permeability. In addition, we found that compared to mild/moderate disease, severe COVID-19 had downregulation of epigenetic mechanisms, including DNA and histone H3K4 methylation and chromatin remodelling regulation. Furthermore, we found 11-ranked miRNAs that may explain miRNA-dependent regulation of histone methylation, some of which share seed sequences with SARS-CoV-2 miRNAs.
Conclusion: Our results may provide novel insights into the epigenetic mechanisms mediating the clinical course of SARS-CoV-2 infection.

Key messages

- Severe COVID-19 was associated with altered gene expression pathways involved in mitochondrial function, membrane permeability, ciliogenesis, and downregulation of epigenetic mechanisms.
- Gene expression signatures mediated by miRNAs (either from the host, including miRNA-548 or the SARS-CoV2) may explain different clinical outcomes.
- Gene-drug interaction analysis showed diethylstilbestrol, trichostatin A, hydralazine hydrochloride, and chlorambucil as significantly predicted drugs associated with severe COVID-19.

1. Introduction

Cumulative and consistent evidence based on the records of patients with COVID-19 demonstrated that the disease course is characterized by a complex clinical spectrum, including asymptomatic infection-which is often not fully captured by cross-sectional epidemiological surveys-, mild symptoms, critical illness, and even the so-called long-COVID-19

☆ SS and CJP should be considered joint senior authors.

* Corresponding authors. Instituto de Investigaciones Médicas, IDIM-CONICET Combatientes de Malvinas 3150, CABA-1427, Argentina
E-mail addresses: ssookoian@intramed.net (S. Sookoian), pirola.carlos@conicet.gov.ar (C.J. Pirola).

https://doi.org/10.1016/j.compbiomed.2022.105895
Received 14 April 2022; Received in revised form 21 June 2022; Accepted 16 July 2022
Available online 30 July 2022

Abbreviations

BP	biological process
GSEA	Gene set enrichment analysis
GO	gene ontology
miRNA	microRNA
scRNA-seq	single-cell RNA sequencing

[1–4].

While progress has been made to understand the mechanism/s behind the evolution from mild to severe clinical outcomes, the molecular processes associated with the worst COVID-19 phenotypes, including severe pneumonia, acute respiratory distress syndrome, and even systemic organ failure, are not entirely understood.

SARS-CoV-2 disseminates by exposure to infectious respiratory fluids, and before symptoms onset, the viral load in cells of nasopharyngeal epithelium is usually high [5]. Therefore, the functional role of the nasal epithelial cells not only as the first physical barrier against viral entry but also in controlling the local immune response is of utmost importance in determining the disease course.

Single-cell RNA sequencing (scRNA-seq) profiling has allowed the identification of genes and molecular pathways associated with diverse human conditions, including pathways that define how a given cell can modify its phenotype after specific stimuli like a viral infection. Thus, scRNA-seq is a state-of-the-art approach for exploring changes in cellular regulatory programs, including activation or repression of genes that may explain the pathophysiology of human disease traits. For example, Ahn and co-workers recently demonstrated that SARS-CoV-2 is massively replicated in a distinct subset of human nasal epithelial cells [6]. Furthermore, by scRNA-seq and *in situ* mappings of SARS-CoV-2 entry–related host molecules, the investigators showed multiciliate cells but not secretory or basal cells of human nasal respiratory epithelium are the primary target for SARS-CoV-2 replication [6]. Ziegler et al., by performing elegant scRNA-seq profiling on nasopharyngeal cells from healthy and COVID-19 participants, revealed vital mechanisms that may define protective or detrimental responses to SARS-CoV-2 [7]. Specifically, the authors demonstrated impaired immune response and anti-viral immunity in cells of the nasal epithelium along 18 different clusters -from inflammatory macrophages to epithelial cell identities that underline severe COVID-19 [7]. Likewise, Ziegler and co-workers mapped the developmental transitions among nasal epithelial cells from healthy epithelium to the epithelium of patients with severe disease [7]. The authors also provided compelling evidence on the mechanisms of host-SARS-CoV-2 interaction in terms of phenotypic differences among cell subtypes and the anti-viral and interferon-mediated response curbing the intracellular levels of viral replication [7]. However, the molecular processes of gene transcriptional regulation in severe COVID-19, particularly those under the influence of miRNAs, have not been fully elucidated. Here, we leveraged scRNA-seq open data to expand the analysis of biological processes and gene regulation programs associated with severe COVID-19.

2. Methods

2.1. Biological samples

Our study comprises data from samples of the nasopharyngeal epithelium of 50 individuals from the University of Mississippi Medical Centre recruited between April and September 2020 [7]. The sample included 15 healthy subjects with negative SARS-CoV-2 PCR tests, denoted by healthy control (HC), and 35 individuals diagnosed with COVID-19, and nasopharyngeal swabs were collected within the first three days following admission to the hospital. Based on the World Health Organization (WHO) guidelines for COVID-19 severity stratification, these 35 patients with positive SARS-CoV-2 PCR tests were divided into two subgroups: 14 individuals with mild/moderate COVID-19 (WHO scores equal to 1–5, COVID-19 M), and 21 individuals with severe COVID-19 (WHO scores equal to 6–8, COVID-19 S). We have taken scRNAseq data from Single Cell Portal: https://singlecell.broadinstitute.org/.

2.2. Single-cell RNA-seq data analysis

We used the single-cell RNA-seq data from nasopharyngeal swabs publicly available for download and visualization via the Single Cell Portal: https://singlecell.broadinstitute.org/, under the identification SCP1289. Specifically, we focused our analyses on cells of the respiratory epithelium located in the nasal mucosa; the transcriptomic profile of 32,871 genes in 32,588 cells was obtained. Owing to the typical dropout of single cells, many of these genes have zero counts in many cells. The analysis of these data is a real challenge because the number of ciliated cells varies from patient to patient, and some cell types are rarely found in some individuals. Thus, we selected those patients with a representative number of ciliated cells. In this sense, only those patients with more than 45 ciliated cells were considered for the analysis.

In addition, our study considered only those genes with a non-null record in at least 650 cells (2% of all cells). Thus, with this restriction, the present study included 11,870 filtered genes. The original research by Ziegler et al. [7] classified cells into 18 cell types (basal cells, B cells, ciliated cells, dendritic cells, deuterosomal cells, developing ciliated cells, developing secretory and goblet cells, enteroendocrine cells, erythroblasts, goblet cells, ionocytes, macrophages, mast cells, mitotic basal cells, plasmacytoid DCs, secretory cells, squamous cells, T cells). In the present study, we only considered the groups of ciliated cells, which include “ciliated cells,” “deuterosomal cells,” and “developing ciliated cells,” representing a set of 12,725 cells. Unlike previous studies where gene expressed genes are determined considering “each cell” as an independent sample or replicate, we thought “each individual or patient” as an independent replicate. Thus, we used the arithmetic mean of normalized gene expression data of ciliated cell for each patient. Specifically, we calculated the mean expression level of each gene in each individual by averaging the normalized counts of all cells of the specific ciliated cell type belonging to that individual. In this manner, we avoided inflating statistical differences as the whole single-cell data taking individual cells as biological replicates might result in an overestimation of significant gene expression changes. The number of cells of a given type varies from patient to patient; some are rare in some individuals. Only those patients with more than 45 ciliary cells were considered in this study.

2.3. Pathway enrichment analysis and data visualization

The scRNA-seq data derived from the different groups of patients were contrasted in three different ways: 1) functional analysis of biological processes (BPs) comparing HC vs. COVID-19 M, and COVID-19 S vs. COVID-19 M using GSEA, as described below; 2) analysis of regulatory target gene sets associated with miRNAs contrasting COVID-19 S vs. COVID-19 M using GSEA; and 3) gene-set enrichment tools by the ToppGene Suite available https://toppgene.cchmc.org/.

