Study on Pathological Mechanism of Pneumonia Infected by Coronavirus Based on Time-Series Gene Co-expression Network Analysis

XingCheng Yi
School of Pharmaceutical Sciences
Jilin University
Changchun, China
yixc18@mails.jlu.edu.cn

Yan Zhang
College of Information Science and Technology
Dalian Maritime University
Dalian, China
jlmuzy@126.com

Tong Xu
Jilin Prochance Biomedical Co., Ltd.
Changchun, China
xtn66@163.com

Xiaoyun Su
School of Pharmaceutical Sciences
Jilin University
Changchun, China
suxy@jlu.edu.cn

Cong Fu
Key Laboratory of Organ Regeneration & Transplantation of the Ministry of Education
The First Hospital of Jilin University
Changchun, China
fucong@jlu.edu.cn

2021 IEEE 9th International Conference on Bioinformatics and Computational Biology (ICBCB) | 978-1-6654-1500-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICBCB52223.2021.9459223

***Abstract*—Recently, the epidemic of COVID-19 infection broke out in Wuhan, China. To explore the pathological mechanism of pneumonia infected by coronavirus, we built a bioinformatics pipeline based on time-series gene co-expression network analysis to analyze the gene expression profile of lung cells in mice infected by SARS-Cov (GSE19137). In this study, Pearson correlation analysis was performed to construct a gene co-expression network. Time-ordered gene network modules were digged out by BFS algorithm . PageRank algorithm was used to explore HUB genes related to pneumonia infected by coronavirus. Based on the information we got, we think that cell lines infected by coronavirus might go through 5 stages, and 10 HUB genes(AKT1, CD68, CTSS, FCGR3A, HSPA8, PTPRC, UBC, VCP, PRPF31, ITPKB) might play a key role in coronavirus infection. This might provide some hints for coronavirus related research.**

Keywords—coronavirus, time-series gene co-expression network, PageRank algorithm, BFS algorithm

I. INTRODUCTION

Recently, the epidemic of COVID-19 infection broke out in Wuhan, China [1]. As of February 23, 2020, 77269 confirmed, 3434 deaths and 24922 cured cases have been reported in China. So far, there hasn't been any effective drugs or vaccines reported [2]. Studies have shown that COVID-19 and severe acute respiratory syndrome coronavirus virus (SARS-Cov) have 79.5% similarity at sequence level, and both have ACE2 as the same cell entry receptor [3-4]. In this study, we took SARS-Cov as research object, explored the mechanism of coronavirus infection from the perspective of transcriptomics, and expected to provide some theoretical support for the treatment of coronavirus.

At present, SARS-Cov has become a hot-spot on the research of respiratory diseases. Hu *et al* found that 24 hours after SARS-Cov infection, IFN-α/β-inducible and cathepsin/proteasome genes were down-regulated and reactive oxygen-related genes were up-regulated. It was inferred that SARS-Cov infection could regulate immunity in monocytes/macrophages [5]. Wang *et al* found that SARS-CoV PLpro stimulated TGF-β1-dependent expression of Type I collagen via activating STAT6 pathway [6]. Shao *et al* found that PBMC could be used to identify gene expression profiles of patients in recovery and track the response of SARS-Cov to host infections. They also found that genes encoding mitochondria were significantly up-regulated in patients with SARS-Cov during recovery, and it was closely related to the antiviral immune response [7]. It can be seen that the researches of coronavirus are mainly focused on the interference of SARS-Cov on the immune system, and the analysis of time series gene expression profile after coronavirus infection has not been studied thoroughly.

With the popularization of sequencing technology and the rapid development of RNA-sequencing, gene co-expression network analysis has become a research hotspot. Among them, WGCNA algorithm is one of the most popular algorithms in the field of gene co-expression network analysis [8-11]. The process of biological activity was a dynamically changed process, and static functional modules analysis cannot satisfy the request of dynamic processes searching. Therefore, the analysis of time-series gene expression datasets has always been a concern of researchers. Wu *et al* performed Gene Ontology (GO) enrichment analysis to analyze differentially expression genes (DEGs) of time series dataset (GSE37069). And it was found that immune-related processes were significantly enriched at all time points, and samples at later time points were closely related to cell activation regulation processes [12]. Park *et al* performed an integrated network recognition algorithm to analyze *GEO* datasets (GSE41714, GSE9593, GSE15299), and found that the cell cycle-specific public network played an important role in regulating aging [13].

