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.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Identifying potential novel insights for COVID-19 pathogenesis and therapeutics using an integrated bioinformatics analysis of host transcriptome

Salem A. El-aarag a,*, Amal Mahmoud b,*, Mahmoud ElHefnawi c,*

a Bioinformatics Department, Genetic Engineering and Biotechnology Research Institute (GEBRI), University of Sadat City, Egypt
b Department of Biology, College of Science, Imam Abdulrahman Ibn Faisal University, P.O. Box. 1982, 31441 Dammam, Saudi Arabia
c Biomedical informatics and chemoinformatics group, Informatics and systems department, National Research Center, Egypt

ARTICLE INFO

Keywords: COVID-19 Systems biology Differentially expressed genes Functional module analysis Protein-protein interaction network Enrichment analysis Host transcriptome and drug repurposing

ABSTRACT

The molecular mechanisms underlying the pathogenesis of COVID-19 have not been fully discovered. This study aims to decipher potentially hidden parts of the pathogenesis of COVID-19, potential novel drug targets, and identify potential drug candidates. Two gene expression profiles were analyzed, and overlapping differentially expressed genes (DEGs) were selected for which top enriched transcription factors and kinases were identified, and pathway analysis was performed. Protein-protein interaction (PPI) of DEGs was constructed, hub genes were identified, and module analysis was also performed. DGIdb database was used to identify drugs for the potential targets (hub genes and the most enriched transcription factors and kinases for DEGs). A drug-potential target network was constructed, and drugs were ranked according to the degree. L1000FDW was used to identify drugs that can reverse transcriptional profiles of COVID-19.

We identified drugs currently in clinical trials, others predicted by different methods, and novel potential drug candidates Entrectinib, Omeprazole, and Exemestane for combating COVID-19. Besides the well-known pathogenic pathways, it was found that axon guidance is a potential pathogenic pathway. Sema7A, which may exacerbate hypercytokinemia, is considered a potential novel drug target. Another potential novel pathway is related to TINF2 overexpression, which may induce potential telomere dysfunction and damage DNA that may exacerbate lung fibrosis. This study identified new potential insights regarding COVID-19 pathogenesis and treatment, which might help us improve our understanding of the mechanisms of COVID-19.

1. Background

Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) is a member of the Coronaviridae family, isolated at the end of 2019. SARS-CoV-2 pandemic has about a 3.06% mortality rate, according to WHO. Nearly the majority of COVID-19 patients, about 80%, have mild symptoms. However, about 20% of cases are severe and develop severe respiratory distress syndrome [1].

Several studies reported that dysregulated immune response (hypercytokinemia) is associated with severe cases [2–7]. Uncontrolled host immune response may be responsible for the fatal outcome rather than unabated viral replication. It was observed that fatal cases had a low viral load [8], so understanding the host response to SARS-CoV-2 is critical to find potential therapeutic targets for severe COVID-19 disease [8].

This study tried to understand the underlying pathological pathways related to SARS-CoV-2 infection and identify potential drug targets and drugs that can combat SARS-CoV-2-induced injuries. Two RNA-seq datasets (GSE147507-GSE153970) were analyzed, and DEGs were identified based on the two datasets. Gene ontology and the Kyoto encyclopedia of genes and genomes pathway enrichment analysis for DEGs were performed. Also, top enriched transcription factors and kinases for DEGs were identified. Protein-protein interaction (PPI) of DEGs was constructed, and hub genes were identified. DGIdb database was used to identify drugs for the potential targets (hub genes and the most enriched transcription factors and kinases for DEGs). A drug-
potential target network was constructed, and drugs were ranked according to the degree. L1000FDW was searched for drugs that can reverse transcriptional profiles of COVID-19. PPI interaction Network of DEGs was expanded, and GO biological process enrichment analysis and KEGG pathway analysis of the top eight modules were performed [Fig. 1] [9–11]. We identified drugs already in use in clinical trials against COVID-19 disease as well as new potential drugs. We also identified new potential drug targets and new pathogenic pathways for COVID-19 disease.

1.1. Datasets

Two gene expression profiles (GSE147507-GSE15397) were obtained from the NCBI Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). The first dataset is stranded RNA-seq data from mock-treated and SARS-CoV-2 (USA-WA1/2020 at MOI of 2 for 24 h.)-infected normal human bronchial epithelial (NHBE) cells. Each group included three independent samples. Each sample has four runs. The second dataset is unstranded RNA-seq data from mock-treated and SARS-CoV-2 (USA-WA1/2020 at MOI of 0.5 for 48 h.)-infected Normal Airway Epithelial cells (NAE). Each group included three independent samples. Each sample has two runs.

1.2. Identifying DEGs

We used Galaxy [https://usegalaxy.org/] as a platform for RNA-seq analysis. First, raw reads were subjected to quality control analysis using FastQC [12] before and after using the trimmomatic tool [13], which was run with its default options. Trimmed reads were aligned against the reference genome (hg19) using HISAT2 [14]. All aligned runs for the same sample were merged using MergeSamFiles [15]. Aligned reads that overlap features in a GFF file (gencode.v19.annotation.gtf.gz) were counted using htsq-count [16]. DESeq2 [17] was used to determine differentially expressed features count tables. Differentially expressed genes (DEGs) were considered using (q. val < 0.05). The intersection of the DEGs was illustrated using VENNY 2.1 (https://bioinfogp.cnb.csic.es/tools/venny/index.html).

1.3. Pathway and GO enrichment Analysis of DEGs

Gene Ontology (GO) term analysis and KEGG pathway enrichment analysis of DEGs were performed using STRING (https://string-db.org/). False discovery rate < 0.05 was considered statistically significant. Transcription factors enrichment analysis and kinases enrichment analysis of DEGs were performed using ChEA3 [18] and KEA3 [19], respectively using the integrated mean rank method. The integrated mean rank method is the mean rank of each TF/kinase across all libraries containing that TF/kinase in which enrichment results from all libraries are integrated to generate this composite rank. This rank provides better prediction of the correct upstream TF/kinase compared to the ranks produced by individual libraries. Integrated mean ranks are sorted in ascending order by score so that TF/Kinase, which has a lower score, is the most relevant.

