Biochemistry and Biophysics Reports 29 (2022) 101219

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

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

Common genetic aspects between COVID-19 and sarcoidosis: A network-based approach using gene expression data

Md. Roman Mogal [a], Sagarika Adhikary Sompa [a], Asadullah Junayed [a], Md. Rashel Mahmod [a], Md. Zainul Abedin [b], Md. Asaduzzaman Sikder [a,*]

[a] Department of Biochemistry and Molecular Biology, Mawlana Bhashani Science and Technology University, Tangail, 1902, Bangladesh
[b] Department of Food Technology and Nutritional Science, Mawlana Bhashani Science and Technology University, Tangail, 1902, Bangladesh

ARTICLE INFO

Keywords:
COVID-19
Sarcoidosis
Protein-protein interaction
Hub genes
Drug network
Vitamin D

ABSTRACT

The pandemic situation of novel coronavirus disease 2019 (COVID-19) is a global threat on our current planet, with its rapid spread and high mortality rate. Sarcoidosis patients are at high risk to COVID-19 severity for having lung injuries as well as treating with immunosuppressive agents. So, physicians are in dilemma whether they should use immunosuppressive agents or not for the patients with sarcoidosis history and COVID-19 infection. Therefore, common factors should be identified to provide effective treatment. For determining the common genes between COVID-19 and sarcoidosis, GSE164805 and GSE18781 were retrieved from the Gene Expression Omnibus (GEO) database. Common upregulated genes were identified by using R language to investigate their involved pathways and gene ontologies (GO). With the aid of the STRING Cytoscape plugin tool, protein-protein interactions (PPIs) network was constructed. From the PPIs network, Hub genes and essential modules were detected by using Cytohubba, and MCODE respectively. For hub genes, TFs, TFs-miRNA, and drug, interaction networks were built through the NetworkAnalyst web platform. A total of 34 common upregulated genes were identified and among them, five hub genes, including TET2, MUC5AC, VDR, NFE2L2, and BCL6 were determined. In addition, a cluster having VDR and NFE2L2 was detected from the PPIs network. Moreover, 32 transcription factors and 9 miRNA were recognized for hub genes. Furthermore, vitamin D and some of its analogous compounds were obtained from the drug interaction network. In conclusion, hub genes identified in this study might have potential roles in modulating COVID-19 infection and sarcoidosis. However, further studies are required to corroborate this study.

1. Introduction

Coronavirus disease 2019 (COVID-19) is a rapidly spreading infectious disease that is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). COVID-19 first reported a cluster of causing pneumonia in Wuhan, Hubei Province on December 31, 2019 and subsequently declared as a pandemic on March 11, 2020 [1]. Pandemics are extensive outbreaks of infectious diseases that pose a serious challenge to medical systems owing to the complexity of estimating the number and severity of cases, the lack of immediate vaccinations, and inadequate antiviral stockpiles to meet demand [2]. According to WHO, until January 13, 2022, the number of confirmed cases in the world were 312, 173,462 including 5,501,000 deaths [3]. WHO also identified highly infected countries including the USA, India, Brazil, UK, France, Russian Federation, Turkey, Italy, Germany, Spain, Argentina, and Iran [3]. Although some vaccines are available in the world, the pandemic will be under control if they are given to enough people [4]. To gain global control of COVID-19, however, approved vaccines yet to achieve some goals: need to be manufactured at high volume, provide affordable price, distribute globally so that they are available where needed, and should be widely implemented in local communities [5]. Therefore, it is suggested to wear the mask, maintain social distance, wash hands frequently, and take extra precaution to pulmonary diseased people to control the severity of COVID-19 pandemic situation.

Sarcoidosis is an inflammatory disease characterized by granuloma formation that can affect multiple organs but dominate the lungs. The etiology of sarcoidosis is still yet to demystify. However, sarcoidosis is thought to be caused by a combination of genetic, host immunologic, and environmental factors [6]. The initial cause of granulomatous inflammation is thought to involve the innate immune response where

* Corresponding author.
E-mail address: sikderma@mbstu.ac.bd (Md.A. Sikder).

https://doi.org/10.1016/j.bbrep.2022.101219
Received 6 November 2021; Received in revised form 18 January 2022; Accepted 21 January 2022

phagocytic cells express pattern recognition receptors such as Toll-like receptors (TLRs) in an attempt to engulf poorly soluble or insoluble material [7]. Even though many people undergo disease remission within the first few years, over 30–50% develop a chronic active disease that requires medication to prevent organ failure and fibrosis [6]. It was reported, in 2010, that the peak age for sarcoidosis in women was 50–69 years old, and for men, it was 40–59 years old [8].

Patients having pulmonary sarcoidosis impose a higher risk for severe illness related to COVID-19 [9,10]. According to Baughman et al. study, in case of sarcoidosis patients, the overall rate of COVID-19 was 2.23% or 22,308 cases per million from April to July 2020 and the rate of hospitalization was 15.8% who required ICU care [10]. Another cohort study reported that the rate of COVID-19 in sarcoidosis patients was 2.1% from March to April 2020 [11]. For several reasons sarcoidosis patients may have severe conditions and death from COVID-19: Involving lungs as a prominent target in sarcoidosis, immunological dysfunction, and dysregulation, providing immunosuppressive medications as primary treatment and having significant relation of sarcoidosis with hypertension, diabetes, and obesity, since they have been identified as independent risk factors for COVID-19 [12]. Studies reported that patients with COVID-19 infection are more likely to be harmed by corticosteroids treatment [13,14]. For all of these reasons, physicians are in predicament to offer effective treatment to patients with COVID-19 infection with sarcoidosis [15]. Therefore, common genetic aspects should be identified to ensure high outcomes in these conditions. This study was designed to find out common genes between COVID-19 and sarcoidosis by utilizing publicly available peripheral blood cells microarray datasets. Initially, upregulated genes were identified, and subsequently GO ontology, PPIs network, hub genes, TFs network, TFs-miRNA network, and drugs network were analyzed.

2. Methods

2.1. Acquisition of datasets

Gene expression profiles of microarray data for COVID-19 (GSE164805) and sarcoidosis (GSE18781) were obtained from Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/gds). Both datasets are dedicated to peripheral blood cells. GPL26963 (Agilent-085982 Arraystar human lncRNA V5 microarray) platform was used for the GSE164805 dataset where samples were got from five severe COVID-19 patients, five mild COVID-19 patients, and five healthy people. On the contrary, for the GSE18781 dataset, GPL570 (Affymetrix Human Genome U133 plus 2.0 Array) platform was adopted where samples were collected from twelve sarcoidosis patients and twelve control subjects.

2.2. Identification of common upregulated genes

For identifying common differentially expressed genes (DEGs) from both datasets, GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/) was applied. GEO2R is used to identify DEGs, using GEOquery and limma R packages from the Bioconductor project, having compared two or more groups of samples [16]. In order to control the false discovery rate, Benjamini-Hochberg was conducted [17]. Common DEGs from both datasets were downloaded from GEO2R as table format and subsequently imported in RStudio for further analysis. Datasets were filtered by setting adjusted p-value<0.01 and Log2-fold change < -1. A Venn diagram was constructed by using Bioinformatics and Evolutionary Genomics (http://bioinformatics.psb.ugent.be/webtools/Venn/) web tool followed by getting common upregulated genes.

