Computers in Biology and Medicine 159 (2023) 106885

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

The pathogenetic influence of smoking on SARS-CoV-2 infection: Integrative transcriptome and regulomics analysis of lung epithelial cells

Md. Ali Hossain [a,b,1], Tania Akter Asa [c,1], Md. Rabiul Auwul [d], Md. Aktaruzzaman [e], Md. Mahfizur Rahman [f], M. Zahidur Rahman [a], Mohammad Ali Moni [g,*]

[a] Department of Computer Science and Engineering, Jahangirnagar University, Dhaka, Bangladesh
[b] Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh
[c] Department of Electrical and Electronics Engineering, Islamic University, Kushtia, Bangladesh
[d] Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur, Bangladesh
[e] Department of Computer Science and Engineering, Islamic University, Kushtia, Bangladesh
[f] Biotechnology and Genetic Engineering, Islamic University, Kushtia, Bangladesh
[g] Artificial Intelligence & Data Science, Faculty of Health and Behavioral Sciences, The University of Queensland, Australia

ARTICLE INFO

Keywords:
COVID-19
SMOKING
Comorbidity
Protein–protein interaction
Drugs
Pathway

ABSTRACT

Corona virus disease (COVID-19) has been emerged as pandemic infectious disease. The recent epidemiological data suggest that the smokers are more vulnerable to infection with COVID-19; however, the influence of smoking (SMK) on the COVID-19 infected patients and the mortality is not known yet. In this study, we aimed to discern the influence of SMK on COVID-19 infected patients utilizing the transcriptomics data of COVID-19 infected lung epithelial cells and transcriptomics data smoking matched with controls from lung epithelial cells. The bioinformatics based analysis revealed the molecular insights into the level of transcriptional changes and pathways which are important to identify the impact of smoking on COVID-19 infection and prevalence. We compared differentially expressed genes (DEGs) between COVID-19 and SMK and 59 DEGs were identified as consistently dysregulated at transcriptomics levels. The correlation network analyses were constructed for these common genes using WGCNA R package to see the relationship among these genes. Integration of DEGs with network analysis (protein–protein interaction) showed the presence of 9 hub proteins as key so called "candidate hub proteins" overlapped between COVID-19 patients and SMK. The Gene Ontology and pathways analysis demonstrated the enrichment of inflammatory pathway such as IL-17 signaling pathway, Interleukin-6 signaling, TNF signaling pathway and MAPK1/MAPK3 signaling pathways that might be the therapeutic targets in COVID-19 for smoking persons. The identified genes, pathways, hubs genes, and their regulators might be considered for establishment of key genes and drug targets for SMK and COVID-19.

1. Introduction

The latest human coronavirus is SARS-CoV-2 causes respiratory illness termed COVID-19. This COVID-19 was emerged in the Wuhan city of China in 2019. Research has shown that SARS-CoV-2 viruses currently referred to as COVID-19 may cause health issues such as fever, vomiting, and fatigue in patients infected with this virus. In severe cases, viral respiratory infections may cause the death of a patient due to serious acute respiratory syndrome (SARS) [1–3]. The pandemic has been steadily growing since its first emersion in December 2019. There are some studies reported in the literature on the association of COVID-19 and smoking [4–10]. Among them, Vardavas et al. [4] concluded with their research study that, smoking most certainly contributes to COVID-19's bad course and unfavorable results. In 2021, Umnuaypornlert et al. [5] conducted a meta-analysis among 1248 studies and concluded that smoking greatly raises the chance of COVID-19 severity and mortality. A recent work conducted by Ram Poudal et al. [9] reported that smoking increases the risk of severe COVID-19, including deaths. Findings have also been reported with some evidence of the correlation of variations in incidence and severity of COVID-19 diseases with sex and of smoking with increased ACE2 expression (the recipient for extreme SARS-CoV-2), which may also be a cause of this type [11].

It has been hypothesized from the above reports that smoking might make more vulnerable to COVID-19 infection and smoking increased

* Corresponding author.
E-mail address: m.moni@uq.edu.au (M.A. Moni).
[1] Equally contributing authors.

https://doi.org/10.1016/j.compbiomed.2023.106885
Received 12 June 2022; Received in revised form 26 February 2023; Accepted 30 March 2023
Available online 31 March 2023