1.4. Hub genes identification

PPI interaction network was obtained by submitting DEGs to the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) with interaction score > 0.4 and analyzed by Cytoscape software (version 3.8.1) [20]. STRING main sources of interactions are co-expression, text-mining, gene fusion, neighbourhood, and experimental data. Hub genes were selected based on the intersection of three centrality methods: subgraph, degree, and maximal clique centrality (MCC). Subgraph and degree were calculated using CytoNCA [21]. However, maximal clique centrality (MCC) was calculated using CytoHubba [22].

Degree centrality is the number of interactions that a node has. The nodes with the highest number of interactions are often nodes of interest [23]. Subgraph centrality measures the participation of each node in all subgraphs in a network, with smaller subgraphs having higher importance [24]. CytoHubba employed several topological algorithms to identify significant nodes in a given network. We select MCC algorithm for hub genes identification as it was reported that MCC performed better than 11 centrality methods [22].

1.5. Identifying potential drugs

1.5.1. Drug-potential target network

Using the Drug-Gene Interaction Database (DGIdb; http://www.dgidb.org/search_interactions), drugs were selected based on the hub genes and the most enriched transcription factors and kinases for DEGs that served as promising drug targets [11]. We accessed the DGIdb web site on 08/01/2021. Cytoscape software was used to draw drug-potential target network. Drugs are ranked according to the degree using Cyto-NCA. Only FDA-approved drugs were selected. We neglected drugs that are not for internal use, like topical cream and eye drop.

1.5.2. L1000FDW analysis

The L1000FDW web-based (http://amp.pharm.mssm.edu/L1000FDW) utility was used to identify potential drugs that can reverse the gene expression signature of COVID-19. L1000FDW computes the similarity between input gene expression signatures given as input of up and down gene sets and data from the Library of Integrated Network-based Cellular Signatures (LINCS)-L1000 data [25]. DEGs were used as input genes. Only FDA-approved drugs were selected.

1.6. Functional modules analysis

PPI interaction Network of DEGs was expanded using STRING and analyzed by Cytoscape software (version 3.8.1). Functional modules were identified via the MCODE (Molecular Complex Detection) using ClusterViz plugin, an APP of Cytoscape for cluster analysis and visualization [26–27], with the default parameters as follows: degree cutoff = 2, node score cutoff = 0.2, k-score = 2, and max. Depth = 100. GO biological process enrichment analysis and KEGG pathway analysis of genes in each functional module were done using STRING.

Fig. 1. Flow chart of integrated bioinformatics approach to identify potential pathogenic pathways, drug targets, and drug candidates for COVID-19. It consists of the following steps: identifying DEGs, pathway and GO enrichment analysis of DEGs, hub genes identification, drug-potential target network, L1000FDW analysis, and functional modules analysis. DEGs, Differentially Expressed Genes; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology.
2. Results

2.1. DEGs

We identified 422 DEGs between SARS-CoV-2 infected NHBE cells and mock-treated NHBE cells and 1829 DEGs between SARS-CoV-2 infected NAE and mock-treated cells. In addition, 187 overlapping genes (48 downregulated and 139 upregulated genes) were identified from the two datasets (Fig. 2 and Table S1).

2.2. Pathway and GO enrichment Analysis of DEGs

KEGG pathway analysis of DEGs was enriched to IL-17 signaling pathway, TNF signaling pathway, cytokine-cytokine receptor interaction, NF-kappa B signaling pathway, rheumatoid arthritis, and NOD-like receptor signaling pathway (Fig. 3a). Receptor ligand activity, cytokine activity, molecular function regulator, molecular function regulator, protein binding and cytokine receptor binding were the most relevant biological functions (Fig. 3b). DEGs were significantly involved in the immune process, inflammatory response, response to an external stimulus, response to biotic stimulus, defense response, and response to cytokine (Fig. 3c), and correlated to extracellular space, secretory granule, cytoplasmic vesicle, and vesicle (Fig. 3d).

RELB, ARNTL2, GRHL3, PLSCR1, SNAI1, CSRNP1, ZNF750, OVOL1, FOSL1, and ZNF267 were the top-ranked transcription factors for DEGs (Table 1), whereas EGFR, RIPK1, CHUK, IKBK, SRC, TBK1, PAK1, IRAK2, FYN, and LYN were the most enriched kinases (Table 2). We considered the top enriched transcription factors and kinases for DEGs as potential drug targets [11].

2.3. Hub genes identification

The PPI network included 186 nodes and 654 edges (Fig. 4). We used the three centrality methods (maximal clique centrality (MCC), degree, and subgraph) for identifying hub genes. The top 20 hub genes were identified from each method. Additionally, 18 overlapping hub genes (IL6, CXCL8, CXCL1, CCL20, CXCL2, CXCL3, IL1B, IL1A, MMP9, CXCL5, CSF3, ICAM1, IL1RN, CSF2, PTGS2, SAA1, NFkB1, and TNFAIP3) were selected based on the intersection of the three methods (Table 3). Finally, the PPI network of the hub genes was constructed using STRING (Fig. 5). It consisted of 18 nodes and 142 edges. This showed that these hub genes highly interact with each other.

2.4. Drug-potential target network

Potential targets consisted of 18 hub genes, top 10 enriched kinases, and transcription factors for DEGs were searched for promising drugs using DGIdb. We did not find drugs for all potential transcription factors.