2.3. Enrichment analysis of common upregulated genes

Gene set enrichment analysis process is used to evaluate collective behaviors of genes in terms of health and diseases [18]. Gene ontologies (GO) and associated pathways of common upregulated genes were retrieved through Enrichr (https://maayanlab.cloud/Enrichr/). Enrichr is a web-based enrichment analysis platform that is simple to use and provides various collective functions of gene lists [19,20]. For pathways analysis, KEGG, BioPlanet, Reactome, and MSigDB databases were selected.

2.4. Network analysis

Networks analysis is an essential part of system biology to understand the molecular and cellular interaction of proteins [21]. In addition, insightful knowledge can be gained about gene sets from network-based studying without studying individual genes [22]. For constructing the network, STRING protein quarry Cytoscape plugin tool was adopted. Search Tool for the Retrieval of Interacting Genes (STRING) (https://string-db.org/) is utilized to provide information about direct (physical) and indirect (functional) associations of protein-protein interaction that covers information more than 2000 organisms [23]. Upregulated genes were inserted in the search box where the confidence score was set up at 0.400. After retrieving the relevant network, it was visualized and modified by Cytoscape [24] and other associated plugin tools.

2.5. Hub genes identification and module analysis

Network has a lot of nodes and edges, among them, those nodes that have higher interconnection are considered hub genes. Having a higher degree of interconnection, hub genes tend to be essential for maintaining biological processes. From our PPI networks, hub genes were determined by using cytoHubba, which is a Cytoscape plugin software. CytoHubba is a user-friendly tool for exploring key nodes in biological networks by providing 11 topological analysis methods [25]. Degree topological method was applied for current research since it works based on the number of interactions for each gene in the PPIs network. Molecular Complex Detection (MCODE), another Cytoscape plugin tool, was exploited to detect highly interconnected portions in the PPIs network. MCODE makes visualization manageable by extracting the dense regions around a protein of interest [26].

2.6. Transcriptional factor regulatory network of hub genes

In mammalian development and adult tissue homeostasis, transcription factors (TFs) networks are a central determinant of cell fate decisions, and they are often corrupted in disease [27]. TFs gene interaction network for hub genes was constructed by NetworkAnalyst (https://www.networkanalyst.ca/NetworkAnalyst/uploads/ListUploadView.xhtml). NetworkAnalyst is a comprehensive web platform that provides a visual network for gene expression analysis [28,29]. TFs gene interaction network was retrieved from JASPAR (http://jaspar.genereg.net/) which is included in NetworkAnalyst platform.

2.7. TFs-miRNA regulatory network analysis

TFs can activate or repress transcription at a pre-transcriptional level [30], while microRNAs can control gene expression at the post-transcriptional stage [31]. TFs-miRNA regulatory network, for hub genes, was built by RegNetwork [32] repository through NetworkAnalyst platform. The network was filtered at 1° cutoff value. Finally, the network was downloaded from the NetworkAnalyst and visualized in Cytoscape software.

2.8. Protein drug interaction network

Lastly, the drug interaction network, for our hub genes, was constructed to know the potential drugs for both COVID-19 and sarcoidosis diseases. The network was made based on DrugBank (https://go.drug

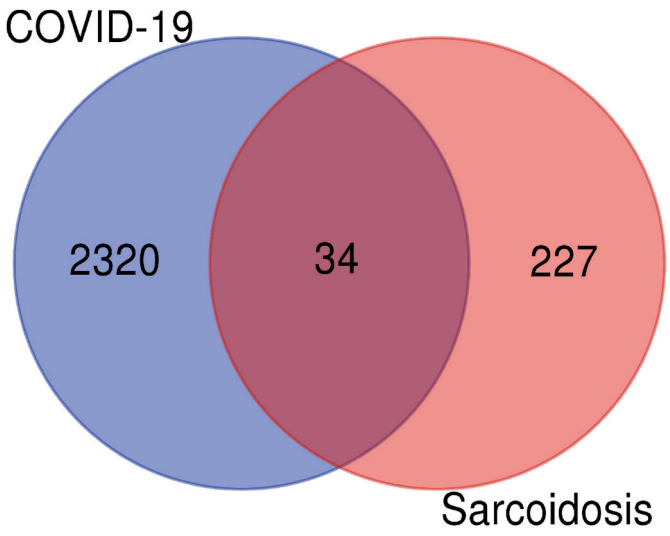

Fig. 1. A Venn diagram of commonly differentially expressed upregulated genes. Common 34 upregulated genes were identified from 2320 upregulated genes of COVID-19 infection and 227 upregulated genes of sarcoidosis.

bank.com/) database [33] through the NetworkAnalyst web platform. The network file was downloaded from NetworkAnalyst and visualized by using Cytoscape.

3. Results

3.1. Identification of common upregulated genes between COVID-19 and sarcoidosis

Selected datasets (GSE164805 and GSE18781) were analyzed to sort out common upregulated genes in between COVID-19 and sarcoidosis patients. 34 common upregulated genes were obtained from GSE164805 and GSE18781 datasets. These common upregulated genes were included as SCARF1, GNAS, ASS1, KLHL6, KIAA1109, VCAN, TBC1D22B, ACTR2, ST3GAL2, KDM4A, TCF7L2, TET2, CSAD, ASAP1, NXPH3, SDCCAG8, SEZ6L2, MALAT1, SLC11A2, NFE2L2, CEACAM1, SMCHD1, BCL6, TMEM80, C7orf50, PRR14L, SAMD9L, VDR, NECTIN2, GTDC1, MUC5AC, GTPBP2, XK, and KATNBL1. The Venn diagram in Fig. 1 indicates the comparison of common upregulated genes.

3.2. Enrichment analysis of common upregulated genes

GO and associated pathways of upregulated genes were analyzed by the Enrichr web platform. Significant ($p < 0.05$) ontologies were listed in Table 1 from three GO subsections (biological process, molecular function, and cellular component). This table indicates, the biological process was enriched with cellular response to laminar fluid shear stress, regulation of natural killer cells, and so on. Moreover, transmembrane transporter activity, beta-galactoside (CMP) alpha-2,3-sialyltransferase activity, and so forth were prominent in molecular function. Furthermore, for cellular components, zonula adherens, t-cell receptor complex were putative components. On the other hand, pathways related to inflammatory response were predominantly found from pathway databases. Significant ($p < 0.05$) pathways from different pathways databases are described in Table 2. Finally, depending on the combined score, a collection of GO terms and associated pathways were depicted in Fig. 2. The combined score is defined by Enrichr based on p-values and z-score.

3.3. Network analysis

The corresponding network was retrieved through the STRING protein quarry plugin tool in Cytoscape software. This network was stored, further to identify hub genes and subsequently for suggesting common drug molecules for both COVID-19 and sarcoidosis. Getting the common synergistic genetic factors and suggesting therapeutic targets for COVID-19 and sarcoidosis were the primary purposes of this study. The network consisted of 32 nodes and 102 edges (Fig. 3).

3.4. Hub genes identification and module analysis

For sorting out the highly interconnected genes in the PPIs network, CytoHubba was utilized. Hub genes were determined by applying the Degree algorithm method. Five genes including TET2, MUC5AC, VDR, NFE2L2, and BCL6 were identified as hub genes (Fig. 4). On the other

Table 1
Gene ontologies, their corresponding p-values and common upregulated genes.