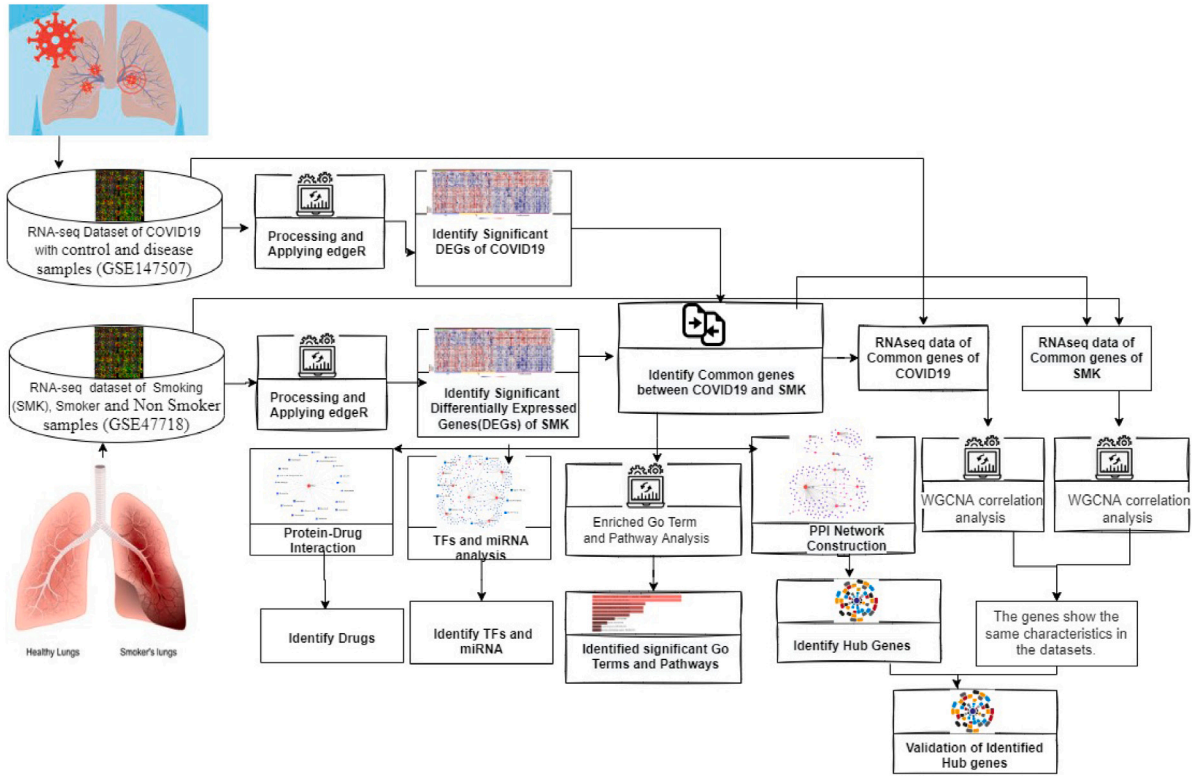

Fig. 1. Flowchart used in this study. Using edgeR package on RNAseq data of COVID19 and SMK, DEGs of these diseases were identified, After that, common significant DEGs between diseases were identified. Correlation analysis using WGCNA for the common genes to see the characteristics of common genes. Then, identify common significant pathways and GO analysis through the pathway and go analysis on the common significant DEGs. Then, to identify the hub proteins, PPI network was constructed around the common genes, and to identify regulatory TFs and miRNA, DEGs–TFs and DEGs–miRNA networks were also constructed. Finally, the protein drug interaction network was constructed around the hub genes.

the high expression of ACE2 expression making it more prone to COVID-19 infection. In addition, a study of meta-analysis observed the significant effect between smokers and no-smokers, where they suggest an increased risk for viral binding and entry of SARS-CoV-2 in lungs epithelial cells of smokers due to significantly and substantially increased pulmonary ACE2 expression [11]. Rao et al. [12] reported some significant genes associated with smoking, alcohol and COVID-19. But, the molecular agents (differentially expressed genes) and pathways related to smoking and COVID-19 have not been identified yet. Thus, identifying the molecular associations at the level of transcriptional changes and pathways is important to identify the impact of smoking on COVID-19 infection prevalence.

Lately, transcriptional signatures of lung epithelial cells infected with COVID-19 have been identified [13]. We have taken this data and analyzed it to identify differentially expressed genes (DEGs) in infected lung epithelial and non-infected cells. Then, we identified DEGs in smokers' epithelial lungs compared to non-smokers. The weighted gene coexpression network analysis (WGCNA) is used to detect the correlated genes in a cluster [14]. It has been widely used for checking the strengthened relationship among the genes [15]. In this research, the network biology approach for the identification of commonly deregulated pathways and molecular signatures in smokers and COVID-19 was adopted. We utilized gene expression data of COVID-19-infected lung epithelial and another dataset of smoking lung epithelial cells to identify common DEGs and pathways overlapped in COVID-19 and smokers. We also constructed the correlation network analysis for these common genes using WGCNA to see the relationship among these genes.

2. Materials and methods

In this research, we employed a series of analysis methods (see Fig. 1). At first, DEGs of these diseases were identified using edgeR package on RNAseq data of COVID19 and SMK. After that, we identified the common significant genes between COVID19 and SMK to find out the disease–gene associations among them. Next, we applied WGCNA correlation analysis on common genes of COVID-19 and SMK datasets separately to see the characteristics pattern of common genes. Considering these common significant DEGs, we performed signaling pathways and GO analysis and then we constructed PPI network and used the topological analyses to identify hub proteins. To identify regulatory TFs and miRNA, DEGs–TFs and DEGs–miRNA networks were also constructed. Finally, using the identified hub genes obtained from PPI network analysis, a protein–drugs interaction network was constructed.

2.1. Datasets employed and statistical methods

RNAseq datasets of COVID-19 and SMK with accession numbers GSE147507 (COVID-19) and GSE47718(SMK) were downloaded from NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) database and were analyzed them to identify the shared significant genes between COVID-19 and SMK [16,17]. The GSE147507 (COVID-19) was RNA-seq transcriptomics data from lung cells infected with SARS-CoV-2 and controls (called mock). The gene expression data was profiled from A549 and NHBE cell within 24-h of infection. Then the samples infected for 24-h were considered and matched mock (controls). While the dataset GSE47718 is RNAseq data derived from the airway epithelium of healthy nonsmokers ($n = 10$) and smokers ($n = 7$) patients. We used edgeR package on RNAseq data of COVID19 and SMK to identify the DEGs of these diseases. We considered $|logFC| >= 1$, and p-value < 0.05 was considered statistically significant to identify DEGs.

2.2. Analysis methods

For comparing gene expression data of COVID-19 with the SMK dataset, global transcriptomic analysis was applied using RNAseq technologies. By comparing disease to normal, these datasets were produced to identify DEGs associated with their respective pathology. COVID19 and SMK datasets are RNAseq data. We applied edgeR [18]-R Bioconductor package-in raw RNAseq data to identify the DEGs of COVID19 and SMK. After that, by comparing two datasets of DEGs, we identified common DEGs between COVID-19 and SMK. For getting more shared information at the molecular level between COVID-19 and SMK, pathway and GO analyses were performed on shared genes between them by using online visual bioinformatics tools Enrichr (http://amp.pharm.mssm.edu/Enrichr/). We used Reactome and KEGG pathway databases for pathway analysis. To identify the significant pathways, we considered $p\text{-}value < 0.05$ as statistically significant. Then, using these common genes, we reconstructed a PPI network utilizing the STRING database [19] via Network Analyst [20]. To identify hub genes from the PPI network topological analysis was applied where degree and betweenness centrality were used.

We know that the TFs and miRNAs affect the expression of transcript levels. So, to see these affect, TF-common significant genes network from the JASPAR database [21] and miRNA-common significant genes network from TarBase [22] and miRTarBase [23] were constructed using NetworkAnalyst tools [20]. For finding significant TFs and miRNAs, the degree and betweenness centrality (BC) filters were used and identified top 10 TFs and miRNAs. Degree centrality can be defined as the following equation:

$$DC(v) = \sum_{j \in G} \frac{a_{vj}}{n-1}$$

Here, n indicates the number of nodes in the network and a_{vj} represents that node v and j are directly connected. BC is also defined as follows:

$$BC(v) = \sum_{i \neq j \neq v \in V} \frac{\sigma_{ivj}}{\sigma_{ij}}$$

Where σ_{ij} = total number of shortest paths from node i to node j, and σ_{ivj} = total number of paths through node v.