---

Table 1

| Rank | Transcription factors | Mean rank |
|------|-----------------------|-----------|
| 1    | RELB                  | 15        |
| 2    | ARNTL2                | 19.33     |
| 3    | GRHL3                 | 21.33     |
| 4    | PLSCR1                | 22.5      |
| 5    | SNAI1                 | 27        |
| 6    | CSRNP1                | 29        |
| 7    | ZNF750                | 33.25     |
| 8    | OVOL1                 | 33.67     |
| 9    | FOSL1                 | 33.8      |
| 10   | ZNF267                | 34        |

Table 2

| Rank | Kinases | Mean Rank |
|------|---------|-----------|
| 1    | EGFR    | 12.8      |
| 2    | RIPK1   | 24.33     |
| 3    | CHUK    | 25.18     |
| 4    | IKBK    | 25.18     |
| 5    | SRC     | 29.09     |
| 6    | TBK1    | 29.73     |
| 7    | PAK1    | 38.64     |
| 8    | IRAK2   | 44        |
| 9    | FYN     | 50        |
| 10   | LYN     | 54.36     |
Fig. 4. Unexpanded PPI network of DEGs using the STRING online database. It consisted of 186 nodes and 654 edges.
except FOSL1. However, all hub genes and potential kinases have drugs except CXCL1, CCL20, CXCL3, CXCL5, and IRAK2. We included only two FDA unapproved botanical-derived drugs and obtained 215 potential drugs (Supplementary data 2). We concentrated on drugs acting on three potential targets or more (with degree $\geq 3$) as Fostamatinib, Entrectinib, Ponatinib, Dasatinib, Bosutinib, Ibrutinib, and Acalabrutinib were the top interactive drugs that act on multi-drug targets (Table 4, Fig. 6).

### 2.5. L1000FDW analysis

Drugs with low similarity scores may potentially reverse the transcriptional profile of COVID-19 and are potential drugs for COVID-19. Sorafenib, Sirolimus, Selumetinib, Fostamatinib, and Dasatinib were the top drugs (Table 5).

#### Table 3

| Rank | Cytohubba (MCC) | CytoNCA-Degree | CytoNCA-Subgraph | Common hub genes | Description |
|------|-----------------|----------------|------------------|------------------|-------------|
| 1    | IL6             | IL6            | IL6              | IL6              | Interleukin 6 (IL6) |
| 2    | CXCL8           | CXCL8          | CXCL8            | CXCL8            | C-X-C motif chemokine ligand 8 (CXCL8) |
| 3    | CXCL1           | IL1B           | IL1B             | CXCL1            | C-X-C motif chemokine ligand 1 (CXCL1) |
| 4    | CCL20           | MMP9           | MMP9             | CCL20            | C-C motif chemokine ligand 20 (CCL20) |
| 5    | CXCL2           | CXCL1          | CXCL1            | CXCL2            | C-X-C motif chemokine ligand 2 (CXCL2) |
| 6    | CXCL3           | PTGS2          | PTGS2            | CXCL3            | C-X-C motif chemokine ligand 3 (CXCL3) |
| 7    | IL1B            | ICAM1          | ICAM1            | IL1B             | Interleukin 1Beta (IL1B) |
| 8    | IL1A            | CSF2           | CCL20            | IL1A             | Interleukin 1 Alpha (IL1A) |
| 9    | MMP9            | CCL20          | CSF2             | MMP9             | Matrix Metalloproteinase 9 (MMP9) |
| 10   | CXCL5           | IL1A           | IL1A             | CXCL5            | Colony Stimulating Factor 3 (CSF3) |
| 11   | CSF3            | NFKB1          | CXCL2            | CSF3             | Interleukin 1 Receptor Antagonist (IL1R) |
| 12   | ICAM1           | CXCL2          | CXCL3            | ICAM1            | Colony Stimulating Factor 2 (CSF2) |
| 13   | IL1RN           | IL1RN          | IL1RN            | IL1RN            | Interleukin 1 Receptor Antagonist (IL1R) |
| 14   | CSF2            | CSF3           | CSF3             | CSF2             | Colony Stimulating Factor 2 (CSF2) |
| 15   | PTGS2           | SAA1           | SAA1             | PTGS2            | Prostaglandin-Endoperoxide Synthase 2 (PTGS2) |
| 16   | SAA1            | CXCL3          | CXCL5            | SAA1             | Serum Amyloid A1 (SAA1) |
| 17   | CXCL6           | TNFAIP3        | NFKB1            | TNFAIP3          | Nuclear Factor Kappa B Subunit 1 (NFKB1) |
| 18   | NFKB1           | PI3            | TNFAIP3          | TNFAIP3          | TNF Alpha Induced Protein 3 (TNFAIP3) |
| 19   | TNFAIP3         | NFKB2          | CXCL6            |                 |             |

#### Table 4

| Rank | Drugs          | Degree | Potential targets (DGIdb) |
|------|----------------|--------|---------------------------|
| 1    | Fostamatinib   | 8      | EGFR-RIPK1-IKBKB-SRC-TBK1-PACK1-FYN-LYN |
| 2    | Entrectinib    | 4      | SRC-TBK1-FYN-LYN          |
| 3    | Ponatinib      | 4      | EGFR-RIPK1-SRC-LYN       |
| 4    | Dasatinib      | 4      | EGFR-SRC-FYN-LYN        |
| 5    | Bosutinib      | 4      | EGFR-SRC-FYN-LYN        |
| 6    | Ibrutinib      | 4      | EGFR-SRC-FYN-LYN        |
| 7    | Acalabrutinib  | 4      | EGFR-SRC-FYN-LYN        |
| 8    | Resveratrol    | 4      | IL1B-PTGS2-NFKB1-RIPK1  |
| 9    | Nintedanib     | 3      | SRC-FYN-LYN              |
| 10   | Vandetanib     | 3      | EGFR-SRC-FYN             |
| 11   | Gefitinib      | 3      | EGFR-RBBB-LYN           |
| 12   | Sulfasalazine  | 3      | PTGS2-CHUK-IKBKB        |
| 13   | Mesalamine     | 3      | PTGS2-CHUK-IKBKB        |
| 14   | Curcumin       | 3      | MMP9-PTGS2-EGFR         |
| 15   | Riluzole       | 3      | IL1B-IL1A-ILRN          |
| 16   | Alteplase      | 3      | CXCL2-IL1B-FOSL1        |
| 17   | Aspirin        | 3      | CXCL8-IL1B-PTGS2       |
| 18   | Bevacizumab    | 3      | CXCL8-MMP9-EGFR         |
| 19   | Omeprazole     | 3      | CXCL8-IL1B-CSF2         |
| 20   | Naproxen       | 3      | CXCL8-PTGS2-SAA1       |
| 21   | Clopaplatin     | 3      | IL6-EGFR-SRC           |

### 2.6. Functional modules analysis

To understand the underlying pathologic pathways, we expanded the PPI network of DEGs to include 386 nodes and 5280 edges (Fig. S1). Then, the significant modules were identified via the MCODE using ClusterViz. The top eight functional clusters of modules were selected. Then, KEGG pathway analysis and biological process enrichment analysis of each module were performed by STRING.