To explore the pathological mechanism of pneumonia infected by coronavirus, we used the gene co-expression network as background, combining with the network analysis algorithm and time series analysis algorithm to mine mice lung tissue transcriptome datasets at 1, 2, 3, 5, 14, and 28 days after SARS-Cov infection (GSE19137). Firstly, gene expression profiles were downloaded from the *GEO* database (https://www.ncbi.nlm.nih.gov/geo/), and DEGs were identified using DESeq2 algorithm [14]. Then, Pearson correlation analysis was used to construct gene co-expression network (GCN), and Breadth First Search (BFS) algorithm [15] was performed to mine time series modules. Furthermore, we performed GO enrichment analysis to explore the biological significance of each module. Later, we queried Search Tool for the Retrieval of Interacting Genes (*STRING*) database (https://string-db.org/) [16] to construct the protein-protein interaction network (PPI network) of module 4 which was found closely related to the regulation of the immune system and module 7 which was closely associated with material metabolism. Then HUB genes in the networks were identified by using PageRank algorithm [17]. Finally, we explored the functions of HUB genes in *DisGeNet* database (http://www.disgenet.org/)[18].

II. Materials and Method

To study the pathological mechanism of pneumonia infected by coronavirus, we took SARS-Cov as the research object and built a bioinformatics pipeline based on time-series gene co-expression network analysis to analyze the gene expression profile of lung cells in mice infected by SARS-CoV.

A. Data Sources

The original dataset of this study were derived from *GEO* database (GSE19137) which included samples of non-SARS-Cov infected mice (GSM474505-GSM474507) and samples of mice infected by SARS-Cov after 1, 2, 3, 5, 14, 28 days (GSM474508-GSM474525), , there were 3 replicates at each time points. There are 3 samples in the control group and 18 samples in the infection group. The expression values of 45,101 genes were included.

B. Analysis of Differentially Expressed Genes

DESeq2 algorithm[14] was used to identify the DEGs. The cut-off value was set at *Fold change* > 2.5 or < 0.4 and *P value* < 0.05. 6 groups of SARS-Cov infected samples were compared to the control group respectively to identify DEGs, and all DEGs were merged together to obtain the DEG dataset for further analysis.

C. Construction of Gene Co-expression Network

Chang *et al* claimed that when the expression levels of two genes in different samples were analyzed by Pearson correlation analysis, if the absolute value of the correlation coefficient was greater than a certain threshold and was in line with statistical significance, two genes could be considered to have a co-expression relationship [19].

The expression levels of DEGs in 18 SARS-Cov infected samples of mice were analyzed by Pearson correlation analysis, and Pearson correlation coefficient (*PCC*) and its *P value* between pairs of genes were obtained. The cut-off value was set at $|PCC| >= 0.85$ and $P\ value < 0.05$ to screen the interactions between gene pair. Two genes that met this qualification were considered to be co-expressed, and a gene GCN was constructed based on these interactions.

D. Mining of Time-Ordered Gene Modules

Since connected genes in a GCN share similar patterns of up-regulation or down-regulation in the time course, we decided to infer the expression time order for all genes in the GCN by BFS algorithm. Firstly, the initial node of BFS should be determined. The gene located in the first time-ordered gene module should be selected as the initial node, which was the gene with peak expression in Day01 (the first day after infection) and monotonously decreased until Day28. Further, the expression time-ordered levels of all genes in the GCN could be inferred by traversing the GCN using BFS algorithm. Finally, the time-ordered levels were merged according to results of GO enrichment analysis to obtain the time-ordered modules.

E. Construction of PPI Network

All the genes in a time-ordered gene module were put into *STRING* database to construct a PPI network. Where"multiple proteins by names/identifiers" were selected for query mode; and the minimum required interaction score was set to "medium confidence (0.400)".

F. Identification of HUB Genes

PageRank algorithm is an algorithm proposed by Google founders Larry Page and Sergey Brin to calculate the importance of nodes in a complex network [17]. In this study, PageRank algorithm was used to score the importance of nodes (genes) in PPI network (the scoring criteria were based on topology principles), and 5 genes with the highest scores in a module were defined as HUB genes. Furthermore, the homologs of these hub genes in human genome were collected.

G. Use of DisGeNet Database

DisGeNet Database contains the information between diseases and genes. In this study, *DisGeNet* database was used to query diseases related to HUB genes to explore the functions of HUB genes.

