



OPEN

SARS-CoV-2 potential drugs, drug targets, and biomarkers: a viral-host interaction network-based analysis

Asmaa Samy^{1,4}, Mohamed A. Maher^{1,4}, Nehal Adel Abdelsalam^{1,2} & Eman Badr^{1,3}✉

COVID-19 is a global pandemic impacting the daily living of millions. As variants of the virus evolve, a complete comprehension of the disease and drug targets becomes a decisive duty. The Omicron variant, for example, has a notably high transmission rate verified in 155 countries. We performed integrative transcriptomic and network analyses to identify drug targets and diagnostic biomarkers and repurpose FDA-approved drugs for SARS-CoV-2. Upon the enrichment of 464 differentially expressed genes, pathways regulating the host cell cycle were significant. Regulatory and interaction networks featured hsa-mir-93-5p and hsa-mir-17-5p as blood biomarkers while hsa-mir-15b-5p as an antiviral agent. MYB, RRM2, ERG, CENPF, CIT, and TOP2A are potential drug targets for treatment. HMOX1 is suggested as a prognostic biomarker. Enhancing HMOX1 expression by neem plant extract might be a therapeutic alternative. We constructed a drug-gene network for FDA-approved drugs to be repurposed against the infection. The key drugs retrieved were members of anthracyclines, mitotic inhibitors, anti-tumor antibiotics, and CDK1 inhibitors. Additionally, hydroxyquinone and digitoxin are potent TOP2A inhibitors. Hydroxyurea, cytarabine, gemcitabine, sotalol, and amiodarone can also be redirected against COVID-19. The analysis enforced the repositioning of fluorouracil and doxorubicin, especially that they have multiple drug targets, hence less probability of resistance.

Owing to its unpredicted prognosis and disease course, COVID-19 is considered one of the most peculiar challenges facing the human race. In less than three years, SARS-CoV-2 infected more than 300 million people globally and caused death to more than 5 million¹. It began in December 2019 when a novel coronavirus (CoV) was identified in Wuhan, China². It belongs to the betacoronavirus genus that includes severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV)³. Accordingly, the International Committee on Taxonomy of Viruses (ICTV) announced the official name of the virus as severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) and the disease it causes as COVID-19⁴. Furthermore, due to the rapid human-to-human transmission, the World Health Organization (WHO) declared COVID-19 as a global public health emergency in January 2020 and a pandemic in March the same year⁵. Lately, the WHO declared the Omicron variant a variant of concern due to mutations in the viral genome that may affect the virus spread⁶. It has been verified in 155 different locations around the globe⁷.

Although SARS-CoV-2 infection causes similar symptoms to SARS-CoV and MERS-CoV, the pathophysiology of COVID-19 and the molecular basis of the disease severity are not fully understood yet. Recent studies highlighted the role of host immunity and inflammation in disease progression. Sudden storms of cytokines and chemokines may lead to lung injury and acute respiratory distress syndrome (ARDS), which is considered the clinical hallmark of SARS-CoV-2 infection. It has also been shown that viral replication results in cellular apoptosis that may cause diffuse alveolar damage, leading to ARDS as well^{8,9}. Cellular apoptosis can be attributed to the fact that viruses lack the needed machinery for replication. Therefore, they rely entirely on the host cells to express the required viral proteins for this process. To efficiently control host cell resources, viruses may either inhibit or induce cell cycle progression. With such interference, some viral infections may eventually cause cancer. Previous studies confirmed that many CoVs induce host cell cycle arrest at different phases, such as G0/G1 by SARS-CoV¹⁰, porcine epidemic diarrhea virus¹¹, and murine hepatitis virus¹². Bronchitis virus¹³ and

¹University of Science and Technology, Zewail City, Giza 12578, Egypt. ²Faculty of Pharmacy, Cairo University, Cairo 11562, Egypt. ³Faculty of Computers and Artificial Intelligence, Cairo University, Giza 12613, Egypt. ⁴These authors contributed equally: Asmaa Samy and Mohamed A. Maher. ✉email: emostafa@zewailcity.edu.eg

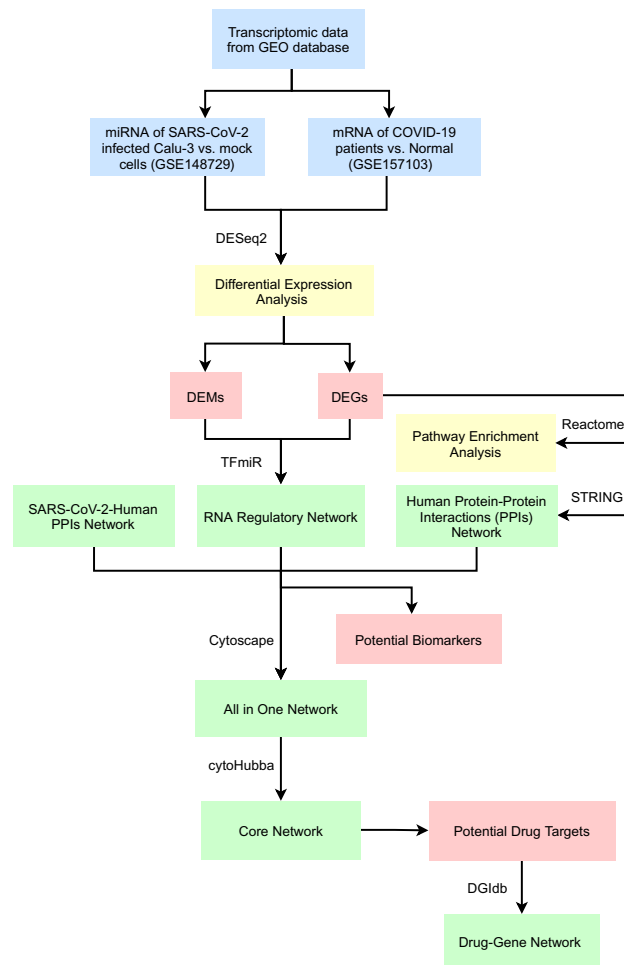


Figure 1. A block diagram depicting the workflow for identifying potential biomarkers and drug targets for COVID-19 infection.

transmissible gastroenteritis virus¹⁴ induce cell cycle arrest at S and G2/M phases. Understanding the cellular pathways associated with viral infection can reveal potential targets for therapy.

One therapeutic approach that has been widely investigated to fight the pandemic is the microRNA (miRNA)-based approach¹⁵. miRNAs are small non-coding RNA molecules of ~21 nucleotides that regulate gene expression at post-transcriptional level¹⁶. They are involved in the regulation of different cellular pathways, including viral infection and host cell response¹⁷. A recent review by Abedi et al.¹⁵ summarized the role of miRNAs in SARS-CoV-2 life cycle and pathogenesis. For instance, hsa-miR-125a-5p and hsa-miR-200 families target the angiotensin-converting enzyme 2 and consequently block viral attachment and entry¹⁸. hsa-let-7e-5p¹⁸ and hsa-miR-98-5p¹⁹ target transmembrane protease, serine 2 (TMPRSS2). Other miRNAs have been reported to inhibit viral replication and protein synthesis. Additionally, hsa-miR-323a-5p and hsa-miR-20b-5p were predicted to have an antiviral effect by reducing inflammatory responses and preventing lung injury²⁰. Besides being potential antiviral agents, miRNA might be diagnostic biomarkers as reported by a recent study that the disease severity and mortality in aged patients may have resulted from the lower expression levels of host miRNAs²¹.

As SARS-CoV-2 is an ongoing health hazard with multiple variants emerging in a short time interval, accelerated research is urgently needed to develop prevention, diagnosis, and treatment strategies against the pandemic. The development of new antiviral agents is a challenging, costly long-term process, especially with the rapid evolution of viral variants. Likewise, resistance adds another layer of complexity to the discovery and design of antivirals²². Drug repurposing or repositioning is a promising strategy in this situation to save both time and cost amidst a pandemic²³. In this study, we applied an integrated computational pipeline of transcriptomic profiling and network analysis, as shown in Fig. 1. Our analysis aimed to (i) enhance the understanding of the molecular mechanism behind SARS-CoV-2 pathogenesis, (ii) identify candidate key targets and their repurposed FDA-approved drugs, and (iii) identify candidate biomarkers for the disease diagnosis. Hence, our study proposes MYB, TOP2A, RRM2, ERG, CENPF, CIT, and HMOX1 as potential drug targets. Moreover, hsa-miR-93-5p and hsa-miR-17-5p are two miRNAs that have been identified as potential biomarkers for COVID-19 infection. hsa-miR-15b-5p is a promising natural miRNA-based antiviral agent. Anthracyclines, mitotic inhibitors, anti-tumor antibiotics, and CDK1 inhibitors are key classes for repositioning against coronavirus infection. Fluorouracil and