Category	GO ID	GO terms	p-value	Genes
GO Biological process 2018	GO:0071499	Cellular response to laminar fluid shear stress.	0.00005859	ASS1, NFE2L2
	GO:0002858	Regulation of natural killer cell mediated cytotoxicity directed against tumor cell target.	0.0001252	CEACAM1,NECTIN2
	GO:0030856	Regulation of epithelial cell differentiation.	0.0004219	CEACAM1, KIAA1109
	GO:0002889	Regulation of immunoglobulin mediated immune response.	0.01016	NECTIN2
	GO:0006526	Arginine biosynthetic process.	0.01016	ASS1
	GO:0000052	Citrulline metabolic process.	0.01016	ASS1
	GO:0060312	Regulation of blood vessel remodeling	0.01016	CEACAM1
	GO:0071803	Positive regulation of podosome assembly.	0.01016	ASAP1
GO Molecular Function 2018	GO:0005384	Manganese ion transmembrane transporter activity.	0.01016	SLC11A2
	GO:0005381	Iron ion transmembrane transporter activity.	0.01016	SLC11A2
	GO:0003836	Beta-galactoside (CMP) alpha-2,3-sialyltransferase activity.	0.01184	ST3GAL2
	GO:0032052	Bile acid binding.	0.01352	VDR
	GO:0051864	Histone demethylase activity.	0.01520	KDM4A
	GO:0046965	Retinoid X receptor binding.	0.01520	VDR
	GO:0070016	Armadillo repeat domain binding.	0.01520	TCF7L2
	GO:0016706	Oxidoreductase activity.	0.001265	KDM4A,TET2
	GO:0031005	Filamin binding.	0.01687	CEACAM1
	GO:0031406	Carboxylic acid binding.	0.00152	VCAN,ASS1
GO Cellular Component 2018	GO:0005915	Zonula adherens.	0.01016	NECTIN2
	GO:1990907	Beta-catenin-TCF complex.	0.01855	TCF7L2
	GO:0042101	T cell receptor complex.	0.02687	CEACAM1
	GO:0097431	Mitotic spindle pole.	0.04167	KATNBL1
	GO:0005657	Replication fork.	0.04330	BCL6
	GO:0031526	Brush border membrane.	0.04655	SLC11A2
	GO:0005796	Golgi lumen.	0.01204	VCAN, MUC5AC
	GO:0005775	Vacuolar lumen.	0.03053	ACTR2, VCAN

Table 2
Top pathways for common upregulated genes from KEGG, BioPlanet, Reactome and MSigDB databases.

Category	Pathways	p-value	Genes
KEGG 2019 Human	Taurine and hypotaurine metabolism.	0.01855	CSAD
	Endocrine and other factor-regulated calcium reabsorption.	0.003013	VDR, GNAS
	Mineral absorption.	0.003395	VDR, SLC11A2
	Arginine biosynthesis.	0.03512	ASS1
	Adherens junction.	0.006655	TCF7L2, NECTIN2
	Mucin type O-glycan biosynthesis	0.05142	ST3GAL2
	Melanogenesis	0.01275	TCF7L2, GNAS
	Parathyroid hormone synthesis, secretion and action	0.01398	VDR,GNAS
BioPlanet 2019	Influenza factor interactions with host	0.01016	XK
	Degradation of cysteine and homocysteine	0.01352	CSAD
	Vitamin D biosynthesis	0.01352	VDR
	Termination of O-glycan biosynthesis	0.0008886	ST3GAL2, MUC5AC
	Rapid glucocorticoid receptor pathway	0.01520	GNAS
	Urea cycle	0.01520	ASS1
	Ganglio sphingolipid metabolism	0.01687	ST3GAL2
	Taurine and hypotaurine metabolism	0.01687	CSAD
	Dermatan sulfate biosynthesis	0.01687	VCAN
	Attenuation of GPCR signaling	0.0185	GNAS
Reactome 2016	Fibronectin matrix formation	0.01016	CEACAM1
	Scavenging by Class F Receptors	0.01016	SCARF1
	Binding of TCF/LEF:CTNNB1 to target gene promoters.	0.01184	TCF7L2
	Defective CHST3 causes SEDCJD	0.01184	VCAN
	Defective CHSY1 causes TPBS	0.01184	VCAN
	Defective CHST14 causes EDS, musculocontractural type	0.01184	VCAN
	Termination of O-glycan biosynthesis	0.0008886	ST3GAL2, MUC5AC
	Degradation of cysteine and homocysteine	0.01520	CSAD
	Urea cycle	0.01687	ASS1
	Dermatan sulfate biosynthesis	0.01855	VCAN
MSigDB Hallmark 2020	TNF-alpha Signaling via NF-Kb	0.04528	BCL6, NFE2L2
	Apical Junction	0.04528	VCAN, NECTIN2
	Inflammatory Response	0.04528	SLC11A2, SCARF1
	Heme Metabolism	0.04528	XK,SLC11A2

hand, a highly clustered region of PPIs networks was obtained by using MCODE. In this clustering network two hub genes, VDR and NFE2L2 were obtained (Fig. 5).

3.5. Transcriptional factor regulatory network of hub genes

For hub genes, TFs regulatory network was constructed through the NetworkAnalyst platform having 48 interactions, 32 TFs, and 4 hub genes. BCL6 was found to be regulated by 17 TFs, TET2 was regulated by 13 TFs, NFE2L2 was regulated by 12 TFs and VDR was regulated by 6 TFs. Nine TFs were detected in the TF regulatory network with a connectivity degree of $\geq$2. Fig. 6 depicts TF regulatory network.

3.6. TFs-miRNA regulatory network analysis

TFs-miRNA regulatory Network provides information about miRNAs and TFs interaction with the hub genes. TF-miRNA coregulatory network was studied by NetworkAnalyst that had 39 nodes and 61 edges. Among 39 nodes, 4 for hub genes, 9 for miRNA, and 26 for TFs. Furthermore, all miRNAs were found to have 2° connectivity except has-miR-544. Fig. 7 illustrates TFs-miRNA regulatory network.

3.7. Protein drug interaction network

Protein drug interaction networks provide essential insight for providing proper treatment to the patients. Protein drug interaction network collected from DrugBank where VDR dominantly connected to 14 different kinds of drug. This network indicates, vitamin D and some of its analogous compounds might play crucial roles in modulating both COVID-19 and sarcoidosis conditions. The drug network is presented in Fig. 8.

4. Discussion

The COVID-19 pandemic is a concern for all clinicians, who are taking care of sarcoidosis patients; many of them require immunosuppressive drugs, which can put them at risk for severe and opportunistic infections as compared to the general population [34]. Therefore, common molecular signatures should be identified so that effective treatment can be provided. This research aids in the description of Bioinformatics approaches for the analysis of peripheral blood cells response during COVID-1 infection and sarcoidosis. Two datasets, GSE164805 and GSE18781, were chosen from GEO to identify common upregulated genes between COVID-19 and sarcoidosis. After obtaining 34 common upregulated genes, the study was continued with the analysis of GO, pathways, PPIs, TF-gene interactions network, TF-miRNA coregulatory network, and drug network.