Module 1, with MCODE score = 32.789, consisted of 77 nodes and 1246 edges (Fig. S2). The KEGG pathway analysis of this module indicated that the IL-17 signaling pathway, axon guidance, TNF signaling pathway, cytokine-cytokine receptor interaction, NF-kappa B signaling pathway, and NOD-like receptor signaling pathway were the most...
related pathways (Table S3). Biological process enrichment analysis identified that cell-surface-receptor signaling pathway, signal transduction, response to cytokine, ephrin receptor signaling pathway, and cytokine-mediated signaling pathway were the top enriched process for the first top module (Table S3).

Module 2, with MCODE score = 16.049, consisted of 42 nodes and 329 edges (Fig. S3). The KEGG pathway analysis of this module indicated that NF-kappa B signaling pathway, response to cytokine, ephrin receptor signaling pathway, and cytokine-mediated signaling pathway were the top related pathways (Table S3). Biological process enrichment analysis identified that cytokine-mediated signaling pathways, response to cytokine, and cellular response to the organic substance were the top enriched processes for module 2 (Table S3).

Module 3, with MCODE score = 10.2, consisted of 11 nodes and 51 edges (Fig. S4). The KEGG pathway analysis of this module indicated that homologous recombination, non-homologous end-joining, and cellular senescence were the most related pathways (Table S3). Biological Process enrichment analysis identified that telomere capping, regulation of telomere maintenance, and negative regulation of telomere maintenance are the top enriched process for module 3 (Table S3).

Module 4, with MCODE score = 10, consisted of 10 nodes and 45 edges (Fig. S5). The KEGG pathway analysis of this module indicated that focal adhesion, regulation of actin cytoskeleton, leukocyte transendothelial migration, tight junction, vascular smooth muscle contraction, and platelet activation were the most related pathways (Table S3). Biological Process enrichment analysis identified muscle contraction and muscle organ development as the top enriched process for module 4 (Table S3).

Module 5, with MCODE score = 9.33, consisted of 10 nodes and 42 edges (Fig. S6). The KEGG pathway analysis of this module indicated that hepatitis C, NOD-like receptor signaling pathway, influenza A, herpes simplex infection, and measles were the most related pathways (Table S3). Biological Process enrichment analysis identified that type I interferon signaling pathway, innate immune response, defense response to the virus, and cellular response to cytokine stimulus were the top enriched process for module 5 (Table S3).

Module 6, with MCODE score = 9, consisted of 6 nodes and 36 edges

---

**Table 5**

| Rank | Drugs       | Similarity scores            |
|------|-------------|------------------------------|
| 1    | Sorafenib   | 0.19496855345912             |
| 2    | Sirolimus   | 0.19496855345912             |
| 3    | Selumetinib | -0.182389937106918           |
| 4    | Fostamatinib| -0.176100628930818           |
| 5    | Dasatinib   | -0.168811320754717           |
| 6    | Vemurafenib | -0.157232704402516           |
| 7    | Neratinib   | -0.15094396262415            |
| 8    | Exemestane  | -0.144654088050314           |
| 9    | Bosutinib   | -0.138364779874214           |
herpes simplex infection, and measles. Apoptosis and RIG-I-like receptor signaling pathways were significantly enriched for module 2. This indicates that COVID 19 induces these important pathogenic pathways [53-57]. Apoptosis is a promising target for combating SARS-CoV-2 virus infection [41,53-56]. It was found that Telomere capping was the top enriched biological process for enriched module 3 (Table S3). TIN2F2 (TIN2) was upregulated in our final DEGs. TIN2 is induced by NF-κB transcription factor [58] and is a core part of the Shelterin complex, which protects telomeres and chromosomes. High TIN2 expression can produce telomere dysfunction-induced DNA damage and aberrant chromosomal alterations [59]. Telomere dysfunction limits the proliferative capacity of human cells by activation of DNA damage responses, inducing senescence or apoptosis [60-63]. Short telomere is manifested as pulmonary fibrosis-ephysmesia in adults [61]. Telomere dysfunction may provide a novel pathway for the pathogenesis of COVID-19. Homologous recombination (HR), non-homologous end-joining (NHEJ), and cellular senescence enriched for module 3 may play a vital role in COVID-19 pathogenesis [64-67]. NHEJ or HR machinery are involved in the interactions between viruses and the host DNA damage response (DDR) machinery [64-65]. Lai et al. proposed that senolytic targeting of virus-infected cells is considered as a novel treatment option against SARS-CoV-2 [66]. Cellular senescence induced by virus is accompanied by a senescence-associated secretory phenotype (SASP) with increased pro-inflammatory cytokines, extracellular matrix-active factors and pro-coagulatory mediators [67].

