ELSEVIER

GENE REPORTS

Check for updates

Using bioinformatics and systems biology to discover common pathogenetic processes between sarcoidosis and COVID-19

Lanyi Fu [1], Maolin Yao [1], Xuedong Liu [*], Dong Zheng [*]

Laboratory of Genetics and Molecular Biology, College of Wildlife and Protected Area, Northeast Forestry University, Harbin 150040, China

ARTICLE INFO

ABSTRACT

The coronavirus disease (COVID-19) pandemic caused by SARS-CoV-2 is ongoing. Individuals with sarcoidosis tend to develop severe COVID-19; however, the underlying pathological mechanisms remain elusive. To determine common transcriptional signatures and pathways between sarcoidosis and COVID-19, we investigated the whole-genome transcriptome of peripheral blood mononuclear cells (PBMCs) from patients with COVID-19 and sarcoidosis and conducted bioinformatic analysis, including gene ontology and pathway enrichment, protein-protein interaction (PPI) network, and gene regulatory network (GRN) construction. We identified 33 abnormally expressed genes that were common between COVID-19 and sarcoidosis. Functional enrichment analysis showed that these differentially expressed genes were associated with cytokine production involved in the immune response and T cell cytokine production. We identified several hub genes from the PPI network encoded by the common genes. These hub genes have high diagnostic potential for COVID-19 and sarcoidosis and can be potential biomarkers. Moreover, GRN analysis identified important microRNAs and transcription factors that regulate the common genes. This study provides a novel characterization of the transcriptional signatures and biological processes commonly dysregulated in sarcoidosis and COVID-19 and identified several critical regulators and biomarkers. This study highlights a potential pathological association between COVID-19 and sarcoidosis, establishing a theoretical basis for future clinical trials.

1. Introduction

On March 11, 2020, the World Health Organization declared coronavirus disease (COVID-19) as a pandemic, which is caused by SARS-CoV-2, a positive-stranded RNA virus (Gorbalenya et al., 2020). Recent reports have demonstrated that individuals with pre-existing health conditions (e.g., cancer, heart disease, diabetes) have a higher risk of SARS-CoV-2 infection compared with the general population (Al-Quteimat and Amer, 2020; Peric and Stulnig, 2020; Radke et al., 2020). Researchers worldwide are striving to unravel therapeutic strategies that are effective for this population (A. Nashiry et al., 2021; M.A. Nashiry et al., 2021; Satu et al., 2021; Taz et al., 2021). Sarcoidosis is a risk factor for COVID-19 (Manansala et al., 2021). However, the pathological mechanisms for this increased risk have not been identified. Identifying the mechanisms might help to develop therapies for the

treatment of COVID-19 in patients with sarcoidosis.

Sarcoidosis is a systemic granulomatous disease that generally causes severe pulmonary dysfunction (Bargagli and Prasse, 2018). Chest discomfort, dyspnea, and dry cough are the most common clinical symptoms (Al-Kailany et al., 2021). According to reports from the United States, sarcoidosis has a mortality rate of 7.6% (Li et al., 2021). Individuals with sarcoidosis are at high-risk of COVID-19, which is difficult to treat owing to the lack of effective drugs and therapeutic regimens (Kondle et al., 2021).

Understanding the pathological mechanisms interlinking sarcoidosis and COVID-19 would provide a clear pathophysiological basis for treatment. Toward this end, in this study, we collected data of qualified samples from the Gene Expression Omnibus (GEO) database, which were combined with transcriptome data to explore unique transcriptional signatures of sarcoidosis and COVID-19, and further investigated
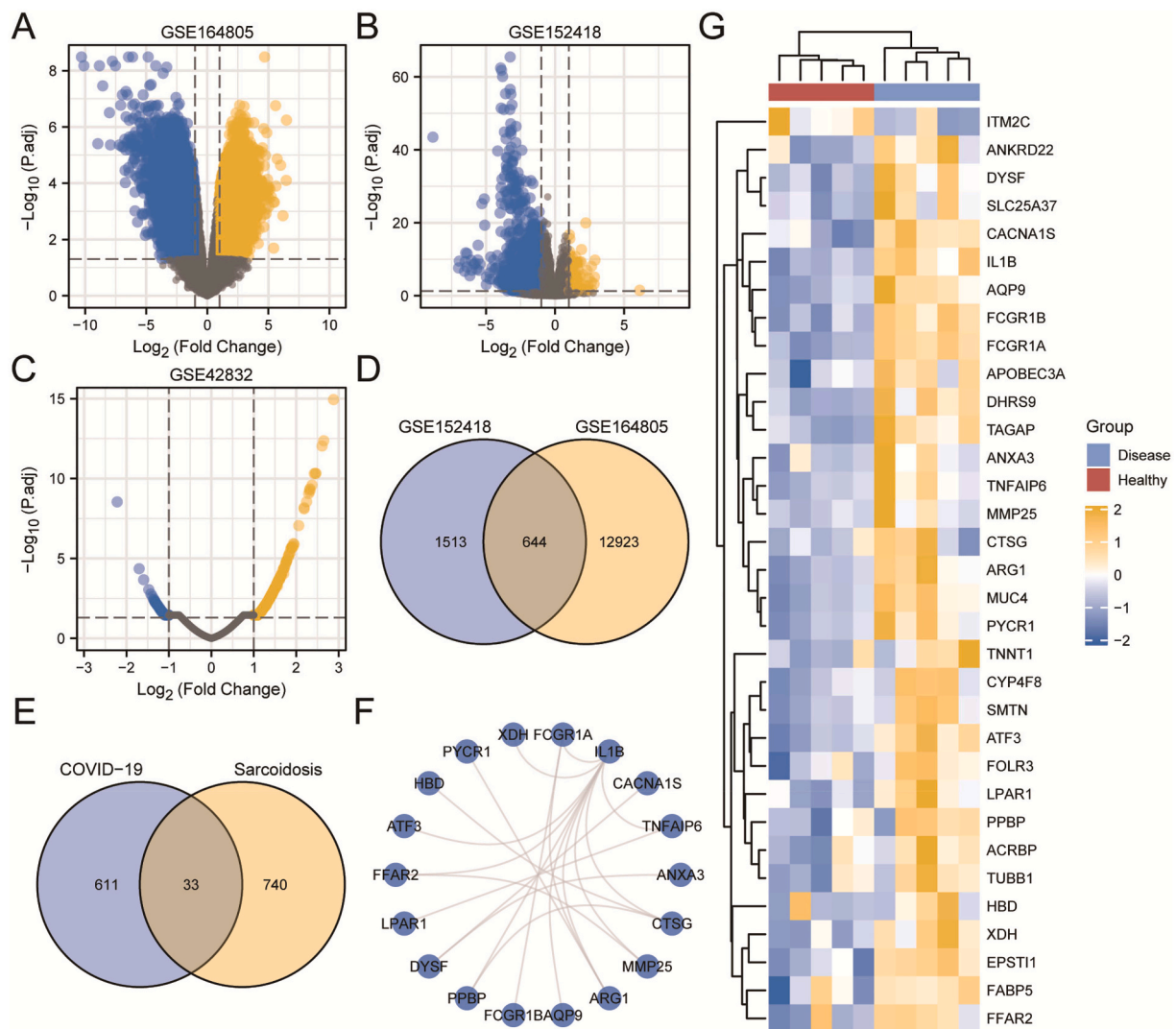
Fig. 1. Identification of common differentially expressed genes between COVID-19 and sarcoidosis. DEGs identified in COVID-19 and sarcoidosis from GSE164805 (A), GSE152418 (B) and GSE42832 (C) are presented as volcano plots. The blue dots represent down-regulated genes, and yellow plots represent up-regulated genes. (D) Venn diagram showing the 644 common DEGs between GSE152418 and GSE164805 datasets. (E) Venn diagram showing the 33 DEGs shared between COVID-19 and sarcoidosis. (F) Protein-protein interaction network constructed using common DEGs shared between COVID-19 and sarcoidosis. (G) Heatmap showing the expression landscape of shared DEGs in sarcoidosis dataset. Abbreviation: DEG, differentially expressed genes.

the common pathological pathways that underlie both diseases. These screened transcriptional signatures and enriched pathways may provide key targets for the treatment of COVID-19 in patients with sarcoidosis.