III. Result

A. Identification of DEGs Using DESeq2 Algorithm

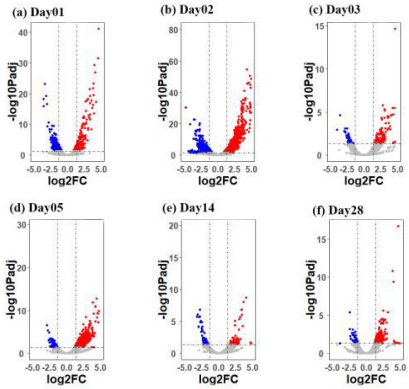

Fig. 1. The DEGs in each group. (a) ~ (f) are the DEG identification results of Day1, 2, 3, 5, 14, 28.

We performed DESeq2 algorithm to identify the DEGs in each group (Fig. 1 & Supplementary Material, Table S1), and got 3,978 DEGs which met appropriate thresholds (*Fold change* > 2.5 or *Fold change* < 0.4 and *P value* < 0.05) after merging all DEGs in 6 groups (Supplementary Material, Table S2).

B. Construction of GCN Using Pearson Correlation Analysis

We performed Pearson correlation analysis to the expression values of DEGs (3,978) in 18 SARS-Cov infected samples, and there were 15,824,484 interactions among these DEGs. Then, interactions failed to meet the chosen threshold (|*PCC*| < 0.85 and *P value* >= 0.05) were removed, and 3,871 genes and 1,291,080 interactions were kept.

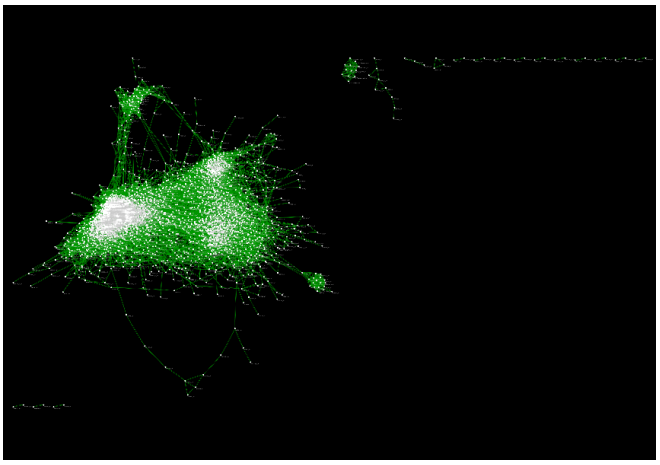

Fig. 2. GCN and some small nets.

The 1,291,080 interactions were visualized using *Cytoscape* software (Fig. 2). A large network and several small networks were built. The large network contained 3,818 genes, while smaller networks only included less than 20 genes. We removed all of small ones and only kept the large one (GCN) for further research (Supplementary Material, Table S3).

C. Mining of Time-Ordered Gene Modules

In this study, gene AV354139 (which had a peak expression in Day01 and monotonously decreased until Day28) was selected as the initial node, and the GCN was traversed by BFS algorithm to infer the expression time order for all genes in the GCN.

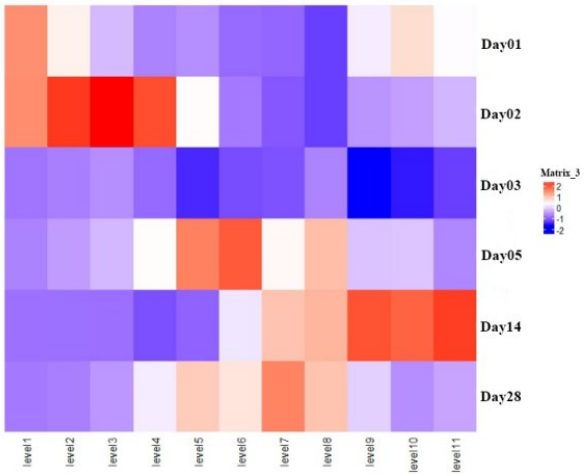

Fig. 3. The heatmap of average normalized gene expression values (z scores) at each time point of DEGs at each level of the GCN. For each DEG, the gene expression values over time points are normalized to z scores first. For each level, the z scores of all DEGs are averaged in the heatmaps.