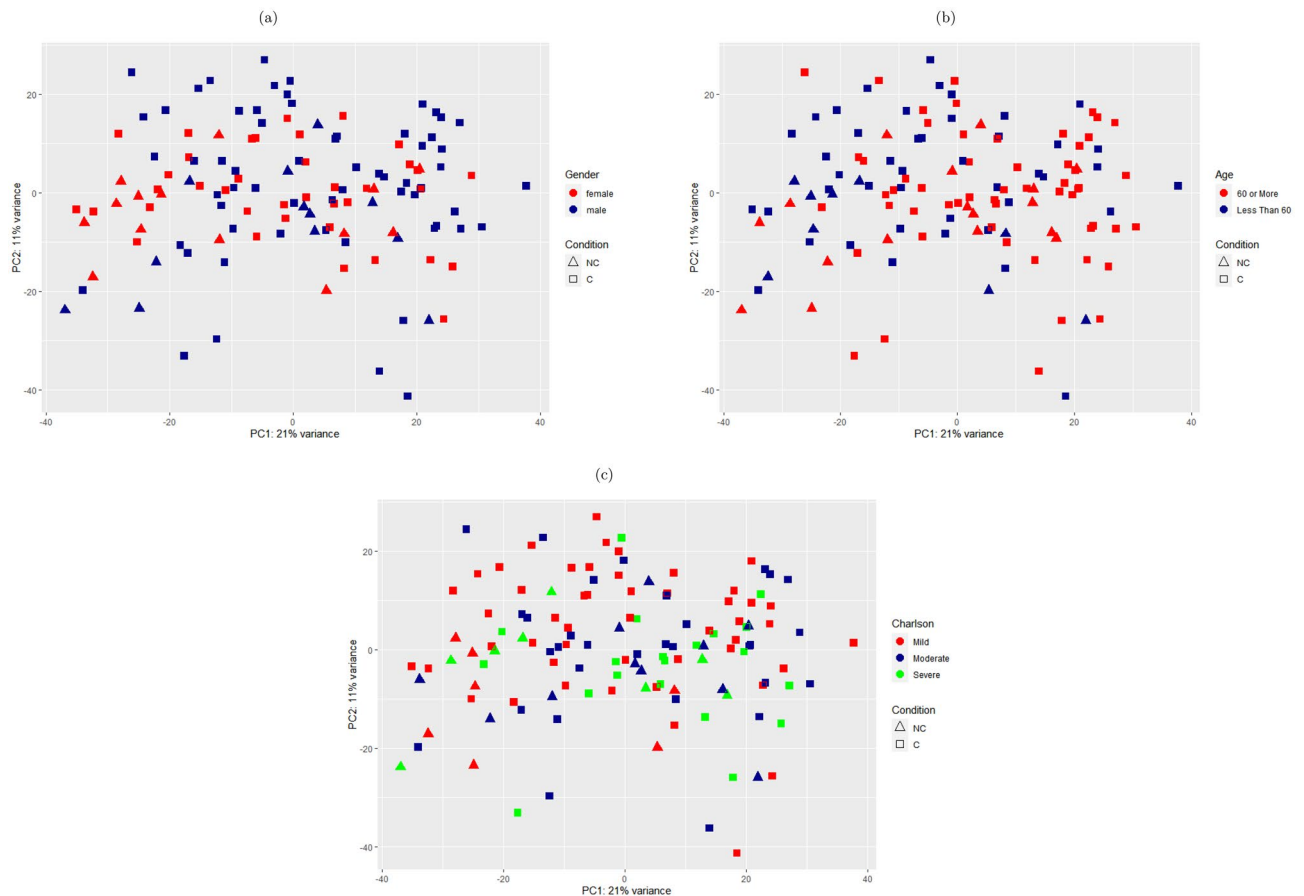


Figure 2. PCA plots showing that (a) gender, (b) age, and (c) Charlson score have no confounding effects on gene expression between COVID-19 patients and non-COVID-19 samples (GSE157103).

doxorubicin are two FDA-approved anticancer drugs and have multiple target genes. They could be repurposed to treat COVID-19 upon thorough experimentation and validation.

Results

Differential expression analysis and biological insights. Differential expression analysis was performed on GSE157103⁶⁶ and GSE148729⁶⁷ to identify differentially expressed genes (DEGs) in plasma and leukocyte samples and differentially expressed microRNAs (DEMs) in lung cancer cell lines, respectively. The dataset GSE157103 holds expression profiles for blood samples from patients tested positive for SARS-CoV-2 and other tested negative. From GSE148729 dataset, only small RNA expression profiles were retrieved for analysis. These samples were SARS-CoV-2 infected and non-infected Calu-3 human epithelial cell line.

For GSE157103, we utilized the Principal Component Analysis (PCA) plots for gender, age, and Charlson severity score to ensure that those factors are not sources of variation. For age, data were split into two groups, below 60 years and 60 years or above. The age groups had 53 samples below 60 years and 73 with 60 years or above. The Charlson score was categorized into three groups: below 3 for mild, between 3 and 6 is moderate, and 6 or more for severe cases. The groups had 56, 42, and 28 samples, respectively. The samples were split into 74 males and 52 females. As Fig. 2 demonstrates, there was no defined separation or clustering between the two cohorts based on the abovementioned characteristics.

Based on statistical thresholds of adjusted P -value < 0.05 and $|\text{Log}_2\text{foldchange}(FC)| > 0.5$, we identified 464 DEGs (389 upregulated and 75 downregulated genes) from GSE157103 dataset (Fig. 3 and Supplementary Table 1). We also identified 37 DEMs (15 upregulated and 22 downregulated miRNAs) from GSE148729 dataset (Fig. 3 and Supplementary Table 2). Same statistical thresholds were set for DEMs and DEGs.

Reactome pathway analysis results showed that the DEGs of patients with COVID-19 were significantly enriched in 97 pathways with maximum False Discovery Rate (FDR) of 0.01 (Supplementary Table 3). The top listed pathways with lowest FDR values were cell cycle, including different phases of mitotic cell cycle and cell cycle checkpoints, DNA replication, and RHO GTPase effectors for signaling as given in Table 1.

Identification of biomarkers and drug targets. Different interactions were retrieved and studied for a wide-scope analysis of the DEGs and DEMs. First, RNA regulatory network was constructed to identify interactions between DEGs and DEMs. The network consists of 96 nodes, including 81 DEGs (5 Transcription Factors (TFs);

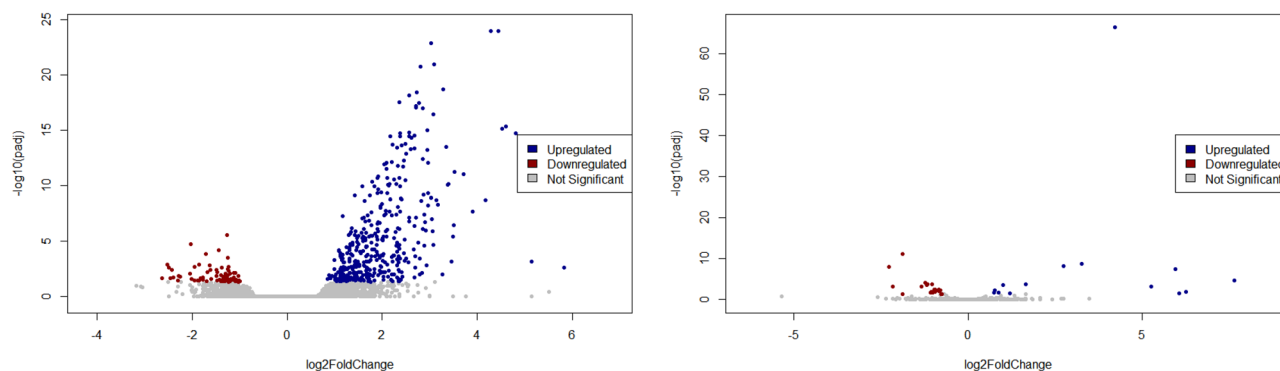


Figure 3. Volcano plots of (a) differentially expressed genes and (b) differentially expressed miRNAs by analyzing GSE157103 and GSE148729 datasets, respectively. Blue circles: upregulated DEGs/DEMs; red circles: downregulated DEGs/DEMs; grey circles: non-significantly expressed mRNAs/miRNAs.

Pathway name	Entities P value	Entities FDR
G1/S transition	1.11E-16	1.75E-14
G1/S-specific transcription	1.11E-16	1.75E-14
Cell cycle, mitotic	1.11E-16	1.75E-14
Cell cycle	1.11E-16	1.75E-14
Mitotic G1 phase and G1/S transition	1.11E-16	1.75E-14
Cell cycle checkpoints	1.11E-16	1.75E-14
Polo-like kinase mediated events	4.04E-14	5.46E-12
Amplification of signal from unattached kinetochores via a MAD2 inhibitory signal	1.03E-13	1.08E-11
Amplification of signal from the kinetochores	1.03E-13	1.08E-11
Resolution of sister chromatid cohesion	1.77E-13	1.69E-11
Mitotic spindle checkpoint	5.63E-13	4.84E-11
Mitotic prometaphase	4.85E-12	3.63E-10
G2/M checkpoints	4.97E-12	3.63E-10
EML4 and NUDC in mitotic spindle formation	2.20E-11	1.47E-09
G0 and Early G1	4.48E-11	2.82E-09
M phase	8.15E-11	4.81E-09
RHO GTPase effectors	2.16E-10	1.19E-08
RHO GTPases activate formins	3.14E-10	1.63E-08
DNA replication	3.90E-10	1.95E-08
Activation of ATR in response to replication stress	6.63E-10	3.12E-08

Table 1. Top 20 enriched pathway terms of DEGs in SARS-CoV-2 patients with lowest FDR values.