Common upregulated genes were used to know GO and their associated pathways. GO and pathways were retrieved by using Enrichr based on their p-value. For biological processes, laminar fluid shear stress, natural killer cell-mediated cytotoxicity, epithelial cell differentiation, immunoglobulin mediated immune response, and arginine biosynthetic process were top prominent. Fluid shear stress has significant effects on endothelial biology, inflammatory response, thrombotic and fibrinolytic homeostasis [35]. NK cell-mediated cytotoxicity has a great impact on the severity of COPD that also has responsibility for self-promoting emphysema [36]. Besides, manganese ion transporter activity, iron transporter activity, beta-galactoside (CMP) alpha-2, 3-sialyltransferase activity, bile acid-binding, and histone demethylase activity were the most GO terms in terms of molecular function. It is demonstrated that a high iron level in the body is associated with many infectious diseases and inflammatory responses, as exemplified by malaria, viral infection, and neurodegeneration [37]. Furthermore, top GO terms regarding the cellular component are zonula adherents, beta-catenin-TCF complex, T cell receptor complex, mitotic spindle pole, and replication fork.

Similarly, upregulated genes were also used to get insights into associated pathways. Prominent pathways from the KEGG database were included taurine and hypotaurine metabolism, endocrine and other factor-regulated calcium reabsorption, mineral absorption, arginine biosynthesis, and adherens' junction. In humans, taurine and taurine derivatives are considered possible medicines against infectious and chronic inflammatory diseases [38]. Moreover, results from BioPlanet showed influenza factor interactions with host, degradation of cysteine and homocysteine, vitamin D biosynthesis, termination of O-glycan biosynthesis, and rapid glucocorticoid receptor pathway. Furthermore, Reactome provided most interacted pathways were fibronectin matrix formation, scavenging by Class F receptors, binding of TCF/LEF: CTNNB1 to target gene promoters, defective CHST3 causes SEDCJD, and defective CHSY1 causes TPBS.

PPIs network was constructed and visualized with the aid of STRING and Cytoscape tools. Five hub genes, including MUC5AC, BCL6, TET2, NFE2L2, and VDR were detected from the PPIs network. MUC5AC, BCL6, NFE2L2, and VDR have been documented as having an

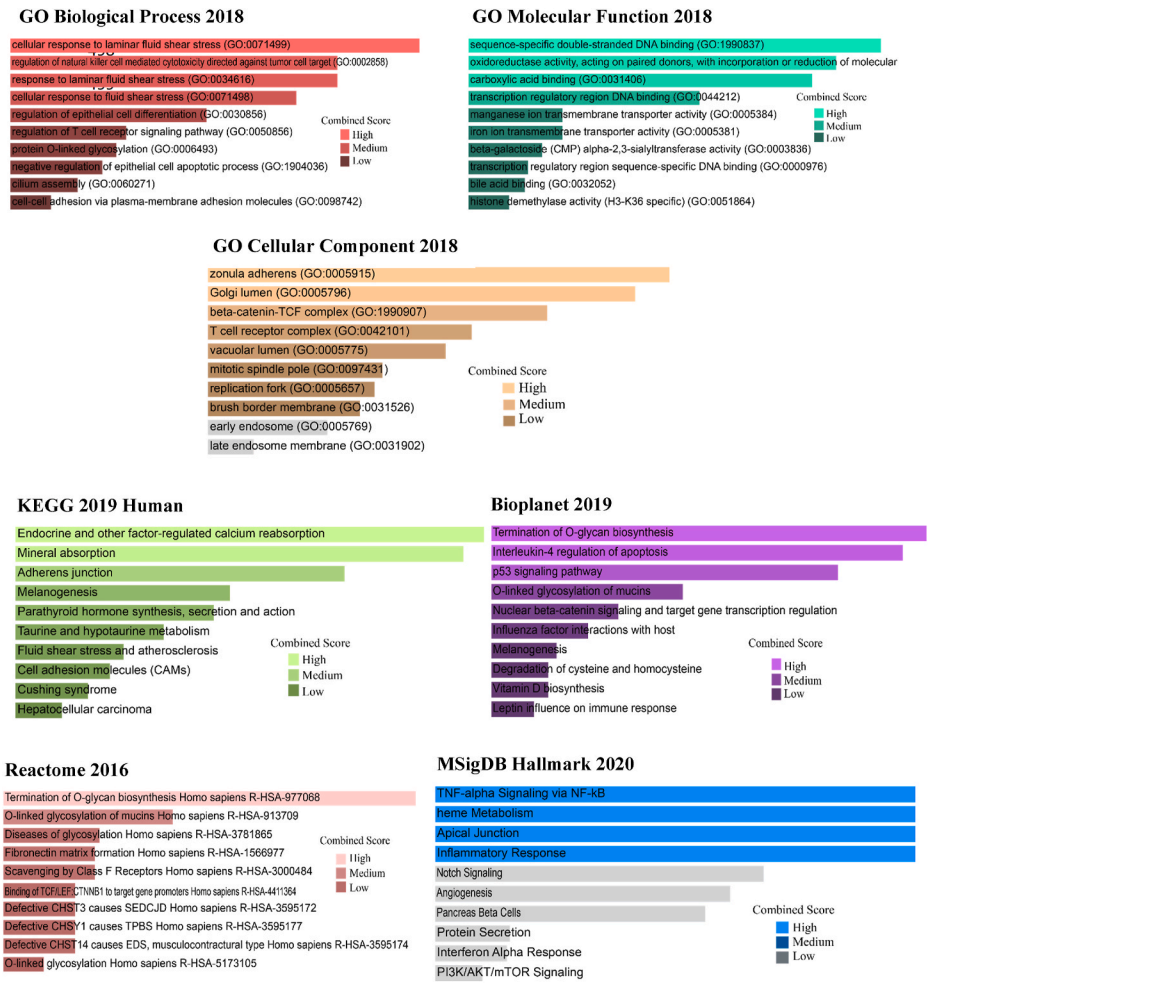

Fig. 2. Gene ontologies (Biological process, Molecular function and Cellular component) and pathways (KEGG, BioPlanet, Reactome and MSigDB) for common upregulated genes based on combined score.

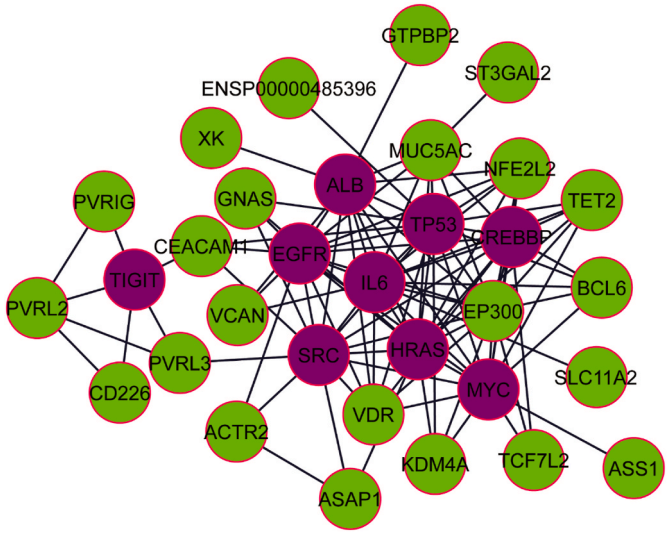

Fig. 3. Protein-protein interactions (PPIs) network for common upregulated genes from COVID-19 and sarcoidosis. The light green color nodes indicate common upregulated genes. Network consists of 32 nodes and 102 edges. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