First, we analysed the GO (gene ontology) terms for BPs associated with gene expression changes using GSEA (Gene Set Enrichment Analysis software, https://www.gsea-msigdb.org) [8]). GSEA is a computational method that determines whether a ranked gene list shows statistically significant, concordant differences between two phenotypes. Here, the rank of genes was established by the Signal2Noise metric, which is defined as the difference between the means expression in each phenotypic class divided by the sum of the deviations as implemented in the GSEA platform for dichotomic phenotypes.

More specifically, we started with the normalized counts as a mea-

sure of the expression level of genes for all cells. First, we selected the ciliated cells of each individual. We considered individuals whose number of ciliated cells was greater than 45 cells. We calculated the mean expression level of gene i in this individual, E_i^j, by averaging the expression level of all its ciliated cells, that is:

$$E_i^j = N_j^{-1} \sum_k^{N_j} e_{ik}^j,$$

where e_{ik}^j is the normalized counts of gene i in the cells k belonging to individual j and N_j is the number of ciliated cells for individual j. This define an average expression profile of ciliated cells for each patient, that will be used for dichotomic comparisons in further steps by means of GSEA.

The matrix E (genes x patients) feeds the GSEA software to perform the pathway enrichment analysis. The rank of genes is established by GSEA for each dichotomic phenotype by using the Signal2Noise metric (other metrics are also available in GSEA), which is defined as:

$$Signal2Noise = \frac{\mu^A - \mu^B}{\sigma^A + \sigma^B},$$

where μ^A and σ^A are the mean and standard deviation, respectively, of gene expression levels observed in the phenotypic class A. More specifically for a gen i

$$\mu_i^A = N_A^{-1} \sum_j^{N_A} E_i^j,$$

$$\sigma_i^A = \sqrt{\frac{\sum \left(E_i^j - \mu_i^A\right)^2}{N_A - 1}},$$

where N_A is the number of individuals in the condition A.

We performed the comparisons considering a subset of the C5 gene set (version 7.4, from MSigDB collection) that includes the BP ontology and a gene set size of 10–500 genes. GO sets are based on ontologies and do not necessarily comprise co-regulated genes. In all cases, we used 1000 permutations over genes to compute the statistical significance of enrichment scores. Next, we performed a cluster analysis of significant BPs by Enrichment Map Cytoscape App 3.3 [9] to obtain a BP network, where nodes represent BPs and edges represent gene overlap between gene sets associated with the connected nodes. Edges with a similarity score lower than 0.5 did not included in the BP network plot. Network

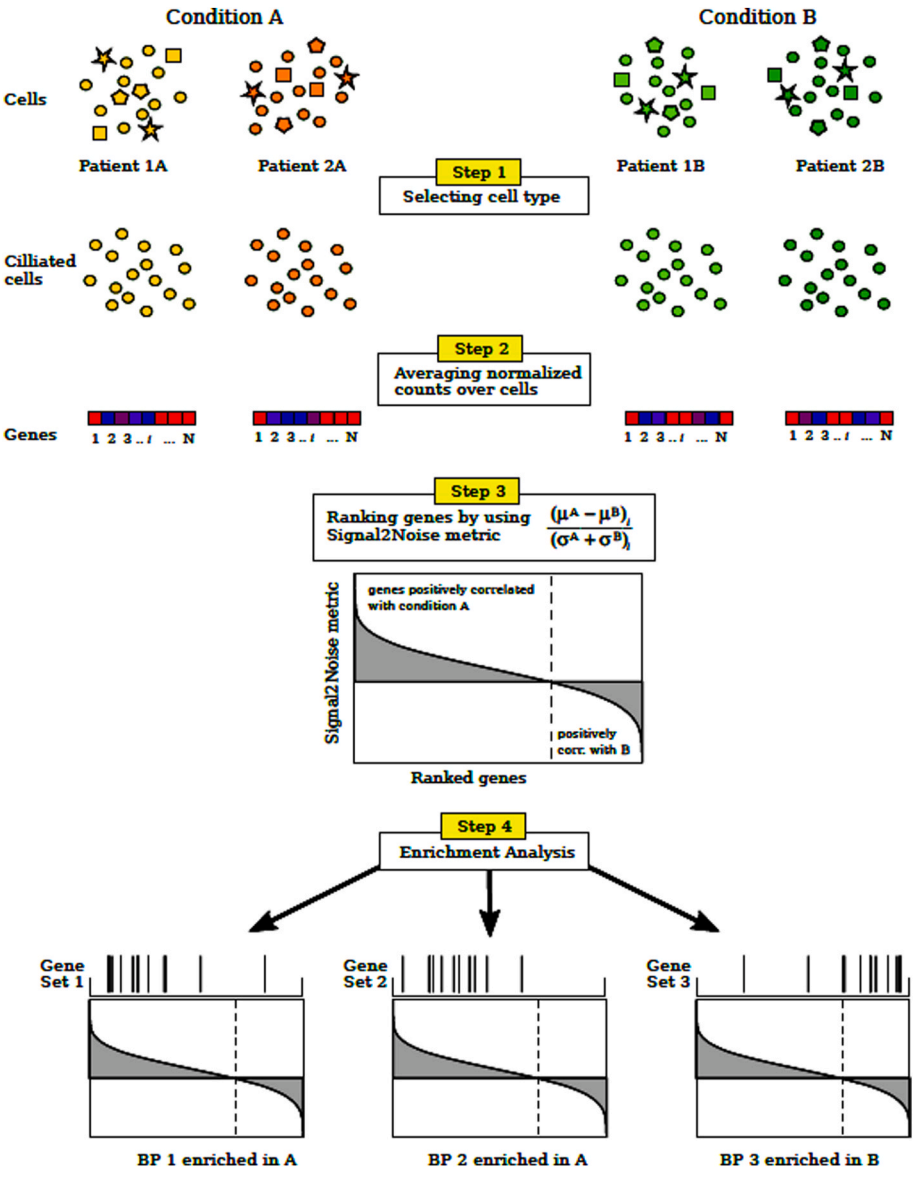

Fig. 1. Methodology flow chart
Step 1: The specific type of cells is selected from each patient (ciliated cells in our case).
Step 2: Normalized counts for each gene are averaged (arithmetic mean) over the selected cells population of each patient, obtaining the average expression profile of ciliated cells for each patient.
Step 3: Signal2Noise metric is computed among the patients of the two conditions of interest. μA(μB) and σA(σB) are the mean and standard deviation of gene expression levels computed in the previous step for patients under condition A (B). Then, the genes are ranked by decreasing order of the metric value.
Step 4: Finally, it is determined if the ranked gene set, associated with a given biological process (BP), shows statistically significant, concordant differences between the two conditions. This last step can be performed with different BPs or even with gene set associated with molecular functions or miRNAs.

clustering and annotation were performed by MCL (Markov Cluster algorithm), and Word Cloud (normalization factor 0.5) using the Cytoscape plug-in auto annotate [10].

Second, we used the GSEA algorithm to prioritize candidate miRNA-target interactions in patients with COVID-19 across the two primary clinical outcomes of interest, including patients with severe vs. patients with mild/moderate disease. To this end, we used the subset of the C3 gene sets (version 7.4, from MSigDB collection) that includes miRNA-targets genes as catalogued by the miRDB v6.0 algorithm [11]. We considered only gene sets with a size of 300–1000 genes.