RRM2, ERG, MYB, MYBL2, and IRF4) and 15 DEMs. It has 150 edges, representing different types of interactions; one TF-miRNA, 24 TF-gene, 117 miRNA-gene, and 8 miRNA-TF (Fig. 4 and Supplementary Table 4). Among the five transcription factors, MYB had the highest interaction by regulating transcription of 17 genes and being regulated by 4 miRNAs. The network revealed 4 upregulated DEMs (hsa-mir-93-5p, hsa-mir-17-5p, hsa-mir-103a-3p, and hsa-mir-107) that could be considered as potential biomarkers for the disease diagnosis.

Superimposing virus-human interactome network on the regulatory network highlighted three common nodes. The two upregulated DEGs, CENPF and CIT, interact with SARS-CoV-2 nsp13. Moreover, HMOX1 is a downregulated DEG that interacts with SARS-CoV-2 ORF3a. On the regulatory level, CENPF is regulated by MYB-TF, CIT is regulated by two upregulated DEMs (hsa-mir-93-5p and hsa-mir-17-5p), and HMOX1 is regulated by ERG-TF.

Merging the two previous networks resulted into 452 nodes (of which 410 are human proteins). Physical and functional interactions of 275 human proteins were retrieved from STRING (1, 062 queries) and overlaid onto the merged network, resulted into the All in One network consists of 452 nodes connected by 1, 544 edges represent three types of interactions: regulatory interactions among miRNAs-TFs-genes, virus-host Protein-Protein Interactions (PPIs), and human PPIs (Supplementary Table 5).

To narrow down the All in One network and explore the influential nodes, we defined the Core network that consists of 86 nodes and 457 edges. It includes 5 TFs, the 5 previously identified potential biomarkers, and their 70 directly connected nodes. Additionally, we included 6 viral proteins that interact with any of the previously selected nodes as shown in Fig. 5. The network nodes were ranked for more investigation based on the

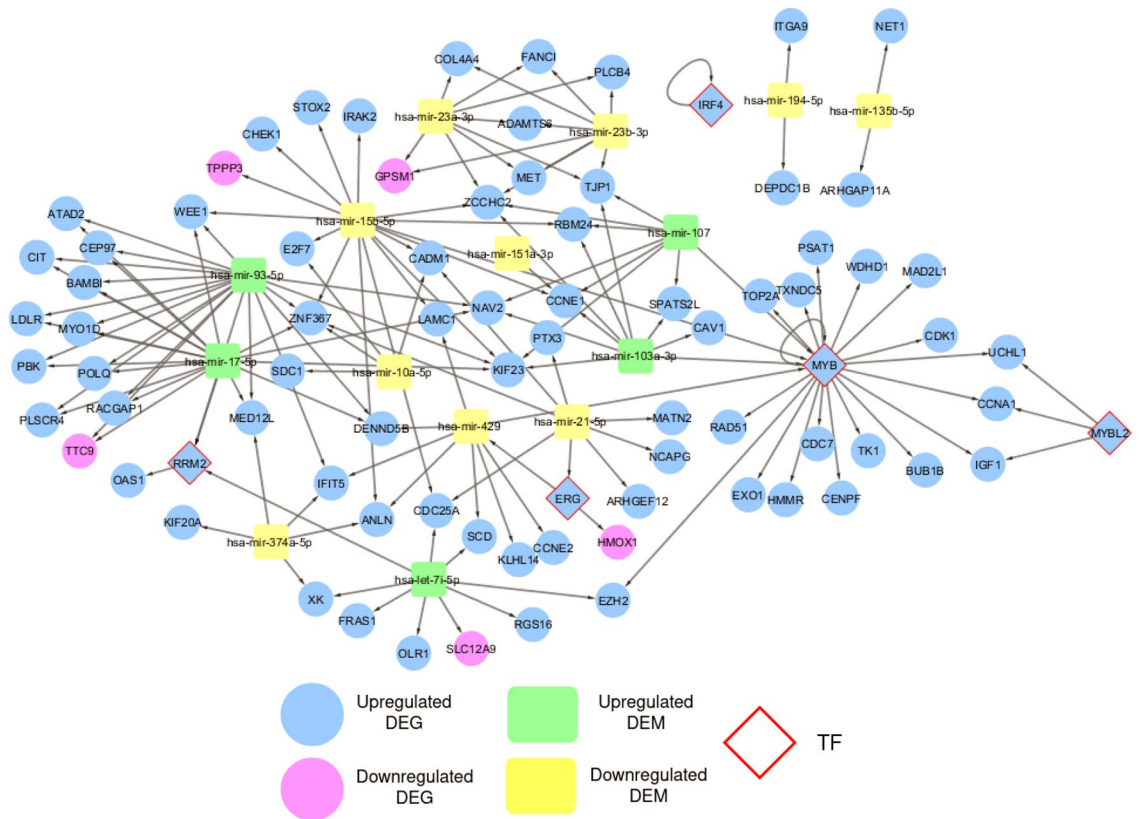


Figure 4. RNA regulatory network that consists of 96 nodes and 150 edges. Node's color and shape are mapped to node type as labeled in the figure. The network is a directed graph to indicate the interaction direction.

Maximal Clique Centrality (MCC) method (Supplementary Table 6). We found that the first 21 genes have the highest MCC score of $9.22E+13$ (Table 2). Twenty of them are upregulated DEGs such as CDK1, HMMR, EXO1, BUB1B, and TOP2A connected to TFs (MYB and RRM2), CENPF (connected to nsp13). PBK is also connected to the previously mentioned nodes and two upregulated miRNAs (hsa-mir-93-5p and hsa-mir-17-5p). Table 2 also listed PRIM1 as a top ranked gene but it does not have Log_2FC and p_{adj} values as it was retrieved from the SARS-CoV-2-human protein-protein interactions network in Gordon et al. study²⁴. It is connected to SARS-CoV-2 nsp1, 17 upregulated DEGs such as CDK1, TOP2A, BUB1B, CENPF, and RRM2-TF.

Screening of FDA-approved drugs with probable potential for COVID-19 treatment. Drug Gene Interaction Database (DGIdb) web resource⁴³ was used to retrieve FDA-approved drugs that interact with 25 genes including the top ranked nodes of the core network (Table 2) and the identified TFs. We found 10 out of the 25 genes (7 upregulated DEGs and 3 TFs) have 64 interactions with 52 approved drugs (Fig. 6 and Supplementary Table 7). TOP2A has the highest number of inhibitory interactions among all genes with 17 drugs such as etoposide, exrazoxane, valrubicin, teniposide, and amsacrine. RRM2-TF can be inhibited by 7 drugs such as gallium nitrate and cladribine, while 5 drugs can block ERG-TF such as sotalol hydrochloride and amiodarone hydrochloride. We also found 10 drugs that could be potential multi-target drugs, such as fluorouracil that targets four genes (HMMR, EXO1, TOP2A, and MYB-TF). Additionally, each of the four anthracycline antitumor antibiotics (doxorubicin, daunorubicin, idarubicin, and epirubicin) can target two genes.

Validation of hub genes and core DEMs. Based on mRNA expression profiles of COVID-19 patients and healthy controls of GSE196822, we evaluated the top 10 ranked genes from the core network (Table 2) and also were identified at the previous section as drug targets (Fig. 6). Five hub genes showed high potential to discriminate COVID-19 patients group from the controls (RRM2, ERG, TOP2A, EXO1, and CDK1) with area under the curve (AUC) values > 0.8 as shown in Fig. 7. AUC values of the other five genes ranged from 0.64 to 0.75. Similarly, the diagnostic potential of the eight identified DEMs of the Core network (Fig. 5) was determined based on miRNA profiles of COVID-19 vs. healthy controls (GSE176498⁷⁹). AUC values showed that the upregulated DEMs (hsa-mir-17-5, hsa-mir-107, hsa-mir-93-5p, and hsa-mir-103a-3p) may be considered as candidate biomarkers for SARS-CoV-2, especially the first two miRNAs with AUC values of 0.89 and 0.82 (Fig. 8). Additionally, the downregulation of hsa-mir-21-5p is significantly associated with SARS-CoV-2 infection with AUC value of 0.92 (Fig. 8).

Node	Log ₂ FC	padj	MCC
TOP2A	3.032766241	1.23E-23	9.22E+13
PBK	3.286974214	1.87E-19	9.22E+13
RRM2	2.563244695	7.40E-19	9.22E+13
BUB1B	2.848843114	9.48E-18	9.22E+13
EXO1	2.569288856	1.65E-15	9.22E+13
HMMR	2.671877231	4.34E-14	9.22E+13
NCAPG	2.319663731	3.83E-14	9.22E+13
CENPF	2.207653	7.83E-13	9.22E+13
CDK1	2.444515359	1.88E-12	9.22E+13
CHEK1	1.63245883	2.46E-09	9.22E+13
RAD51	1.970628548	9.30E-09	9.22E+13
KIF20A	2.410368374	1.81E-09	9.22E+13
ANLN	2.372484524	2.12E-08	9.22E+13
ATAD2	1.296402031	3.01E-06	9.22E+13
WDHD1	1.334236601	2.11E-06	9.22E+13
KIF23	1.315400626	6.74E-05	9.22E+13
ARHGAP11A	1.243130514	3.70E-05	9.22E+13
RACGAP1	0.994416067	0.000535663	9.22E+13
FANCI	1.133952503	0.000318433	9.22E+13
MAD2L1	1.157533489	0.003931285	9.22E+13
PRIM1	N/A	N/A	9.22E+13

Table 2. Top ranked nodes of the Core network based on MCC algorithm.