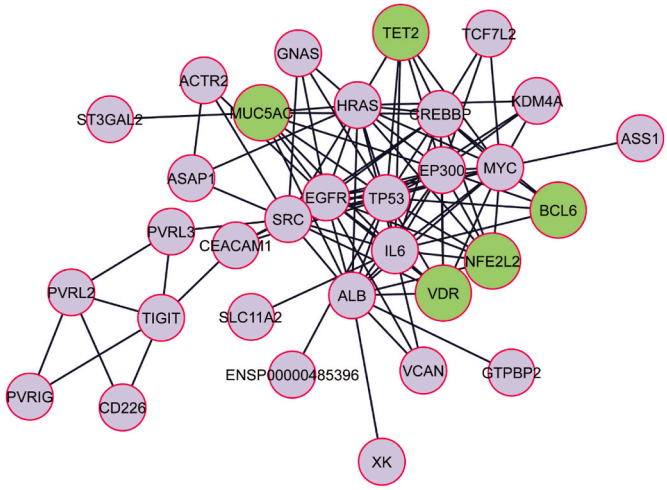

Fig. 4. Identification of hub genes from PPIs network for common upregulated genes. Five hub genes were detected based on their degree value. Light green color nodes (TET2, MUC5AC, BCL6, NFE2L2 and VDR) indicate common hub genes. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

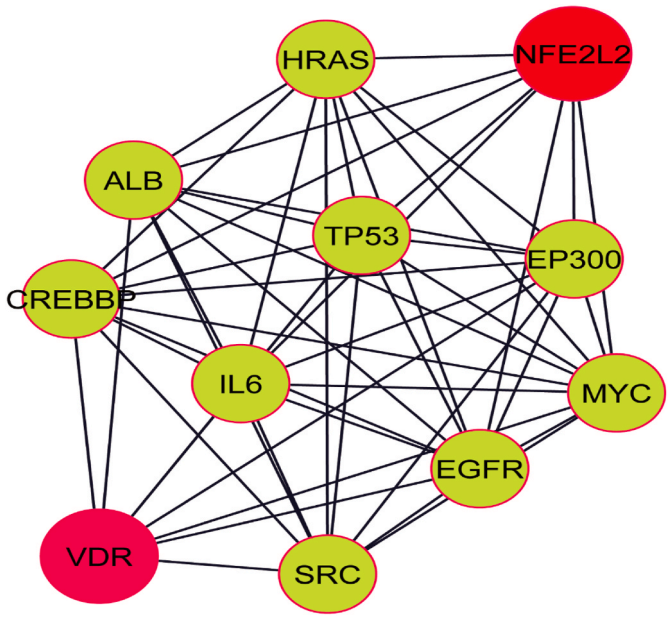

Fig. 5. Cluster analysis network extracted from PPIs network. This network represents highly interconnected network in PPIs network. From five hub genes NFE2L2 and VDR obtained in cluster network. Two hub genes highlighted in red color. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

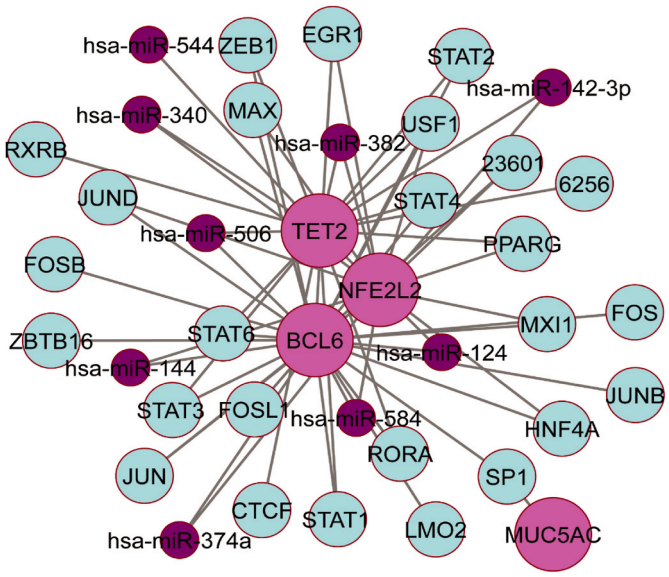

Fig. 7. Network indicates TF-miRNA coregulatory network for common hub genes. This network embodied with 39 nodes and 61 edges where 4 noes for hub genes, 9 nodes for miRNA and 26 nodes for TFs. Magenta color denotes hub genes; Maroon color denotes miRNA Cyan color denotes TFs. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

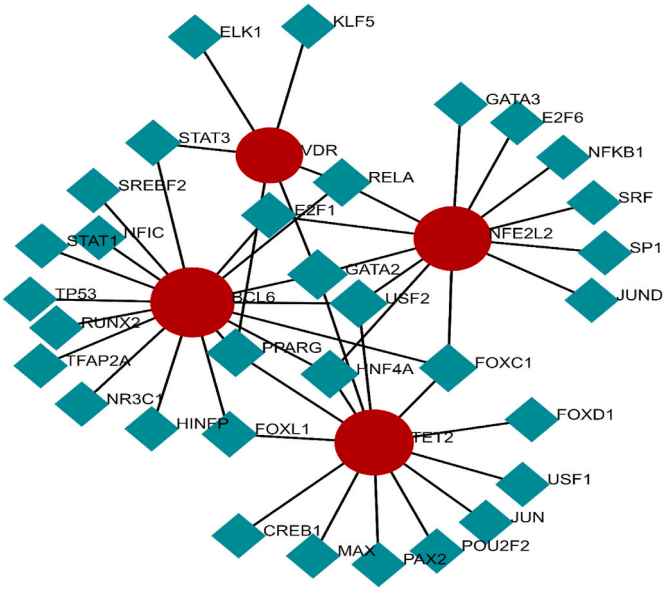

Fig. 6. TF-interaction network for common hub genes. Red color nodes denote hub gene and other nodes represent TF genes. This network contains 32 TFs and four hub genes. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

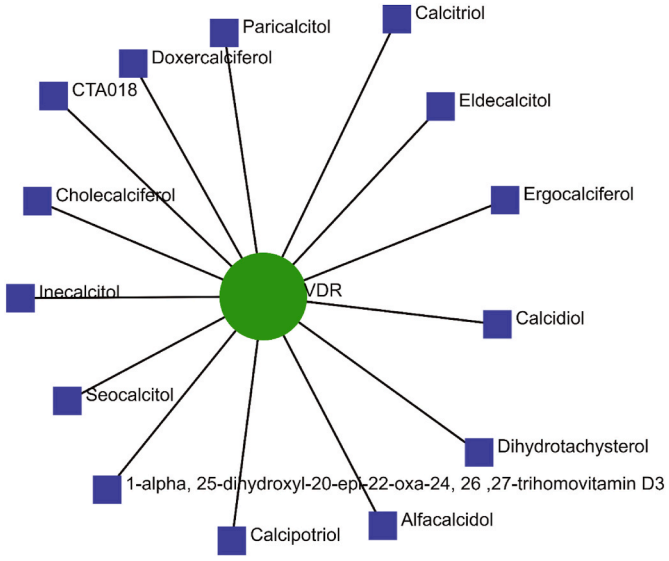

Fig. 8. Drugs protein interaction network.

association with the pathogenesis of lung diseases [39–42]. In module analysis, VDR and NFE2l2 were clustered together which was performed to suggest more effective drug compounds.

