In Silico Investigation of Potential COVID-19-Associated MicroRNA Signatures

Athanasia Pavlopoulou¹, Seyedehsadaf ASFA²

¹ Dökuz Eylül University
² Izmir Biomedicine and Genome Center (IBG)

Funding: No specific funding was received for this work.
Potential competing interests: No potential competing interests to declare.

Abstract

**Background:** The global pandemic COVID-19, caused by the coronavirus SARS-CoV-2, is persistent despite the increasing vaccination rates, with new cases being reported per week. MicroRNAs, non-coding RNA species that regulate gene expression at the post-transcriptional level, play a pivotal role in the SARS-CoV-2 life cycle, pathophysiology and host's anticoronaviral responses.

**Materials and methods:** In the present study, an integrative bioinformatics approach was employed – including database searching, gene set enrichment analysis, and network-based and microRNA target prediction methods – towards discovering epigenetic determinants of COVID-19.

**Results:** An intricate microRNA-target gene network was constructed, and a set of eight highly interacting microRNAs that potentially co-target and co-regulate key COVID-19-related genes was detected. These miRNAs and their corresponding genes are likely involved in the response to SARS-CoV-2 infection.

**Conclusion:** The eight functionally associated miRNAs could represent the components of a signature for COVID-19 diagnosis.
Seyedehsaad AFSA\textsuperscript{1,2,a}, Athanasia PAVLOPOULOU\textsuperscript{1,2,b,}\textsuperscript{*}

\textsuperscript{1} Izmir Biomedicine and Genome Center (IBG), 35340 Balcova, Izmir, Turkey.
\textsuperscript{2} Izmir International Biomedicine and Genome Institute, Genomics and Molecular Biotechnology Department, Dokuz Eylül University, 35340 Balcova, Izmir, Turkey.

\textsuperscript{a} ORCID iD: 0000-0001-9134-7517
\textsuperscript{b} ORCID iD: 0000-0002-0815-3808

\textsuperscript{*} Correspondence: athanasia.pavlopoulou@ibg.edu.tr

Keywords: COVID-19, bioinformatics, microRNAs.

1. Introduction

The ongoing global pandemic COVID-19, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), persists despite the vaccination efforts, with thousands of COVID-19 cases and a great number of deaths to be reported weekly (WHO Coronavirus (COVID-19) Dashboard). Of note, a novel SARS-CoV-2 variant BA.2.86 (‘Pirola’) has recently emerged, which is spreading in the USA and UK (Mahase, 2023). SARS-CoV-2-mediated infection is characterized by activation of the innate immune response (the first line of defense) and hyperinflammation (Diamond and Kanneganti, 2022; Gustine and Jones, 2021; Tufan et al., 2020). COVID-19 is associated with numerous complications and comorbidities, including common cold, pneumonia, blood coagulation, lymphopenia, acute respiratory distress syndrome (ARDS), tissue injury, multiple organ dysfunction, etc. (Chen et al., 2020; Krynytska et al., 2021; Nalbandian et al., 2021; Sanyaolu et al., 2020).

MicroRNAs (miRNAs) are endogenous, short (22-nucleotide) single-stranded non-protein-coding RNA molecules that can regulate gene expression at the post-transcriptional level through base-pairing with the 3’ untranslated region (3’UTRs) of the target mRNAs, and subsequently suppress mRNA expression by mRNA degradation or translational inhibition. A single miRNA can potentially target multiple genes, and a gene can be targeted by many miRNAs (Ambros, 2004; Bartel, 2004).