Discussion

Effective diagnostic and therapeutic strategies against viruses rely heavily on comprehending the mutual interplay between the virus and its host. Transcriptomic and network analyses are the key players in understanding those interactions, thus SARS-CoV-2-host interactions are currently under research²⁵. We conducted an *in silico* analysis of DEGs, DEMs, and human PPIs to infer enriched pathways and potential drug targets in patients and lung cell lines infected with SARS-CoV-2. Although we expect the immune pathways to be the most significantly enriched in the infected samples, to our surprise, pathways involved in regulating the cell cycle progression and DNA replication and repair were ranked the top Reactome terms. As viruses lack the cellular machinery required for genome replication and proteins synthesis, they hijack host cell machinery at the site of infection. Viruses either promote or arrest cell cycle progression for their benefit²⁶. Hence, viruses may promote host cell cycle arrest at the active metabolic state to utilize all the cellular resources of nucleotides and proteins and produce new viral cells²⁷. In SARS-CoV-2 infected cell lines, cell cycle arrest patterns were observed²⁸. However, in peripheral blood samples in our study as shown in Table 1, host cell cycle progression was induced. This is further supported by other studies on peripheral blood samples of infected patients in which cell cycle progression was the highest signal among other cellular, metabolic, and signaling pathways^{29–31}. This may be attributed to the fact that virally infected B cells signal cell cycle progression³². Intriguingly, cell cycle progression pathways are also enriched in SARS-CoV-2-infected human gastrointestinal cells³³. In addition to our study, the aforementioned studies confirm that the impact of SARS-CoV-2 on host cell cycle regulation is yet not fully understood³² and requires further *in vivo* testing and validation.

miRNAs are key players in gene expression regulation through interaction with genes, transcription factors, and other regulatory miRNAs. Additionally, they serve as biomarkers for the diagnosis of various diseases³⁴. In our study, hsa-mir-93-5p and hsa-mir-17-5p were found to be significantly upregulated. In a study on DEGs in peripheral blood samples of patients with SARS-CoV-2, these two miRNAs were deduced as potentially interacting miRNAs³⁰. hsa-mir-93-5p was deduced as a signature that may distinguish patients with COVID-19 from healthy individuals³⁵. These findings may nominate both miRNAs as candidate biomarkers for the viral infection using only blood samples from the patients. On the other hand, hsa-mir-21-5p, hsa-mir-10a-5p and hsa-mir-15b-5p were significantly downregulated in our analysis with AUC values of 0.9, 0.85, and 0.6, respectively. On a study on patients with SARS-CoV-2 infection, hsa-mir-21-5p were found to be downregulated in peripheral blood samples of both moderate and severe infections and correlated with immune hyperactivation³⁶. hsa-mir-10a-5p is involved in regulating host immune responses. For protection from host immune attacks, SARS-CoV-2 suppresses human hsa-mir-10a-5p and dysregulates its transcription¹⁵. hsa-mir-15b-5p was downregulated in our analysis and it was experimentally proven to bind to RNA-dependent RNA polymerase structure of SARS-CoV-2 leading to inhibition of viral replication⁴⁰. Through targeting this gene, hsa-mir-15b-5p acts as a natural antiviral agent and thus it may serve as a miRNA-based treatment for SARS-CoV-2³⁹. Both hsa-mir-107 and hsa-mir-103a-3p with AUC of 0.815 and 0.671, respectively were predicted to be bound via three binding sites to SARS-CoV-2 with perfect matching to the seed region³⁷. hsa-mir-103a-3p specifically binds to the viral nucleocapsid protein coding region³⁸.

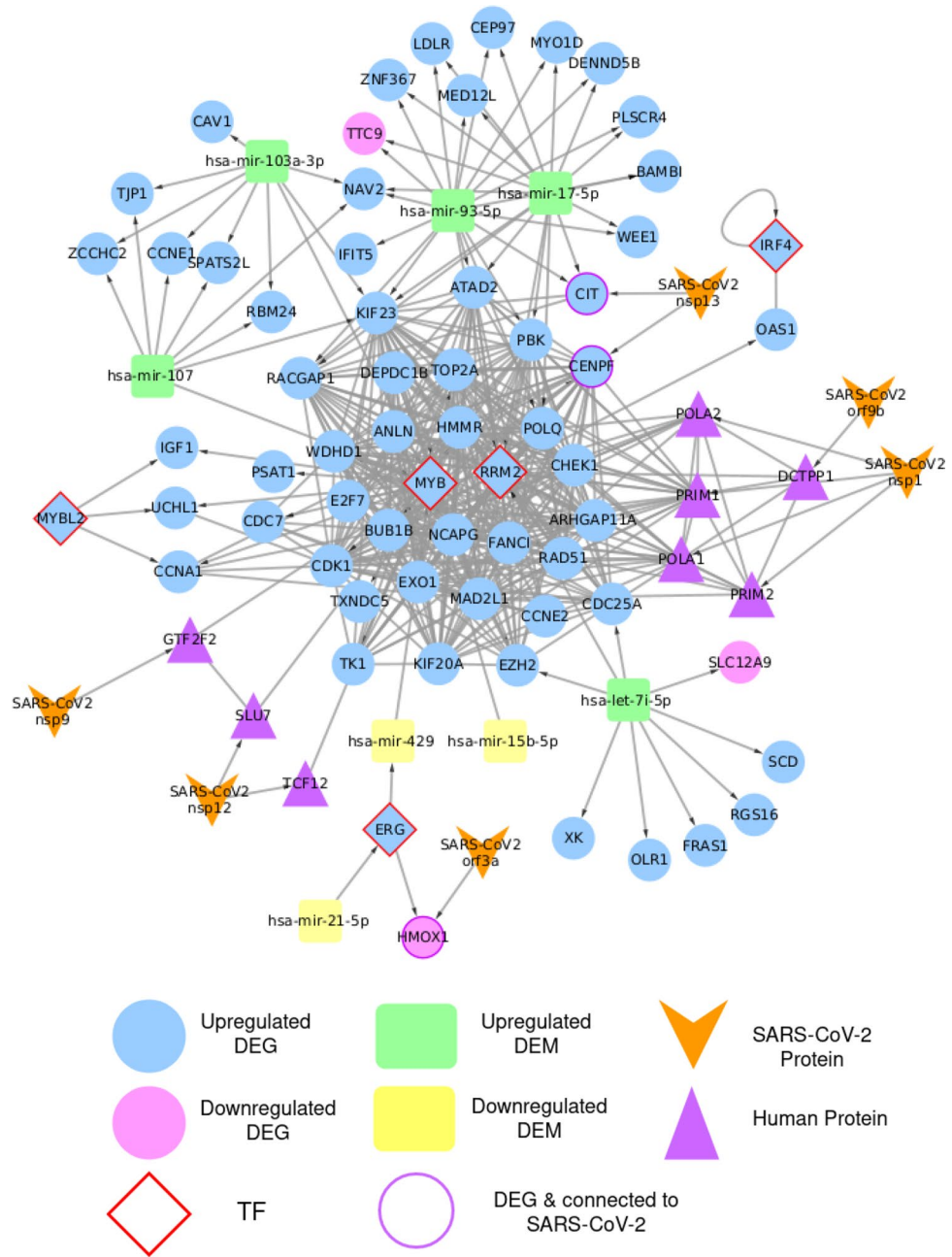


Figure 5. Core network consists of 86 nodes and 457 edges. It contains TFs, potential biomarkers, and their first neighbors. It also includes viral proteins connected to any of the selected nodes. Node’s color and shape are mapped to node type as labeled in the figure. Border is colored to highlight special cases of nodes. Directed edges represent regulatory and virus-host interactions while undirected edges represent human PPIs.

Transcription factors are regulators that govern the gene expression process in living organisms. Through binding to specific sites on the DNA, direct interaction with miRNAs, or both mechanisms, TFs play a vital role in cell differentiation and even human diseases⁴¹. Based on TFmiR results, five TFs (MYB, RRM2, ERG, MYBL2, and IRF4) were highlighted in the RNA regulatory network (Fig. 4). MYB is especially interesting because it was found to regulate 17 genes (TOP2A, CCNA1, CENPF, CDK1, BUB1B, CDC7, MAD2L1, RAD51, TK1, TXNDC5, PSAT1, HMMR, EXO1, UCHL1, TXNDC5, EZH2, WDHD1, IGF1) and be regulated by miRNAs (hsa-mir-107, hsa-mir-15b-5p, hsa-mir-103a-3p). Hence, we propose MYB as a golden drug target for COVID-19 treatment. In a study on transcription regulation in leukocytes of patients infected with SARS-CoV-2, MYB regulation mechanisms were upregulated in patients with poor prognosis and related to the severity of the disease⁴². Since many of the genes regulated by MYB are involved in the cell cycle, anticancer drugs might serve as potential repurposed drugs for COVID-19. According to DGIdb, drugs such as doxorubicin, paclitaxel and fluorouracil have an inhibitory effect on MYB⁴³. RRM2 (AUC=0.857) is one of the principal genes responsible for the de