Hub genes were considered to construct a TFs gene interaction network through the NetworkAnalyst web platform. This network consisted of 36 nodes and 48 edges. Among hub genes, BCL6 had high interaction with 17° value in the TFs genes network. On the other hand, among transcription factor genes GATA2 gene was revealed as a prominent transcription factor having 4° value. GATA2 deficiency has disclosed the association with pulmonary alveolar proteinosis and fibrosis [43].

Gene expression is an intricate process regulated by TFs and miRNAs. TFs and miRNAs can be useful as biomarkers in complex diseases. Identification of miRNAs in sarcoidosis is a great strategy for better understanding the disease's pathogenesis and identifying potential biomarkers [44]. For hub genes, TFs-miRNA coregulatory network was built through NetworkAnalyst tools where the network was filtered at 2° value. As a result, a network containing 39 nodes and 61 edges was gained. Among network interaction, BCL6 had a higher degree of interaction value of 27. On the other side, all miRNA contained 2° value except hsa-miR-544 having 1° value. Finally, drug interaction network was also retrieved through NetworkAnalyst. The drug interaction network provided 14 potential interactors with the VDR gene.

From cluster and drug networks, this study hypothesized that vitamin D and its analogous compound may have a significant role in regulating COVID-19 and sarcoidosis conditions. Several studies documented that vitamin D deficiency is associated with lung diseases and predispose to respiratory infection by virus and bacteria [45–49]. A systemic review study reported, vitamin D can evaluate the risk of infection, seriousness, and mortality from COVID-19 as well as also suggest public to maintain an appropriate level of vitamin D to cope with the pandemic [50]. Likely, hypovitaminosis D tends to be linked to increased sarcoidosis disease activity, and thus can be considered a risk factor for sarcoidosis disease activity [51,52]. However, the exact mechanism of how vitamin D modulate immune and inflammation response in COVID-19 and sarcoidosis conditions is still unknown. Therefore, the more wet-lab research study is demanded to use vitamin D as a supplement in COVID-19 and sarcoidosis.

Even though rigorous bioinformatics analysis was executed, some flaws remain. First, our study is limited to the number of data, where just two datasets were selected. Second, our data is limited to peripheral blood cells. Augmenting the number and types of data might increase the accuracy of our analyzed results. Third, this study cannot explain how upregulation of five hub genes can play role in COVID-19 and sarcoidosis.

5. Conclusion

So far there has been no other study on COVID-19 and sarcoidosis in the sense of peripheral blood cells transcriptomic analysis. In this circumstance, this study concludes that five hub genes, including, BCL6, MUC5AC, TET2, NFE2L2, and VDR might have synergistic effects and could be used as a therapeutic target on COVID-19 patients with sarcoidosis history. Since COVID-19 is a recent discovery infection and sarcoidosis etiology is not fully understood, therefore, more wet laboratory-based studies are necessary to validate the findings of this study.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contributions

Md. Roman Mogal and Md. Asaduzzaman Sikder planned and designed the research. Md. Roman Mogal and Md. Rashel Mahmod conducted bioinformatics analyses. Sagarika Adhikary Sompa and Asadullah Junayed interpreted the results and performed statistical analysis. Md. Roman Mogal wrote the manuscript. Md. Asaduzzaman Sikder and Md. Zainul Abedin thoroughly edited and revised the manuscript. Md. Asaduzzaman Sikder supervised the whole research. All authors read and approved the manuscript.

Declaration of competing interest

Authors declared no competing interest among them.

Data availability

No data was used for the research described in the article.

Acknowledgment

Authors are grateful to the Department of Biochemistry and Molecular Biology, Mawlana Bhashani Science and Technology University, for providing technical support.