2. Materials and methods

2.1. Datasets

We acquired three gene expression datasets (GSE164805, GSE152418, and GSE42832) from GEO (https://www.ncbi.nlm.nih.gov/geo/; Barrett et al., 2013). The GSE164805 and GSE152418 datasets provide the whole-genome transcriptome of peripheral blood mononuclear cells (PBMCs) from healthy controls and patients with COVID-19. Ten samples, five samples each from healthy controls and from patients with severe COVID-19, were extracted from the GSE164805 dataset to identify differentially expressed genes (DEGs; Zhang et al., 2021). The GSE152418 dataset contains 34 PBMC samples, with 17 samples each from healthy controls and 17 samples from patients with COVID-19. Among them, two patients were sampled twice in succession; thus, these repeated-measurement samples were excluded

before analysis. The RNA sequences of the GSE152418 dataset were obtained using the Illumina NextSeq 500 platform (Blanco-Melo et al., 2020). The GSE42832 dataset contains the transcriptional profiles of PBMCs from healthy controls and patients with sarcoidosis. Ten samples, including five from healthy controls and five from patients with sarcoidosis, were extracted from this dataset to identify DEGs (Bloom et al., 2013).

2.2. Identification of differentially expressed genes

The R package "limma" was used (with default parameters) to identify DEGs associated with COVID-19 and sarcoidosis from the GSE164805 and GSE42832 datasets, respectively. The limma package uses linear models and moderated t-statistics to analyze microarray data. It provides normalization functions, supporting features that are especially useful for a linear modeling approach (Smyth, 2005). We screened the DEGs of controls and COVID-19 patients in the GSE152418 dataset using the R package "DESeq2," which uses shrinkage estimation for dispersions and fold changes to improve estimate stability and interpretability (Love et al., 2014). A log2-fold change (absolute) > 1.0

Table 1

Statistical measurement of common genes between sarcoidosis and COVID-19 datasets.

Gene	logFC	*P*.value	Adj.*P*.val	Gene	logFC	*P*.value	Adj.*P*.val
FCGR1A	2.878674	3.29E−20	1.14E−15	PPBP	1.165549	0.00018746	0.02113713
FCGR1B	2.647999	2.49E−17	4.32E−13	TUBB1	1.160117	0.00019872	0.02153726
DHRS9	2.608846	7.91E−17	9.14E−13	MUC4	1.156990	0.00019919	0.02153726
AQP9	2.184951	2.67E−12	6.63E−09	FABP5	1.145486	0.00023993	0.02448231
IL1B	1.897941	1.14E−09	1.98E−06	CTSG	1.112332	0.00037636	0.03065154
DYSF	1.713539	3.91E−08	3.77E−05	CACNA1S	1.105433	0.000379	0.03072175
TAGAP	1.628856	1.70E−07	0.000131	ANXA3	1.086968	0.00048986	0.03517862
ANKRD22	1.613975	2.32E−07	0.000167	PYCR1	1.077254	0.00053426	0.03517862
APOBEC3A	1.591804	3.42E−07	0.000215	FFAR2	1.057344	0.00069084	0.03517862
FOLR3	1.551470	6.83E−07	0.000382	LPAR1	1.047851	0.00075782	0.03517862
MMP25	1.352393	1.54E−05	0.004017	ATF3	1.041001	0.00081275	0.03517862
TNFAIP6	1.296651	3.17E−05	0.006799	CYP4F8	1.035995	0.00087245	0.03517862
SLC25A37	1.254662	5.63E−05	0.009862	XDH	1.033177	0.00089368	0.03517862
ARG1	1.228584	7.86E−05	0.012342	TNNT1	1.015741	0.00112162	0.03517862
EPSTI1	1.225689	8.12E−05	0.012472	SMTN	1.014436	0.00111006	0.03517862
HBD	1.176368	0.000185	0.021099	ITM2C	−1.00671	0.00122408	0.03517862
ACRBP	1.168952	0.000177	0.020864				

The first and fifth columns represent the gene symbol of the common genes between sarcoidosis and COVID-19 datasets; the second and sixth columns represent the log fold change between the groups; the third and seventh columns represent the t-statistics *P*-value; the fourth and eighth columns represent the adjusted *P*-values (*P*. adjust), which means *P*-value has been adjusted using Benjamini and Hochberg.
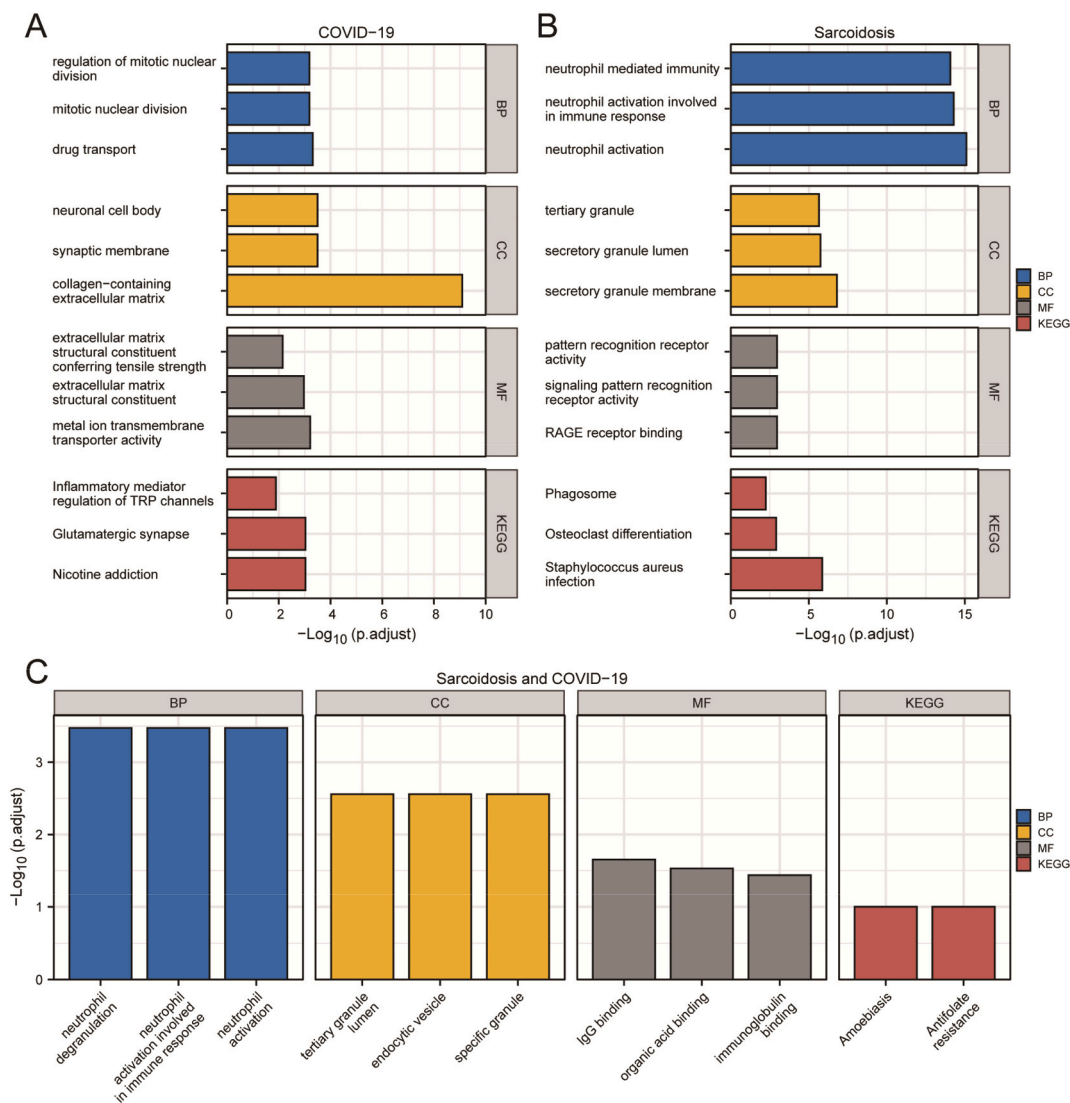
Fig. 2. Functional insights into differentially expressed genes. We performed GO and KEGG analysis to uncover the molecular function of identified DEGs, including COVID-19-specific DEGs (A), sarcoidosis-specific DEGs (B) and the common DEGs between sarcoidosis and COVID-19 (C). Abbreviations: BP, biological process; CC, cellular component; DEGs, differentially expressed genes; GO, gene ontology; MF, molecular functions; KEGG, Kyoto encyclopedia of genes and genomes.