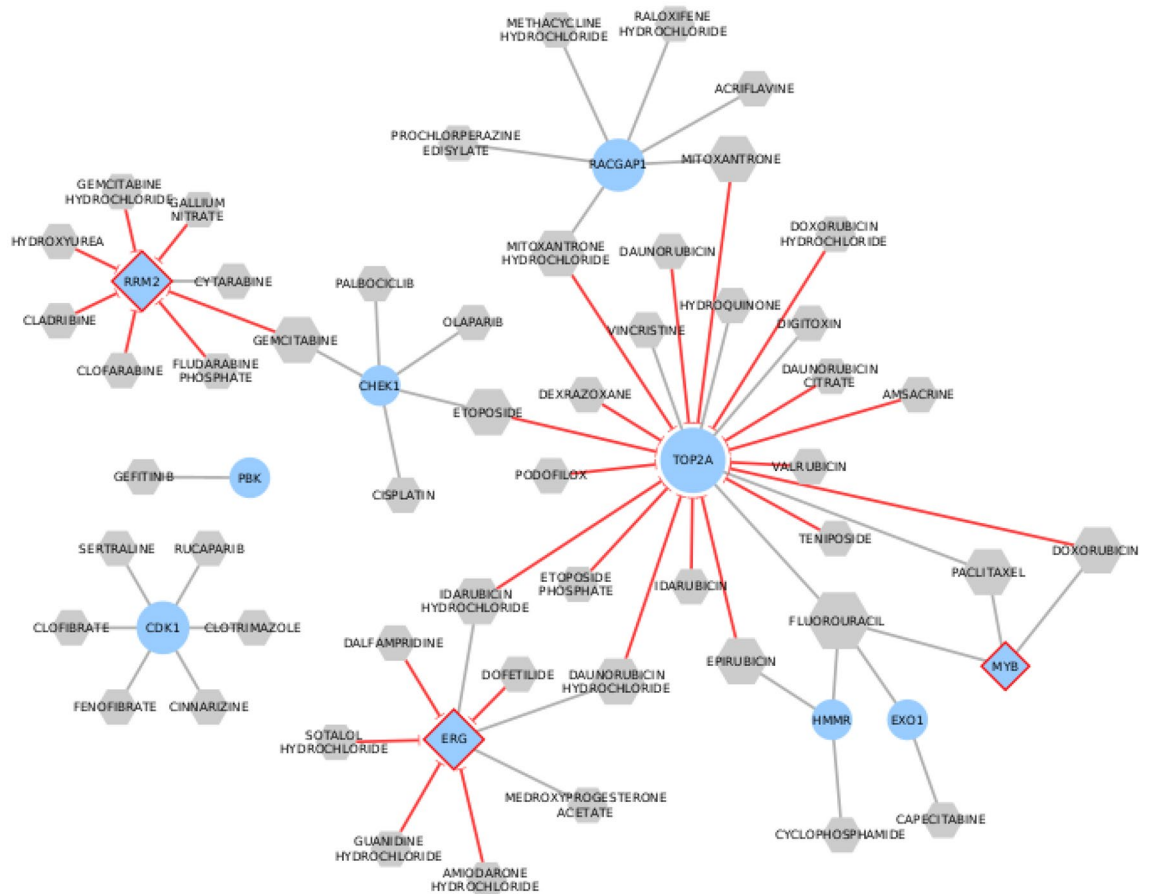


Figure 6. Drug-gene network consists of 62 nodes connected by 64 edges. It contains 7 upregulated DEGs (cyan ellipses), 3 TFs (cyan diamonds with red border), and 52 approved drugs (gray hexagons). Red edges with T shape arrows represent inhibitor or blocker interaction types while gray edges have missing interaction type. Top five targeted genes are represented by large nodes as well as the multi-targeted drugs.

novo synthesis of nucleotides during DNA replication. In addition to its role as a tumor promoter in various cancers, RRM2 was found to be actively interacting with proteins of RNA viruses such as mouse hepatitis virus and hepatitis C virus^{44,45}. The RRM2-viral proteins interactions further boost viral replication⁴⁵. Moreover, the E7 protein of the human papillomavirus induces the overexpression of RRM2; thus, it promotes human cervical cancer⁴⁶. Therefore, RRM2 is a promising target for repurposed anti-cancers as it was found to be overexpressed in our study and probably interact with SARS-CoV-2 proteins. Gemcitabine, cytarabine, and hydroxyurea are among the FDA-approved anti-cancer drugs that target RRM2 and are suggested by the drug-gene network. ERG (AUC=0.844) is a transcription factor that plays a role in embryo development. It is also an oncogene that contributes to leukemia and Ewing's sarcoma. Lately, ERG has been associated with prostate cancer through gene fusion to TMPRSS2⁴⁷. TMPRSS2 has gained particular interest since the beginning of the pandemic as it facilitates the entry of SARS-CoV-2 into host cells and viral activation. Ergo, patients with prostate cancer might experience unusual pathogenesis of COVID-19⁴⁸. In our analysis, ERG was overexpressed and notably interacting with multiple genes. It can be targeted by many drugs, such as the two FDA-approved antiarrhythmic drugs, sotalol, and amiodarone.

Coronaviruses proteins might directly interact with the host cell proteins involved in the cell cycle to recruit the necessary machinery for viral replication²⁷. CENPF and CIT are two proteins involved in the stages of cell cycle progression. In our analysis as illustrated in Fig. 5, they were differentially expressed and bound to nsp13 of SARS-CoV-2. Nsp13 helicase is highly conserved in the coronaviruses family and pivotal in viral replication. Further studies are needed to investigate the effect of targeting CENPF and CIT as a possible therapeutic strategy to treat COVID-19. It is worth mentioning that the overexpression of CENPF and CIT, as is the case in this study, has been associated with different types of cancer such as cervical cancer, hepatocellular carcinoma, osteosarcoma, bladder cancer, and esophageal squamous cell carcinoma⁴⁹⁻⁵³. Targeting MYB may be also an alternative strategy to inhibit CENPF and indirectly control viral replication. Another SARS-CoV-2 protein found to bind to HMOX1 host protein is ORF3a. ORF3a is a transmembrane protein that aids in translocating the replicated virions outside the infected cells. ORF3a might also induce cellular apoptosis, and it is involved in the host immune response to the virus^{54,55}. HMOX1 encodes for heme oxygenase protein (HO-1) that catabolizes human heme. Lungs are one of the main sites of HO-1 expression for protective purposes. In viral infections, the

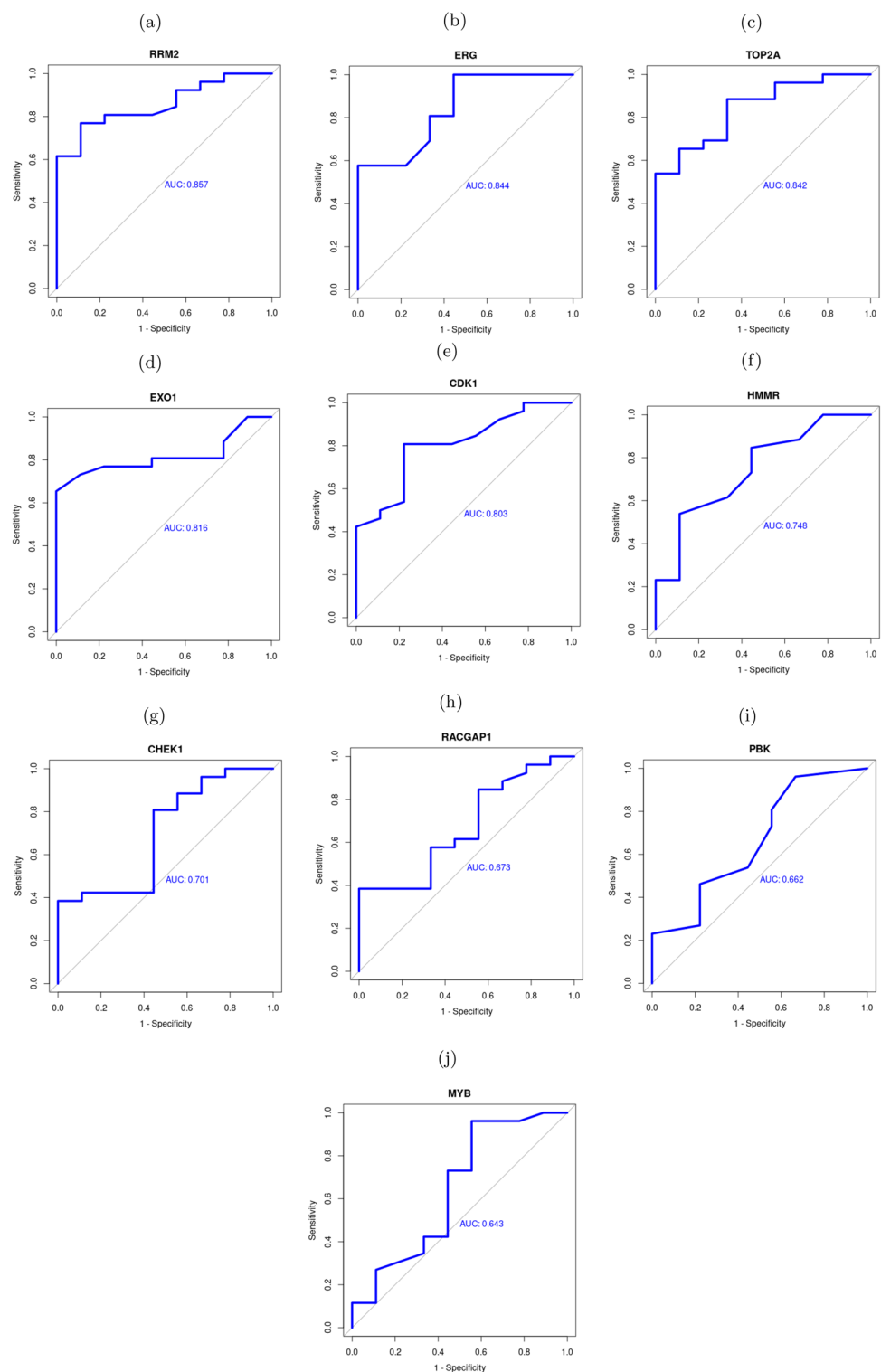


Figure 7. ROC curves of 10 hub genes between COVID-19 patients and healthy controls.

expression of HO-1 decreases, which is further supported by our study as a differentially expressed downregulated protein. HMOX1 expression might be a prognostic biomarker for COVID-19 disease⁵⁶. In addition, the HMOX1 pathway regulates platelets aggregation and poses anti-inflammatory effects. As a defense mechanism, ORF3a binds to HMOX1 to inhibit mechanisms hindering viral replication. Thus, drugs targeting HMOX1-ORF3a binding or enhancing expression of HMOX1 gene can be suggested as new alternatives for COVID-19 treatment. A study proposed the use of the leaf extract of the neem plant to overexpress HMOX1 gene and target SARS-CoV-2 infection⁵⁷.