Finally, we performed gene list functional enrichment of a selected input list (see the results section for further details) to predict gene-drug interactions and co-expression analysis that was compiled from various data sources and used to a build training set gene profile by the ToppGene/ToppCluster resource (https://toppgene.cchmc.org/) [12,13]. The same platform was used to interpret BPs, molecular functions, family genes, and human traits associated with particular subsets of genes or prioritize putative viral miRNA targets.

A detailed flow chart to explain all steps of the methodology is shown in Fig. 1.

3. Results

We found that the number of cells of a given type varied considerably from patient to patient, and some cell types are rare in some individuals. Therefore, we only considered those patients that presented more than 45 ciliary cells. Consequently, we included 13 healthy controls that were SARS-CoV-2 negative by PCR, 13 patients with mild/moderate COVID-19, and 14 patients with severe COVID-19.

3.1. Gene set enrichment analysis associated with COVID-19 across all levels of the disease severity identifies distinctive biological processes

As explained earlier, we performed gene set enrichment analysis (GSEA) of BPs associated with SARS-CoV-2 infection in cells of the respiratory epithelium located in the nasal mucosa. Hence, we analysed a total number of 12,725 ciliated cells. We first compared BPs in infected patients presenting mild/moderated COVID-19 versus control (healthy) subjects. Besides the GO terms associated with innate immune response and TNF-mediated signalling, genome replication and viral release, viral budding and symbiont-host interaction, granulocyte migration, and lymphocyte homeostasis that were already described by Ziegler et al. [7], we found some other BPs and or sub-networks that deserve further analysis (Fig. 2). For example, we found that sub-networks associated with mitochondrial membrane permeabilization, protein targeting to mitochondria, and lysosome and protein targeting to the membrane were significantly enriched in mild/moderate COVID-19. In addition, sub-networks related to protein folding and misfolding are highlighted. This last observation could be related to a previous characteristic described for SARS-CoV2 codon composition of highly expressed genes, which could influence the tRNAs pools of the infected cells with relevant implications in host protein folding and expression [14]. The entire list of BPs, including the GO terms, is shown in Supplementary Table 1.

We next asked whether the differential pattern of BPs observed in ciliated cells may distinguish mild and moderate cases from severe COVID-19. The most relevant sub-networks identified by GSEA are shown in Fig. 3. Compared to mild/moderate COVID-19 infection, patients with the severe disease showed an overrepresentation of sub-networks linked to fatty acid metabolism and eicosanoid synthesis, acute inflammatory processes, keratinocyte differentiation, and humoral immune response, which were all described by Ziegler et al. [7]. However, we noted that cells derived from severe COVID-19 patients had some remarkable underrepresented sub-networks linked to

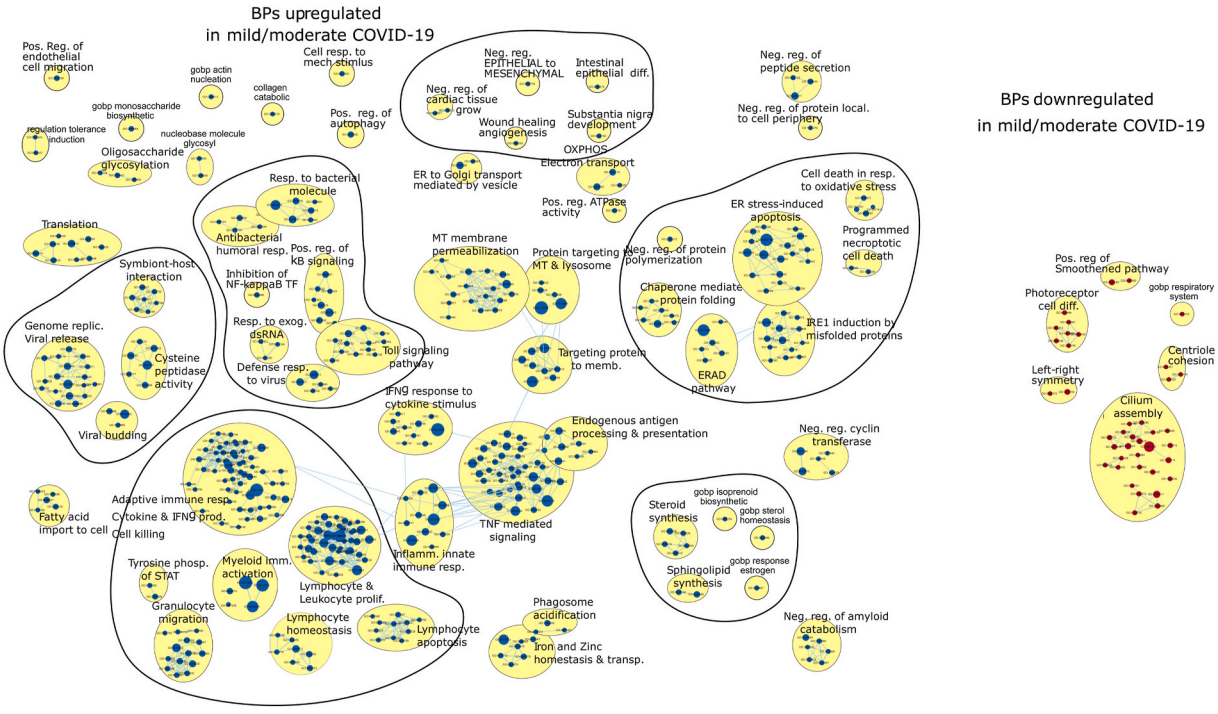

Fig. 2. Gene set enrichment analysis derived from ciliated cells in mild/moderate COVID-19 patients vs. healthy controls
The Figure illustrates a network of GO (Gene Ontology) terms corresponding to biological processes (BPs) significantly enriched in genes upregulated (blue nodes) and downregulated (red nodes) in mild-moderate COVID-19 (COVID-M) patients concerning to healthy control (HC) individuals (Bonferroni adjusted p-value, FDR 0.05). The node's size is proportional to the number of genes associated with that BP. The edges represent gene overlap between gene sets related to different GO terms. Connected nodes are organized in clusters of interconnected BPs obtained by the MCL algorithm, which considers similarity among gene sets to assign the edges, with a similarity score threshold of 0.5.