Table 2
GO items and KEGG pathways enriched by common genes between COIVD-19 and sarcoidosis datasets.

Ontology	ID	Description	Count	Pvalue	P.adjust
BP	GO:0043312	Neutrophil degranulation	8	9.13E−07	0.000336
BP	GO:0002283	Neutrophil activation involved in immune response	8	9.56E−07	0.000336
BP	GO:0042119	Neutrophil activation	8	1.11E−06	0.000336
BP	GO:0002446	Neutrophil mediated immunity	8	1.13E−06	0.000336
BP	GO:0006692	Prostanoid metabolic process	3	4.89E−05	0.009682
BP	GO:0006693	Prostaglandin metabolic process	3	4.89E−05	0.009683
BP	GO:0008643	Carbohydrate transport	4	0.000101	0.01714
BP	GO:0070942	Neutrophil mediated cytotoxicity	2	0.000119	0.017465
BP	GO:0002367	Cytokine production involved in immune response	3	0.000637	0.029095
BP	GO:0002369	T cell cytokine production	2	0.002001	0.048849
CC	GO:1904724	Tertiary granule lumen	3	8.70E−05	0.002774
CC	GO:0030139	Endocytic vesicle	5	0.000102	0.002774
CC	GO:0042581	Specific granule	4	0.000111	0.002774
CC	GO:0034774	Secretory granule lumen	5	0.000133	0.002774
CC	GO:0060205	Cytoplasmic vesicle lumen	5	0.000169	0.002774
CC	GO:0031983	Vesicle lumen	5	0.000172	0.002774
CC	GO:0035578	Azurophil granule lumen	3	0.000389	0.005394
CC	GO:0005766	Primary lysosome	3	0.001822	0.018216
MF	GO:0019864	IgG binding	2	0.000162	0.022157
MF	GO:0043177	Organic acid binding	4	0.000431	0.029493
MF	GO:0019865	Immunoglobulin binding	2	0.000801	0.036541
MF	GO:0015144	Carbohydrate transmembrane transporter activity	2	0.001904	0.065201
KEGG	hsa05146	Amoebiasis	3	0.001391	0.099746
KEGG	hsa01523	Antifolate resistance	2	0.002099	0.099746

The first column represents the category for each item. Gene ontology (GO) contains three aspects: cellular component (CC), molecular functions (MF), and biological process (BP); the second column represents the serial number of each item; the third column shows the full name of each item; the fourth column represents the number of genes enriched in each item; the fourth column represents the *P*-values; the sixth column represents the adjusted *P*-values (*P*.adjust), which means *P*-value has been adjusted using Benjamini and Hochberg.

and adjusted *P*-value < 0.05 were used as the cut-off criteria to identify DEGs in all three datasets. The Benjamini-Hochberg procedure was used to control the false discovery rate. The common DEGs among the GSE164805, GSE152418, and GSE42832 datasets were identified and visualized using the R packages.

2.3. Functional analysis of differentially expressed genes

We performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis to investigate the biological functions of the identified DEGs. GO analysis contains three aspects: cellular component (CC), molecular functions (MF), and biological process (BP) (Doms and Schroeder, 2005). First, we used the R package "org.Hs.eg.db" to convert the IDs of overlapping DEGs. Subsequently, we used the R packages "clusterprofiler" (Yu et al., 2012) and "GOplot" (Walter et al., 2015) for the GO and KEGG analysis. The species was limited to *Homo sapiens*. GO terms and pathways were declared to be significantly enriched when the adjusted *P*-value was <0.05.

2.4. Protein-protein interaction network analysis

Protein-protein interaction (PPI) networks help to untangle complex molecular mechanisms. The PPI networks of sarcoidosis-specific genes, COVID-19-specific genes, and common genes between sarcoidosis and COVID-19 were respectively constructed using the Search Tool for Retrieval of Interacting Genes/Proteins (STRING: https://string-db.org/) database (Szklarczyk et al., 2019). All DEGs were input into the STRING database as seed genes, and their neighboring interacting partners were also included in the network. The generated PPI networks were exported to Cytoscape software (Shannon et al., 2003), and the hub genes in three networks were identified by the cytoHubba plugin (Chin et al., 2014). We selected genes in the top 10 based on maximal clique centrality as hub genes. The diagnostic potential of hub genes in COVID-19 or sarcoidosis was tested using receiver operating characteristic (ROC) curve analysis. An area under the ROC curve (AUC) of >0.5 was considered statistically significant.

2.5. Gene regulatory network analysis

Gene regulatory networks (GRNs) describe regulatory relationships between transcription factors (TFs) and microRNAs (miRNAs) and their target genes (Wang et al., 2018). We built DEG-TF and DEG-miRNA interaction networks using the JASPAR (Khan et al., 2018) and Tarbase (Sethupathy et al., 2006) databases in the Network Analyst website (Zhou et al., 2019) to explore the effects of TFs and miRNAs on DEGs. The list of the generated DEG-TF and DEG-miRNA pairs was reintroduced into Cytoscape software for further visualization. Significant hub miRNAs and TFs were detected using the cytoHubba plugin in Cytoscape and ranked by the Matthews correlation coefficient (MCC) value.

3. Results

3.1. Transcriptomic signature

The DEGs identified in COVID-19 and sarcoidosis from the three GEO datasets are presented as volcano plots in Fig. 1A–C, with 13,567, 2157 and 773 DEGs were identified in the GSE164805, GSE152418, and GSE42832 datasets, respectively, of which 33 DEGs were common to the three datasets, indicating a potentially common molecular mechanism contributing to COVID-19 and sarcoidosis pathology (Fig. 1D–E, Table 1). The interactions among the 33 common DEGs based on the STRING database are shown in Fig. 1F. The heatmap in Fig. 1G shows that most of the DEGs were upregulated in the disease group compared with the corresponding expression in the normal controls.