2.2.1. Weighted gene co-expression networks construction

To see the clustering nature of the detected common gene between COVID-19 and Smoking datasets, we used weighted gene co-expression network analysis (WGCNA) package [14] of *R* to find the weighted gene co-expression network among these genes. For this purpose, firstly we remove the outlier samples (if there exist) by constructing the sample cluster dendrogram by hclust *R* function for both of the datasets. Then using this *R* in this study, we used the pickSoftThreshold function for finding numerous soft thresholding powers β over R^2 and picking the value of β for which the R^2 value is higher. Then we construct the adjacency matrix and Topological Overlap Matrix (TOM) by using the transformed gene expression matrix. The dissimilarity of TOM (dissTOM) was also conducted to construct a network heatmap plot and for further analysis.

3. Results

3.1. Identification of differentially expressed genes

To identify the influence of smoking and COVID-19, we analyzed RNA seq datasets of COVID-19 and SMK obtained from the NCBI-GEO database. We detected there were 739 DEGs (353 up-regulated and 386 down-regulated) in COVID-19 and 3866 DEGs (1916 up-regulated and 1950 down-regulated) in SMK. Then, we performed the cross-comparative analysis to identify common DEGs between COVID-19 and SMK, identified 33 upregulated DEGS common in COVID-19 and SMK, and 26 downregulated DEGs common in COVID-19 and SMK. These common DEGs indicated that they are comorbid.

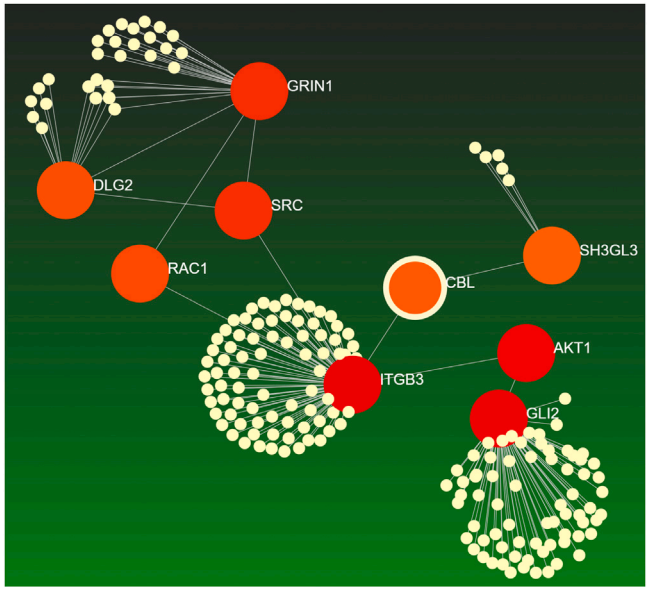

Fig. 2. Protein–protein interaction network around 59 common significant genes in COVID-19 and SMK.

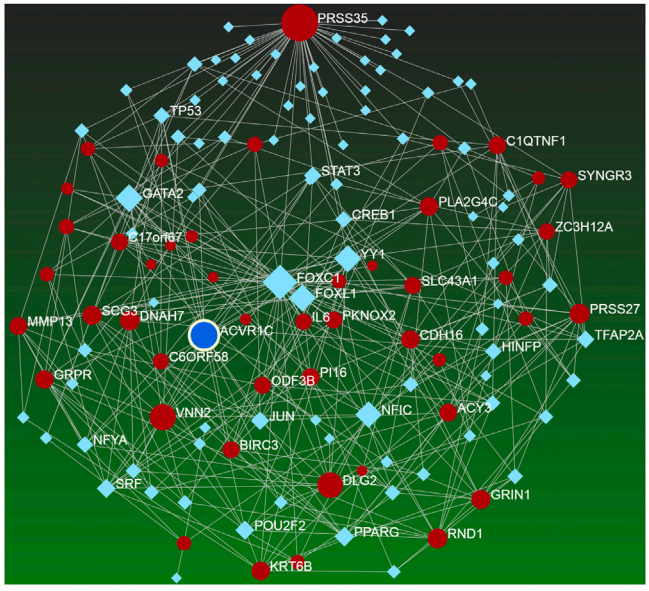

Fig. 3. The gene–TFs interaction network obtained from JASPAR database.

3.2. Functional enrichment of differentially expressed genes common to COVID-19 and SMK

Because of the underlying molecular or biological mechanisms, different diseases are related to each other that can be understood by pathway-based analysis [24]. For these reasons, we find out the common pathways using the common significant genes between COVID-19 and SMK from Enrichr, where we used KEGG and Reactome data as preferred data annotation for enrichment analysis. Significant common pathways are summarized in Table 1. To get further insight into the biological significance of the DEGs, we performed gene ontology analysis (Biological process, Molecular Function and Cellular component) on commonly significant genes for getting the more biological significance of these genes. Identified gene ontologies are summarized in Table 2.

Table 1
Molecular pathways enriched by the common differentially expressed genes shared by COVID-19 and SMK diseases. These include significant pathways common to COVID-19 and SMK.

	Pathways	P-value	Genes in Pathways
KEGG	Hippo signaling pathway	1.07E-02	DLG2;GLI2;BIRC3
	Amyotrophic lateral sclerosis	1.89E-02	DNAH7;DNAH17;PFN4;GRIN1
	Rap1 signaling pathway	2.10E-02	ITGB3;PFN4;GRIN1
	Hypertrophic cardiomyopathy	2.64E-02	IL6;ITGB3
	IL-17 signaling pathway	2.86E-02	IL6;MMP13
	Hematopoietic cell lineage	3.14E-02	IL6;ITGB3
	Salmonella infection	3.26E-02	IL6;PFN4;BIRC3
	Amoebiasis	3.32E-02	SERPINB4;IL6
	TNF signaling pathway	3.93E-02	IL6;BIRC3
	Glutamatergic synapse	4.06E-02	PLA2G4C;GRIN1
	Pathways of neurodegeneration	4.38E-02	IL6;DNAH7;DNAH17;GRIN1
	Platelet activation	4.73E-02	ITGB3;PLA2G4C
Reactome	GLI proteins bind promoters of Hh responsive genes to promote transcription	1.94E-02	GLI2
	Sema4D mediated inhibition of cell attachment and migration	2.22E-02	RND1
	MAPK1 (ERK2) activation	2.49E-02	IL6
	MAPK3 (ERK1) activation	2.77E-02	IL6
	MAPK1/MAPK3 signaling	3.00E-02	IL6;ITGB3;GRIN1
	Interleukin-6 signaling	3.04E-02	IL6
	Gastrin-CREB signalling pathway via PKC and MAPK	3.27E-02	ITGB3;GPR4;GRPR;GRIN1
	PECAM1 interactions	3.31E-02	ITGB3
	Platelet degranulation	3.50E-02	ITGB3;SCG3
	Response to elevated platelet cytosolic Ca2+	3.81E-02	ITGB3;SCG3
	GRB2:SOS provides linkage to MAPK signaling for Integrins	4.12E-02	ITGB3
	p130Cas linkage to MAPK signaling for integrins	4.12E-02	ITGB3
	MAPK family signaling cascades	4.53E-02	IL6;ITGB3;GRIN1
	Ras activation uopn Ca2+ infux through NMDA receptor	4.66E-02	GRIN1