The BFS algorithm assigned 3818 genes into 11 time-ordered levels (Supplementary Material, Table S4). Furthermore, we made a heatmap of the average normalized gene expression values (*z scores*) (Fig. 3). As can be seen from Fig. 3, red squares (highly expressed) are roughly distributed around the main diagonal. Further, GO enrichment analysis was used to explore the biological significance of each time-ordered level, and the adjacent levels with similar functions were merged to obtain a total of 8 time-ordered gene modules (Fig. 4 & Supplementary Material, Table S5).

The biological process (BP) results of 8 time-ordered modules are shown in Supplementary Material, Table S6. It could be seen that module 1 was closely related to the cell's stress protection; module 2 was closely related to the regulation of the γ-interferon response and the defense response to virus; module 3 was closely related to the regulation of β-interferon response and negative regulation of viral processes; module 4 was closely related to the activation of T cells and neutrophils; module 5 was closely related to the metabolism of reactive oxygen species and the regulation of muscle tissue development and cell growth; module 6 was closely related to the regulation of cell matrix adhesion and cell chemotaxis; module 7 was closely related to the cell's hypoxic response and cell migration; module 8 was closely related to the apoptotic process of motor neurons.

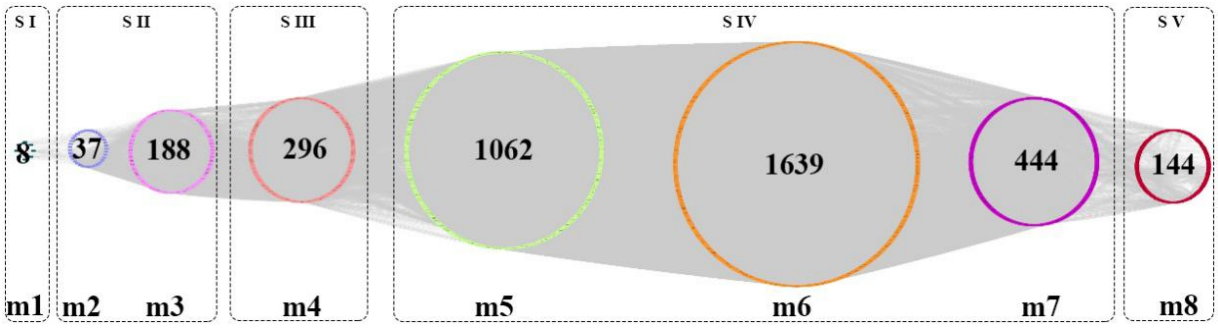

Fig. 4. Time-ordered gene modules. The GCN structure with DEGs as nodes. The number in a circle indicates the DEG number at that module.

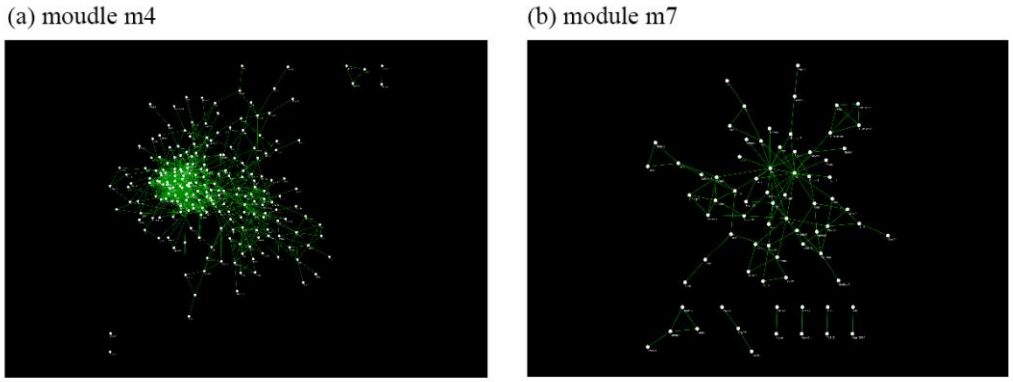

Fig. 5. The PPI networks of key modules. (a) ~ (b) are PPI networks for module 4 and module 7.

From above analysis results, it could be found that the 8 time-ordered modules could be roughly divided into 5 stages: 1. stage I (module 1): the main biological processes were related to the body's own stress protection response; 2. stage II (module 2, module 3): the main biological processes were related to the regulation of the interferon response and the defense response to the virus; 3. stage III (module 4): the main biological processes were related to response to immune cell-mediated immune regulation; 4. stage IV (module 5, module 6, module 7): the main biological processes were related to metabolism of reactive oxygen species, tissue development and cell differentiation; stage V (module 8): the main biological processes were related to motor neuron apoptotic process.