It has been demonstrated that miRNAs play a critical role in coronaviral life cycle, pathogenesis and host antiviral responses (Abedi et al., 2021; Arghiani et al., 2021). SARS-CoV-2 infection-induced changes in the expression patterns of the host miRNAs can lead to the down-regulation of key genes involved in the immune response and, consequently, to immunosuppression or attenuated host immune surveillance, thereby enabling the coronavirus to evade the host’s immune system (Farr et al., 2021; Liang et al., 2023; Yang et al., 2022). SARS-CoV-2 can also co-opt host miRNAs
implicated in immune response suppression (e.g., hsa-miR-939 and hsa-miR-146b), in order to facilitate coronaviral replication and subvert host immune responses (Arghiani et al., 2021; Panda et al., 2022). Moreover, miRNAs encoded by the coronaviral genome can target genes involved in host immune response/inflammation, like IFN1-mediated signaling (Khan et al., 2020; Singh et al., 2022). SARS-CoV-2 non-coding RNAs, by acting as miRNA sponges, can also deplete the host’s miRNA pool (Bartoszewski et al., 2020; Li et al., 2022).

According to Fayyad-Kazan and colleagues (2021), differentially expressed circulating miRNAs (i.e., miR-19a-3p, miR-19b-3p, and miR-92a-3p) are considered powerful biomarkers for the timely and accurate diagnosis of COVID-19 (Fayyad-Kazan et al., 2021). In addition, deregulated plasma miRNAs were shown to have predictive and discriminatory potential for COVID-19 severity and mortality (Fernandez-Pato et al., 2022). In a study by Farr et al. (2021), it was shown that three differentially expressed miRNAs (miR-423-5p, miR-23a-3p and miR-195-5p) could accurately discriminate COVID-19 patients from healthy controls (Farr et al., 2021).

In this study, we have employed a bioinformatics approach towards the investigation of potential miRNA determinants of the epigenetic regulation of genes that are prominently associated with COVID-19, in order to decipher the co-regulation activities of miRNAs exerted upon those genes in coronavirus infection.

2. Materials and Methods

2.1. Acquisition of COVID-19-related genes

To obtain a comprehensive list of genes significantly linked to COVID-19, the GeneCards database (https://www.genecards.org/) (Rebhan et al., 1998; Stelzer et al., 2016) [accessed March 2023] was searched using the keywords “COVID-19” and “SARS-CoV-2”, and the genes with relevance score ≥ 9 were extracted.

2.2. Protein-Protein Interaction Network

The physical and functional associations among the protein products of the retrieved COVID-19-related genes were investigated and visualized using STRING (Search Tool for Retrieval of Interacting Genes/Proteins) v11.5 (Szklarczyk et al., 2021), a database of both experimentally supported or predicted, functional and/or physical, association data among genes/proteins derived collected from diverse resources. A relatively high confidence interaction score (≥0.7) was set as a cutoff, and only those associations based on text mining, database and experimental evidence were considered, in order to enhance the reliability of the given interactions. The associations were further visualized and analyzed through the open-source platform Cytoscape (v.3.10.0) (Shannon et al., 2003). Moreover, the Cytoscape plugin cytoHubba (Chin et al., 2014), which allows network topological analysis by twelve local and global ranking algorithms (Betweenness (BC), BottleNeck (BN), Clustering Coefficient (CC), Closeness (Clo), Degree method (Deg), Density of Maximum Neighborhood Component (DMNC), EcCentricity (EC), Edge Percolated Component (EPC), Maximum Clique Centrality (MCC), Maximum Neighborhood Component (MNC), Radiality (Rad) and Stress (Str)) was utilized to select the top 25 nodes, by
using the node-degree filter.

2.3. Functional Enrichment Analysis

Gene set enrichment analysis was conducted with the online tool WebGestalt (WEB-based GEne SeT AnaLysis Toolkit) (Kirov et al., 2014; Liao et al., 2019), for the identification of statistically significant over-represented terms in the COVID-19 genes under study. The WebGestalt parameters selected were “Organism of Interest”: Homo sapiens, “Method of Interest”: Over-Representation Analysis (ORA), “Functional database”: pathway/Reactome for biological paths, “Select gene ID type”: gene symbol, “Select Reference set”: genome; the default advanced parameters were chosen, and only those pathways with a Benjamini-Hochberg-adjusted p-value (Benjamini and Hochberg, 1995) ≤ 0.05 were included in the analysis. Affinity propagation was applied to reduce the terms and cluster them into representative categories.