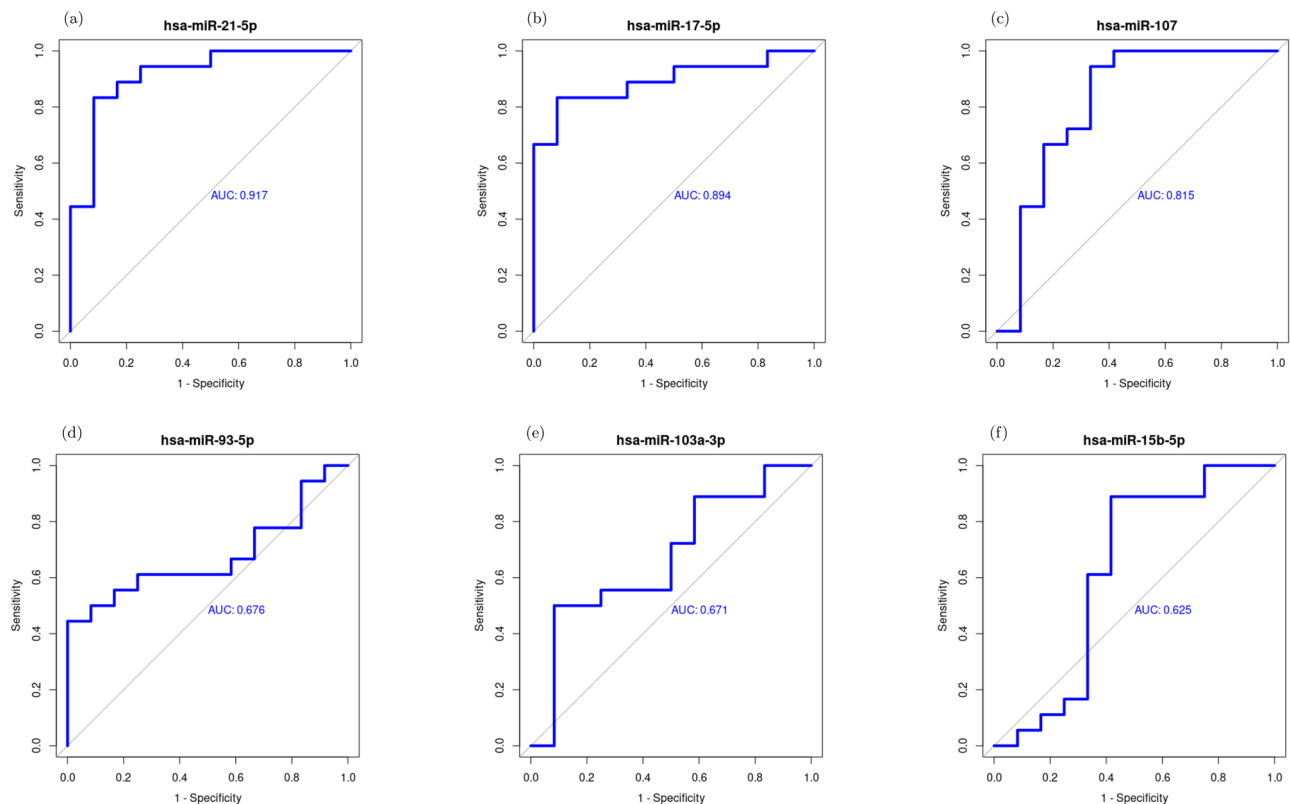


Figure 8. ROC curves of the identified DEMs of the Core network between COVID-19 patients and healthy controls.

To conclude the potential repurposed drugs for SARS-CoV-2, we constructed a drug-gene network (Fig. 6) based on the potential drug targets identified in the Core network (Fig. 5). Since the cell cycle progression pathways were significantly enriched in our study, most of the suggested drugs that target genes involved in the cell cycle. Multiple studies also supported the repurposing of anticancer drugs to treat COVID-19^{26,30,58–63}. This can be explained by the similar influence that pathogens and cancers have on the biologically active biological processes in the human body²⁶.

Using DGIdb, all sources of drug-gene interactions were included to ensure a comprehensive thorough search. Among the 25 top ranked genes and TFs, only 10 genes showed interactions with 52 FDA-approved drugs: CDK1, HMMR, EXO1, PBK, RRM2, CHEK1, RACGAP1, MYB, ERG, and TOP2A. TOP2A (AUC = 0.842) is a crucial gene in the DNA replication process during the cell cycle. It was identified as a central hub gene in a study on comparing the gene expression profiles of blood samples from patients with COVID-19 with healthy controls⁶⁴. According to the drug-gene network constructed, it was inhibited by different classes of anti-cancers such as anthracyclines, mitotic inhibitors, anti-tumor antibiotics, and topoisomerase II inhibitors. Etoposide is a topoisomerase II inhibitor that is currently under phase II clinical trial for COVID-19 therapy⁶⁵. Moreover, hydroxyquinone and digitoxin also have a potential inhibitory effect on TOP2A. CDK1 (AUC=0.803) is a cyclin-dependent kinase that modulates cell cycle progression. Several viruses can control the expression of CDKs to create a suitable cellular environment for viral replication. CDK inhibitors have been proposed to limit viral infectivity to the cells in general, and SARS-CoV-2 in particular⁶². CDK1 was an overexpressed hub gene in our study and previous studies³⁰. Subsequently, CDK1 inhibition may offer a therapeutic strategy to fight COVID-19. According to our drug-gene network, this can be achieved by repurposing riviciclib, dinaciclib, milciclib, and others.

Drugs targeting multiple genes are of particular interest within the drug-gene network. This can be explained by the fact that these repurposed drugs, when given as single agents, would be more efficient and have fewer side effects than drug combinations. For example, in our study, fluorouracil targets four genes with a transcription factor included. Carmofur, a fluorouracil-based anti-cancer, and doxorubicin have been proposed for COVID-19 treatment^{30,65}. Despite being promising drug targets, more studies need to be conducted to validate the use of repurposed anti-cancer drugs to fight COVID-19 with the proper dosing. This is due to the detrimental side effects of anti-cancers on normal cells and the quality of life in general.

Conclusion

We sought potential drug targets, biomarkers, and repurposed FDA-approved drugs to help early diagnosis and treatment of COVID-19. Our integrative bioinformatics analyses identified MYB, RRM2, ERG, CENPF, CIT, TOP2A, and HMOX1 as candidate drug targets. Both hsa-mir-93-5p and hsa-mir-17-5p were identified as

candidate biomarkers for COVID-19 infection. In addition, hsa-mir-15b-5p is a promising miRNA-based anti-viral agent. The key drugs retrieved were members of anthracyclines, mitotic inhibitors, anti-tumor antibiotics, and CDK1 inhibitors. Additionally, hydroxyquinone and digitoxin are potent TOP2A inhibitors. The analysis enforced the repositioning of fluorouracil and doxorubicin, especially those with multiple drug targets assuming less probability of resistance. Our study provides promising results that could facilitate developing diagnostic and therapeutic approaches against the COVID-19 pandemic.

Methods

Data retrieval. Two gene expression profiling datasets were retrieved from the Gene Expression Omnibus (GEO) repository. The first RNA-seq data (Accession no. GSE157103)⁶⁶ was generated on Illumina NovaSeq 6000 platform and consisted of 126 blood samples; 100 COVID-19 patients and 26 non-COVID-19 samples. The dataset included additional metadata about the samples, such as age, gender, and clinical data. It had 74 males and 52 females. The age range was from 21 to 89 years. PCA has been utilized to ensure that factors such as age, gender, or Charlson severity index score do not affect gene expression. The second dataset (Accession no. GSE148729)⁶⁷ was generated on Illumina NextSeq 500 platform. It consisted of different gene expression profiling measurements using bulk and single-cell sequencing. We only retrieved read counts of miRNAs from mock and infected epithelial lung cancer cell lines (Calu-3) 24 h post-infection (two replicates for each group) for the expression analysis.