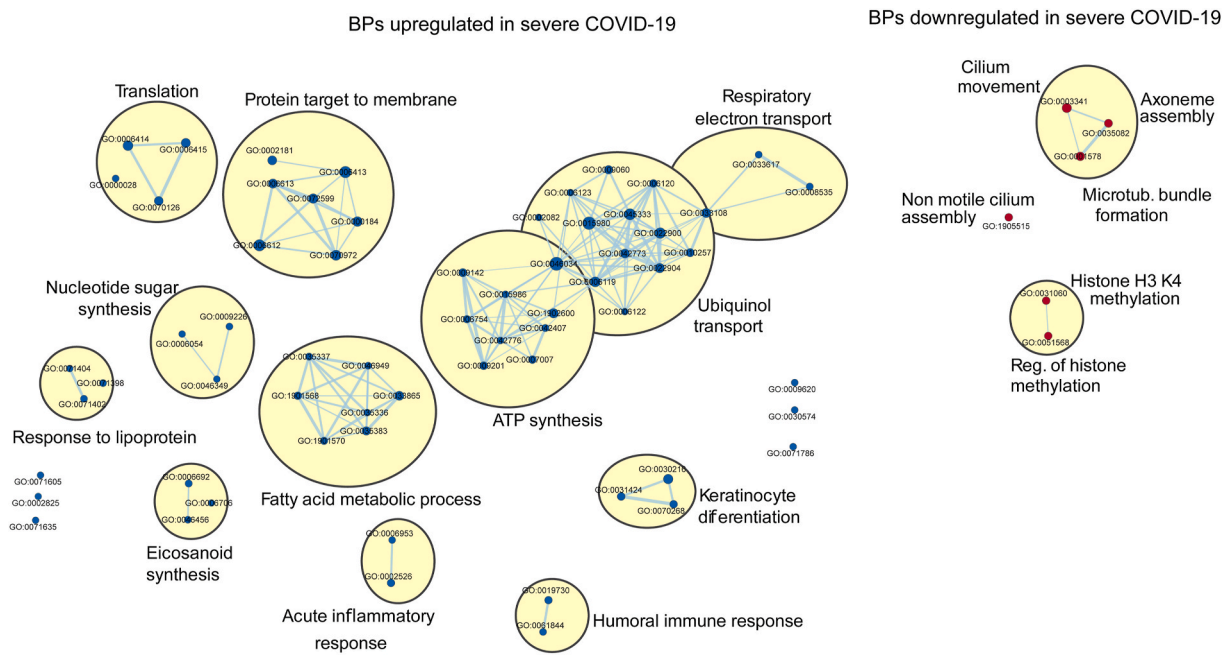

Fig. 3. Gene set enrichment analysis derived from ciliated cells in severe vs. mild/moderate COVID-19 patients
The Figure illustrates enrichment analysis of biological processes (BPs) upregulated (blue nodes) and downregulated (red nodes) in patients with severe COVID-19 vs. patients with mild/moderate disease(Bonferroni adjusted p-value, FDR 0.05). BPs were clustered by using the MCL algorithm with a similarity score threshold of 0.5.

epigenetic mechanisms, including histone H3K4 methylation and regulation of histone methylation (Fig. 3). In addition, GO terms associated with ciliogenesis (GO:0003341, GO:0035082, GO:1905515, and GO:0001578) were significantly downregulated in severe COVID-19 (Fig. 3). The entire list of BPs, including the GO terms, is shown in Supplementary Table 2.

3.2. *Gene expression profile of cells derived from the nasal mucosa in patients with COVID-19 across diverse stages of disease severity: sub-networks analysis focused on organelle function*

To gain further insight into the gene lists generated by the GSEA analysis, we focus on specific pathways connected with impaired organelle function. Specifically, we selected the pathways of mitochondrial membrane permeabilization, protein targeting to mitochondria, and lysosome and protein targeting to the membrane, in which transcript levels of most upregulated genes in COVID-19 M patients concerning to control subjects are in the percentile 0.10. Then, we

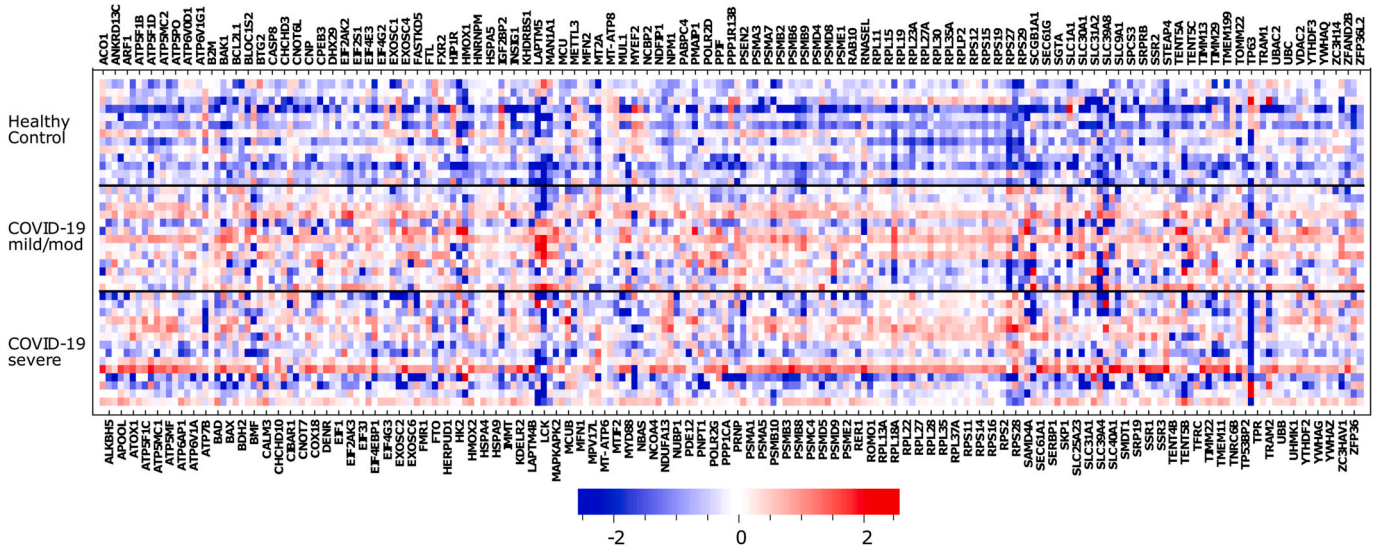

Fig. 4. Heat map of expression genes belonging to the pathways associated with mitochondrial membrane permeabilization, protein targeting to mitochondria, and lysosome and protein targeting the membrane.
Blue squares correspond to downregulated genes and red squares to upregulated genes. We selected only the subset of genes most upregulated (the percentile 0.1 of the ranked genes)in mild/moderate COVID-19 vs. healthy patients. The colour scale represents the RNA abundance relative to the media of transcript levels in all subjects and then log2 transformed.

compared the expression profile of these genes within each experimental group, including HC, COVID-19 M, and COVID-19 S; transcript levels in each group of subjects were referred to the media of all samples. A heat map of the results appears in Fig. 4. Including the COVID-19 S group, the analysis shows significant downregulation of specific genes in cells derived from COVID-19 patients with severe disease. For instance, the expression of *LCK* (Lymphocyte Cell-Specific Protein-Tyrosine Kinase, of which the encoded protein is a key signalling molecule in the selection and maturation of developing T-cells), *TP63* (a member of the p53 family of transcription factors), *TRAM1* (Translocation Associated Membrane Protein 1), and *SLC40A1* (Solute Carrier Family 40 Member 1) were significantly decreased in severe COVID-19. Upon inspecting the expression pattern in members of the ribosomal protein family (RPL and RPS) that encode structural constituents of ribosomes, we found that patients with COVID-19 had significant upregulation regardless of the disease severity. Interestingly, these ribosomal protein members are associated with other viral diseases, particularly influenza [15].

3.3. The regulatory network of miRNAs associated with severe COVID-19

We next sought to explore the putative regulatory mechanisms associated with more severe illnesses. Specifically, we asked whether miRNA-mediated gene expression deregulation may explain severe COVID-19. Hence, using the ranked list of genes obtained by GSEA from the comparison of COVID-19 S vs. COVID-19 M as input, we detected 29 miRNAs potentially affecting the disease severity.

Some of these miRNAs share the same so-called seed region, which ultimately defines the mRNA target, and in that case, have the same gene targets. Therefore, we curated the list of miRNAs by an in-house script to avoid redundancy in the subsequent analyses. The correlations among the targets of the 29 miRNAs mentioned above are depicted in Supplementary Fig. 1. Upon inspection of redundancy, 11 ranked miRNAs (miR-16-5p, miR-130-3p, miR548h-3p, miR-1283, miR548an, miR-19a-3p, miR144-3p, miR9983, miR101-3p, miR181a-5p, and miR8485) were used in the search for their targets among the whole set of putative miRNA target genes of the human genome (hsa-miRNA targets as catalogued by the miRDB v6.0 algorithm). As a result, we found a total of 3076 target genes. The complete list of target genes associated with the 11 miRNAs is shown in Supplementary Table 3.