2.4. miRNA Regulators of COVID-19 Genes

The miRNAs potentially regulating the COVID-19 genes were investigated. To this end, both the experimentally verified and predicted miRNAs targeting the COVID-19 genes were obtained by applying four state-of-the-art software tools: i) microT_CDS (http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=microT_CDS/index) [accessed May 2023], is based on the DIANA-microT-CDS algorithm to predict miRNA target genes (Paraskevopoulou et al., 2013); ii) TargetScan (https://www.targetscan.org/vert_80/) [accessed May 2023] searches for the presence of conserved sites matching the miRNA seed region, including the k-mers: 8mer, 7mer, and 6mer (Agarwal et al., 2015); iv) PITA (https://genie.weizmann.ac.il/pubs/mir07/) recognizes thermodynamically favorable miRNA-mRNA interactions (Kertesz et al., 2007); iii) miRDB (http://mirdb.org/) [accessed May 2023] predicts functionally annotated miRNA gene targets by machine learning approaches (Chen and Wang, 2020; Liu and Wang, 2019). The miRNA-mRNA interactions were retrieved from each tool combined, and the duplicates were removed. To enhance the accuracy of the prediction, only those miRNA-gene targets predicted by more than three methods were considered in this study.

2.5. Pairwise miRNA associations

The miRNA-miRNA relationships were collected from two different sources, MiRGOFS (Y. Yang et al., 2018) and GOSemSim (Yu et al., 2010), which infer functional similarities between miRNA pairs based on the degree of functional relatedness of their corresponding target genes. The pairwise miRNA interactions from both sources were merged, and the duplicates were removed. Only those miRNA-miRNA interactions with a weight score above 0.85 (where “1.0” is the highest score) were chosen, so as to enhance robustness.

2.6. Independent Validation

The findings of this study were further compared against an independent transcriptomic dataset available in the comprehensive online resource COVID19db (Zhang et al., 2022), by using the “differential expression” module, which provides gene expression profiling in whole blood derived from COVID-19 patients and healthy controls.
3. Results and Discussion

Collectively, 149 COVID-19-relevant genes were obtained from GeneCards (Table S1). A functional network of the products of those genes was generated, and 128 nodes appear to be highly interconnected (Figure 1), suggesting physical and functional associations among the corresponding proteins. By examining the topological properties of the network, we identified those key nodes that are more relevant to the overall function of the network and, therefore, biologically meaningful (Barabasi et al., 2011; Kontou et al., 2016). The 25 top nodes were detected in the protein-protein interaction network (Figure 1 and Table S1) based on the combined output of the twelve algorithms in cytoHubba.

Figure 1. Interaction network of the COVID-19 proteins. The nodes denote genes/gene products, and the edges represent functional associations. The nodes corresponding to the top-ranking gene products are indicated by light red fill colors.
The terms over-represented in the genes, the products of which are involved in the protein association network (Figure 1), are mainly associated with immune signaling pathways (Figure 2), e.g., cytokine, interleukin, interferon, Toll-like receptor, MYD88, NF-κF, T-cell receptor (TCR)-mediated signaling cascades. Notably, 109 (out of 128) and all 25 top-ranking genes are found in the over-represented Reactome pathways, further highlighting their prominent role in COVID-19 pathophysiology.

Figure 2. Bar plot illustrating the over-represented Reactome pathways in the COVID-19 genes. The width of bar plots is proportional to the number of genes in each pathway.