The enriched pathways for module 4 as Focal adhesion, regulation of actin cytoskeleton, leukocyte transendothelial migration, tight junction, vascular smooth muscle contraction, and platelets activation are potential targets for COVID-19 [68-70]. MYLK, in module 4, is found to be downregulated, and its role in the pathogenesis of COVID-2 is unclear. Besides, Genes of module 5 (MX2, IFI27, and GBP5) were upregulated and highlighted the role of type I interferon signaling pathway and innate immune response in COVID-19 pathogenesis [27,71]. SARS-CoV-2-infected upregulated genes (TGM1, SPRR1B, LOR, SPRR2E, SPRR2D, SPRR1A, IFL, FLG, and SPRR2A) involved in cornification and peptide cross-linking, which were enriched for module 6. This may lead to cell death and lung fibrosis [74-76]. Module 7 genes (TUBA1B, TUBB4B, TUBA1C, and TUBA1A) were upregulated and involved in the microtubule-based process and cytoskeleton organization, which may explore the potential significance of gap junction, phagosome, and tight junction in COVID-19 pathogenesis. Colchicine, which acts on tubulin, was repurposed for COVID-19 [77]. Genes of module 8 (AAA1, DXDS, SARPIN2, F3, TNFRSF10A, BID, PLAU, HBGF, ANXA1, and PLAIR) were upregulated. Four (F3, PLAU, PLAIR, and SARPIN2) were involved in complement and coagulation cascades and considered potentially important in coagulation disorders associated with COVID-19 infection. Many papers explored the significance of the Jak-STAT signaling pathway and natural killer cell-mediated cytotoxicity [78-79].

From the unexpanded PPI network of DEGs, 18 hub genes (IL6, CXCL8, CXCL1, CCL20, CXCL2, CXCL3, IL1B, IL1A, MMP9, CXCL5, CSF3, Icam1, IL1RN, CSF2, PTGS2, SAA1, NFKB1, and TNFAIP3) were identified by the overlap of three methods, and it was found that these hub genes were highly interactive with each other. Hub genes and the top enriched transcription factors and kinases for DEGs were used to construct a drug potential target network. Drugs are ranked according to the degree. Though targeting a single potential target can be considered as a therapeutic option for COVID-19, drugs targeting three potential targets or more (with degree ≥ 3) were prioritized. The top potential drugs as Fostamatinib (Degree = 8), Ibrutinib (Degree = 4), Acalabrutinib (Degree = 4), Bevacizumab (Degree = 3), and Nintedanib (Degree = 3) were already suggested to be repurposed for COVID-19 in recent papers [80-84]. Neha Tabassum et al. (2020) showed that Fostamatinib (an inhibitor of spleen tyrosine kinase (SYK)) was able to reduce MUC1 in a relevant pre-clinical model and may improve serious outcomes of COVID-19, including acute respiratory distress syndrome (ARDS) and...
acute lung injury (ALI) [80]. Treon SP et al. (2020) showed that 420 mg/day of Ibrutinib (Bruton tyrosine kinase (BTK) inhibitor) improved 5 COVID-19 patients’ outcomes, especially oxygenation, and did not require hospitalization [81]. Acalabrutinib (a selective Bruton tyrosine kinase (BTK) inhibitor) was administered to 19 patients with severe COVID-19, and it was found that it improved oxygenation and clinical status in a majority of patients, often within 1–3 days [82]. Moreover, 26 patients with severe COVID-19 were administered Bevacizumab, 24 patients showed improvement in oxygenation, 17 patients were discharged, and none died [83]. Ogata H et al. (2021) found that Nintedanib improved serious outcomes for a 78-year-old Japanese woman with post-COVID-19 fibrosis [84]. Naproxen has antiviral and anti-inflammatory effects. It inhibited replication of SARS-CoV-2 and protected against virus-induced damage pulmonary epithelium in vitro [85]. Clinical trials were conducted to study the safety and efficacy of Fostamatinib, Acalabrutinib, Ibrutinib, and Bevacizumab, Alteplase, Aspirin, and Naproxen for hospitalized patients with COVID-19 (ClinicalTrials.gov) (NCT04579392-NCT04497948-NCT04375397-NCT04275414-NCT04357730-NCT04365309-NCT04322565). Another top potential drug, Nintedanib with degree = 3, is under phase III to treat SARS-CoV-2 Induced Pulmonary Fibrosis (NCT04541680). Another Study (Phase IV) has been launched to evaluate the effect of Nintedanib in slowing lung fibrosis in patients with COVID-19 (NCT04619680). Galimberti S et al. (2020) considered Bosutinib (degree = 4) as a potential drug for COVID-19 and discouraged the use of Dasatinib (degree = 4) due to its toxicity [86]. Ponatinib was found to inhibit cytokine release in myeloid cells in vitro and lung inflammation in the mouse model [87]. O’Donovan S.M. et al. (2021) identified Dasatinib and Vandetanib (degree = 3) as candidate drugs to manage COVID-19 using a signature-based approach [88]. Other top drugs as Entrectinib, Resveratrol, Gefitinib, Curcumin, Rilonacept, Omeprazole, Naproxen, and Ribavirin are considered potential drugs for COVID-19. Some papers addressed the beneficial effects of the two botanical-derived drugs, Resveratrol and Curcumin, in COVID-19 [89–93]. Rilonacept as Anti-IL-1 inhibitor therapies is considered as a potential drug for severe COVID-19 patients with cytokine release syndrome (CRS) due to its rapid and sustained blockade of inflammation [94]. Omeprazole inhibits activation of the nuclear factor-κB, release of proinflammatory cytokines (tumor necrosis factor (TNF-α), interleukin-1β (IL-1β) and interleukin-6 (IL-6)), and chemotaxis of neutrophils and reduces the oxidative stress [95–96]. So, Omeprazole may offer an attractive treatment option for COVID-19. There is a controversy about using Ribavirin in COVID-19 [97–100]. The potential of gefitinib to attenuate pulmonary fibrosis remained controversial [101–106]. Though Sulfasalazine was considered a potential drug [88], Mesalamine and Sulfasalazine are discouraged as they may induce lung injury [107]. Paclitaxel has an induction effect on interleukin-8 (IL-8), for which it worsen COVI-19 complications [108]. Cisplatin is discouraged as it may cause lung fibrosis [109]. Along with the degree, we added another dimension for prioritization. We selected those novel potential drugs which act on the same targets as those under clinical investigation. It was found that targets of Entrectinib (SRC-TK1-FYN-LYN), Ponatinib (EGFR-RIPK1-SRC-LYN) shared three of the targets of Acalabrutinib and Ibrutinib (EGFR-SRC-FYN-LYN), and Omeprazole (CXCL8-IL1B-CSF2) shared two of the targets of Atelplase (CXCL2-IL1B-FOSL1) and aspirin (CXCL8-IL1B-PDGSS) (Table 4). L1000FDW analysis added new potential drugs, as shown in Table 5. Fostamatinib, Dasatinib, and Bosutinib were common between the two methods (drug-potential target network and L1000FDW analysis). Other potential drugs revealed by L1000FDW analysis were Sorafenib, Sirolimus, Selumetinib, Neratinib, Vemurafenib, Exemestane, and Bosutinib. Some papers identified Sirolimus, Sorafenib, Selumetinib, Vemurafenib, and Neratinib as potential drugs for COVID-19 [88,110–113]. Bischof et al. (2021) proposed that targeting the hallmarks of aging by mTORC1 inhibitors as Sirolimus is a novel potential strategy for combating COVID-19. It may diminish inflammation and immunosenescence [112]. Another study (Phase II) was planned to illustrate the efficacy and safety of Sirolimus in COVID-19 infection (NCT04461340). Vemurafenib was found to significantly modulate ACE2 expression [113]. The inhibition of host growth-factor receptor signaling by Sorafenib prevents SARS-CoV-2 replication in vitro [114]. Sorafenib and Selumetinib (Kinase inhibitors) were identified as potential candidates for combating COVID-19, using a signature-based approach [88]. Neratinib is found to have antiviral activity due to its putative inhibition of SARS-CoV-2 virus main protease M-pro [115]. Finally, Entrectinib, Omeprazole, and Exemestane are considered potential novel candidates for combating COVID-19. Entrectinib is a tyrosine kinase inhibitor that was promising in COVID-19 [116]. Exemestane is an aromatase inhibitor that was found to decrease radiation-induced lung fibrosis scores [117].