References

[1] WHO, WHO Timeline COVID-19, Who, 2020, 2020, https://www.who.int/news/item/27-04-2020-who-timeline—covid-19. (Accessed 1 April 2021).
[2] N. Madhav, B. Oppenheim, M. Gallivan, P. Mulembakani, E. Rubin, N. Wolfe, Pandemics: risks, impacts, and mitigation, in: Dis. Control Priorities, third ed., ume 9, Improv. Heal. Reducing Poverty, 2017, pp. 315–345, https://doi.org/10.1596/978-1-4648-0527-1_ch17.
[3] WHO coronavirus (COVID-19) dashboard | WHO coronavirus (COVID-19) dashboard with vaccination data, n.d. https://covid19.who.int/?gclid=CjwKCAjwnK36BRBVEiwAsMT8WJ3y00_BUzvrLsvbl3uthuoTH_Occ45gyEUbpYRyEqAzll3aZB6TYxoCcM0QAvD_BwE. (Accessed 13 January 2022).
[4] S. Mallapaty, Can COVID vaccines stop transmission? Scientists race to find answers, Nature (2021), https://doi.org/10.1038/d41586-021-00450-z.
[5] O.J. Wouters, K.C. Shadlen, M. Salcher-Konrad, A.J. Pollard, H.J. Larson, Y. Teerawattananon, M. Jit, Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment, Lancet 397 (2021) 1023–1034, https://doi.org/10.1016/S0140-6736(21)00306-8.
[6] E.S. Chen, D.R. Moller, Etiologies of sarcoidosis, Clin. Rev. Allergy Immunol. 49 (2015) 6–18, https://doi.org/10.1007/s12016-015-8481-z.
[7] H. Clay, J.M. Davis, D. Beery, A. Huttenlocher, S.E. Lyons, L. Ramakrishnan, Dichotomous role of the macrophage in early Mycobacterium marinum infection of the zebrafish, Cell Host Microbe 2 (2007) 29–39, https://doi.org/10.1016/j.chom.2007.06.004.
[8] P. Ungprasert, E.M. Carmona, J.P. Utz, J.H. Ryu, C.S. Crowson, E.L. Matteson, Epidemiology of sarcoidosis 1946-2013: a population-based study, Mayo Clin. Proc. 91 (2016) 183–188, https://doi.org/10.1016/j.mayocp.2015.10.024.
[9] B. Southern, Patients with interstitial lung disease and pulmonary sarcoidosis are at high risk for severe illness related to COVID-19: posted April 18, 2020, Cleve, Clin. J. Med. 87 (2020) 1–7, https://doi.org/10.3949/CCJM.87A.CCC026.
[10] R.P. Baughman, E.E. Lower, M. Buchanan, P. Rottoli, M. Drent, J. Sellares, M. Terwiel, M. Elfferich, J. Francesqui, M.R.B. Cabrerizo, N. Sweiss, F. Martone, T. Al-Hakim, M.A. Judson, Risk and outcome of COVID-19 infection in sarcoidosis patients: results of a self-reporting questionnaire, Sarcoidosis Vasc, Diffus. Lung Dis. 37 (2020), https://doi.org/10.36141/svdld.v37i4.10726.
[11] M. Manansala, C. Ascoli, A.G. Alburquerque, D. Perkins, M. Mirsaedi, P. Finn, N. J. Sweiss, Case series: COVID-19 in African American patients with sarcoidosis, Front. Med. 7 (2020), https://doi.org/10.3389/fmed.2020.588527.
[12] N.J. Sweiss, P. Korsten, H.J. Syed, A. Syed, R.P. Baughman, A.M.F. Yee, D. A. Culver, T. Sosenko, A. Azuma, F. Bonella, U. Costabel, W.P. Drake, M. Drent, E. E. Lower, D. Israel-Biet, R.L.M. Mostard, H. Nunes, P. Rottoli, P. Spagnolo, A. U. Wells, W.A. Wuyts, M.A. Judson, When the game changes: guidance to adjust sarcoidosis management during the coronavirus disease 2019 pandemic, Chest 158 (2020) 892–895, https://doi.org/10.1016/j.chest.2020.04.033.
[13] C.D. Russell, J.E. Millar, J.K. Baillie, Clinical evidence does not support corticosteroid treatment for 2019-nCoV lung injury, Lancet 395 (2020) 473–475, https://doi.org/10.1016/S0140-6736(20)30317-2.
[14] A.H. Györfi, M. Kopp, M. May, M. Vetter, M. Uder, A.E. Kremer, G. Schett, T. Harrer, J.H.W. Distler, Glucocorticoid-induced relapse of COVID-19 in a patient with sarcoidosis, Ann. Rheum. Dis. (2020), https://doi.org/10.1136/annrheumdis-2020-218258.
[15] M. Manansala, A. Chopra, R.P. Baughman, R. Novak, E.E. Lower, D.A. Culver, P. Korsten, W.P. Drake, M.A. Judson, N. Sweiss, COVID-19 and sarcoidosis, readiness for vaccination: challenges and opportunities, Front. Med. 8 (2021), https://doi.org/10.3389/fmed.2021.672028.
[16] D. Sean, P.S. Meltzer, GEOquery, A bridge between the gene expression Omnibus (GEO) and BioConductor, Bioinformatics (2007), https://doi.org/10.1093/bioinformatics/btm254.
[17] Y. Benjamini, Y. Hochberg, Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Stat. Soc. Ser. B. 1995 (1995) 289–300. J. R. Stat. Soc. B.
[18] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 15545–15550, https://doi.org/10.1073/pnas.0506580102.
[19] E.Y. Chen, C.M. Tan, Y. Kou, Q. Duan, Z. Wang, G.V. Meirelles, N.R. Clark, A. Ma'ayan, Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool, BMC Bioinf. 14 (2013), https://doi.org/10.1186/1471-2105-14-128.
[20] M.V. Kuleshov, M.R. Jones, A.D. Rouillard, N.F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S.L. Jenkins, K.M. Jagodnik, A. Lachmann, M.G. McDermott, C. D. Monteiro, G.W. Gundersen, A. Ma'ayan, Enrichr: a comprehensive gene set enrichment analysis web server 2016 update, Nucleic Acids Res. 44 (2016) W90–W97, https://doi.org/10.1093/nar/gkw377.
[21] A. Chaiboonchoe, W. Jurkowski, J. Pellet, E. Glaab, A. Kolodkin, A. Rausell, A. Le Béchec, S. Ballereau, L. Meyniel, I. Crespo, H. Ahmed, V. Volpert, V. Lotteau, N. Baliga, L. Hood, A. Del Sol, R. Balling, C. Auffray, On different aspects of network analysis in systems biology, in: Syst. Biol. Integr. Biol. Simul. Tools, Springer, 2013, pp. 181–207, https://doi.org/10.1007/978-94-007-6803-1_6.
[22] A. Chaiboonchoe, S. Samarasinghe, D. Kulasiri, K. Salehi-Ashtiani, Integrated analysis of gene network in childhood leukemia from microarray and pathway databases, BioMed Res. Int. 2014 (2014), https://doi.org/10.1155/2014/278748.
[23] D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K.P. Tsafou, M. Kuhn, P. Bork, L.J. Jensen, C. Von Mering, STRING v10: protein-protein interaction networks, integrated over