In parallel, we selected the 10% top-ranked genes downregulated in COVID-19 S vs. COVID-19 M according to the ranking based on the Signal2Noise metric, which yielded 1167 genes (Fig. 5a). From the intersection of the 11 miRNAs-targeted genes (3076) and the above described 1167 downregulated genes in severe COVID-19, we found 336 deregulated transcripts also targets of the 11 miRNAs (Fig. 5b). The expression pattern of these 336 genes within each group of patients relative to the media of all groups (healthy controls, COVID-19 M, and COVID-19 S) is shown in Fig. 5c. A comprehensive network of the BPs

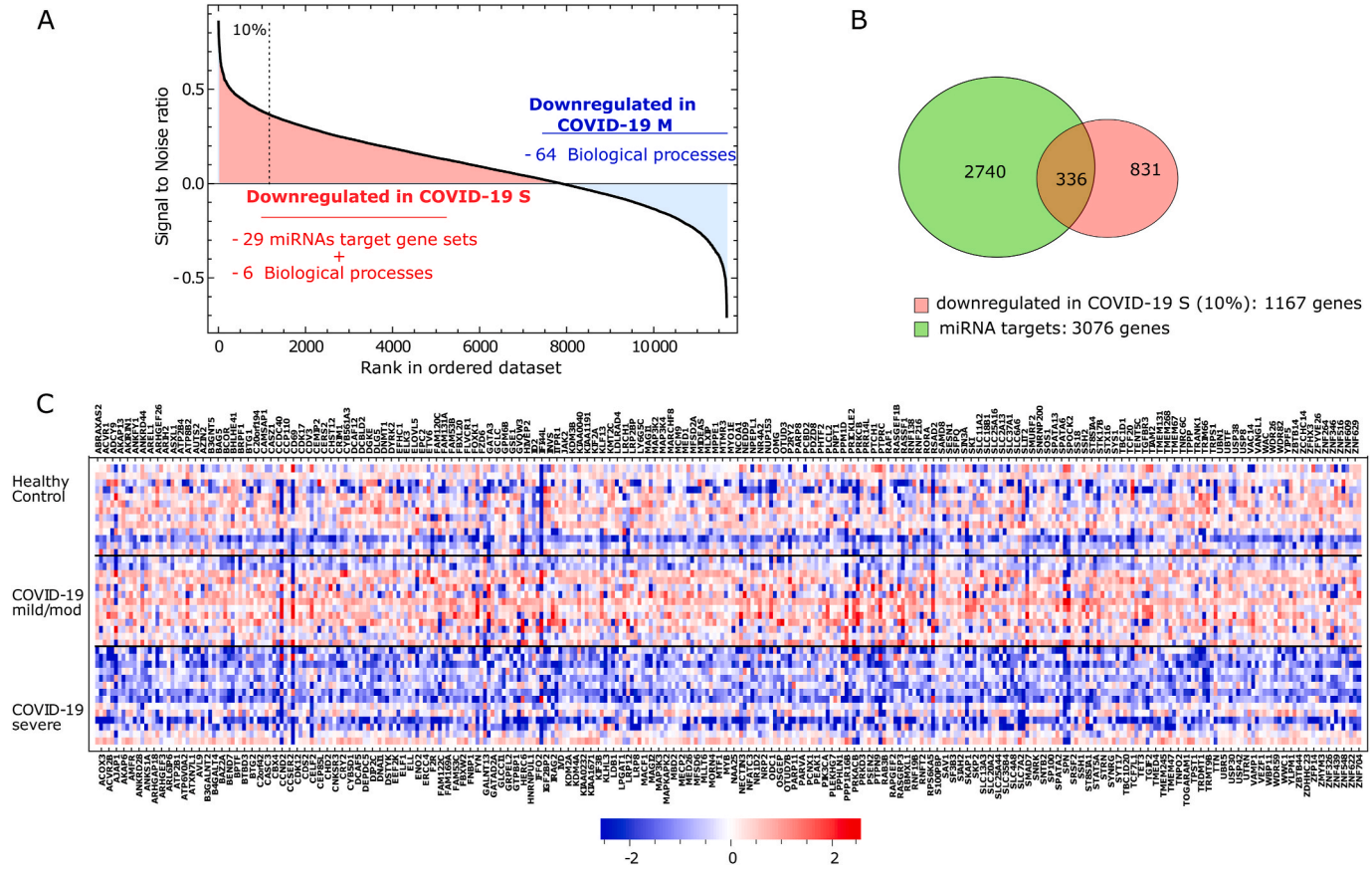

Fig. 5. The role of epigenetic mechanisms in COVID-19: a regulatory network of miRNAs-target genes

A. Rank of genes according to the Signal2Noise metric as implemented in the GSEA platform (COVID-19 S vs. COVID-19 M) using the media of the normalized counts of the whole cell population in each individual. The top 10% of downregulated genes (rank over the dashed line) was used in the subsequent analysis shown in panels B and C.

B. Venn Diagram showing the overlap between the 3076 gene targets of the 11 miRNAs associated with the target sets enriched in COVID-19 S vs. COVID-19 M and the 1167 top-ranked genes downregulated in COVID-19 S (as explained in panel A). The intersection shows the 336 genes used in panel C.

C. A heat map of the 336 genes that result from the intersection mentioned in panel B. Squares in blue and red represent genes with lower and higher RNA abundance relative to the media of transcript level of all subjects and then log2 transformed, respectively.

COVID-19 S: severe disease; COVID-19 M: mild/moderate disease.

represented by these downregulated genes in patients with severe COVID-19 is shown in Fig. 6.

We then focused our analysis on those miRNA-target pairs involved in epigenetic mechanisms, including BPs linked to H3K4 methylation and regulation of histone methylation. These 11 miRNAs together regulate most of the histone methylation pathway genes that were downregulated in severe COVID-19 (*MYB*, BCOR, *DNMT1*, *KMT2C*, *TET3*, *KMT2A*, *KMT2C*, *WDR82*, *GATA3*, *AUTS2*, *MECP2*, *and MLLT6*) (Supplementary Fig. 2).

3.4. Gene list enrichment analysis based on functional annotations: insights into drug-target and co-expression functional enrichments

We finally sought to explore potential druggable targets in the list of differentially expressed miRNA-target genes associated with epigenetic mechanisms.

The transcripts in the input list (*MYB*, BCOR, *DNMT1*, *KMT2C*, *TET3*, *KMT2A*, *KMT2C*, *WDR82*, *GATA3*, *AUTS2*, *MECP2*, *and MLLT6*) were then screened for gene-drug interactions using the ToppGene Suite resource.