3.2. Functional enrichment of differentially expressed genes

The COVID-19-specific DEGs were most strongly associated with the GO terms and KEGG pathways "regulation of mitotic nuclear division," "neuronal cell body," "extracellular matrix structural constituent conferring tensile strength," and "inflammatory mediator regulation of TRP channels" (Fig. 2A), whereas the sarcoidosis-specific DEGs were most strongly associated with "neutrophil mediated immunity," "tertiary granule," "pattern recognition receptor activity," and "phagosome"
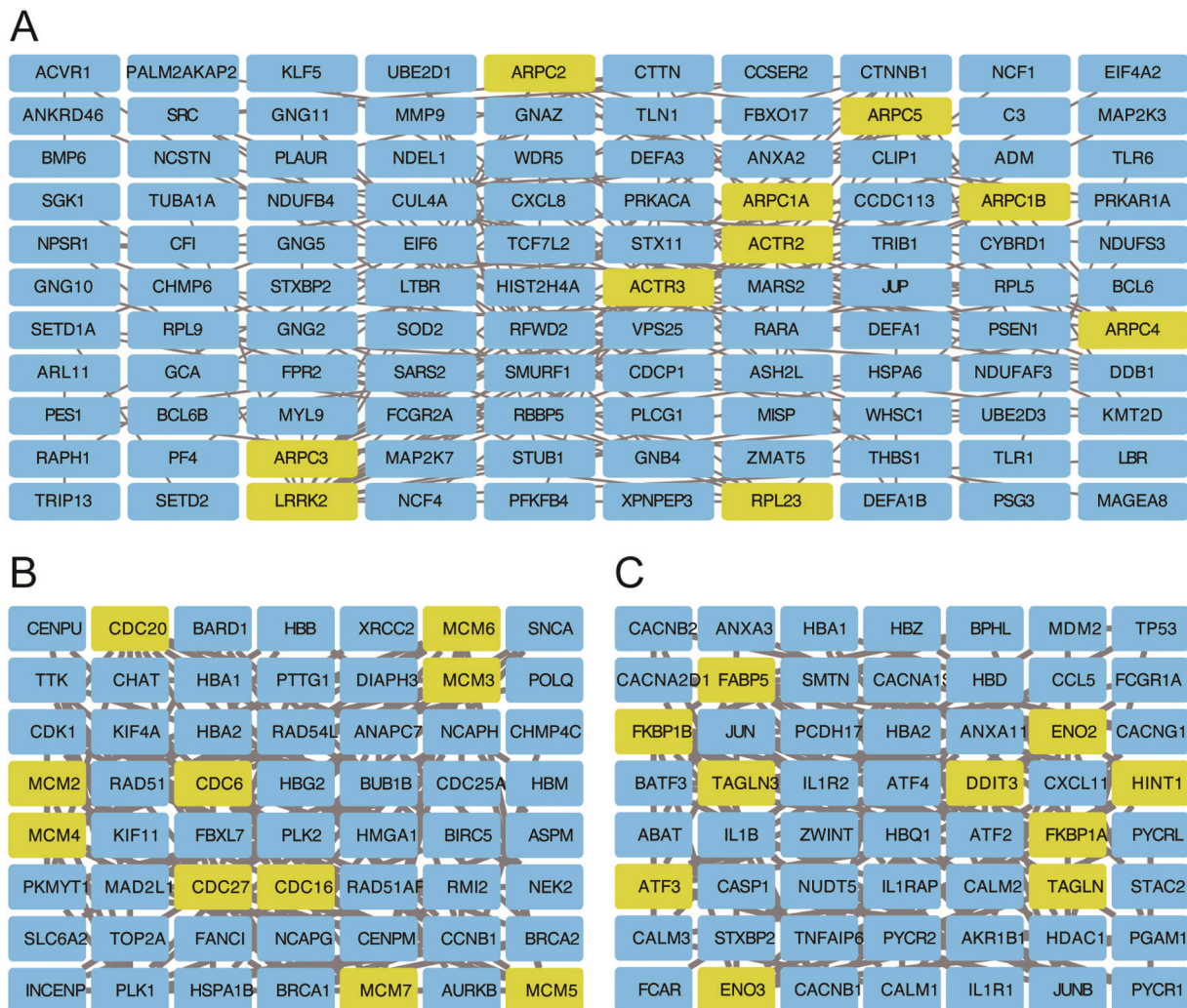
Fig. 3. Identification of proteomic signatures. Protein-protein interaction networks was constructed using sarcoidosis-specific DEGs (A), COVID-19-specific DEGs (B), common DEGs (C) and their neighboring interacting partners. The yellow nodes represent the identified hub genes. Abbreviation: DEGs, differentially expressed genes.

(Fig. 2B). Most of the 33 shared between sarcoidosis and COVID-19, were associated with the BP term "neutrophil degranulation", the CC term "tertiary granule lumen," the MF term "IgG binding", and the "amoebiasis" and "antifolate resistance" KEGG pathways (Fig. 2C). Table 2 summarizes the top six GO items and KEGG pathways associated with the common dysregulated genes in the two diseases.

3.3. Proteomic signature

The PPI network of sarcoidosis-specific DEGs contained 110 nodes and 146 edges (Fig. 3A); the PPI network of COVID-19-specific DEGs contained 56 nodes and 115 edges (Fig. 3B); and the PPI network of the 33 common DEGs contained 56 nodes and 109 edges (Fig. 3C). The MF GO terms of the 10 hub genes (*FKBP1B, ATF3, FABP5, TAGLN3, ENO3, DDIT3, ENO2, FKBP1A, TAGLN,* and *HINT1*) are listed in Table 3. The ROC curves for genes with significant AUC values for predicting disease are shown in Fig. 4. The ROC curves identified *MCM3, MCM5, CDC27, CDC16, FABP5, ENO2, TAGLN,* and *ATF3* as having diagnostic potential for COVID-19, and *ACTR2, ACTR3, ARPC1B, ARPC1A, FABP5, ENO2, TAGLN,* and *ATF3* as having diagnostic potential for sarcoidosis. These results revealed that hub genes identified from the PPI network of common DEGs had higher diagnostic potential in both COVID-19 and sarcoidosis than hub genes identified from the PPI network of disease-specific DEGs.

3.4. Regulatory signature

TFs and miRNAs regulate various biological processes, including pathological processes in sarcoidosis and COVID-19. We constructed miRNA-DEG and TF-DEG regulatory networks and identified several hub regulatory factors for each network. For the 33 common DEGs, the TF-DEG network contained 36 nodes and 123 edges and the miRNA-DEG network contained 33 nodes and 98 edges (Fig. 5A). For COVID-19-specific DEGs, the TF-DEG network contained 24 nodes and 54 edges and the miRNA-DEG network contained 44 nodes and 299 edges (Fig. 5B). For sarcoidosis-specific DEGs, the TF-DEG network contained 46 nodes and 198 edges and the miRNA-DEG network contained 40 nodes and 244 edges (Fig. 5C).

4. Discussion

A previous study investigated the cellular pathways involved in sarcoidosis and COVID-19; however, only autophagy-related pathways were considered (Calender et al., 2020). With a broader systems biology approach, our study identified pivotal genes, pathways, regulators, and biomarkers which were associated with COVID-19 and sarcoidosis, individually and collectively. These findings therefore provide novel potential therapeutic targets for patients with sarcoidosis and COVID-19. Moreover, to the best of our knowledge, our study is the first to

Table 3
Overview of the hub genes obtained from the PPI network consisted of common DEGs.