Conclusion

In this study, we tried to decipher potentially hidden parts of the pathogenesis of COVID-19 as well as repurposing FDA-approved drugs using an integrative bioinformatics approach. Besides the well-known pathogenic pathways like IL-17 signaling pathway, TNF signaling pathway, Cytokine-cytokine receptor interaction, NF-kappa B signaling pathway, NOD-like receptor signaling pathway, Toll-like receptor signaling pathways, platelets activation, and leukocyte transendothelial migration, axon guidance is a potential pathogenic pathway. Sema7A and its receptor, PlexinC1, is a potential novel drug target. Sema7A may exacerbate cytokines production. Another potential novel pathway is related to TNF2 overexpression, which may induce potential telomere dysfunction and cause DNA damage that may exacerbate lung fibrosis. We identified repurposed drugs that are under investigation for hospitalized patients with COVID-19. We also repurposed new potential drugs, such as Entrectinib and Omeprazole, which were prioritized based on two factors: the degree and the similarity of targets with those drugs which were under clinical investigation. In addition, L1000FDW analysis identified Exemestane as a potential novel drug. Our findings provide new insight into the mechanisms of COVID-19 complications as well as potential new drugs. Further validation of these findings is needed. Although we did not do an experimental validation, finding known pathogenic pathways and drugs currently in clinical trials for COVID-19 in our analysis is one way of validation for our method.

Availability of data and materials

The data sets is publicly available at Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/), under data source access (GSE147507-GSE153970). Galaxy, VENNY 2.1, STRING, ChEA3, KEA3 and L1000FDW are online tools. Cytoscape software (version 3.8.1) was downloaded. Data supporting the conclusions of this article are included within the article.

CRedit authorship contribution statement

Salem A. El-aarag: Conceptualization, Methodology, Formal analysis, Writing—original draft preparation, Writing- Reviewing and Editing, and Visualization.

Amal Mahmoud: Conceptualization, Writing- Reviewing and Editing, Funding acquisition, Project administration.

Mahmoud ElHefnawi: Conceptualization, Writing- Reviewing and Editing, Visualization. Supervision.

All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by Deanship of Scientific Research, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia, Grant No. Covid19-2020-001-BASRC. We acknowledge the support from Science Technology and Innovation Funding Authority (STIFA) project number...
The authors declare no conflict of interest.

Acknowledgements

We are grateful for the GEO, KEGG, STRING, L1000FDW, and DGIib databases. Also, we wish to thank Dr. Ahmed Mostafa (National Research Center, Egypt) for his support and suggestions.

Appendix A. Supplementary data

Supplementary data to this article can be found at https://doi.org/10.1016/j.ijbiomac.2021.11.124.

References

[1] 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. Xu, X. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Guo, Q. Jin, J. Wang, B. Cao. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet 395 (10223) (2020) 497–506.

[2] D. Blanco-Melo, B.E. Nilsson-Payant, W.C. Liu, Imbalanced host response to SARS-CoV-2 during development of COVID-19, Cell 181 (5) (2020) 1036–1045, https://doi.org/10.1016/j.cell.2020.04.026, e9.

[3] A. Vanderheiden, P. Ralfs, T. Chirkova, A.A. Upadhaya, M.G. Zimmerman, S. Bedoya, H. Aeud, G.M. Tharp, K.L. Pellegrini, C. Misra, E. Teschner, B. Mainou, J.L. Lobey, J.E. Kohleim, A.C. Lowen, F.-Y. Shi, V.D. Menachery, L. J. Anderson, A. Gakouli, S.E. Bosinger, M.S. Suthar, Type I and type III interferons restrict SARS-CoV-2 infection of human airway epithelial cultures, J. Virol. 94 (2020), e00965-20, https://doi.org/10.1128/JVI.00965-20.

[4] Peipei Song, Wei Li, Jianqin Xie, Yanlong Hou, Chongge You, Cytokine storm induced by SARS-CoV-2, Clin. Chim. Acta 509 (2020) 280–287.