We found several compounds as putative ligands of the genes enriched in the input list; however, some specific drugs were ranked among the top ones. These drugs included diethylstilbestrol (FDR = 2.904E-2), trichostatin A (FDR = 1.902E-2), hydralazine hydrochloride (FDR = 1.902E-2), and chlorambucil (FDR = 1.902E-2). In addition, we found that a large proportion of the predicted tissue co-expression of the target genes was mainly located in cells of the immune system, including alpha-beta T cells and gamma-delta T cells of the pancreas, thymus, lymph node, spleen, and CD4 positive of the spleen (FDR 1.209E-2). Furthermore, by searching for clinical trial targets' in the Therapeutical target database (TTD, http://db.idrblab.net/ttd/) [16] we found that *MYB* (LR3001, antisense anticancer drug in phase 2), *DNMT1*

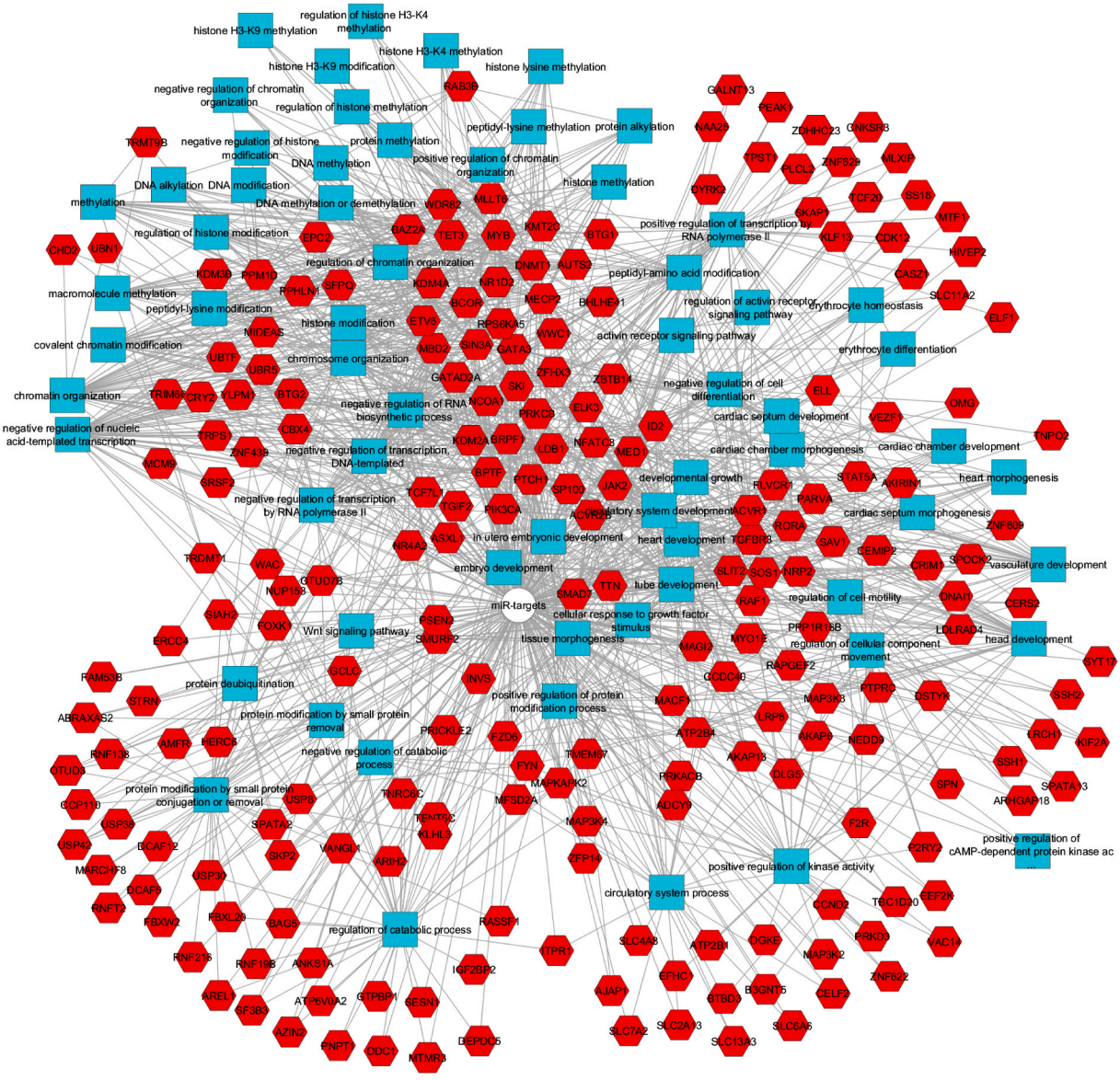

Fig. 6. Network of the 336 genes downregulated in COVID-S and their associated GO terms for biological processes (BP)
The network was constructed using the ToppGene/ToppCluster resource (https://toppgene.cchmc.org/) with an FDR of 0.05. Hexagons in red and squares in light blue stand for genes and BPs, respectively, which were distributed using the edge-weighted Spring embedded layout and slightly modified for readability using Cytoscape v3.4.0. In general, the upper left region shows BP associated with DNA and histone modifications, regulation of transcription, and chromatin or chromosome remodelling. In contrast, the right middle part shows BPs associated with organ morphogenesis, particularly cardiac and vascular. The lower left region shows metabolic processes, primarily catabolic.

(S-110, DNA demethylating agent, phase 2/3), and *KMT2A* (KO-539, Menin-mll interaction inhibitor, phase 1/2), have specific interacting drugs in diverse clinical trial phases.

4. Discussion

This study identified distinctive biological processes related to epigenetic mechanisms in patients with severe COVID-19. Specifically, we analysed scRNA-seq profiles from nasopharyngeal ciliated cells of patients with COVID-19 and performed pathway enrichment analysis. Our research was based on two distinctive strategies, including the analysis of scRNA-seq data at an individual-patient level rather than at the bulk cell-level and the prioritization of molecular mechanisms associated with severe COVID-19.

Patients with severe COVID-19 presented deregulation of gene expression in pathways linked to mitochondrial function and membrane permeability. The expression of certain genes, including *LCK*- a tyrosine kinase involved in T-cells maturation and expansion, the tumour encoding protein *TP63* that plays a role in the regulation of epithelial morphogenesis, *TRAM1* that encodes for a multi-pass membrane protein that is part of the mammalian endoplasmic reticulum, and *SLC40A1* that mediates iron ion transmembrane transporter activity-among other transcripts, were consistently decreased across ciliated cells from all severe COVID-19 patients. Collectively, these findings reinforce the importance of ciliated cells differentiation, maturation, and immune response in these specific cells of the nasopharyngeal epithelium as highlighted by Ziegler et al. [7]. In addition to the relevance of mitochondrial dynamics, impairment of mitochondrial membrane potential, and cell death in any viral infection, including SARS-CoV-2 infection, one may speculate that some other molecular mechanisms are relevant to determine the worst disease outcome. For instance, Singh et al. reported that SARS-CoV-2 interferes with mitochondrial function to evade host cell immunity and facilitate virus replication and COVID-19 development [17]. Likewise, disturbances in iron transport may affect mitochondrial function and induce oxidative stress. Earlier reports demonstrated that hyperferritinemia is a predictor of increased mortality of the disease [18]. Besides, ferroptosis, an iron-related cell death program, may be involved in COVID-19-induced multiple organ failure [19].

Furthermore, our results highlight the importance of ciliated cells in tissue homeostasis. Nasopharyngeal ciliated cells of patients with severe COVID-19 presented significant downregulation of pathways associated with cilium movement, microtubule bundle formation, and axoneme assembly. This finding has implications for understanding impaired physiological processes of the nasal mucosa, for instance, olfaction-related disorders and systemic functioning deregulation of ciliated cells across the body, including primary cilia in testicular cells. Recent clinical evidence suggests that the spermatogenic function of the testis is impaired in patients with COVID-19 [20]. It is then plausible to speculate that the dysfunction of ciliated cells in severe COVID-19 underlies pleiotropic effects in many different tissues and organs of the body. Here, we found that in this phenomenon, the dynein protein family may play a crucial role, a finding already reported by others [21]. Therefore, the associated signalling pathways resemble the syndromic disorders known as ciliopathies.