D. Construction of PPI Networks in Key Modules

According to current literature reports, the PLpro in SARS-Cov had the additional function of stripping ubiquitin and ISG15 from host cell proteins to help coronaviruses evade the innate immune response of host [20]. SARS-Cov 3a protein activated NLRP3 inflammasomes in macrophages induced by lipopolysaccharide, and K+ efflux and mitochondrial reactive oxygen species play important roles in activation of SARS-Cov 3a-induced NLRP3 inflammasome [21]. Thus, combining the results of GO enrichment analysis, we defined modules 4 (stage III) and 7 (stage IV) as the key modules of pneumonia infected by coronavirus. Furthermore, genes of module 4 and 7 were put into the *STRING* database respectively to construct the PPI network (Fig. 6).

E. Identification of HUB Genes in Key Modules and Exploration of Their Functions

We used *PageRank* algorithm to score the importance of nodes (genes) in the two large, major PPI networks built based on modules 4 and 7 respectively, and defined 5 genes with the highest scores in one module as HUB genes respectively. After mapping, a total of 10 HUB genes of human (AKT1, CD68, CTSS, FCGR3A, HSPA8, ITPKB, PRPF31, PTPRC, UBC, VCP) were obtained.

To explore the functions of these HUB genes, we put them into the *DisGeNet* database to query diseases related. From the results returned by the *DisGeNet* database, 10 HUB genes were mainly related to 3 types of diseases, namely virus diseases, lung diseases, immune system diseases and inflammatory diseases. There were 8 genes related to viral diseases: AKT1, CD68, CTSS, FCGR3A, HSPA8, PTPRC, UBC, VCP, and the corresponding diseases included HIV Infections, Kaposi Sarcoma, Coinfection, Hepatitis B, Merkel cell carcinoma, *etc.*; There are 8 genes related to lung diseases: AKT1, CD68, CTSS, FCGR3A, HSPA8, PRPF31, UBC, VCP, and the corresponding diseases included Carcinoma of lung, Adenocarcinoma of lung, Small cell carcinoma of lung, Abnormal lung lobation, Interstitial lung fibrosis, *etc*.; There were 9 genes related to immune system diseases and inflammatory diseases: AKT1, CD68, CTSS, FCGR3A, HSPA8, ITPKB, PTPRC, UBC, VCP, and the corresponding diseases included Rhumatoid Arthritis, Lupus Vasculitis, Central Nervous System, IGA Glomerulonephritis , HIV Infections, B-Cell Lymphomas, Immunologic Deficiency Syndromes, *etc.*

IV. DISCUSSION

A. Some Problems about the Bioinformatics Algorithm

To explore the pathological mechanism of pneumonia infected by coronavirus, this study used apply a bioinformatics pipeline based on analysis of time-series gene co-expression networks to analyze the gene expression profile of lung cells in mice infected by SARS-CoV. BFS algorithm is an important search algorithm in graph theory. It is a non-heuristic search method [15]. In this study, BFS algorithm was performed to infer the expression time order for all genes in the GCN. As for the initial node of BFS algorithm, we believed that the initial node must be located in the first time-ordered gene module. That was, the gene with peak expression in Day01 and monotonously decreased until Day28 should be selected as the initial node[20]. In the data analysis, some parameters need to be explained. If the standard setting of DEG identification was stricter, the reliability of the identified genes would be higher. But, this might prevent some important genes away from the research scope; On the contrary, if the DEG identification standard was set too low, it might cause lots of false positives. In this study, we believed that it was reasonable to control the number of DEGs between 3000-4000. Similarly, if the threshold selection of the GCN construction was higher, the GCN result would be more accurate. However, this would not only reduce the number of genes contained in GCN, but also reduce the network density of GCN, and even failed to construct an effective GCN, which would affect subsequent analysis results. Zu *et al* considered that when $|PCC| >= 0.5$,

the requirements for GCN construction could be met [22]. We believe that if two genes met the condition $|PCC| >= 0.85$, they could be considered to have a co-expression relationship.