[5] Z. Z lou, L. Ren, L. Zhang, J. Zhong, Y. Xiao, J. Liu, G. Yang, J. Wang, S. Jiang, D. Jiang, Z. Zhang, H. Li, F. Chen, Y. Xu, M. Chen, Z. Gao, J. Dong, B. Liu, X. Zhang, W. Wang, K. He, Q. Jin, M. Li, J. Wang, Z. Zhou, L. Ren, L. Zhang, J. Zhong, Y. Xiao, J. Liu, G. Yang, C. Wang, S. Jiang, D. Yang, G. Zhang, H. Li, F. Chen, Y. Xu, M. Chen, Z. Gao, J. Dong, B. Liu, X. Zhang, W. Wang, K. He, Q. Jin, M. Li, J. Wang, Heightened innate immune responses in the respiratory tract of COVID-19 patients, Cell Host Microbe 27 (6) (2020) 883–890.

[6] R.M. Golonka, P. Saha, B.S. Yeoh, S. Chattopadhyay, A.T. Gewertz, B. Joe, M. Vijay Kumar, Harnessing innate immunity to eliminate SARS-CoV-2 and ameliorate COVID-19 disease, Physiol. Genomics 52 (5) (2020) 217–221.

[7] K. Prasad, F. Khatoon, S. Rashid, N. Ali, A.P. Alasmri, M.Z. Ahmed, A. S. Alqattan, M.S. Alqattan, V. Kumar, Targeting hub genes and pathways of innate immune response in COVID-19: A network biology perspective, Int. J. Biol. Macromol. 163 (2020 Nov 15) 1–8, https://doi.org/10.1016/j.ijbiomac.2020.06.228, Epub 2020 Jun 26, PMID: 32599245; PMCID: PMC7371964.

[8] M. Wu, Y. Chen, H. Xia, C. Wang, C.Y. Tan, X. Cai, Y. Liu, F. Ji, P. Xiong, R. Liu, Y. Guan, D. Duan, D. Kuang, S. Xu, H. Cai, Q. Xia, D. Yang, M.W. Wang, L.M. Chiu, C. Cheng, P.P. Ahern, L. Liu, G. Wang, N.K. Surana, T. Y. D. Karger, Transcriptomic and proteomic insights into the host response in fatal COVID-19 cases, Proc. Natl. Acad. Sci. U. S. A. 117 (40) (2020 Nov 20) 28336–28343, https://doi.org/10.1073/pnas.2008301117, Epub 2020 Oct 20, PMID: 33088228; PMCID: PMC6768053.

[9] Dengchuan Wang, Shenghou Liu, Shi Xu, Identification of hub genes, key pathways, and therapeutic agents in Hutchinson-Gilford Progeria syndrome using bioinformatics analysis, Medicine 99 (7) (February 2020), e19222, https://doi.org/10.1097/MD.00000000000019222.

[10] Y. Gao, X. Li, Y. He, J. Yang et al., Integrated bioinformatics analysis for the screening of hub genes and therapeutic drugs in ovarian cancer, J. Ovarian Res. 13 (1) (2020) 10, https://doi.org/10.1186/s13059-020-00613-9.

[11] S.A. El-aarag, A. Mahmoud, M.H. Hashem, et al., In silico identification of potential key regulatory factors in smoking-induced lung cancer, BMC Med. Genet. 10 (2017) 40, https://doi.org/10.1186/s12881-017-0284-z.

[12] S. Andrews (n.d.). FastQC a quality control tool for high throughput sequence data. Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

[13] A.M. Bolger, M. Lohse, B. Usadel, Trimmmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics 30 (15) (2014), https://doi.org/10.1093/bioinformatics/btu170, 21144f7-2120. Retrieved from.

[14] D. Kim, B. Langmead, S.L. Salzberg, HISAT: a fast spliced aligner with low memory requirements, Nat. Methods 12 (4) (2015), https://doi.org/10.1038/nmeth.3311, 357af4-360. Retrieved from.

[15] Institute, B. (n.d.). Picard. Broad Institute, GitHub repository. Retrieved from http://broadinstitute.github.io/picard/Picard, B. (n.d.). Picard. Broad Institute, GitHub repository. Retrieved from http://broadinstitute.github.io/picard/.}

International Journal of Biological Macromolecules 194 (2022) 770–780

© 2022 Published by Elsevier B.V.

778
A. Kovacs, M. Reverta, H. Wang, X. Jing, Y. Zhang, Y. Chen, Y. Cao, Efficacy and tolerability of bevacizumab in patients with severe COVID-19. Nat. Commun. 12 (1) (2021 Feb 5) 841. https://doi.org/10.1038/s41467-020-18992-8. PMID: 33547300; PMCID: PMC7684918.

[84] H. Ogata, T. Nakagawa, S. Sakoda, et al., Nintedanib treatment for pulmonary fibrosis after coronavirus disease 2019, Respiratory Care Res. 9 (5) (2021 May), e60744, https://doi.org/10.1097/01.rcr.2.44.

[85] L. Li, L. Cai, L. Zheng, H. Yu, W. Yuan, Z. Guo, W. Li, Gefitinib inhibits bleomycin-induced pulmonary fibrosis via alleviating the oxidative damage in mice, Oxidative Med. Cell. Longev. 12 (2018) (2018 Apr) 8249693. https://doi.org/10.1155/2018/8249693; PMCID: PMC592497.

[86] Y. Manoharan V. Haridas K. Vasanthakumar S. Muthu F.F. Thavoorullah P. Shetty T. Loganathan, S. Ramachandran, P. Shankaran, D. Nagarajan, S.S. Mohan, Host learning identifies molecular regulators and therapeutics for targeting SARS-CoV2-induced cytokine release, Mol. Syst. Biol. 17 (9) (2021), e10426, https://doi.org/10.15252/msb.202110426.

[87] S. Galimberti, O. Terrier, S. Dilly, A. Pizzorno, et al., Antiviral properties of the NSAID drug naproxen targeting the nucleoprotein of SARS-CoV-2 Coronavirus, Molecules 26 (9) (2021) 2593. https://doi.org/10.3390/26092593. Published 2021 Apr 29.