We also noted that cells derived from severe COVID-19 patients had downregulation of sub-networks linked to epigenetic mechanisms, including histone H3K4 methylation and regulation of histone methylation. As shown in Fig. 6, genes in the pathways of histone modifications, histone lysine methylation, DNA methylation or demethylation, DNA modification, chromatin remodelling, chromosome organization, negative regulation of RNA biosynthetic process, negative regulation of the cellular biosynthetic process, and even cardiac morphogenesis -among many other pathways, are all dramatically downregulated in cells derived from severe COVID-19 patients. Some genes in this network belong to the Methyl-CpG binding domain-containing NuRD complex, Zinc fingers C2HC-type PHD finger proteins, Lysine acetyltransferases, and Zinc fingers CXXC-type gene families. Collectively, these findings might explain the differences in developmental and maturation trajectories of ciliated cells in severe COVID-19 patients reported by Ziegler and co-workers [7].

To expand the understanding of the potential molecular processes involved in the regulation of gene expression, we examined the enrichment of miRNAs in the list of genes deregulated in cells of patients with severe vs. mild/moderate COVID-19. We found some miRNA-gene-target pairs that might be relevant to the disease biology. For example, we highlight the miRNA-548, which has been shown to down-regulate host anti-viral response via direct targeting of IFN-λ1 [22]. Notably, miRNA-548 seems to participate in numerous signalling pathways, such as the Wnt, the MAPK and TGF-β pathways, and regulation of the immune system in the transition from immune tolerance to immune activation in chronic hepatitis B [23]. Notably, predictions of putative small ligands of genes enriched in the input list of predicted miRNAs showed some interesting gene-drug interactions that might explain differences in clinical outcomes mediated by epigenetic mechanisms or help treat the disease. For example, we predicted a significant interaction between five of the target genes of the list of enriched miRNAs with diethylstilbestrol. It appears that oestrogen treatment silences the inflammatory reactions and decreases virus titters leading to an improved survival rate [24]. Previous experimental studies showed that 17β-oestradiol protects females against influenza by recruiting neutrophils and increasing virus-specific CD8 T cell responses in the lungs [25]. Perets and colleagues showed that 17β-oestradiol reduces influenza A virus replication in primary human nasal epithelial cells derived from females [26]. Some other gene-drug target predictions suggest that compounds used to treat haematological conditions, such as leukaemia and lymphomas, or the antihypertensive agent hydralazine hydrochloride could be tested for drug repurposing in the treatment of severe COVID-19.

As a putative mechanistic explanation, we observed some overlap between the seeds of gene expression changes associated with miRNAs and those of the recently reported SARS-CoV-2-encoded viral miRNAs [27–29], as shown in Supplementary Table 4. Thus, these findings could indicate that SARS-CoV-2 has miRNAs-like sequences [29], and -like many other viruses-, hitchhikes the regulatory processes of the host cells by targeting epigenetic modulators.

We are aware that our research may have limitations. First, the nature of retrieved metadata and the limited number of patients do not allow us to examine relevant demographic information of the infected patients. Therefore, we did not examine the correlation between the patients' features, such as age and gender, with deregulated gene expression pathways. Second, the analysed dataset contains scRNA-seq data from the nasopharyngeal swabs collected within the first three days following admission to the hospital [7], which does not allow any assessment of patient's follow-up. Hence, further validation of our observations using experimental models, either in cell lines or animal models is required.

Together, our results may provide novel insights into the cellular and molecular mechanisms that modulate the clinical course of SARS-CoV-2 infection. Based on predicted biological processes and functional enrichment analysis, we detected some relevant mechanisms that may explain the pathogenesis of severe COVID-19. Among these molecular explanations, we found some clues on the importance of post-transcriptional regulation of gene expression by small-non coding RNAs and epigenetic mechanisms of gene transcription regulation, such as DNA and histone methylation and ultimately chromatin remodelling.

Authors contribution

All authors should have made substantial contributions to all of the following: (Diambra L, Sookoian S, Pirola CJ) the conception and design of the study, acquisition of data, analysis and interpretation of data; (Diambra L and Alonso A) scRNA-seq data acquisition and analysis;

(Diambra L, Sookoian S, Pirola CJ) drafting the article; (Diambra L, Alonso A, Sookoian S, Pirola CJ) final approval of the version to be submitted.

Data availability

Data supporting the findings of this study are available within the article and Supplementary Information files. Derived data supporting the methodology of this study are available from the corresponding authors on request.

Declaration of competing interest

None Declared

Acknowledgment

This study was partially supported by grants PICT 2018-889 and PICT 2019-0528 (SS), PICT 2016-0135 and PICT 2018-0620 (CJP), PICT 2018–03713 (LD) (Agencia Nacional de Promoción Científica y Tecnológica, FONCyT), and CONICET Proyectos Unidades Ejecutoras 2017, PUE 0055 (SS, CJP).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2022.105895.

References

[1] B. Blomberg, K.G. Mohn, K.A. Brokstad, F. Zhou, D.W. Linchausen, B.A. Hansen, S. Lartey, T.B. Onyango, K. Kuwelker, M. Saevik, H. Bartsch, C. Tondel, B. R. Kittang, R.J. Cox, N. Langeland, Long COVID in a prospective cohort of home-isolated patients, Nat. Med. 27 (2021) 1607–1613, https://doi.org/10.1038/s41591-021-01433-3.

[2] W.J. Guan, Z.Y. Ni, Y. Hu, W.H. Liang, C.Q. Ou, J.X. He, L. Liu, H. Shan, C.L. Lei, D. S.C. Hui, B. Du, L.J. Li, G. Zeng, K.Y. Yuen, R.C. Chen, C.L. Tang, T. Wang, P. Y. Chen, J. Xiang, S.Y. Li, J.L. Wang, Z.J. Liang, Y.X. Peng, L. Wei, Y. Liu, Y.H. Hu, P. Peng, J.M. Wang, J.Y. Liu, Z. Chen, G. Li, Z.J. Zheng, S.Q. Qiu, J. Luo, C.J. Ye, S. Y. Zhu, N.S. Zhong, Clinical characteristics of coronavirus disease 2019 in China, N. Engl. J. Med. (2020), https://doi.org/10.1056/NEJMoa2002032.

[3] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet 395 (2020) 497–506, https://doi.org/10.1016/S0140-6736(20)30183-5.

[4] J.T. Wu, K. Leung, M. Bushman, N. Kishore, R. Niehus, P.M. de Salazar, B. J. Cowling, M. Lipsitch, G.M. Leung, Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China, Nat. Med. 26 (2020) 506–510, https://doi.org/10.1038/s41591-020-0822-7.

[5] E.A. Meyerowitz, A. Richterman, R.T. Gandhi, P.E. Sax, Transmission of SARS-CoV-2, Ann. Intern. Med. 174 (2021) 1037, https://doi.org/10.7326/L21-0166.

[6] J.H. Ahn, J. Kim, S.P. Hong, S.Y. Choi, M.J. Yang, Y.S. Ju, Y.T. Kim, H.M. Kim, M. D.T. Rahman, M.K. Chung, S.D. Hong, H. Bae, C.S. Lee, G.Y. Koh, Nasal ciliated cells are primary targets for SARS-CoV-2 replication in the early stage of COVID-19, J. Clin. Invest. 131 (2021), https://doi.org/10.1172/JCI148517.