B. Biological Significance of the Time-Ordered Gene Modules and the HUB Genes

From the results of GO enrichment analysis, we found that the cell line infected by SARS-Cov mainly underwent 5 stages: 1. stage I(module 1): the main biological processes were related to the body's own stress protection response (regulation of response to endoplasmic reticulum stress *etc*.); 2. stage II(module 2, module 3): the main biological processes were related to the regulation of the interferon response and the defense response to the virus (response to interferon-gamma etc.); 3. stage III (module 4): the main biological processes were related to response to immune cell-mediated immune regulation (T cell mediated immunity etc.); 4. stage IV (module 5, module 6, and module 7): the main biological processes were related to metabolism of reactive oxygen species, and tissue development and cell differentiation (positive regulation of reactive oxygen species metabolic process and muscle tissue development *etc.*); Stage V(module 8): the main biological processes were related to motor neuron apoptotic process. From the results of HUB genes analysis of key modules, 10 HUB genes (AKT1, CD68, CTSS, FCGR3A, HSPA8, PTPRC, UBC, VCP, PRPF31, ITPKB) were found playing key roles in module 4 and 7, and HUB genes were closely related to 3 types of diseases: viral diseases, lung diseases, and inflammatory diseases. It could be inferred that these HUB genes played a key role in pneumonia infected by coronavirus.

C. A Review of Research on HUB Genes

AKT1 is a serine protein kinase involved in cell metabolism, cell growth, and proliferation [23]. Recent studies showed that AKT1-mediated mitochondrial phagocytosis contributed to apoptosis of alveolar macrophage, and it was a necessary biological process for the development of pulmonary fibrosis [24]. CD68 is a syrup protein and is the most reliable marker of macrophages [25-26]. CTSS is a cysteine protease with higher concentrations in BAL fluid and plasma in subjects with chronic obstructive pulmonary disease[27]. Weldon *et al* found that the miR-31/IRF-1/CTSS pathway might play a key role in the pathogenesis of pulmonary cystic fibrosis [28]. FcγRs played important regulatory roles in humoral immune response. Single nucleotide polymorphisms (SNPs) of FCGR2A and FCGR3A could affect the expression of CD32 and CD16 [29]. Arce *et al* found that treatment of inflamed tumors could enhance anti-CTLA-4 activity in inflamed tumors by enhancing FcγR binding[30]. HSPA8 was a heat shock homologous protein that bound to nascent peptides to promote proper folding, played an important role in transporting membrane protein-coated vesicles through cell membrane components[31]. HSPA8 mutation played an important role in the detection of lung tumors [32]. PTPRC was a member of the protein tyrosine phosphatase family and played an important role in biological processes such as cell growth, differentiation, metabolism, and immune response etc.[33]. Li *et al* found that 14 target genes such as PTPRC were significantly correlated with survival of patients with lung adenocarcinoma[34]. Ubiquitin was a small molecule protein that was widely present in eukaryotic cells. It could be linked to each other through enzymatic reactions to mediate the degradation of target proteins[35].Tang *et al* found that knock-down of ubiquitin by mixed shRNAs targeting its coding genes ubiquitin B (UBB) and ubiquitin C (UBC) suppressed the growth and increased the radiosensitivity in NSCLCH1299 cells[36]. Song *et al* found that VCP could promote proteasome degradation of polyubiquitinated proteins through its unfolding activity [37]. Bäder *et al* found that ITPKB negatively controlled transmigration of H1299 cells in vitro by blocking Ins (1,4,5) P3-mediated calcium release [38].

D. Conclusion

In summary, we found that individuals infected by coronavirus might go through 5 stages, and 10 HUB genes might play key roles in the pneumonia infected by coronavirus. This discovery might provide some new clues for coronavirus related research.

ACKNOWLEDGMENT

Supported by National Natural Science Foundation of China (No. 31401080).

REFERENCES

[1] Chan JF, Kok KH, Zhu Z, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerg Microbes Infect. 2020;9(1):221‑236. doi:10.1080/22221751.2020.1719902.

[2] Tian X, Li C, Huang A, et al. Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. Emerg Microbes Infect. 2020;9(1):382 - 385. doi:10.1080/22221751.2020.1729069.

[3] Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin [published online ahead of print, 2020 Feb 3]. Nature. 2020;10.1038/s41586-020-2012-7. doi:10.1038/s41586-020-2012-7.

[4] Kirchdoerfer RN, Ward AB. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. Nat Commun. 2019;10(1):2342. Published 2019 May 28. doi:10.1038/s41467-019-10280-3.