Among the 25 key genes, there are mostly genes encoding components of the immune system (CD4, CD8A, HLA-B, IRF3, MYD88 etc.), including pro- and anti-inflammatory factors, such as chemokines and cytokines (CXCL8, CXCL10, IL1B, IL2, IL4, IL6, IL10, IL17A, IL18, IFNG) and the tumor necrosis factor TNF. In many COVID-19 cases, exacerbated inflammatory responses ("cytokine storm") are observed, which result from the acute increase in the levels of circulating pro-inflammatory cytokines and chemokines, and their uncontrolled release, both at local and systemic levels (Coperchini et al., 2020; Hu et al., 2021; Montazersaheb et al., 2022). Several studies, though, suggest that the human host's immune system is rather compromised upon SARS-CoV-2 infection and is not capable of eliciting a sufficient immune response.
(Ozbek et al., 2022; Remy et al., 2020).

The NF-kB subunits, NFKB1 and RELA (Table S1), are tightly connected to several pro-inflammatory agents. NF-kB can modulate the host’s immediate innate immune response to SARS-CoV-2 infection. In particular, SARS-CoV-2-mediated activation of NF-kB was found to induce the expression of the genes IL1, IL2, IL6, IL8 and TNF (Hariharan et al., 2021). Human miRNAs, by targeting and suppressing the gene coding for NFKB1, could attenuate the effects (e.g., organ injuries) of the COVID-19-induced inflammatory cascade (Abedi et al., 2021).

The effector gene OAS1 affects susceptibility to SARS-CoV-2 and has a protective effect against COVID-19 severity (Asgari and Pousaz, 2021; Huffman et al., 2022; Zhou et al., 2021). TLR4 (Toll-like receptor 4) is a sensor for innate immunity. SARS-CoV-2 is suggested to bind and activate TLR4, leading to increased expression of ACE2, the major receptor for SARS-CoV-2 cell entry (Scialo et al., 2020), thereby facilitating coronavirus entry and amplifying the inflammatory response (Aboudounya and Heads, 2021; Mukherjee, 2022). CSF2 (colony-stimulating factor 2) was also found to be over-expressed in SARS-CoV-2-infected human lung epithelial cells (Chandrashekar et al., 2021).

The differential expression profiles of several COVID19-related top-25 genes (Table S1) between SARS-CoV-2-infected patients and healthy controls were further investigated through the “differential expression” module of COVID19db (Zhang et al., 2022). Among those, the cytokines IL10, IL18 and CXCl0 are up-regulated in the COVID-19 group (Figure 4), consistent with their pro-inflammatory effects (Callahan et al., 2021; Dinarello, 2000; Lauw et al., 2000).
Overexpression of the cytokine IL6 has been reported to be highly associated with the mortality risk of COVID-19, and IL6 acts as a perpetrator of the SARS-CoV-2-induced cytokine storm (Chen et al., 2020; Geronikolou et al., 2022; Santa Cruz et al., 2021). However, IL6 appears to be downregulated in the blood of COVID-19 patients (Figure 4), consistent with the findings of a previous bioinformatics study by Özbek and colleagues (2023), wherein IL6 was not found significantly dysregulated in diverse SARS-CoV-2-infected tissues (Ozbek et al., 2022). This is probably due to the dual role of IL6 as
a pro- and anti-inflammatory cytokine (Borsini et al., 2020; Scheller et al., 2011), as well as the diverse immune response patterns observed in COVID-19 patients (Galbraith et al., 2021; Ozbek et al., 2022). The pleiotropic anti-inflammatory cytokine IL4 (Chatterjee et al., 2014) is also down-regulated in COVID-19 blood samples (Figure 4), in agreement with previous studies suggesting cytokine storm suppression during SARS-CoV-2 infection (Ozbek et al., 2022; Remy et al., 2020).