Gene	Description	Related diseases and pathways	Reference
FKBP1B	FKBP prolyl Isomerase 1B	This gene is a member of the immunophilin protein family. Increased FKBP1b reversed calcium dysregulation and memory impairment in aging rats	PMID: 26224869
ATF3	Activating transcription factor 3	This gene is associated with various steps of tumorigenesis. Decreased expression of ATF3 contributes to gastric cancer progression via increasing β-catenin and CEMIP	PMID: 34728784
FABP5	Fatty acid binding protein 5	This gene encodes the fatty acid binding protein, which promotes lymph node metastasis in cervical cancer	PMID: 32550890
TAGLN3	Transgelin 3	This gene contributes to the sequential progression of vertebrate neurogenesis	PMID: 25565981
ENO3	Enolase 3	This gene is a metalloenzyme that acts as a tumor inhibitor in hepatocellular carcinoma development	PMID: 35004693
DDIT3	DNA damage inducible transcript 3	This gene is associated with many mineralization processes and inflammatory diseases and suppresses cementoblast differentiation via the NF-κB pathway	PMID: 30488444
ENO2	Enolase 2	This gene encodes enolase isoenzymes that promotes cell proliferation and glycolysis in acute lymphoblastic leukemia	PMID: 29689546
FKBP1A	FKBP prolyl isomerase 1A	This gene plays an important role in intercellular communication between endocardium and myocardium via regulating Notch1	PMID: 23571217
TAGLN	Transgelin	This gene encodes an actin cross-linking/gelling protein that mediates ovarian cancer progression via RhoA/ROCK pathway	PMID: 34538264
HINT1	Histidine triad nucleotide binding protein 1	This gene mediates activation of MITF transcriptional activity in human melanoma cells	PMID: 28394346

The first and second columns show the gene symbol and full name of hub proteins. The third and fourth columns describe the related disease and pathway that hub gene involved in, and relevant references.

collectively analyze the host transcriptional response to COVID-19 and sarcoidosis.

We performed differential expression analysis to identify the common DEGs between sarcoidosis and COVID-19 followed by enrichment analysis to reveal the potential common mechanisms. We identified 33 common DEGs, which were largely associated with cytokine-related signaling pathways. This finding implies that cytokines have a strong biological relevance with the pathogenesis of COVID-19 and sarcoidosis. Consistent with our findings, a previous study found that patients with COVID-19 have high cytokine signatures, which appeared to be the driving feature of COVID-19 (Blanco-Melo et al., 2020). Similarly, multiple studies have reported exuberant inflammatory cytokine production in patients with sarcoidosis (Beirne et al., 2009; Gerke and Hunninghake, 2008). Cytokines are soluble proteins that are often responsible for many infection-related symptoms, such as myalgia, fever, and headache (Slifka and Whitton, 2000), which are the common symptoms of COVID-19.

PPI networks are fundamental in systems biology and help to detect critical molecules among proteins (Ewing et al., 2007). We identified 10

hub proteins in the PPI network of common DEGs. ROC analysis showed that most of the hub proteins had high diagnostic potential (AUC > 0.7), suggesting their potential as biomarkers of COVID-19 and sarcoidosis. Among the hub genes, FKBP prolyl isomerase 1B (*FKBP1B*) is a member of the immunophilin family. Increased *FKBP1b* levels have been reported to reverse calcium dysregulation and memory impairment in aging rats (Gant et al., 2015). Activating transcription factor 3 (*ATF3*) is associated with various steps of tumorigenesis, and its decreased expression contributes to gastric cancer progression via increasing β-catenin and CEMIP levels (Xie et al., 2021). Fatty acid binding protein 5 (*FABP5*) encodes a fatty acid binding protein, which has been shown to promote lymph node metastasis in cervical cancer (Zhang et al., 2020). Transgelin 3 (*TAGLN3*) contributes to the sequential progression of vertebrate neurogenesis (Ratié et al., 2014). Enolase 3 (*ENO3*) is a metalloenzyme that acts as a tumor inhibitor in hepatocellular carcinoma development (Cui et al., 2021). DNA damage inducible transcript 3 (*DDIT3*) is associated with many mineralization processes and inflammatory diseases, and suppresses cementoblast differentiation via the NF-κB pathway (Liu et al., 2019). Enolase 2 (*NEO2*) encodes enolase isoenzymes that promote cell proliferation and glycolysis in acute lymphoblastic leukemia (Liu et al., 2018). FKBP prolyl isomerase 1A (*FKBP1A*) plays an important role in intercellular communication between the endocardium and myocardium via regulating Notch1 expression (Liu et al., 2018). Transgelin (*TAGLN*) encodes an actin cross-linking/gelling protein that mediates ovarian cancer progression via the RhoA/ROCK pathway (Wei et al., 2021). Histidine triad nucleotide binding protein 1 (*HINT1*) mediates the activation of MITF transcriptional activity in human melanoma cells (Motzik et al., 2017).

GRNs, which describe the regulatory relationships between the TFs and miRNAs and their target genes, are an essential component of bioinformatics. In this study, GRN analysis revealed that 36 TFs and 33 miRNAs could regulate the common DEGs. Among those, hsa-mir-16-5p and *FOXC1* showed the strongest interactions with the common DEGs. A previous study demonstrated that hsa-mir-16-5p can regulate ACE2-related networks (Wicik et al., 2020). ACE2 is one of the main cellular receptors for SARS-CoV-2: the virus binds to ACE2 receptor to infiltrate host cells by binding the receptor and then rapidly replicates and spreads, ultimately causing organ damage (Renhong et al., 2020). Thus, this study suggests that hsa-mir-16-5p affects the clinical outcomes of patients with COVID-19. Forkhead box C1 (*FOXC1*) mediates chemotherapeutic resistance and the metastasis of gastric adenocarcinoma via the epithelial-to-mesenchymal transition and the hedgehog pathway (Jun et al., 2021).

This study has some limitations. We conducted an in-depth analysis of the transcriptomic data of COVID-19 and sarcoidosis to identify pivotal DEGs, pathways, regulators, and biomarkers. However, our findings have not been validated in vitro, ex vivo, and in vivo, and we hope to work on this in the future. Furthermore, we did not perform transcriptomic analyses in patients with sarcoidosis having COVID-19 owing to the current lack of relevant sample data. However, with continuous accumulation of samples in the future, this analysis is likely to become possible.