Gene expression and pathway enrichment analyses. Differential expression analysis (DEA) was performed to identify the differentially expressed genes (DEGs) in the COVID-19 patients and the differentially expressed miRNAs (DEMs) in the SARS-CoV-2 infected human cell lines, respectively. In both cases, DESeq2 Bioconductor package⁶⁸ was used to run the analysis. Messenger RNAs (mRNAs) and miRNAs with adjusted P -value < 0.05 and $|Log_2FC| > 0.5$ were considered as DEGs and DEMs, respectively. To reveal biological insight into the resulting DEGs and their association with the virus infection, pathway enrichment analysis was performed using Reactome analysis tool⁶⁹. False discovery rate (FDR) cutoff of 0.01 was used to highlight the significantly enriched pathways.

Network construction and analysis. Multiple networks were constructed for an integrated analysis of the identified DEGs and DEMs as follows:

1. An RNA regulatory network between the identified DEGs and DEMs was created using the TFmiR web server⁷⁰. TFmiR highlights transcription factors (TFs) if they exist in the given input of DEGs and provides their interactions with the given miRNAs and genes in the network.
2. For SARS-CoV-2-human protein-protein interactions (PPIs) retrieval, we used Gordon et al. study²⁴. Using affinity-purification mass spectrometry, the authors reported the human proteins that were physically associated with SARS-CoV-2 expressed proteins in human cells. The dataset is publicly available at STRING v11.5 database⁷¹. It contains 332 human proteins connected to 27 SARS-CoV-2 proteins.
3. STRING v11.5 database was used to retrieve PPIs of the identified DEGs, the human proteins identified in the second network, and between both categories. Co-expression, experimental validation, and information from databases were the only interaction evidences included in the network with a minimum confidence score of 0.4.

All the aforementioned different types of interactions were integrated into one complete network, referred to as the “All in One network” using Cytoscape v3.9.0 software⁷². It has also been utilized in visualization and analysis. The “Core network” was extracted from the All in One network to understand the complex interactions better. It includes TFs, upregulated DEMs identified from the RNA regulatory network, and their first neighboring nodes. We also added the interacting viral proteins with any of the mentioned nodes. Then, Maximal Clique Centrality (MCC) algorithm was employed to identify and prioritize the hub genes for further investigation using Cytoscape plugin cytoHubba⁷³. Among the 11 topological analysis methods provided by cytoHubba, we selected MCC for ranking as it showed more precise performance over other centrality measures such as degree, closeness, and betweenness⁷³.

Drug repurposing. Drug repurposing is considered an effective alternative strategy for the emerging pandemic. Accordingly, we chose the Drug Gene Interaction Database (DGIdb) v4.2.0⁴³ to suggest drugs that target both the top ranked genes of the Core network and the previously identified TFs. DGIdb is an expert-curated database that integrates 100, 273 drug-gene interactions from 22 source databases such as DrugBank⁷⁴, Drug Target Commons⁷⁵, PharmGKB⁷⁶, and ChEMBL⁷⁷. It has more than 40, 000 genes and 10, 000 drugs from different drug classes. To filter the interacting drugs with the query genes, we chose only approved drugs filter to make sure we included all approved drug classes and not only antineoplastics or immunotherapies. We also included all drug interactions databases, all gene categories, and all drug-gene interaction types.

ROC curve analysis. The receiver operating characteristic (ROC) curve analysis was performed to assess the ability of the identified hub genes and DEMs of the Core network to distinguish between COVID-19 patients and healthy controls groups. pROC R package⁷⁸ was used to generate ROC curves and calculate the area under the curve (AUC) values for the hub genes and DEMs from GEO datasets GSE196822 and GSE176498⁷⁹, respectively.

Data availability

The read counts of gene expression profiling datasets used for differential expression analysis and ROC analysis are available in the GEO repository. GSE157103 for Overmyer et al.⁶⁶, GSE148729 for Wyler et al.⁶⁷, GSE196822 for Banerjee U et al., and GSE176498 for Gutmann C et al.⁷⁹. All results generated by the study analyses are included in this published article (and its Supplementary Information file).

Code availability

The code for reproducing the transcriptomic analysis results will be made available at the Github repository See Code <https://github.com/ComputationalBiologyLab/Transcriptomic-analysis-of-SARS-CoV-2>.

Received: 11 February 2022; Accepted: 30 June 2022

Published online: 13 July 2022

References

1. Organization, W. H. Who coronavirus (COVID-19) dashboard. <https://covid19.who.int/> (2022).
2. Hui, D. S. *et al.* The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int. J. Infect. Dis.* **91**, 264–266 (2020).
3. Zhou, P. *et al.* Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. *BioRxiv* (2020).
4. Gorbalenya, A. E. *et al.* The species Severe acute respiratory syndrome-related coronavirus: Classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **5**, 536 (2020).
5. Organization, W. H. Who director-general's opening remarks at the media briefing on COVID-19-11 march 2020. <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020> (2020).
6. Organization, W. H. Update on omicron. <https://www.who.int/news/item/28-11-2021-update-on-omicron> (2021).
7. CDC, C. D. T. Global variants report. <https://covid.cdc.gov/covid-data-tracker/#global-variant-report-map> (2022).
8. Mason, R. J. Pathogenesis of COVID-19 from a cell biology perspective (2020).
9. Wu, Z. & McGoogan, J. M. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72 314 cases from the Chinese center for disease control and prevention. *Jama* **323**, 1239–1242 (2020).
10. Yuan, X., Shan, Y., Zhao, Z., Chen, J. & Cong, Y. G0/G1 arrest and apoptosis induced by SARS-CoV 3b protein in transfected cells. *Virology* **421**, 1–5 (2005).
11. Sun, P. *et al.* Porcine epidemic diarrhea virus through P53-dependent pathway causes cell cycle arrest in the G0/G1 phase. *Virus Res.* **253**, 1–11 (2018).
12. Chen, C.-J. & Makino, S. Murine coronavirus replication induces cell cycle arrest in G0/G1 phase. *J. Virol.* **78**, 5658–5669 (2004).
13. Li, F. Q., Tam, J. P. & Liu, D. X. Cell cycle arrest and apoptosis induced by the coronavirus infectious bronchitis virus in the absence of p53. *Virology* **365**, 435–445 (2007).
14. Ding, L. *et al.* Transmissible gastroenteritis virus infection induces cell cycle arrest at S and G2/M phases via P53-dependent pathway. *Virus Res.* **178**, 241–251 (2013).
15. Abedi, F., Rezaee, R., Hayes, A. W., Nasiripour, S. & Karimi, G. MicroRNAs and SARS-CoV-2 life cycle, pathogenesis, and mutations: Biomarkers or therapeutic agents?. *Cell Cycle* **20**, 143–153 (2021).
16. Bushati, N. & Cohen, S. M. microRNA functions. *Annu. Rev. Cell Dev. Biol.* **23**, 175–205 (2007).
17. Bernier, A. & Sagan, S. M. The diverse roles of microRNAs at the host-virus interface. *Viruses* **10**, 440 (2018).
18. Nersisyan, S., Shkurnikov, M., Turchinovich, A., Knyazev, E. & Tonevitsky, A. Integrative analysis of miRNA and mRNA sequencing data reveals potential regulatory mechanisms of ACE2 and TMPRSS2. *PLoS One* **15**, e0235987 (2020).
19. Matarese, A., Gambardella, J., Sardu, C. & Santulli, G. miR-98 regulates TMPRSS2 expression in human endothelial cells: Key implications for COVID-19. *Biomedicines* **8**, 462 (2020).
20. Khan, M. *et al.* Epigenetic regulator miRNA pattern differences among SARS-CoV, SARS-CoV-2, and SARS-CoV-2 world-wide isolates delineated the mystery behind the epic pathogenicity and distinct clinical characteristics of pandemic COVID-19. *Front. Genet.* **11**, 765 (2020).
21. Fulzele, S. *et al.* COVID-19 virulence in aged patients might be impacted by the host cellular MicroRNAs abundance/profile. *Aging Dis.* **11**, 509 (2020).
22. Bagga, S. & Bouchard, M. J. Cell cycle regulation during viral infection. *Cell Cycle Control* 165–227 (2014).
23. Pushpakom, S. *et al.* Drug repurposing: Progress, challenges and recommendations. *Nat. Rev. Drug Discov.* **18**, 41–58 (2019).
24. Gordon, D. E. *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).
25. Suryawanshi, R. K., Koganti, R., Agelidis, A., Patil, C. D. & Shukla, D. Dysregulation of Cell Signaling by SARS-CoV-2. *Trends Microbiol.* (2020).
26. Tutuncuoglu, B. *et al.* The landscape of human cancer proteins targeted by SARS-CoV-2. *Cancer Discov.* **10**, 916–921 (2020).
27. Su, M. *et al.* A mini-review on cell cycle regulation of coronavirus infection. *Front. Vet. Sci.* **7**, 943 (2020).
28. Bouhaddou, M. *et al.* The global phosphorylation landscape of SARS-CoV-2 infection. *Cell* **182**, 685–712 (2020).
29. Gao, X. *et al.* Genome-wide screening of SARS-CoV-2 infection-related genes based on the blood leucocytes sequencing dataset of patients with COVID-19. *J. Med. Virol.* (2021).
30. Shahjaman, M., Rahman, M. R. & Auwul, M. R. A network-based systems biology approach for identification of shared gene signatures between male and female in COVID-19 datasets. *Inform. Med. Unlocked* **25**, 100702 (2021).
31. Mo, S. *et al.* Comprehensive analysis of the systemic transcriptomic alternations and inflammatory response during the occurrence and progress of COVID-19. *Oxid. Med. Cell. Longev.* **2021** (2021).
32. Vázquez-Jiménez, A. *et al.* On deep landscape exploration of COVID-19 patients cells and severity markers. *Front. Immunol.* **12** (2021).
33. Chen, J.-C. *et al.* Identification of key pathways and genes in SARS-CoV-2 infecting human intestines by bioinformatics analysis. *Biochem. Genet.* 1–19 (2021).
34. O'Brien, J., Hayder, H., Zayed, Y. & Peng, C. Overview of MicroRNA biogenesis, mechanisms of actions, and circulation. *Front. Endocrinol.* **9**, 402 (2018).
35. Farr, R. J. *et al.* Detection of sars-cov-2 infection by microRNA profiling of the upper respiratory tract. *Plos ONE* **17**, e0265670 (2022).
36. Tang, H. *et al.* The noncoding and coding transcriptional landscape of the peripheral immune response in patients with COVID-19. *Clini. Transl. Med.* **10**, e200 (2020).
37. Jafarinejad-Farsangi, S., Jazi, M. M., Rostamzadeh, F. & Hadizadeh, M. High affinity of host human microRNAs to sars-cov-2 genome: An in silico analysis. *Non-coding RNA Res.* **5**, 222–231 (2020).