[88] S. Chanchal, U. Mahajan, S. Siddharth, et al., In vivo and in vitro protective effects of nimesulide against COVID-19 infection, PeerJ 8 (2020) e9357. https://doi.org/10.7717/peerj.9357.

[89] A. Pizzorno, et al., Antiviral properties of the NSAID drug naproxen targeting the nucleoprotein of SARS-CoV-2 Coronavirus, Molecules 26 (9) (2021) 2593. https://doi.org/10.3390/26092593. Published 2021 Apr 29.

[90] T. Hoang, An approach of fatty acids and resveratrol in the prevention of COVID-19 severity, Phytother. Res. (2020 Nov 17), https://doi.org/10.1002/ptr.6956. Epub ahead of print. PMID: 32008039; PMCID: PMC7753401.

[91] K. Klann, D. Bojkova, G. Tascher, S. Ciesek, C. Münch, J. Cinatl, Growth factor receptor signaling inhibition prevents SARS-CoV-2 replication, e4, Mol. Cell 80 (2020 Jul 23), 102571, https://doi.org/10.1016/j.autrev.2020.102571. Published 2020 Jan 23. PMID: 32440966; PMCID: PMC7518749.

[92] K. Li, W. Huang, K. Zhang, C. Lu, L. Wang, H. Chen, F. Sun, Y. He, Metformin attenuates gefitinib-induced exacerbation of pulmonary fibrosis in vivo, Life Sci. 177 (2018) 339-347, https://doi.org/10.1016/j.lfs.2017.11.006. Epub 2017 Nov 29. PMID: 30830136; PMCID: PMC6055208.

[93] L. Que, C.S. Lau, K.H. Chan, K.K. To, K.Y. Yuen, Triple combination of interferon therapy in patients infected with the Middle East respiratory syndrome coronavirus 2019, Cell Discov. 6 (2020) 14, https://doi.org/10.1038/s41421-020-00214-x. Epub 2020 Jul 23. PMID: 32712134; PMCID: PMC7777727.

[94] L.F. Lee, J.S. Haskill, N. Mukaida, K. Matsushima, J.P. Ting, Identification of tumor-specific paclitaxel (Taxol)-responsive regulatory elements in the interleukin-8 promoter, Mol. Cell Biol. 17 (9) (1997) 5070-5105, https://doi.org/10.1128/MCB.17.9.5070.

[95] H. Ishii, S. Fujimoto, T. Fukuda, Gefitinib prevents bleomycin-induced lung fibrosis in mice, Am. J. Respir. Crit. Care Med. 174 (5) (2006 Sep) 550-556, https://doi.org/10.1164/rccm.200509-1540OC. Epub 2006 Jan 1 PMID: 16741154.

[96] S. Tong, X. Wang, Y. Zhang, Y. Chen, Y. Cao, Efficacy and tolerability of bevacizumab in patients with severe COVID-19. Nat. Commun. 12 (1) (2021 Feb 5) 841. https://doi.org/10.1038/s41467-020-18992-8. PMID: 33547300; PMCID: PMC7684918.

[97] Y. Su, Y. Yu, W. Yuan, Z. Guo, W. Li, Gefitinib inhibits bleomycin-induced pulmonary fibrosis via alleviating the oxidative damage in mice, Oxidative Med. Cell. Longev. 12 (2018) (2018 Apr) 8249693. https://doi.org/10.1155/2018/8249693; PMCID: PMC592497.

[98] S. Galimberti, C. Baldini, C. Barati, F. Ricci, S. Baldacci, S. Grassi, The CoV-2 outbreak: how hematologists could help to fight Covid-19, Pharmcol. Res. (2020), 104866, https://doi.org/10.1016/j.phrs.2020.104866.

[99] B. Zheng, Y. Wang, H. Xu, W. Han, Z. Li, S. Zhang, L. Zhang, Y. He, Metformin attenuates gefitinib-induced exacerbation of pulmonary fibrosis in vivo, Life Sci. 177 (2018) 339-347, https://doi.org/10.1016/j.lfs.2017.11.006. Epub 2017 Nov 29. PMID: 30830136; PMCID: PMC6055208.

[100] Nuran Bese, Aromatase inhibitors decrease radiation-induced lung fibrosis: results of a pilot study, Cancer Lett. 46 (2014 Oct) 207-212, https://doi.org/10.1016/j.canlet.2014.06.001. Epub 2014 Jun 27. PMID: 25061555; PMCID: PMC4433624.

[101] L. Zhou, Y. Hou, J. Shen, et al., Network-based drug repurposing for novel coronavirus 2019-nCoV, Cell Discov. 6 (2020) 14, https://doi.org/10.14321/2020-01421-01425. Epub 2020 Apr 14. PMID: 32406576; PMCID: PMC7110882.

[102] S. Sinha, K. Cheng, A.A. Schaffer, K. Aldape, E. Schiff, E. Ruppin, In vitro and in vivo identification of clinically approved drugs that modify ACE2 expression, Mol. Syst. Biol. 16 (7) (2020 Jul), e6628. https://doi.org/10.15252/msb.202006628. PMCID: PMC7229748. https://doi.org/10.15252/msb.202006628.

[103] S. Galimberti, M. Petrini, C. Barati, F. Ricci, S. Baldacci, S. Grassi, F. Guerini, E. Ciabatti, S. Mechelli, A. Di Paolo, C. Baldini, L. Baglietto, L. Macera, P. G. Spezia, F. Maggi, Tyrosine kinase inhibitors play an antiviral action in patients affected by chronic myeloid leukaemia: a possible model supporting their use in the fight against SARS-CoV-2, Front. Oncol. 10 (2020) 1429. https://doi.org/10.3389/foncot.2020.01426. PMID: 33014780; PMCID: PMC7493657.

[104] A. Ayse Altn, Sukru Vildirim, T. Altug, N. Sut, A. Ober, E.M. Ozsahin, D. Azria, Nuran Bese, Aromatase inhibitors decrease radiation-induced lung fibrosis results of an animal experimental study, Breast 28 (2016) 174-177. https://doi.org/10.1016/j. breast.2016.04.003.