the tree of life, Nucleic Acids Res. 43 (2015) D447–D452, https://doi.org/10.1093/nar/gku1003.
[24] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software Environment for integrated models of biomolecular interaction networks, Genome Res. 13 (2003) 2498–2504, https://doi.org/10.1101/gr.1239303.
[25] C.H. Chin, S.H. Chen, H.H. Wu, C.W. Ho, M.T. Ko, C.Y. Lin, cytoHubba: identifying hub objects and sub-networks from complex interactome, BMC Syst. Biol. 8 (2014), https://doi.org/10.1186/1752-0509-8-S4-S11.
[26] G.D. Bader, C.W.V. Hogue, An automated method for finding molecular complexes in large protein interaction networks, BMC Bioinf. 4 (2003), https://doi.org/10.1186/1471-2105-4-2.
[27] A.C. Wilkinson, H. Nakauchi, B. Göttgens, Mammalian transcription factor networks: recent advances in interrogating biological complexity, Cell Syst 5 (2017) 319–331, https://doi.org/10.1016/j.cels.2017.07.004.
[28] J. Xia, M.J. Benner, R.E.W. Hancock, NetworkAnalyst - integrative approaches for protein-protein interaction network analysis and visual exploration, Nucleic Acids Res. 42 (2014), https://doi.org/10.1093/nar/gku443.
[29] G. Zhou, O. Soufan, J. Ewald, R.E.W. Hancock, N. Basu, J. Xia, NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis, Nucleic Acids Res. 47 (2019) W234–W241, https://doi.org/10.1093/nar/gkz240.
[30] K.S. Zaret, J.S. Carroll, Pioneer transcription factors: establishing competence for gene expression, Genes Dev. 25 (2011) 2227–2241, https://doi.org/10.1101/gad.176826.111.
[31] I. Slezak-Prochazka, D. Selvi, B.J. Kroesen, A. Van Den Berg, MicroRNAs, macrocontrol: regulation of miRNA processing, Rna 16 (2010) 1087–1095, https://doi.org/10.1261/rna.1804410.
[32] Z.P. Liu, C. Wu, H. Miao, H. Wu, RegNetwork: an integrated database of transcriptional and post-transcriptional regulatory networks in human and mouse, Database 2015 (2015) 1–12, https://doi.org/10.1093/database/bav095.
[33] D.S. Wishart, C. Knox, A.C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang, J. Woolsey, DrugBank: a comprehensive resource for in silico drug discovery and exploration, Nucleic Acids Res. 34 (2006), https://doi.org/10.1093/nar/gkj067.
[34] A. Duréault, C. Chapelon, L. Biard, F. Domont, L. Savey, B. Bodaghi, V. Pourcher, M.R. Rigon, P. Cacoub, D. Saadoun, Severe infections in sarcoidosis: incidence, predictors and long-term outcome in a cohort of 585 patients, Med. (United States) (2017) 96, https://doi.org/10.1097/MD.0000000000008846.
[35] B. Kim, Z. Arany, Could shear stress mimetics delay complications in COVID-19? Trends Cardiovasc. Med. (2021) https://doi.org/10.1016/j.tcm.2021.01.004.
[36] B. Hervier, J. Russick, I. Cremer, V. Vieillard, NK cells in the human lungs, Front. Immunol. 10 (2019), https://doi.org/10.3389/fimmu.2019.01263.
[37] J. Kim, The role of iron metabolism in lung inflammation and injury, J. Allergy Ther. 1 (2012), https://doi.org/10.4172/2155-6121.s4-004.
[38] J. Marcinkiewicz, E. Kontny, Taurine and inflammatory diseases, Amino Acids 46 (2014) 7–20, https://doi.org/10.1007/s00726-012-1361-4.
[39] J. Li, Z. Ye, The potential role and regulatory mechanisms of MUC5AC in chronic obstructive pulmonary disease, Molecules 25 (2020), https://doi.org/10.3390/molecules25194437.
[40] B. Zhu, R. Zhang, C. Li, L. Jiang, M. Xiang, Z. Ye, H. Kita, A.M. Melnick, A.L. Dent, J. Sun, BCL6 modulates tissue neutrophil survival and exacerbates pulmonary inflammation following influenza virus infection, Proc. Natl. Acad. Sci. U. S. A. 116 (2019) 11888–11893, https://doi.org/10.1073/pnas.1902310116.
[41] K. Mizumura, S. Maruoka, T. Shimizu, Y. Gon, Role of Nrf2 in the pathogenesis of respiratory diseases, Respir. Investig. 58 (2020) 28–35, https://doi.org/10.1016/j.resinv.2019.10.003.
[42] I.K. Sundar, I. Rahman, Vitamin D and susceptibility of chronic lung diseases: role of epigenetics, Front. Pharmacol. AUG (2011), https://doi.org/10.3389/fphar.2011.00050.
[43] A. Ballerie, S. Nimubona, C. Meunier, F.L. Gutierrez, B. Desrues, P. Delaval, S. Jouneau, Association of pulmonary alveolar proteinosis and fibrosis: patient with GATA2 deficiency, Eur. Respir. J. 48 (2016) 1510–1514, https://doi.org/10.1183/13993003.00252-2016.
[44] B. Pattnaik, P.B. Sryma, S. Mittal, A. Agrawal, R. Guleria, K. Madan, MicroRNAs in pulmonary sarcoidosis: a systematic review, Respir. Investig. 58 (2020) 232–238, https://doi.org/10.1016/j.resinv.2020.02.008.
[45] C.R. Gilbert, S.M. Arum, C.M. Smith, Vitamin D deficiency and chronic lung disease, Cancer Res. J. 16 (2009) 75–80, https://doi.org/10.1155/2009/829130.
[46] P.E. Pfeffer, C.M. Hawrylowicz, Vitamin D and lung disease, Thorax 67 (2012) 1018–1020, https://doi.org/10.1136/thoraxjnl-2012-202139.
[47] A.J. Ghosh, M. Moll, L.P. Hayden, J. Bon, E. Regan, C.P. Hersh, J.D. Crapo, E. K. Silverman, B.J. Make, E.A. Regan, T. Beaty, F. Begum, P.J. Castaldi, M. Cho, D. L. Demeo, A.R. Boueiz, M.G. Foreman, E. Halper-Stromberg, L.P. Hayden, C. P. Hersh, J. Hetmanski, B.D. Hobbs, J.E. Hokanson, N. Laird, C. Lange, S.M. Lutz, M.L. McDonald, M.M. Parker, D. Qiao, E.A. Regan, E.K. Silverman, E.S. Wan, S. Won, P. Sakornsakolpat, D. Prokopenko, M. Al Qaisi, H.O. Coxson, T. Gray, M.L. K. Han, E.A. Hoffman, S. Humphries, F.L. Jacobson, P.F. Judy, E.A. Kazerooni, A. Kluiber, D.A. Lynch, J.D. Newell, E.A. Regan, J.C. Ross, R.S.J. Estepar, J. Schroeder, J. Sieren, D. Stinson, B.C. Stoel, J. Tschirren, E. Van Beek, B. Van Ginneken, E. Van Rikxoort, G. Washko, C.G. Wilson, R. Jensen, D. Everett, J. Crooks, C. Moore, M. Strand, C.G. Wilson, J.E. Hokanson, J. Hughes, G. Kinney, S.M. Lutz, K. Pratte, K.A. Young, S. Bhatt, J. Bon, M.L.K. Han, B. Make, C. Martinez, S. Murray, E. Regan, X. Soler, C.G. Wilson, R.P. Bowler, K. Kechris, F. Banaei-Kashani, J.L. Curtis, C.H. Martinez, P.G. Pernicano, N. Hanania, P. Alapat, M. Atik, V. Bandi, A. Boriek, K. Guntupalli, E. Guy, A. Nachiappan, A. Parulekar, D. L. Demeo, C. Hersh, F.L. Jacobson, G. Washko, R.G. Barr, J. Austin, B. D'Souza, G. D.N. Pearson, A. Rozenshtein, B. Thomashow, N. MacIntyre, H.P. McAdams, L. Washington, C. McEvoy, J. Tashjian, R. Wise, R. Brown, N.N. Hansel, K. Horton, A. Lambert, N. Putcha, R. Casaburi, A. Adami, M. Budoff, H. Fischer, J. Porszasz, H. Rossiter, W. Stringer, A. Sharafkhaneh, C. Lan, C. Wendt, B. Bell, M.G. Foreman, E. Berkowitz, G. Westney, R. Bowler, D.A. Lynch, R. Rosiello, D. Pace, G. Criner, D. Ciccolella, F. Cordova, C. Dass, G. D'Alonzo, P. Desai, M. Jacobs, S. Kelsen, V. Kim, A.J. Mamary, N. Marchetti, A. Satti, K. Shenoy, R.M. Steiner, A. Swift, I. Swift, M.E. Vega-Sanchez, Vitamin D deficiency is associated with respiratory symptoms and airway wall thickening in smokers with and without COPD: a prospective cohort study, BMC Pulm. Med. 20 (2020), https://doi.org/10.1186/s12890-020-1148-4.
[48] S. Hansdottir, M.M. Monick, Vitamin D Effects on Lung Immunity and Respiratory Diseases, 2011, https://doi.org/10.1016/B978-0-12-386960-9.00009-5.
[49] J.D. Finklea, R.E. Grossmann, V. Tangpricha, Vitamin D and chronic lung disease: a review of molecular mechanisms and clinical studies, Adv. Nutr. 2 (2011) 244–253, https://doi.org/10.3945/an.111.000398.
[50] H. Yisak, A. Ewunetei, B. Kefale, M. Mamuye, F. Teshome, B. Ambaw, G. Y. Yitbarek, Effects of vitamin d on covid-19 infection and prognosis: a systematic review, Risk Manag. Healthc. Pol. 14 (2021) 31–38, https://doi.org/10.2147/RMHP.S291584.
[51] L.S. Kamphuis, F. Bonte-Mineur, J.A. Van Laar, P.M. Van Hagen, P.L. Van Daele, Calcium and vitamin D in sarcoidosis:is supplementation safe? J. Bone Miner. Res. 29 (2014) 2498–2503, https://doi.org/10.1002/jbmr.2262.
[52] S. Filipovic, V. Violeta, V. Jelica, S. Mihailo, J. Aleksandar, Vitamin D deficiency and activity of sarcoidosis. https://doi.org/10.1183/13993003.congress-2016.pa827, 2016. PA827.