38. Bertolazzi, G., Cipollina, C., Benos, P. V., Tumminello, M. & Coronello, C. mir-1207-5p can contribute to dysregulation of inflammatory response in COVID-19 via targeting sars-cov-2 rna. *Front. Cell. Infect. Microbiol.* **673** (2020).
39. Sardar, R., Satish, D., Birla, S. & Gupta, D. Integrative analyses of SARS-CoV-2 genomes from different geographical locations reveal unique features potentially consequential to host-virus interaction, pathogenesis and clues for novel therapies. *Heliyon* **6**, e04658 (2020).
40. Sato, A., Ogino, Y., Tanuma, S.-I. & Uchiyama, F. Human microRNA hsa-mir-15b-5p targets the rna template component of the rna-dependent rna polymerase structure in severe acute respiratory syndrome coronavirus 2. *Nucleosides Nucleotides Nucleic Acids* **40**, 790–797 (2021).
41. Lambert, S. A. *et al.* The human transcription factors. *Cell* **172**, 650–665 (2018).
42. Lam, M. T. Y. *et al.* Profiling transcription initiation in peripheral leukocytes reveals severity-associated cis-regulatory elements in critical COVID-19. *BioRxiv* (2021).
43. Freshour, S. L. *et al.* Integration of the drug-gene interaction database (DGIdb 4.0) with open crowdsourcing efforts. *Nucleic Acids Res.* **49**, D1144–D1151 (2021).
44. Zhan, Y. *et al.* Inhibiting RRM2 to enhance the anticancer activity of chemotherapy. *Biomed. Pharmacother.* **133**, 110996 (2021).
45. Perrin-Cocon, L. *et al.* The current landscape of coronavirus-host protein-protein interactions. *J. Transl. Med.* **18**, 1–15 (2020).
46. Wang, N. *et al.* Increased expression of RRM2 by human papillomavirus E7 oncoprotein promotes angiogenesis in cervical cancer. *Br. J. Cancer* **110**, 1034–1044 (2014).
47. Adamo, P. & Ladomery, M. The oncogene ERG: A key factor in prostate cancer. *Oncogene* **35**, 403–414 (2016).
48. Afshari, A., Janfeshan, S., Yaghobi, R., Roozbeh, J. & Azarpira, N. COVID-19 pathogenesis in prostatic cancer and TMPRSS2-ERG regulatory genetic pathway. *Infect. Genet. Evol.* **88**, 104669 (2021).
49. Yu, B. *et al.* TOP2A and CENPF are synergistic master regulators activated in cervical cancer. *BMC Med. Genom.* **13**, 1–17 (2020).
50. Huang, Y. *et al.* Centromere Protein F (CENPF) serves as a potential prognostic biomarker and target for human hepatocellular carcinoma. *J. Cancer* **12**, 2933 (2021).
51. Zou, P.-A., Yang, Z.-X., Wang, X. & Tao, Z.-W. Upregulation of CENPF is linked to aggressive features of osteosarcoma. *Oncol. Lett.* **22**, 1–9 (2021).
52. Liu, Z. *et al.* Down-regulation of CIT can inhibit the growth of human bladder cancer cells. *Biomed. Pharmacother.* **124**, 109830 (2020).
53. Lu, W. *et al.* High expression of citron kinase contributes to the development of esophageal squamous cell carcinoma. *Front. Genet.* **12** (2021).
54. Ren, Y. *et al.* The ORF3a protein of SARS-CoV-2 induces apoptosis in cells. *Cell. Mol. Immunol.* **17**, 881–883 (2020).
55. Bianchi, M., Borsetti, A., Ciccozzi, M. & Pascarella, S. SARS-CoV-2 ORF3a: Mutability and function. *Int. J. Biol. Macromol.* **170**, 820–826 (2021).
56. Dattilo, M. The role of host defences in Covid 19 and treatments thereof. *Mol. Med.* **26**, 1–15 (2020).
57. Batra, N., De Souza, C., Batra, J., Raetz, A. G. & Yu, A.-M. The HMOX1 pathway as a promising target for the treatment and prevention of SARS-CoV-2 of 2019 (COVID-19). *Int. J. Mol. Sci.* **21**, 6412 (2020).
58. Borgo, C., D'Amore, C., Sarno, S., Salvi, M. & Ruzzene, M. Protein kinase CK2: A potential therapeutic target for diverse human diseases. *Signal Transduct. Target. Ther.* **6**, 1–20 (2021).
59. Stolfi, P., Manni, L., Soligo, M., Vergni, D. & Tieri, P. Designing a network proximity-based drug repurposing strategy for COVID-19. *Front. Cell Dev. Biol.* **8** (2020).
60. Pillaiyar, T. & Laufer, S. Kinases as potential therapeutic targets for anti-coronaviral therapy. *J. Med. Chem.* (2021).
61. Kim, Y.-J., Witwit, H., Cubitt, B. & de la Torre, J. C. Inhibitors of anti-apoptotic Bcl-2 family proteins exhibit potent and broad-spectrum anti-mammarenavirus activity via cell cycle arrest at G0/G1 phase. *J. Virol.* **95**, e01399-21 (2021).
62. Weisberg, E. *et al.* Repurposing of kinase inhibitors for treatment of COVID-19. *Pharm. Res.* **37**, 1–29 (2020).
63. Krishnamoorthy, P., Raj, A. S., Roy, S., Kumar, N. S. & Kumar, H. Comparative transcriptome analysis of SARS-CoV, MERS-CoV, and SARS-CoV-2 to identify potential pathways for drug repurposing. *Comput. Biol. Med.* **128**, 104123 (2021).
64. Hasan, M. I. *et al.* Systems biology and bioinformatics approach to identify blood based signatures molecules and drug targets of patient with COVID-19. *Inform. Med. Unlocked* **28**, 100840 (2022).
65. Poduri, R., Joshi, G. & Jagadeesh, G. Drugs targeting various stages of the SARS-CoV-2 life cycle: Exploring promising drugs for the treatment of COVID-19. *Cellul. Signal.* **74**, 109721 (2020).
66. Overmyer, K. A. *et al.* Large-scale multi-omic analysis of COVID-19 severity. *Cell Syst.* **12**, 23–40 (2021).
67. Wylter, E. *et al.* Transcriptomic profiling of SARS-CoV-2 infected human cell lines identifies HSP90 as target for COVID-19 therapy. *Science* **24**, 102151 (2021).
68. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
69. Fabregat, A. *et al.* The reactome pathway knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).
70. Hamed, M., Spaniol, C., Nazarieh, M. & Helms, V. TFmiR: A web server for constructing and analyzing disease-specific transcription factor and miRNA co-regulatory networks. *Nucleic Acids Res.* **43**, W283–W288 (2015).
71. Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
72. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
73. Chin, C.-H. *et al.* cytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* **8**, 1–7 (2014).
74. Wishart, D. S. *et al.* Drugbank 5.0: A major update to the drugbank database for 2018. *Nucleic Acids Res.* **46**, D1074–D1082 (2018).
75. Tang, J. *et al.* Drug target commons: A community effort to build a consensus knowledge base for drug-target interactions. *Cell Chem. Biol.* **25**, 224–229 (2018).
76. Thorn, C. F., Klein, T. E. & Altman, R. B. Pharmgkb. *Pharmacogenomics* 179–191 (2005).
77. Gaulton, A. *et al.* The chembl database in 2017. *Nucleic Acids Res.* **45**, D945–D954 (2017).
78. Robin, X. *et al.* Proc: An open-source package for r and s+ to analyze and compare roc curves. *BMC Bioinform.* **12**, 1–8 (2011).
79. Gutmann, C. *et al.* Association of cardiometabolic micRNAs with COVID-19 severity and mortality. *Cardiovasc. Res.* **118**, 461–474 (2022).

Author contributions

A.S. and M.M. developed the pipeline, N.A. performed the biological analysis, E.B. conducted the project conceptualization and administration, A.S. and E.B. worked on the methodology, A.S., M.M., and N.A. worked on the original draft preparation, E.B. conducted the manuscript review and editing.

Funding

Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-15898-w>.

Correspondence and requests for materials should be addressed to E.B.

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.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022