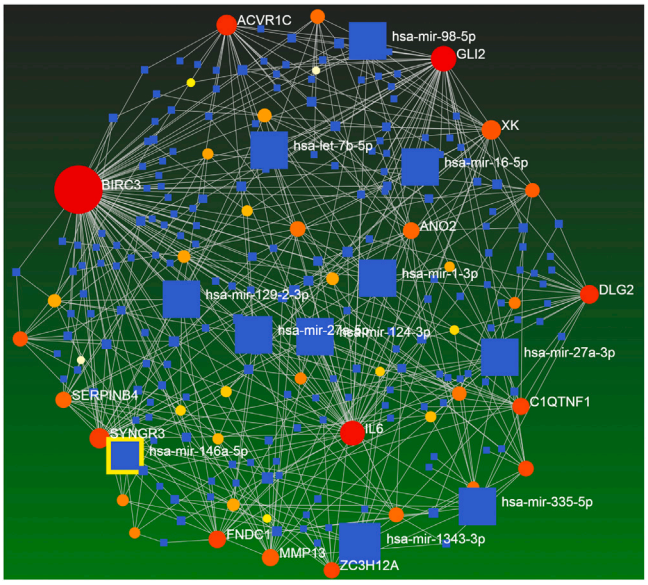

Fig. 4. The gene–microRNAs interaction network obtained from miRTarbase and Tarbase databases.

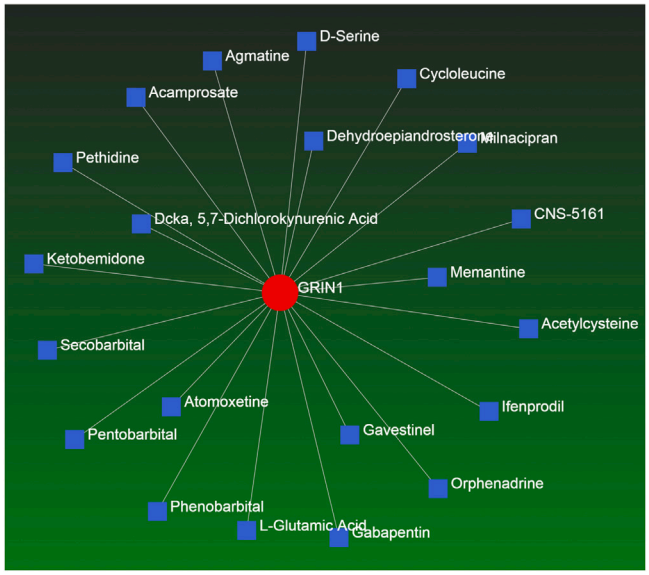

Fig. 5. The Protein–Drug interaction network analysis.

3.3. Hub proteins from protein–protein interaction analysis

We built PPI network around the common 59 significant genes in COVID-19 and SMK (Fig. 2) using the online visual software NetworkAnalyst. We identified nine hub genes namely ITGB3, GLI2, GRIN1, DLG2, SH3GL3, SRC, AKT1, RAC1 and CBL (see Fig. 2).

3.4. Regulators of the DEGs

We constructed the DEG–TFs interactions network (Fig. 3) and the DEG–miRNA interactions network(Fig. 4.). The top ten TFS (PRSS35, FOXC1, ACVR1C, GATA2, YY1, VNN2, FOXL1, NFIC, DLG2 and PRSS27) and miRNA (mir-27a-3p, mir-335-5p, mir-1-3p, mir-27a-5p, let-7b-5p, mir-124-3p, mir-1343-3p, mir-129-2-3p, mir-16-5p and mir-21-3p) were identified based on topological analysis (see Fig. 5).

Table 2
Gene ontology (biological processes) common to COVID-19 and SMK diseases.

	ID	Pathways	P-value	Genes in Pathways
Biological Process	GO:1905953	negative regulation of lipid localization	3.04E-04	IL6;ITGB3
	GO:1904037	positive regulation of epithelial cell apoptotic process	1.00E-03	IL6;ECSCR
	GO:0010888	negative regulation of lipid storage	1.00E-03	IL6;ITGB3
	GO:0045124	regulation of bone resorption	1.90E-03	IL6;ITGB3
	GO:0030262	apoptotic nuclear changes	2.08E-03	ACVR1C;DNASE1L3
	GO:0060285	cilium-dependent cell motility	2.26E-03	DNAH17;DNAH7
	GO:0050919	negative chemotaxis	2.66E-03	ITGB3;SEMA3G
	GO:0032675	regulation of interleukin-6 production	6.83E-03	IL6;ZC3H12A
	GO:0003341	cilium movement	9.20E-03	DNAH7;DNAH17
	GO:0045055	regulated exocytosis	9.54E-03	ITGB3;SCG3;SYNGR3
Molecular Function	GO:0008569	ATP-dependent microtubule motor activity, minus-end-directed	1.13E-03	DNAH17;DNAH7
	GO:1990939	ATP-dependent microtubule motor activity	8.85E-03	DNAH7;DNAH17
	GO:0005518	collagen binding	1.03E-02	C1QTNF1;MMP13
	GO:0004518	nuclease activity	1.31E-02	ZC3H12A;DNASE1L3
	GO:0003777	microtubule motor activity	1.36E-02	DNAH7;DNAH17
	GO:0043184	vascular endothelial growth factor receptor 2 binding	1.76E-02	ITGB3
	GO:0004532	exoribonuclease activity	1.76E-02	ZC3H12A
	GO:0004519	endonuclease activity	1.97E-02	ZC3H12A;DNASE1L3
	GO:0005138	interleukin-6 receptor binding	2.05E-02	IL6
	GO:0017002	activin-activated receptor activity	2.05E-02	ACVR1C
Cellular Component	GO:0005858	axonemal dynein complex	1.00E-03	DNAH17;DNAH7
	GO:0031258	lamellipodium membrane	2.34E-02	ITGB3
	GO:0031528	microvillus membrane	2.34E-02	ITGB3
	GO:0044224	juxtaparanode region of axon	2.62E-02	DLG2
	GO:0005887	integral component of plasma membrane	2.67E-02	IL6;C1QTNF1;ACVR1C;DLG2;ITGB3;SLC43A1;GPR4;GRPR;CDH16
	GO:0031527	filopodium membrane	3.48E-02	ITGB3
	GO:0098878	neurotransmitter receptor complex	3.48E-02	DLG2
	GO:0005788	endoplasmic reticulum lumen	4.55E-02	IL6;SCG3;ARSE
	GO:0031092	platelet alpha granule membrane	4.90E-02	ITGB3