[5] Hu W, Yen YT, Singh S, Kao CL, Wu-Hsieh BA. SARS-CoV regulates immune function-related gene expression in human monocytic cells. Viral Immunol. 2012;25(4):277‑288. doi:10.1089/vim.2011.0099.

[6] Wang CY, Lu CY, Li SW, et al. SARS coronavirus papain-like protease up-regulates the collagen expression through non-Samd TGF-β 1 signaling. Virus Res. 2017;235:58 - 66. doi:10.1016/j.virusres.2017.04.008.

[7] Shao H, Lan D, Duan Z, et al. Upregulation of mitochondrial gene expression in PBMC from convalescent SARS patients. J Clin Immunol. 2006;26(6):546‑554. doi:10.1007/s10875-006-9046-y.

[8] Yang Q, Wang R, Wei B, et al. Candidate Biomarkers and Molecular Mechanism Investigation for Glioblastoma Multiforme Utilizing WGCNA. Biomed Res Int. 2018;2018:4246703. Published 2018 Sep 26. doi:10.1155/2018/4246703.

[9] Tian H, Guan D, Li J. Identifying osteosarcoma metastasis associated genes by weighted gene co-expression network analysis (WGCNA). Medicine (Baltimore). 2018;97(24):e10781. doi:10.1097/MD.0000000000010781.

[10] Mei-Lee Hwang, Yu-Da Lin, Li-Yeh Chuang, and Cheng-Hong Yang, "Determination of the SNP-SNP Interaction between Breast Cancer Related Genes to Analyze the Disease Susceptibility," International Journal of Machine Learning and Computing vol. 4, no. 5, pp. 468-473, 2014.

[11] Nalini Singh, Ambarish G Mohapatra, Biranchi Narayan Rath, and Guru Kalyan Kanungo, "GUI Based Automatic Breast Cancer Mass and Calcification Detection in Mammogram Images using K-means and

Fuzzy C-means Methods," International Journal of Machine Learning and Computing vol. 2, no. 1, pp. 7-12, 2012.

[12] Wu D, Zhou M, Li L, et al. Time series analysis of gene changes and processes after burn with human gene expression profiles. Burns. 2019;45(2):387‑397. doi:10.1016/j.burns.2018.08.022.

[13] Park C, Yun SJ, Ryu SJ, et al. Systematic identification of an integrative network module during senescence from time-series gene expression. BMC Syst Biol. 2017;11(1):36. Published 2017 Mar 15. doi:10.1186/s12918-017-0417-1.

[14] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. doi:10.1186/s13059-014-0550-8.

[15] Silvela J, Portillo J. Breadth-first search and its application to image processing problems. IEEE Trans Image Process. 2001;10(8):1194–1199. doi:10.1109/83.935035.

[16] Szklarczyk D, Franceschini A, Kuhn M, et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res. 2011;39(Database issue):D561‑D568. doi:10.1093/nar/gkq973.

[17] Brin S, Page L. The anatomy of a large-scale hypertextual web search engine(J). Computer networks and ISDN systems, 1998, 30(1-7): 107-117.

[18] Piñero J, Bravo À, Queralt-Rosinach N, et al. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. Nucleic Acids Res. 2017;45(D1):D833‑D839. doi:10.1093/nar/gkw943.

[19] Chang YM, Lin HH, Liu WY, et al. Comparative transcriptomics method to infer gene coexpression networks and its applications to maize and rice leaf transcriptomes. Proc Natl Acad Sci U S A. 2019;116(8):3091‑3099. doi:10.1073/pnas.1817621116.

[20] Báez-Santos YM, St John SE, Mesecar AD. The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. Antiviral Res. 2015;115:21 ‑ 38. doi:10.1016/j.antiviral.2014.12.015.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[21] Chen IY, Moriyama M, Chang MF, Ichinohe T. Severe Acute Respiratory Syndrome Coronavirus Viroporin 3a Activates the NLRP3 Inflammasome. Front Microbiol. 2019;10:50. Published 2019 Jan 29. doi:10.3389/fmicb.2019.00050.

[22] Zu J, Gu Y, Li Y, et al. Topological evolution of coexpression networks by new gene integration maintains the hierarchical and modular structures in human ancestors. Sci China Life Sci. 2019;62(4):594‑608. doi:10.1007/s11427-019-9483-6.

[23] Balasuriya N, McKenna M, Liu X, Li SSC, O'Donoghue P. Phosphorylation-Dependent Inhibition of Akt1. Genes (Basel). 2018;9(9):450. Published 2018 Sep 7. doi:10.3390/genes9090450.