5. Conclusion

In this study, we identified common pathological mechanisms between COVID-19 and sarcoidosis using multiple analytical approaches. We identified 33 common DEGs between COVID-19 and sarcoidosis. A total of 36 TFs and 33 miRNAs regulated the common DEGs, including hsa-mir-16-5p and *FOXC1*. We also identified several hub proteins related to the common DEGs that were identified. The identified hub genes have the potential for use as biomarkers for the diagnosis of COVID-19 in patients with sarcoidosis. We believe our findings can provide a pathophysiological basis explaining why patients with sarcoidosis tend to develop severe COVID-19, which can inform treatment to improve outcomes in this high-risk population.
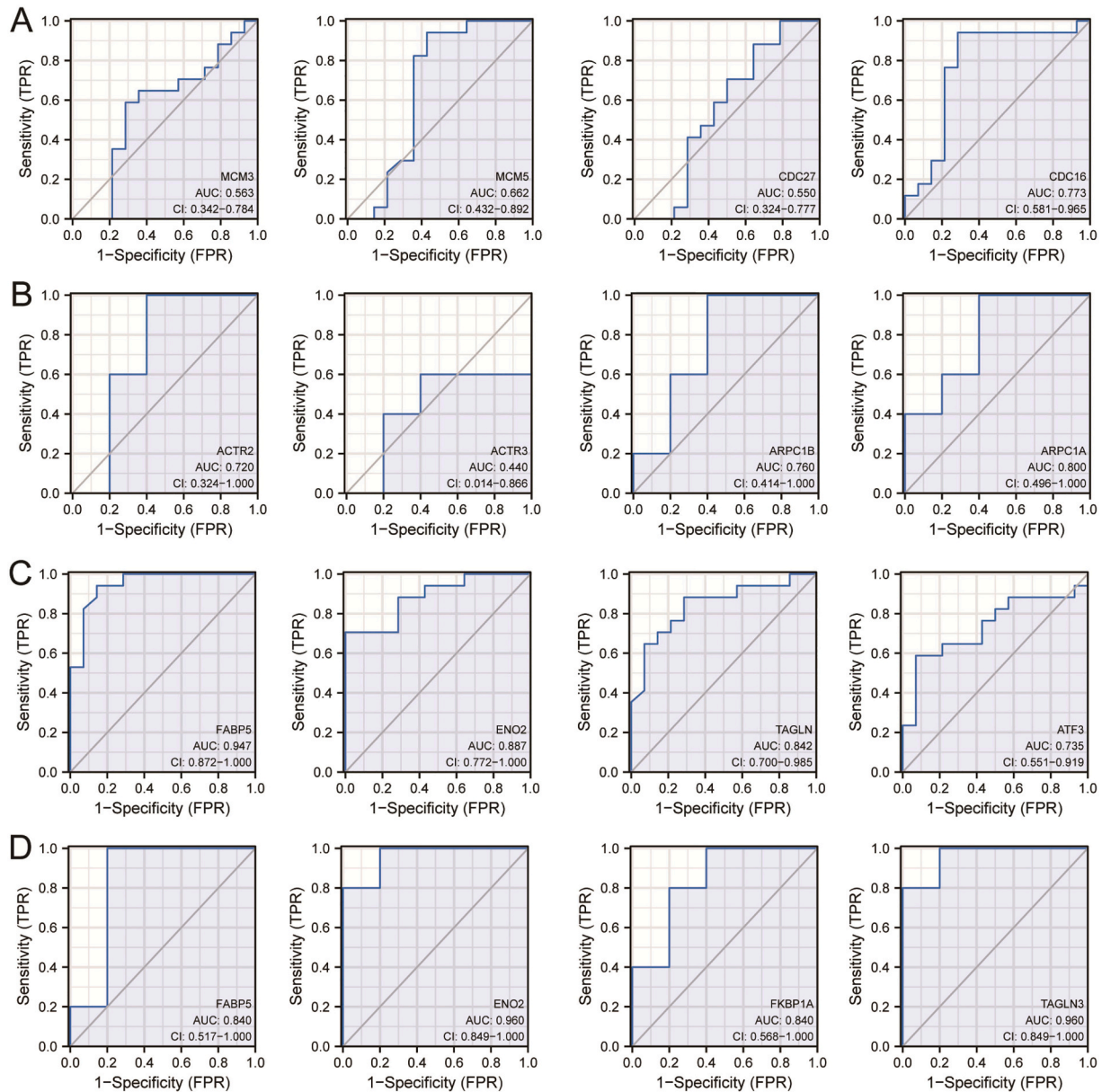
Fig. 4. Receiver-operating characteristic curves to assess the potential diagnostic value of biomarkers for COVID-19 and sarcoidosis. (A) shows the diagnostic potential of *MCM3, MCM5, CDC27* and *CDC16* in COVID-19. (B) shows the diagnostic potential of *ACTR2, ACTR3, ARPC1B* and *ARPC1A* in sarcoidosis. (C) shows the diagnostic potential of *FABP5, ENO2, FKBP1A,* and *TAGLN3* in COVID-19. (D) shows the diagnostic potential of *FABP5, ENO2, TAGLN,* and *ATF3* in sarcoidosis.
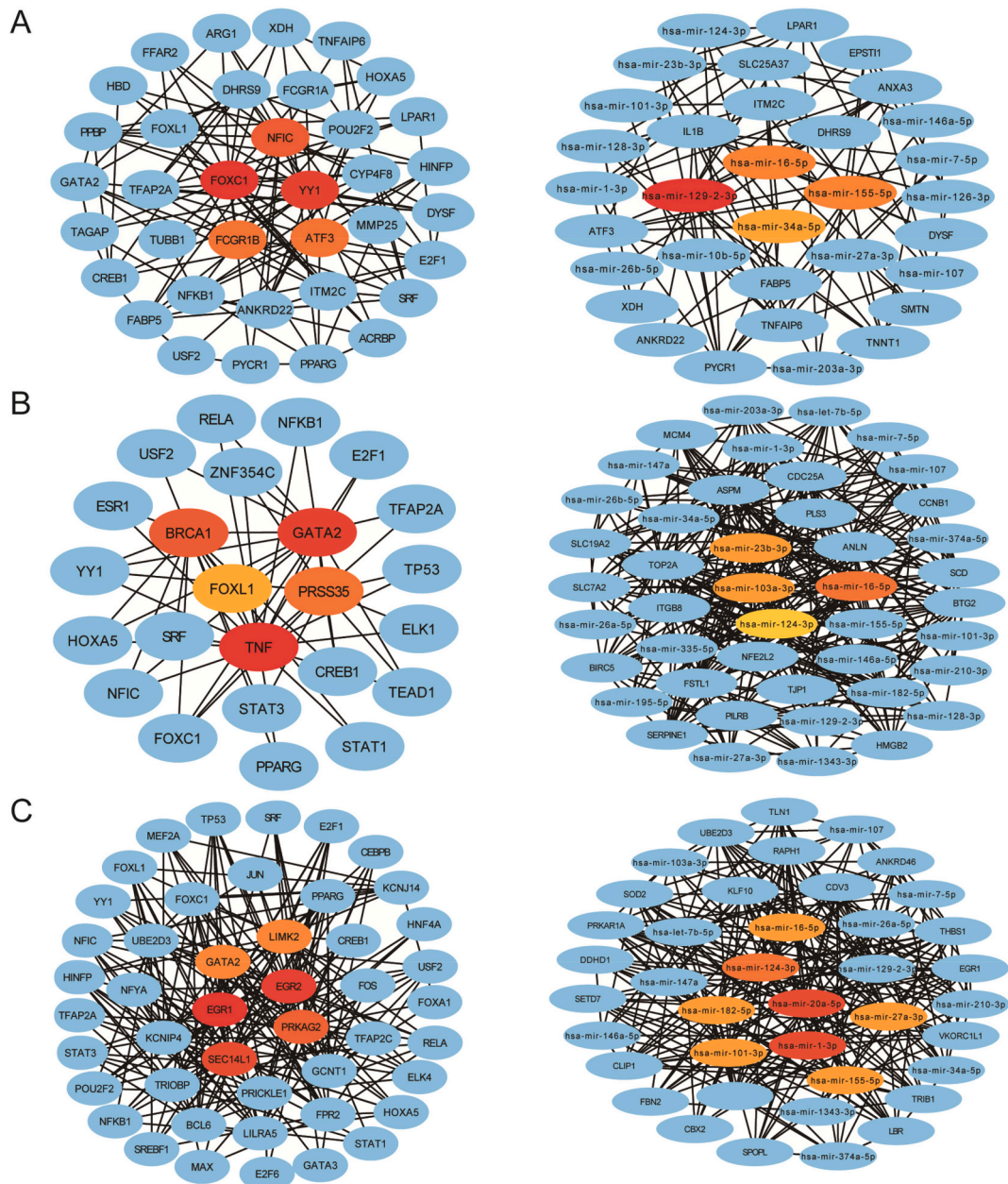
Fig. 5. Construction of gene regulatory networks. (A) Protein-protein interaction network consisted of common DEGs between sarcoidosis and COVID-19. (B) Protein-protein interaction network consisted of COVID-19-specific DEGs. (C) Protein-protein interaction network consisted of sarcoidosis-specific DEGs. The genes in highlight represent hub TFs/miRNA. Abbreviations: DEGs, differentially expressed genes; miRNA, microRNA; TFs, transcription factors.