3.5. Drug protein interaction

To predict possible drugs, we performed a protein–drugs interaction analysis. The analysis showed that GRIN1 protein had interactions with 21 drugs (Acamprosate, Atomoxetine, Gabapentin, L-Glutamic Acid, Memantine, Pentobarbital, Pethidine, Secobarbital, Agmatine, D-Serine, Gavestinel, Orphenadrine, Ifenprodil, Acetylcysteine, CNS-5161, Milnacipran, Dehydroepaindrosterone, Cycloleucine, Dcka, 5, 7-Dichlorokynurenic Acid, ketobemidone and Phenobarbital). The description of the drugs is summarized in Table 3.

3.6. Significant genes identification using WGCNA analysis

In Fig. 6, the gene co-expression network analysis has been executed with 59 DEGs for the COVID-19 dataset including 23 disease samples as well as the Smoking dataset including 7 disease samples. Fig. 6A to Fig. 6C for COVID-19 and Fig. 6D to Fig. 6F for SMK. Fig. 6A visualized the cluster dendrogram of the sample and no outlier samples were detected. To identify the modules through WGCNA, we found the optimized soft thresholding powers $\beta = 6$ as the scale-free topology criteria (Fig. 6B). With this power value, we constructed the co-expression network and detect 2 modules through the Dynamic Tree Cut technique using deepSplit = 2 and minClusterSize = 20 parameters. We found 12 and 47 genes for gray and turquoise modules respectively for COVID-19. The network heatmap of all genes with these 2 modules has been shown in Fig. 6C. Fig. 6D visualized the cluster dendrogram of the sample and no outlier samples were detected. To identify the modules through WGCNA, we found the optimized soft thresholding powers $\beta = 16$ as the scale-free topology criteria (Fig. 6E). With this power value, we constructed the co-expression network and detect 2 modules through the Dynamic Tree Cut technique using deepSplit = 2 and minClusterSize = 20 parameters. We found all 59 genes are in the gray module for SMK. The network heatmap of all genes with this module is shown in Fig. 6F.

4. Discussion

COVID-19 has been emerged as one of the worst pandemics and impacting health worldwide. SMK are more susceptible to COVID-19 infection and higher fatality was reported. In this research, we identified the molecular alterations at the level of transcriptome dynamics that occur in COVID-19-infected lung epithelial and shared transcriptional elements detected in the lung epithelial of SMK. When compared, discovered identified 33 upregulated DEGS common in COVID-19 and SMK and 26 downregulated DEGs common in COVID-19 and SMK. The WGCNA analysis identifies that the common genes showed the exact nature in both COVID-19 and Smoking datasets. The functional annotation of these common transcriptional signatures revealed several molecular pathways enriched by the DEGs. Some of the prominent pathways namely, the Hippo signaling pathway, Amyotrophic lateral sclerosis, Rap1 signaling pathway, Hypertrophic cardiomyopathy, IL-17 signaling pathway, Hematopoietic cell lineage, Salmonella infection, Amoebiasis, TNF signaling, MAPK1/MAPK3 signaling, and Interleukin-6 signaling pathways. It may be possible to consider the drugs that interfere in these pathways to repurpose drugs for COVID-19 infection. Rap1 signaling and MAPK signaling pathways have significant associations with COVID-19 [25]. Another Pathways Hippo signaling has a strong association with COVID-19 infection [26]. One significant pathway common between COVID-19 and SMK is the Interleukin-6 signaling pathway which has a significant impact on the COVID-19 patients as

Table 3
Descriptions of drugs.

Drugs	Status	Descriptions of Drugs uses
Acamprosate	Approved, Investigational	is a drug used for treating alcohol dependence.
Atomoxetine	Approved	used in the treatment of attention deficit hyperactivity disorder (ADHD) in children and adults.
Gabapentin	Approved, Investigational	,it is used in the treatment of Partial-Onset Seizures.
L-Glutamic Acid	Approved, Nutraceutical	helps control alcoholism, schizophrenia and the craving for sugar
Memantine	Approved, Investigational	is used to manage moderate to severe Alzheimer's dementia
Pentobarbital	Approved, Investigational, Vet approved	Used For the short-term treatment of insomnia.
Pethidine	Approved	Used to control moderate to severe pain.
Secobarbital	Approved, Vet approved	Secobarbital is a barbiturate derivative drug with anaesthetic, anticonvulsant, sedative and hypnotic properties.
Agmatine	Experimental	Agmatine is being studied experimentally for several indications such as cardioprotection, diabetes, decreased kidney function, neuroprotection (stroke, severe CNS injuries, epilepsy, glaucoma, and neuropathic pain), and psychiatric conditions (depression, anxiety, schizophrenia, and cognition).
D-Serine	Approved, Experimental	D-Serine is an amino acid that plays a role in cognitive enhancement and schizophrenia treatment.
Gavestinel	Investigational	Excitatory Amino Acid Agents
Orphenadrine	Approved	used to treat drug-induced parkinsonism and to relieve pain from muscle spasm.
Ifenprodil	Approved, Investigational	ithdrawn, Ifenprodil is a selective NMDA receptor (glutamate) antagonist.
Acetylcysteine	Approved, Investigational	is used mainly as a mucolytic and in the management of paracetamol (acetaminophen) overdose.
CNS-5161	Investigational	Investigated for use/treatment in migraine and cluster headaches, neuropathy (diabetic), and pain (acute or chronic).
Milnacipran	Approved, Investigational	may be used for the treatment of major depressive disorder (MDD)
Dehydroepaindrosterone	Approved, Investigational, Nutraceutical	used for the treatment of schizophrenia ; improving the appearance of older people's skin ; improving ability to achieve an erection in men with sexual dysfunction.
Cycloleucine	Experimental	Amino Acids, widely used in biochemical experiments
Dcka,5,7-Dichlorokynurenic Acid	Experimental	
Ketobemidone	Investigational,	Used for the treatment of all types of severe pain, such as postoperative, cancer, kidney stones and fractures.
Phenobarbital	Approved, Investigational,	For the treatment of all types of seizures except absence seizures.