[24] Larson-Casey JL, Deshane JS, Ryan AJ, Thannickal VJ, Carter AB. Macrophage Akt1 Kinase-Mediated Mitophagy Modulates Apoptosis Resistance and Pulmonary Fibrosis. Immunity. 2016;44(3):582 ‑ 596. doi:10.1016/j.immuni.2016.01.001.

[25] Yamashita M, Saito R, Yasuhira S, et al. Distinct Profiles of CD163-Positive Macrophages in Idiopathic Interstitial Pneumonias. J Immunol Res. 2018;2018:1436236. Published 2018 Feb 4. doi:10.1155/2018/1436236.

[26] Chistiakov DA, Killingsworth MC, Myasoedova VA, Orekhov AN, Bobryshev YV. CD68/macrosialin: not just a histochemical marker. Lab Invest. 2017;97(1):4‑13. doi:10.1038/labinvest.2016.116.

[27] Doherty DF, Nath S, Poon J, et al. Protein Phosphatase 2A Reduces Cigarette Smoke-induced Cathepsin S and Loss of Lung Function. Am J Respir Crit Care Med. 2019;200(1):51‑62. doi:10.1164/rccm.201808-1518OC.

[28] Weldon S, McNally P, McAuley DF, et al. miR-31 dysregulation in cystic fibrosis airways contributes to increased pulmonary cathepsin S production. Am J Respir Crit Care Med. 2014;190(2):165 ‑ 174. doi:10.1164/rccm.201311-1986OC.

[29] Paul P, Pedini P, Lyonnet L, et al. FCGR3A and FCGR2A Genotypes Differentially Impact Allograft Rejection and Patients' Survival After Lung Transplant. Front Immunol. 2019;10:1208. Published 2019 Jun 12. doi:10.3389/fimmu.2019.01208.

[30] Arce Vargas F, Furness AJS, Litchfield K, et al. Fc Effector Function Contributes to the Activity of Human Anti-CTLA-4 Antibodies. Cancer Cell. 2018;33(4):649‑663.e4. doi:10.1016/j.ccell.2018.02.010.

[31] Lu Z, Song Q, Yang J, et al. Comparative proteomic analysis of anti-cancer mechanism by periplocin treatment in lung cancer cells. Cell Physiol Biochem. 2014;33(3):859‑868. doi:10.1159/000358658.

[32] Rusin M, Zientek H, Krześniak M, et al. Intronic polymorphism (1541-1542delGT) of the constitutive heat shock protein 70 gene has functional significance and shows evidence of association with lung cancer risk. Mol Carcinog. 2004;39(3):155‑163. doi:10.1002/mc.20009.

[33] Higgs R. PTPRC mutation associated with response to anti-tNF therapy in rheumatoid arthritis. Nat Rev Rheumatol. 2010;6(6):311. doi:10.1038/nrrheum.2010.69.

[34] Li L, Peng M, Xue W, et al. Integrated analysis of dysregulated long non-coding RNAs/microRNAs/mRNAs in metastasis of lung adenocarcinoma. J Transl Med. 2018;16(1):372. Published 2018 Dec 27. doi:10.1186/s12967-018-1732-z.

[35] Grumati P, Dikic I. Ubiquitin signaling and autophagy. J Biol Chem. 2018;293(15):5404–5413. doi:10.1074/jbc.TM117.000117.

[36] Tang Y, Geng Y, Luo J, Shen W, Zhu W, Meng C, Li M, Zhou X, Zhang S, Cao J.Downregulation of ubiquitin inhibits the proliferation and radioresistance ofnon-small cell lung cancer cells in vitro and in vivo. Sci Rep. 2015 Mar 30;5:9476. doi: 10.1038/srep09476. PubMed PMID: 25820571.

[37] Song C, Wang Q, Song C, Rogers TJ. Valosin-containing protein (VCP/p97) is capable of unfolding polyubiquitinated proteins through its ATPase domains. Biochem Biophys Res Commun. 2015;463(3):453‑457. doi:10.1016/j.bbrc.2015.05.111.

[38] Bäder S, Glaubke E, Grüb S, et al. Effect of the actin- and calcium-regulating activities of ITPKB on the metastatic potential of lung cancer cells. Biochem J. 2018;475(12):2057‑2071. Published 2018 Jun 26. doi:10.1042/BCJ20180238.