CRediT authorship contribution statement

Lanyi Fu: Visualization, Writing – original draft, Data curation, Project administration. **Maolin Yao:** Formal analysis, Data curation, Writing – original draft, Project administration. **Xuedong Liu:** Conceptualization, Supervision, Funding acquisition. **Dong Zheng:** Conceptualization, Supervision, Funding acquisition.

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.

Ethical approval

This study does not involve any human or animal testing.

Data availability statement

All data utilized in this manuscript are available online from NCBI GEO database with accession number GSE164805 (COVID-19), GSE152418 (COVID-19) and GSE42832 (sarcoidosis).

References

Al-Kailany, W., Timens, W., Venmans, B., de Jonge, G., van der Werf, T.S., 2021. Sarcoidosis presenting with glazy mucoid sputum and dyspnea: a case report. J. Med. Case Rep. 15, 1–5. https://doi.org/10.1186/s13256-021-02809-2.

Al-Quteimat, O.M., Amer, A.M., 2020. The impact of the COVID-19 pandemic on cancer patients. Am. J. Clin. Oncol. 43, 453–455. https://doi.org/10.1097/COC.0000000000000712.

Bargagli, E., Prasse, A., 2018. Sarcoidosis: a review for the internist. Intern. Emerg. Med. 13, 325–331. https://doi.org/10.1007/s11739-017-1778-6.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C.L., Serova, N., Davis, S., Soboleva, A., 2013. NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res. 41, 991–995. https://doi.org/10.1093/nar/gks1193.

Beirne, P., Pantelidis, P., Charles, P., Wells, A.U., Abraham, D.J., Denton, C.P., Welsh, K. I., Shah, P.L., Du Boise, R.M., Kelleher, P., 2009. Multiplex immune serum biomarker profiling in sarcoidosis and systemic sclerosis. Eur. Respir. J. 34, 1376–1382. https://doi.org/10.1183/09031936.00028209.

Blanco-Melo, D., Nilsson-Payant, B.E., Wen Chun, L., Uhl, S., Hoagland, D., Møller, R., Jordan, T.X., Oishi, K., Panis, M., Sachs, D., Taia, T.W., Schwartz, R.E., Lim, J.K., Albrecht, R.A., TenOever, B.R., 2020. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. Cell 181, 1036–1045. https://doi.org/10.1016/j.cell.2020.04.026.

Bloom, C.I., Graham, C.M., Berry, M.P.R., Rozakeas, F., Redford, P.S., Yuanyuan, W., Zhaohui, X., Wilkinson, K.A., Wilkinson, R.J., Kendrick, Y., Devouassoux, G., Ferry, T., Miyara, M., Bouvry, D., Dominique, V., Gorochov, G., Blankenship, D., Saadatian, M., Vanhems, P., Beynon, H., Vancheeswaran, R., Wickremasinghe, M., Chaussabel, D., Banchereau, J., Pascual, V., Lingpei, H., Lipman, M., O'Garra, A., 2013. Transcriptional blood signatures distinguish pulmonary tuberculosis, pulmonary sarcoidosis, pneumonias and lung cancers. PLoS ONE 8, e70630. https://doi.org/10.1371/journal.pone.0070630.

Calender, A., Israel-Biet, D., Valeyre, D., Pacheco, Y., 2020. Modeling potential autophagy pathways in COVID-19 and sarcoidosis. Trends Immunol. 41, 856–859. https://doi.org/10.1016/j.it.2020.08.001.

Chin, C., Chen, S., Wu, H., Ho, C., Ming, T.K., Lin, C., 2014. CytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst. Biol. 8, S11. https://doi.org/10.1186/1752-0509-8-S4-S11.

Cui, H., Guo, D., Zhang, X., Zhu, Y., Wang, Z., Jin, Y., Guo, W., Zhang, S., 2021. ENO3 inhibits growth and metastasis of hepatocellular carcinoma via Wnt/β-catenin signaling pathway. Front. Cell Dev. Biol. 9, 97–102. https://doi.org/10.3389/fcell.2021.797102.

Doms, A., Schroeder, M., 2005. GoPubMed: exploring PubMed with the gene ontology. Nucleic Acids Res. 33, 783–786. https://doi.org/10.1093/nar/gki470.

Ewing, R.M., Peter, C., Elisma, F., Hongyan, L., Taylor, P., Climie, S., McBroom-Cerajewski, L., Robinson, M.D., O'Connor, L., Michael, L., Taylor, R., Dharsee, M., Yuen, H., Heilbut, A., Moore, L., Shudong, Z., Ornatsky, O., Bukhman, Y.V., Ethier, M., Yinglun, S., Vasilescu, J., Abu-Farha, M., Lambert, J.P., Duewel, H.S., Stewart, I.I., Kuehl, B., Hogue, K., Colwill, K., Gladwish, K., Muskat, B., Kinach, R., Adams, S.L., Moran, M.F., Morin, G.B., Topaloglou, T., Figeys, D., 2007. Large-scale mapping of human protein-protein interactions by mass spectrometry. Mol. Syst. Biol. 3, 1–17. https://doi.org/10.1038/msb4100134.

Gant, J.C., Kuey Chu, C., Kadish, I., Blalock, E.M., Thibault, O., Porter, N.M., Landfield, P.W., 2015. Reversal of aging-related neuronal Ca2+ dysregulation and cognitive impairment by delivery of a transgene encoding FK506-binding protein 12.6/1b to the hippocampus. J. Neurosci. 35, 10878–10887. https://doi.org/10.1523/JNEUROSCI.1248-15.2015.

Gerke, A.K., Hunninghake, G., 2008. The immunology of sarcoidosis. Clin. Chest Med. 29, 379–390. https://doi.org/10.1016/j.ccm.2008.03.014.

Gorbalenya, A.E., Baker, S.C., Baric, R.S., de Groot, R.J., Drosten, C., Gulyaeva, A.A., Haagmans, B.L., Lauber, C., Leontovich, A.M., Neuman, B.W., Penzar, D., Perlman, S., Poon, L.L.M., Samborskiy, D.V., Sidorov, I.A., Sola, I., Ziebuhr, J., 2020. The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat. Microbiol. 5, 536–544. https://doi.org/10.1038/s41564-020-0695-z.

Jun, L., Longlong, C., Yu, X., Xiaoyan, H., Soojeong, C., Ashktorab, H., Smoot, D.T., Ping, L., Chaohui, Z., Kim, J., Ryeom, S.W., Yoon, S.S., Yoon, C., Changming, H., 2021. FOXC1 modulates stem-like cell properties and chemoresistance through hedgehog and EMT signaling in gastric adenocarcinoma. Mol. Ther. 21, 464–470. https://doi.org/10.1016/J.YMTHE.2021.09.009.

Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J.A., Van Der Lee, R., Bessy, A., Chèneby, J., Kulkarni, S.R., Tan, G., Baranasic, D., Arenillas, D.J., Sandelin, A., Vandepoele, K., Lenhard, B., Ballester, B., Wasserman, W.W., Parcy, F., Mathelier, A., 2018. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic Acids Res. 46, D260–D266. https://doi.org/10.1093/nar/gkx1126.

Kondle, S., Hou, T., Manansala, M., Ascoli, C., Novak, R.M., Sweiss, N., 2021. Treatment of COVID-19 in patients with sarcoidosis. Front. Med. 8, 689539 https://doi.org/10.3389/fmed.2021.689539.

Li, Y., Liang, Z., Zheng, Y., Qiao, J., Wang, P., 2021. Pulmonary sarcoidosis: from clinical features to pathology-narrative review. Ann. Palliat. Med. 10, 3438–3444. https://doi.org/10.21037/apm-21-344.