well as the patients of SMK [27,28]. From some studies, it was reported that the L-17 signaling pathway has significant rules with the patients of COVID-19 and SMK [29–31]. Exposing to tobacco smoke contributes to the inflammatory lung phase, increased activation of the mucosa, the release of inflammatory cytokines, tumour necrosis factor α, epithelial permeability, excess mucous and compromised clearing [32]. However, a large number of research has to date shown that the most significant cause of COVID-19 patient death is a moderate or extreme cytokine (excessive immune response) floods. Interleukin-6 (IL-6) has a great role in cytokine excess release. As the signal transduction pathway of IL-6 can be disrupted, it will potentially evolve into a new method of treating serious COVID-19 patients [33,34]. Nasir et al. described the research results from cellular and experimental investigations defining the function of IL-6 as a therapeutic target in COVID-19 [35]. Protein–protein interaction network topological analysis provides critical proteins and signaling molecules, and drug targets. The PPI analysis was performed to determine the hub proteins — ITGB3, GLI2, GRIN1, DLG2, SH3GL3, SRC, AKT1, RAC1 and CBL. These hub proteins might be considered as drug targets. Tin Wang et al. [36] reported RAC1 protein as a therapeutic target in Acute Lung Injury induced by severe pneumonia of COVID-19. Another hub protein CBL was reported as a potential drug target for COVID-19 [37]. The hub gene AKT1 is associated with SMK and COVID-19 [36]. In the replication of viruses, particularly those connected to SARS-CoV-2, several SRC family kinases have been found to be active [38]. Another paper suggested that SRC protein is a potential targeted drug for COVID-19 [39]. The hub protein ITGB3 was declared as a potential therapeutic target for COVID-19-related stroke [40]. GLI2 variants play a role in the pathogenesis of nonsyndromic cleft lip with or without cleft palate [41] and hazra et al. [42] found possible interaction of GLI2 with COVID-19. GRIN1 involved in neurodevelopmental disease [43]. DLG2 is involved in the developmental and intellectual disability [44].

Lack of lab facilities, the qRT-PCR analysis was not performed for the further validation of our identified significant hub proteins which is the limitation of our work. To provide insights into the transcriptional regulations at the level of transcriptional and post-transcriptional levels, we built a DEGs–TFs and DEGs–miRNAs interaction network. We need to test our findings in the laboratory. Otherwise, covid-19 research is ongoing and still, no research finding is concrete or absolute.

5. Conclusions

The present study was intended to decipher the influence of smoking on COVID-19 infection using High-throughput RNA-Seq gene expression data. Our employed methodologies showed the 59 DEG shared between COVID-19 and SMK. Further pathway analysis of the identified DEGs demonstrated the inflammatory pathways are crucial in COVID-19 infections and smoking. The hub genes ITGB3, GLI2, GRIN1, DLG2, SH3GL3, SRC, AKT1, RAC1 and CBL were detected from PPI analysis that might be candidate genes. The network-based further regulomics data showed the TFS (PRSS35, FOXC1, ACVR1C, GATA2, YY1, VNN2, FOXL1, NFIC, DLG2 and PRSS27) and miRNAs (mir-27a-3p, mir-335-5p, mir-1-3p, mir-27a-5p, let-7b-5p, mir-124-3p, mir-1343-3p, mir-129-2-3p, mir-16-5p and mir-21-3p) may regulate the identified DEGs. Since the above results are based on in silico analysis, thus we propose to conduct experimental studies to evaluate the molecular functions of the identified genes and molecules.