[7] C.G.K. Ziegler, V.N. Miao, A.H. Owings, A.W. Navia, Y. Tang, J.D. Bromley, P. Lotfy, M. Sloan, H. Laird, H.B. Williams, M. George, R.S. Drake, T. Christian, A. Parker, C.B. Sindel, M.W. Burger, Y. Pride, M. Hasan, G.E. Abraham III, M. Senitko, T.O. Robinson, A.K. Shalek, S.C. Glover, B.H. Horwitz, J. Ordovas-Montanes, Impaired local intrinsic immunity to SARS-CoV-2 infection in severe COVID-19, Cell 184 (2021) 4713–4733, https://doi.org/10.1016/j.cell.2021.07.023.

[8] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 15545–15550, https://doi.org/10.1073/pnas.0506580102.

[9] D. Merico, R. Isserlin, O. Stueker, A. Emili, G.D. Bader, Enrichment map: a network-based method for gene-set enrichment visualization and interpretation, PLoS One 5 (2010), e13984, https://doi.org/10.1371/journal.pone.0013984.

[10] M. Kucera, R. Isserlin, A. Arkhangorodsky, G.D. Bader, AutoAnnotate: A Cytoscape App for Summarizing Networks with Semantic Annotations, vol. 5, 2016, p. 1717, https://doi.org/10.12688/f1000research.9090.1. F1000Res.

[11] Y. Chen, X. Wang, miRDB: an online database for prediction of functional microRNA targets, Nucleic Acids Res. 48 (2020) D127–D131, https://doi.org/10.1093/nar/gkz757.

[12] J. Chen, E.E. Bardes, B.J. Aronow, A.G. Jegga, ToppGene Suite for gene list enrichment analysis and candidate gene prioritization, Nucleic Acids Res. 37 (2009) W305–W311, https://doi.org/10.1093/nar/gkp427.

[13] V. Kaimal, E.E. Bardes, S.C. Tabar, A.G. Jegga, B.J. Aronow, ToppCluster: a multiple gene list feature analyzer for comparative enrichment clustering and network-based dissection of biological systems, Nucleic Acids Res. 38 (2010) W96–W102, https://doi.org/10.1093/nar/gkq418.

[14] A.M. Alonso, L. Diambra, SARS-CoV-2 codon usage bias downregulates host expressed genes with similar codon usage, Front. Cell Dev. Biol. 8 (2020) 831, https://doi.org/10.3389/fcell.2020.00831.

[15] A.C. Becker, M. Gannage, S. Giese, Z. Hu, S. Abou-Eid, C. Roubaty, P. Paul, L. Buhler, C. Gretzmeier, V.I. Dumit, S. Kaeser-Pebernard, M. Schwemmle, C. Munz, J. Dengjel, Influenza A virus induces autophagosomal targeting of ribosomal proteins, Mol. Cell. Proteomics 17 (2018) 1909–1921, https://doi.org/10.1074/mcp.RA117.000364.

[16] Y. Wang, S. Zhang, F. Li, Y. Zhou, Y. Zhang, Z. Wang, R. Zhang, J. Zhu, Y. Ren, Y. Tan, C. Qin, Y. Li, X. Li, Y. Chen, F. Zhu, Therapeutic target database 2020: enriched resource for facilitating research and early development of targeted therapeutics, Nucleic Acids Res. 48 (2020) D1031–D1041, https://doi.org/10.1093/nar/gkz981.

[17] K.K. Singh, G. Chaubey, J.Y. Chen, P. Suravajhala, Decoding SARS-CoV-2 hijacking of host mitochondria in COVID-19 pathogenesis, Am. J. Physiol. Cell Physiol. 319 (2020) C258–C267, https://doi.org/10.1152/ajpcell.00224.2020.

[18] P. Mehta, D.F. McAuley, M. Brown, E. Sanchez, R.S. Tattersall, J.J. Manson, COVID-19: consider cytokine storm syndromes and immunosuppression, Lancet 395 (2020) 1033–1034, https://doi.org/10.1016/S0140-6736(20)30628-0.

[19] M. Yang, C.L. Lai, SARS-CoV-2 infection: can ferroptosis be a potential treatment target for multiple organ involvement? Cell Death Dis. 6 (2020) 130, https://doi.org/10.1038/s41420-020-00369-w.

[20] Y. He, J. Wang, J. Ren, Y. Zhao, J. Chen, X. Chen, Effect of COVID-19 on male reproductive system - a systematic review, Front. Endocrinol. 12 (2021), 677701, https://doi.org/10.3389/fendo.2021.677701.

[21] G. Nunnari, C. Sanfilippo, P. Castrogiovanni, R. Imbesi, V.G. Li, I. Barbagallo, G. Musumeci, R.M. Di, Network perturbation analysis in human bronchial epithelial cells following SARS-CoV2 infection, Exp. Cell Res. 395 (2020), 112204, https://doi.org/10.1016/j.yexcr.2020.112204.

[22] Y. Li, J. Xie, X. Xu, J. Wang, F. Ao, Y. Wan, Y. Zhu, MicroRNA-548 down-regulates host antiviral response via direct targeting of IFN-lambda1, Protein Cell 4 (2013) 130–141, https://doi.org/10.1007/s13238-012-2081-y.

[23] T.J. Xing, H.T. Xu, W.Q. Yu, B. Wang, J. Zhang, MiRNA-548ah, a potential molecule associated with transition from immune tolerance to immune activation of chronic hepatitis B, Int. J. Mol. Sci. 15 (2014) 14411–14426, https://doi.org/10.3390/ijms150814411.

[24] Z. Suba, Prevention and therapy of COVID-19 via exogenous estrogen treatment for both male and female patients, J. Pharm. Pharmaceut. Sci. 23 (2020) 75–85, https://doi.org/10.18433/jpps31069.

[25] D.P. Robinson, O.J. Hall, T.L. Nilles, J.H. Bream, S.L. Klein, 17beta-estradiol protects females against influenza by recruiting neutrophils and increasing virus-specific CD8 T cell responses in the lungs, J. Virol. 88 (2014) 4711–4720, https://doi.org/10.1128/JVI.02081-13.

[26] J. Peretz, A. Pekosz, A.P. Lane, S.L. Klein, Estrogenic compounds reduce influenza A virus replication in primary human nasal epithelial cells derived from female, but not male, donors, Am. J. Physiol. Lung Cell Mol. Physiol. 310 (2016) L415–L425, https://doi.org/10.1152/ajplung.00398.2015.

[27] Z. Fu, J. Wang, Z. Wang, Y. Sun, J. Wu, Y. Zhang, X. Liu, Z. Zhou, L. Zhou, C. Y. Zhang, Y. Yi, X. Xia, L. Wang, X. Chen, A virus-derived microRNA-like small RNA serves as a serum biomarker to prioritize the COVID-19 patients at high risk of developing severe disease, Cell Discov. 7 (2021) 48, https://doi.org/10.1038/s41421-021-00289-8.

[28] F. Meng, G.K. Siu, B.W. Mok, J. Sun, K.S.C. Fung, J.Y. Lam, N.K. Wong, L. Gedefaw, S. Luo, T.M.H. Lee, S.P. Yip, C.L. Huang, Viral MicroRNAs encoded by nucleocapsid gene of SARS-CoV-2 are detected during infection, and targeting metabolic pathways in host cells, Cells 10 (2021), https://doi.org/10.3390/cells10071762.

[29] G.A. Merino, J. Raad, L.A. Bugnon, C. Yones, L. Kamenetzky, J. Claus, F. Ariel, D. H. Milone, G. Stegmayer, Novel SARS-CoV-2 encoded small RNAs in the passage to humans, Bioinformatics 36 (2021) 5571–5581, https://doi.org/10.1093/bioinformatics/btaa1002.