Liu, C., Wang, H., Wang, W., Wang, L., Liu, W., Wang, J., Geng, Q., Lu, Y., 2018. ENO2 promotes cell proliferation, glycolysis, and glucocorticoid-resistance in acute lymphoblastic leukemia. Cell. Physiol. Biochem. 46, 1525–1535. https://doi.org/10.1159/000489196.

Liu, X., Sun, H., Yu, M., Liu, J., Yang, B., Wu, Y., Wang, J., 2019. DDIT3 regulates cementoblast mineralization by isocitrate dehydrogenase 1 through nuclear factor-κB pathway. J. Cell. Physiol. 234, 11602–11609. https://doi.org/10.1002/jcp.27811.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 1–21. https://doi.org/10.1186/s13059-014-0550-8.

Manansala, M., Chopra, A., Baughman, R.P., Novak, R., Lower, E.E., Culver, D.A., Korsten, P., Drake, W.P., Judson, M.A., Sweiss, N., 2021. COVID-19 and sarcoidosis, readiness for vaccination: challenges and opportunities. Front. Med. 8, 1–7. https://doi.org/10.3389/fmed.2021.672028.

Motzik, A., Amir, E., Erlich, T., Wang, J., Kim, B.G., Han, J.M., Kim, J.H., Nechushtan, H., Guo, M., Razin, E., Tshori, S., 2017. Post-translational modification of HINT1 mediates activation of MITF transcriptional activity in human melanoma cells. Oncogene 36, 4732–4738. https://doi.org/10.1038/onc.2017.81.

Nashiry, A., Sarmin Sumi, S., Islam, S., Quinn, J.M.W., Moni, M.A., 2021. Bioinformatics and system biology approach to identify the influences of COVID-19 on cardiovascular and hypertensive comorbidities. Brief. Bioinform. 22, 1387–1401. https://doi.org/10.1093/bib/bbaa426.

Nashiry, M.A., Sumi, S.S., Sharif Shohan, M.U., Alyami, S.A., Azad, A.K.M., Moni, M.A., 2021. Bioinformatics and system biology approaches to identify the diseasome and comorbidities complexities of SARS-CoV-2 infection with the digestive tract disorders. Brief. Bioinform. 22, 1–16. https://doi.org/10.1093/bib/bbab126.

Peric, S., Stulnig, T.M., 2020. Diabetes and COVID-19: disease-management-people. Wien. Klin. Wochenschr. 132, 356–361. https://doi.org/10.1007/s00508-020-01672-3.

Radke, R.M., Frenzel, T., Baumgartner, H., Diller, G.P., 2020. Adult congenital heart disease and the COVID-19 pandemic. Heart 106, 1302–1309. https://doi.org/10.1136/heartjnl-2020-317258.

Ratié, L., Ware, M., Jagline, H., David, V., Dupé, V., 2014. Dynamic expression of notch-dependent neurogenic markers in the chick embryonic nervous system. Front. Neuroanat. 8, 1–13. https://doi.org/10.3389/fnana.2014.00158.

Renhong, Y., Yuanyuan, Z., Yaning, L., Lu, X., Yingying, G., Qiang, Z., 2020. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science 367, 1444–1448. https://doi.org/10.1126/science.abb2762.

Satu, S., Khan, I., Rahman, R., Howlader, K.C., Roy, S., Roy, S.S., Quinn, J.M.W., Moni, M.A., 2021. Diseasome and comorbidities complexities of SARS-CoV-2 infection with common malignant diseases. Brief. Bioinform. 22, 1415–1429. https://doi.org/10.1093/bib/bbab003.

Sethupathy, P., Corda, B., Hatzigeorgiou, A.G., 2006. TarBase: a comprehensive database of experimentally supported animal microRNA targets. RNA 12, 192–197. https://doi.org/10.1261/rna.2239606.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models. Genome Res. 13, 2498–2504. https://doi.org/10.1101/gr.1239303.metabolite.

Slifka, M.K., Whitton, J.L., 2000. Clinical implications of dysregulated cytokine production. J. Mol. Med. 78, 74–80. https://doi.org/10.1007/s001090000086.

Smyth, G.K., 2005. Limma: linear models for microarray data. In: Bioinformatics And Computational Biology Solutions Using R And Bioconductor, pp. 397–420.

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., Jensen, L.J., Mering, C., 2019. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 47, D607–D613. https://doi.org/10.1093/nar/gky1131.

Taz, T.A., Ahmed, K., Paul, B.K., Kawsar, M., Aktar, N., Mahmud, S.M.H., Moni, M.A., 2021. Network-based identification genetic effect of SARS-CoV-2 infections to idiopathic pulmonary fibrosis (IPF) patients. Brief. Bioinform. 22, 1254–1266. https://doi.org/10.1093/bib/bbaa235.

Walter, W., Sánchez-Cabo, F., Ricote, M., 2015. GOplot: an R package for visually combining expression data with functional analysis. Bioinformatics 31, 2912–2914. https://doi.org/10.1093/bioinformatics/btv300.

Wang, Y., Cho, D.Y., Lee, H., Fear, J., Oliver, B., Przytycka, T.M., 2018. Reprogramming of regulatory network using expression uncovers sex-specific gene regulation in drosophila. Nat. Commun. 9, 4061–4071. https://doi.org/10.1038/s41467-018-06382-z.

Wei, X., Lou, H., Zhou, D., Jia, Y., Li, H., Huang, Q., Ma, J., Yang, Z., Sun, C., Meng, Y., Xu, S., Yang, X., Li, X., Ji, T., Guo, Z., Gao, Q., 2021. TAGLN mediated stiffness-regulated ovarian cancer progression via RhoA/ROCK pathway. J. Exp. Clin. Cancer Res. 40, 292–306. https://doi.org/10.1186/s13046-021-02091-6.

Wicik, Z., Eyileten, C., Jakubik, D., Simões, S.N., Martins, D.C., Pavão, R., Siller-Matula, J.M., Postula, M., 2020. ACE2 interaction networks in COVID-19: a physiological framework for prediction of outcome in patients with cardiovascular risk factors. J. Clin. Med. 9, 3743–3772. https://doi.org/10.3390/jcm9113743.

Xie, G., Dong, P., Chen, H., Xu, L., Liu, Y., Ma, Y., Zheng, Y., Yang, J., Zhou, Y., Chen, L., Shen, L., 2021. Decreased expression of ATF3, orchestrated by β-catenin/TCF3, miR-17-5p and HOXA11-AS, promoted gastric cancer progression via increased β-catenin and CEMIP. Exp. Mol. Med. 53, 1706–1722. https://doi.org/10.1038/s12276-021-00694-9.

Yu, G., Wang, L., Han, Y., He, Q., 2012. ClusterProfiler: an R package for comparing biological themes among gene clusters. Omi. A J. Integr. Biol. 16, 284–287. https://doi.org/10.1089/omi.2011.0118.

Zhang, C., Liao, Y., Liu, P., Du, Q., Liang, Y., Ooi, S., Qin, S., He, S., Yao, S., Wang, W., 2020. FABP5 promotes lymph node metastasis in cervical cancer by reprogramming fatty acid metabolism. Theranostics 10, 6561–6580. https://doi.org/10.7150/thno.44868.

Zhang, Q., Meng, Y., Wang, K., Zhang, X., Chen, W., Sheng, J., Qiu, Y., Diao, H., Li, L., 2021. Inflammation and antiviral immune response associated with severe progression of COVID-19. Front. Immunol. 12, 1–12. https://doi.org/10.3389/fimmu.2021.631226.

Zhou, G., Soufan, O., Ewald, J., Hancock, R.E.W., Basu, N., Xia, J., 2019. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. Nucleic Acids Res. 47, W234–W241. https://doi.org/10.1093/nar/gkz240.