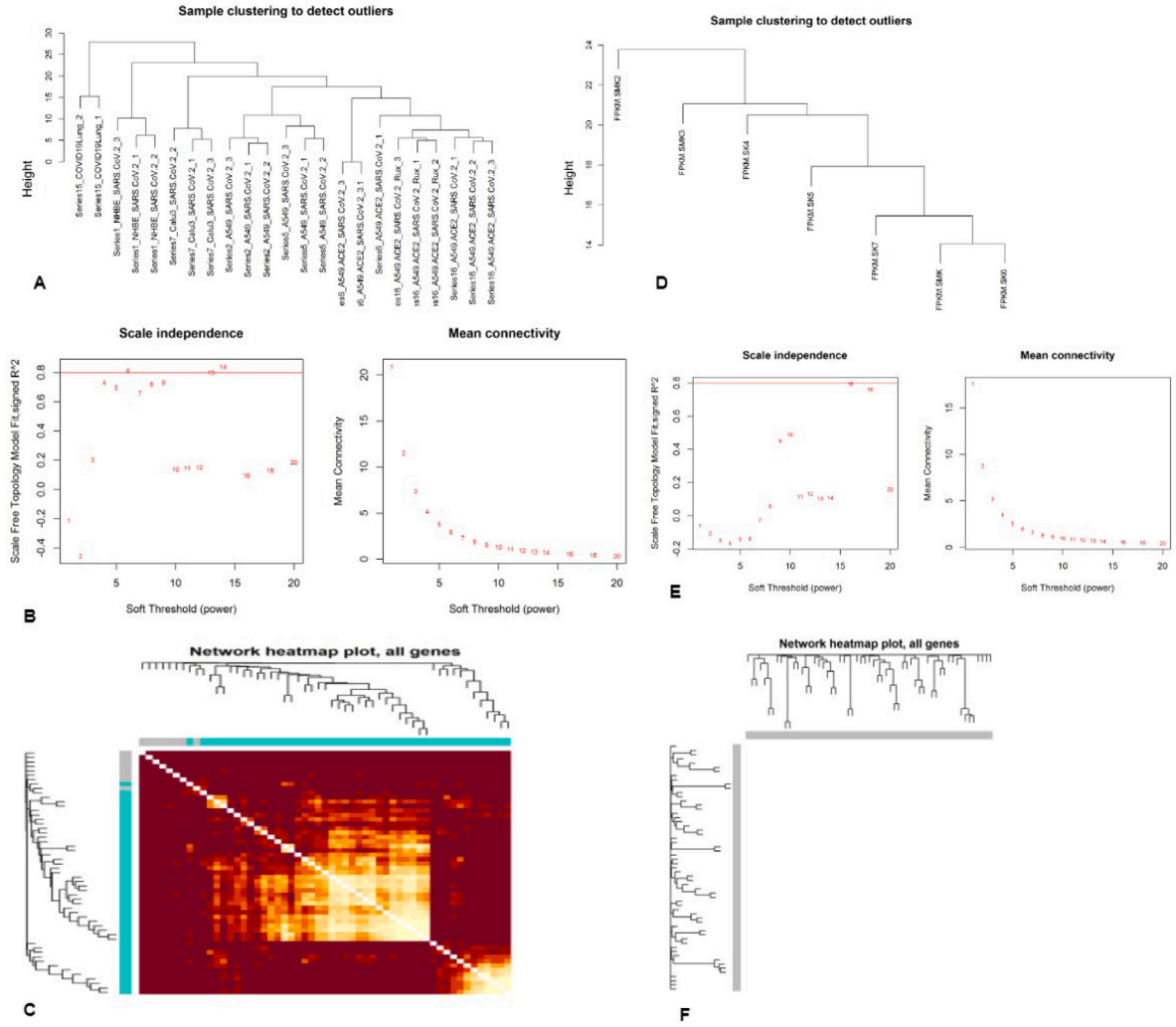

Fig. 6. Correlation analysis using WGCNA for the common significant genes in COVID-19 and SMK.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei, et al., Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study, Lancet 395 (10223) (2020) 507–513.

[2] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, et al., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet 395 (10223) (2020) 497–506.

[3] D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, et al., Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China, JAMA 323 (11) (2020) 1061–1069.

[4] C.I. Vardavas, K. Nikitara, COVID-19 and smoking: A systematic review of the evidence, Tob. Induc. Dis. 18 (March) (2020) http://dx.doi.org/10.18332/tid/119324.

[5] A. Umnuaypornlert, S. Kanchanasurakit, D.E.I. Lucero-Prisno, S. Saokaew, Smoking and risk of negative outcomes among COVID-19 patients: A systematic review and meta-analysis, Tob. Induc. Dis. 19 (2021).

[6] J. Baker, N. Krishnan, L.C. Abroms, C.J. Berg, The impact of tobacco use on COVID-19 outcomes: A systematic review, J. Smok. Cessat. 2022 (2022).

[7] A.K. Clift, A. von Ende, P. San Tan, H.M. Sallis, N. Lindson, C.A. Coupland, M.R. Munafò, P. Aveyard, J. Hippisley-Cox, J.C. Hopewell, Smoking and COVID-19 outcomes: an observational and mendelian randomisation study using the UK Biobank cohort, Thorax 77 (1) (2022) 65–73.

[8] N. Ismail, N. Hassan, M.H.N. Abd Hamid, U.N. Yusoff, N.R. Khamal, M.A. Omar, X.C. Wong, M.D. Pathmanathan, S.M. Zin, F.M. Zin, et al., Association of smoking and severity of COVID-19 infection among 5,889 patients in Malaysia: a multi-center observational study, Int. J. Infect. Dis. 116 (2022) 189–196.

[9] R. Poudel, L.B. Daniels, A.P. DeFilippis, N.M. Hamburg, Y. Khan, R.J. Keith, R.S. Kumar, A.C. Strokes, R.M. Robertson, A. Bhatnagar, Smoking is associated with increased risk of cardiovascular events, disease severity, and mortality among patients hospitalized for SARS-CoV-2 infections, PLoS One 17 (7) (2022) e0270763.

[10] Y. He, Y. He, Q. Hu, S. Yang, J. Li, Y. Liu, J. Hu, Association between smoking and COVID-19 severity: A multicentre retrospective observational study, Medicine 101 (29) (2022).

[11] G. Cai, Y. Bossé, F. Xiao, F. Kheradmand, C.I. Amos, Tobacco smoking increases the lung gene expression of ACE2, the receptor of SARS-CoV-2, Am. J. Respir. Crit. Care Med. (ja) (2020).

[12] S. Rao, A. Baranova, H. Cao, J. Chen, X. Zhang, F. Zhang, Genetic mechanisms of COVID-19 and its association with smoking and alcohol consumption, Brief. Bioinform. 22 (6) (2021) bbab284.

[13] D. Blanco-Melo, B. Nilsson-Payant, W.-C. Liu, R. Møller, M. Panis, D. Sachs, R. Albrecht, et al., SARS-CoV-2 launches a unique transcriptional signature from in vitro, ex vivo, and in vivo systems, BioRxiv (2020).

[14] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinformatics 9 (1) (2008) 1–13.

[15] M.R. Auwul, M.R. Rahman, E. Gov, M. Shahjaman, M.A. Moni, Bioinformatics and machine learning approach identifies potential drug targets and pathways in COVID-19, Brief. Bioinform. 22 (5) (2021) bbab120.

[16] D. Blanco-Melo, B.E. Nilsson-Payant, W.-C. Liu, S. Uhl, D. Hoagland, R. Møller, T.X. Jordan, K. Oishi, M. Panis, D. Sachs, et al., Imbalanced host response to SARS-CoV-2 drives development of COVID-19, Cell (2020).