Increased expression of TLR4 greatly affects heart failure in COVID-19 patients (Choudhury and Mukherjee, 2021; Mukherjee, 2022). TLR4 is known to trigger the activation of pro-inflammatory cytokines (Swanson et al., 2020), including TNF, which is also up-regulated in COVID-19 (Figure 4); increased expression of TNF was found to be a prognostic factor for mortality among COVID-19 patients with comorbidities and disease progression (Mohd Zawawi et al., 2023). MYD88, which plays a central role in TLR/IL1R-mediated signalling in innate and adaptive immunity (Chen et al., 2020), is overexpressed in the COVID-19 group (Figure 4); MYD88 polymorphisms were shown to be tightly associated with COVID19 severity (Martinez-Gomez et al., 2023). Finally, the expression level of the interferon-inducible gene OAS1 is markedly higher in the SARS-CoV-2 infected patients as compared to the healthy samples, in agreement with the findings of a recent study, wherein SARS-CoV-2 infection significantly increased the expression of OAS1 (Assou et al., 2023).

As shown in Figure 3, 24 of the top-ranking genes are potentially targeted by more than eight miRNAs, namely hsa-miR-1276, hsa-miR-3121-3p, hsa-miR-338-5p, hsa-miR-340-5p, hsa-miR-5692a, hsa-miR-570-3p, hsa-miR-664a-3p and hsa-miR-7-5p. All eight miRNAs are also associated in a pairwise fashion (Figure 3), suggesting participants in an intricate regulatory network, where these miRNAs act in a synergistic fashion to co-target and co-regulate the expression of COVID-19 genes.
There is evidence that several of those eight miRNAs are associated with SARS-CoV-2 and COVID-19. Vastrad and colleagues (2020) analyzed gene expression data from whole blood from human patients with COVID-19, and found hsa-miR-5692a among those miRNAs that target genes up-regulated in COVID-19 samples as compared to healthy controls (Vastrad et al., 2020). Moreover, hsa-miR-340-5p was down-regulated in the peripheral blood collected from COVID-19 patients (Li et al., 2020). According to Katopodis et al. (2022), hsa-miR-3121-3p and hsa-miR-570-3p could regulate the expression of mediators of SARS-CoV-2 cell entry (Katopodis et al., 2022). Also, hsa-miR-1276 was shown to have the capacity to bind viral RNA and to be modulated significantly in pneumocytes, suggesting a role in SARS-CoV-2 infection (Milenkovic et al., 2021). Bartoszewski et al. (2020) suggested potential SARS-CoV-2-mediated modulation of host miRNAs, including hsa-miR-664a-3p, so as to perpetuate coronaviral persistence/replication and evade host immune defences (Bartoszewski et al., 2020).

The causal genes of diseases/disorders usually share common regulatory mechanisms in order to ensure their
coordinated regulation. Disease-related miRNAs act in a cooperative manner so as to exert their regulatory effect upon their corresponding target genes (Arshinchi Bonab et al., 2022; Zinani et al., 2022).

The protein-protein interaction and miRNA-gene networks constructed in this study provide a fundamental framework for detecting protein-coding genes and epigenetic regulators (e.g., miRNAs) that likely respond to SARS-CoV-2 infection in a coordinated way. In this study, by applying stringent criteria, we discovered eight highly interacting miRNAs that potentially co-target key COVID-19 genes. This panel of miRNAs could represent potential signature components for COVID-19.

4. Conclusion

Herein, by employing a bioinformatics pipeline, we detected eight functionally related miRNAs that likely co-regulate pivotal COVID-19-associated genes. The differential expression status of the predicted miRNAs, and their corresponding pairwise interactions, merit further experimental validation in the context of COVID-19. These miRNAs could be taken into consideration in the clinical setting for updating and complementing currently used biomarkers towards improving the accuracy of diagnosing SARS-CoV-2-infected patients.

Conflict of interest

The authors report that there are no competing interests to declare.

References


• Katopodis, P., Randeva, H. S., Spandidos, D. A., Saravi, S., Kyrou, I., et al. (2022). Host cell entry mediators implicated in the cellular tropism of SARS-CoV-2, the pathophysiology of COVID-19 and the identification of microRNAs that can modulate the expression of these mediators (Review). International Journal of Molecular Medicine, 49(2). https://doi.org/10.3892/ijmm.2021.5075