[17] D.M. Ryan, T.L. Vincent, J. Salit, M.S. Walters, F. Agosto-Perez, R. Shaykhiev, Y. Strulovici-Barel, R.J. Downey, L.J. Buro-Auriemma, M.R. Staudt, et al., Smoking dysregulates the human airway basal cell transcriptome at COPD risk locus 19q13. 2, PLoS One 9 (2) (2014) e88051.
[18] M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data, Bioinformatics 26 (1) (2010) 139–140.
[19] D. Szklarczyk, J.H. Morris, H. Cook, M. Kuhn, S. Wyder, M. Simonovic, A. Santos, N.T. Doncheva, A. Roth, P. Bork, et al., The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible, Nucleic Acids Res. (2016) gkw937.
[20] J. Xia, E.E. Gill, R.E. Hancock, NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data, Nat. Protoc. 10 (6) (2015) 823.
[21] A. Khan, O. Fornes, A. Stigliani, M. Gheorghe, J.A. Castro-Mondragon, R. van der Lee, A. Bessy, J. Cheneby, S.R. Kulkarni, G. Tan, et al., JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework, Nucleic Acids Res. 46 (D1) (2017) D260–D266.
[22] P. Sethupathy, B. Corda, A.G. Hatzigeorgiou, TarBase: A comprehensive database of experimentally supported animal microRNA targets, Rna 12 (2) (2006) 192–197.
[23] S.-D. Hsu, F.-M. Lin, W.-Y. Wu, C. Liang, W.-C. Huang, W.-L. Chan, W.-T. Tsai, G.-Z. Chen, C.-J. Lee, C.-M. Chiu, et al., miRTarBase: a database curates experimentally validated microRNA–target interactions, Nucleic Acids Res. 39 (suppl_1) (2010) D163–D169.
[24] L. Jin, X.-Y. Zuo, W.-Y. Su, X.-L. Zhao, M.-Q. Yuan, L.-Z. Han, X. Zhao, Y.-D. Chen, S.-Q. Rao, Pathway-based analysis tools for complex diseases: a review, Genom., Proteom. Bioinform. 12 (5) (2014) 210–220.
[25] P. Khanal, Y.N. Dey, R. Patil, R. Chikhale, M.M. Wanjari, S.S. Gurav, B. Patil, B. Srivastava, S.N. Gaidhani, Combination of system biology to probe the anti-viral activity of andrographolide and its derivative against COVID-19, RSC Adv. 11 (9) (2021) 5065–5079.
[26] G. Garcia Jr., Y. Wang, J.I. Irudayam, A.V. Jeyachandran, S.C. Cario, C. Sen, S. Li, Y. Li, A. Kumar, K. Nielsen-Saines, et al., Hippo signaling pathway activation during SARS-CoV-2 infection contributes to host antiviral response, BioRxiv (2022).
[27] G. Magro, SARS-CoV-2 and COVID-19: is interleukin-6 (IL-6) the'culprit lesion'of ARDS onset? What is there besides tocilizumab? SGP130fc, Cytokine: X (2020) 100029.
[28] A. Jamil, A. Rashid, A.K. Naveed, M. Asim, Effect of smoking on interleukin-6 and correlation between IL-6 and serum amyloid A-low density lipoprotein in smokers, J. Postgrad. Med. Inst. (Peshawar-Pakistan) 31 (4) (2017).
[29] O. Pacha, M.A. Sallman, S.E. Evans, COVID-19: a case for inhibiting IL-17? Nat. Rev. Immunol. 20 (6) (2020) 345–346.
[30] K.-H. Lee, C.-H. Lee, J. Woo, J. Jeong, A.-H. Jang, C.-G. Yoo, Cigarette smoke extract enhances IL-17a-induced IL-8 production via up-regulation of IL-17R in human bronchial epithelial cells, Mol. Cells 41 (4) (2018) 282.
[31] M.Z. Hasan, S. Islam, K. Matsumoto, T. Kawai, SARS-CoV-2 infection initiates interleukin-17-enriched transcriptional response in different cells from multiple organs, Sci. Rep. 11 (1) (2021) 1–11.
[32] A. Strzelak, A. Ratajczak, A. Adamiec, W. Feleszko, Tobacco smoke induces and alters immune responses in the lung triggering inflammation, allergy, asthma and other lung diseases: a mechanistic review, Int. J. Environ. Res. Public Health 15 (5) (2018) 1033.
[33] C. Zhang, Z. Wu, J.-W. Li, H. Zhao, G.-Q. Wang, The cytokine release syndrome (CRS) of severe COVID-19 and Interleukin-6 receptor (IL-6R) antagonist tocilizumab may be the key to reduce the mortality, Int. J. Antimicrob. Ag. (2020) 105954.
[34] M. Shirvaliloo, The unfavorable clinical outcome of COVID-19 in smokers is mediated by H3K4me3, H3K9me3 and H3K27me3 histone marks, Epigenomics 14 (3) (2022) 153–162.
[35] A. Nasirzadeh, J. Bazeli, J. Hajavi, N. Yavarmanesh, M. Zahedi, M. Abounoori, A. Razavi, M.M. Maddah, P. Mortazavi, M. Moradi, et al., Inhibiting IL-6 during cytokine storm in COVID-19: Potential role of natural products, 2021.
[36] T. Wang, M. Yegambaram, C. Gross, X. Sun, Q. Lu, H. Wang, X. Wu, A. Kangath, H. Tang, S. Aggarwal, et al., RAC1 nitration at Y32 IS involved in the endothelial barrier disruption associated with lipopolysaccharide-mediated acute lung injury, Redox Biol. 38 (2021) 101794.
[37] G. Selvaraj, S. Kaliamurthi, G.H. Peslherbe, D.-Q. Wei, Identifying potential drug targets and candidate drugs for COVID-19: biological networks and structural modeling approaches, F1000Research 10 (2021).
[38] E. Weisberg, A. Parent, P.L. Yang, M. Sattler, Q. Liu, Q. Liu, J. Wang, C. Meng, S.J. Buhrlage, N. Gray, et al., Repurposing of kinase inhibitors for treatment of COVID-19, Pharm. Res. 37 (9) (2020) 1–29.
[39] M. Zou, X. Su, L. Wang, X. Yi, Y. Qiu, X. Yin, X. Zhou, X. Niu, L. Wang, M. Su, The molecular mechanism of multiple organ dysfunction and targeted intervention of COVID-19 based on time-order transcriptomic analysis, Front. Immunol. (2021) 3366.
[40] G. Cen, L. Liu, J. Wang, X. Wang, S. Chen, Y. Song, Z. Liang, Weighted gene co-expression network analysis to identify potential biological processes and key genes in COVID-19-related stroke, Oxidative Med. Cellular Longev. (2022).
[41] P. Meng, H. Zhao, W. Huang, Y. Zhang, W. Zhong, M. Zhang, P. Jia, Z. Zhou, G. Maimaitili, F. Chen, et al., Three GLI2 mutations combined potentially underlie non-syndromic cleft lip with or without cleft palate in a Chinese pedigree, Molecular Genet. Genom. Med. 7 (9) (2019) e714.
[42] S. Hazra, A.G. Chaudhuri, B.K. Tiwary, N. Chakrabarti, Candidate genes associated with neurological manifestations of COVID-19: Meta-analysis using multiple computational approaches, BioRxiv (2022).
[43] K. Platzer, J.R. Lemke, GRIN1-related neurodevelopmental disorder, in: GeneReviews® [Internet], University of Washington, Seattle, 2019.
[44] C. Reggiani, S. Coppens, T. Sekhara, I. Dimov, B. Pichon, N. Lufin, M.-C. Addor, E.F. Belligni, M.C. Digilio, F. Faletra, et al., Novel promoters and coding first exons in DLG2 linked to developmental disorders and intellectual disability, Genome Med. 9 (1) (2017) 1–20.