European Review for Medical and Pharmacological Sciences 2020; 24: 7497-7505

CXCL10 an important chemokine associated with cytokine storm in COVID-19 infected patients

N. ZHANG, Y.-D. ZHAO, X.-M. WANG

Department of Pulmonary Medicine, First Affiliated Hospital of Dalian Medical University, Dalian, China

Na Zhang and Yadong Zhao contributed equally to this work

Abstract. – OBJECTIVE: The specific mechanism of cytokine storm in COVID-19 infected patients is not clear. This study aims to identify the key genes that cause cytokine storm in COVID-19 infected patients.

MATERIALS AND METHODS: We conducted a difference analysis on the GSE147507 data set. The analysis results are combined with immune genes to obtain immune-related genes among the differential genes. Finally, GO enrichment analysis, PPI analysis, core gene identification, and ssGSEA enrichment analysis were performed on the new gene set.

RESULTS: A total of 232 differential genes were screened out. After merging with immune genes, a total of 29 immune-related genes were obtained. Further analysis revealed that the genes were enriched in 16 pathways, and the protein interaction network had a total of 29 nodes and 139 edges. After screening, the core gene was CXCL10. The ssGSEA results of CXCL10 showed that CD4 and CD8 immune-related signature were significantly enriched in high CXCL10 expression, and the samples with low CXCL10 expression were significantly enriched with monocytes and DC immune-related signature.

CONCLUSIONS: CXCL10 may be a key gene related to the cytokine storm of COVID-19 infection, and it is expected to become the therapeutic target.

Key Words:
CXCL10, COVID-19, Infection, Therapeutic target, Cytokine storm.

Acronyms

COVID-19, Coronavirus Disease 2019; GO, Gene oncology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GEO, Gene Expression Omnibus; CXCL10, Chemokine (C-X-C motif) ligand 10.

Introduction

The Coronavirus Disease 2019 (COVID-19) is an emerging infectious disease caused by severe acute respiratory syndrome Coronavirus 2 (SARS-CoV-2), which broke out at the end of December 2019 in Wuhan, China, is highly contagious and spreads quickly and widely, posing a major threat to world public health security[1]. At present, COVID-19 patients currently take supportive treatment, and were mainly administered hormone drugs to inhibit the inflammatory response induced by virus and cytokines[2-3]; thus, there are no specific antiviral drugs or vaccine against COVID19 infection for therapy of humans.

When external pathogens invade, the body will immediately release cytokines (small molecule proteins) to mediate the immune response to protect itself. However, under the immune dysregulation state, patients will re-lease cytokines explosively, forming a cytokine storm, resulting in systemic inflammatory response[4-5]. There is increasing evidence that cytokine storms are widespread in patients with severe infections, causing diffuse damage to pulmonary capillary endothelial cells and alveolar epithelial cells, which in turn leads to the occurrence of ARDS, which worsens the condition in the short term, with respiratory failure, and eventually multiple organ failure[6-8]. The appearance of cytokine storm marks the critical condition and poor prognosis of patients with COVID-19. Inhibiting cytokine storm is also an important starting point for treatment[9]. However, the wide variety of cytokines and intricate interrelationships restrict their drug development and therapeutic practice. If the key genes that cause the cytokine storm cascade reaction in COVID-19 infected persons are used as clues, the develop-

Corresponding Author: Ximing Wang, MM; e-mail: wangximingdalian@126.com

ment of new drugs or corresponding treatments will be beneficial to the clinical benefit of the patients and the control of the epidemic. However, there is no relevant research so far.

In this study, we performed bioinformatics analysis on the COVID-19 sequencing data uploaded and shared in the GEO database, aiming to reveal the expression changes of immune genes in the body after COVID-19 infection, and try to identify the key chemokines that cause the cytokine storm cascade in the body.

Materials and Methods

Data Source and Difference Analysis

Our data comes from the GEO database, a free data sharing platform. The analysis data set used is GSE147507, which contains 3 MOCK-processed COVID-19 samples and 3 COVID-19-infected samples. The GSE5972 data set was used to verify the expression of core genes. Differential genes (DEGs) are screened through the limma package, and the screening conditions set are $p<0.05$ and logFC>1.

Acquisition of Immune-Related Genes in DEGs

The immune gene set was downloaded from the immport website (https://www.immport.org/) to the local area and intersected with the DEGs of the GSE147507 sample. This process was implemented using R package VennDiagram.

Functional Enrichment Analysis of Immune-Related Genes

Functional analysis is an important part of studying the molecular mechanism of action. In order to explore the possible biological functions of differential immune-related genes in the COVID-19 infection group, we used Metascape (metascape.org/) to perform functional enrichment analysis on differential genes and set the condition to Min overlap>3, p-value<0.01, Min Enrichment>1.5.

Analysis of Protein-Protein Interaction of Immune-Related Genes

As a classic protein interaction database, the String database is often used by us to predict the interaction relationship of immune-related proteins. The prediction result file is uploaded to Cytoscape for subsequent core gene screening and visual analysis. In addition, we also performed K-means clustering analysis on the predicted network.

Core Gene Expression Verification and ssGSEA Enrichment Analysis

GSE5972, a data set containing 10 normal samples and 55 SARS samples, was used to verify the expression of core genes. ssGSEA enrichment analysis was used to explain the causes of differences in phenotypes of different samples.

Statistical Analysis

All data analysis is performed under the operation of R software. The t-test was used to compare differences between groups. All tests are two-tailed tests, $p<0.05$ is defined as statistically significant.

Results

Acquisition of DEGs and Immune-Related Genes in DEGs

The data set GSE147507 containing 3 MOCK-processed COVID-19 samples and 3 COVID-19-infected samples, was downloaded from the GEO database. The unstandardized data baseline is not balanced. After standardized preprocessing, it can be seen that the sample baseline has been balanced **(Figure S1)**. The standardized data is used for subsequent analysis. After the difference analysis, a total of 232 differential genes (DEGs) were screened, including 147 upregulated genes and 85 downregulated genes. Top 20 DEGs are shown in Table I. In order to display the distribution of differential genes more directly, we visualized the results of differential genes. As shown in Figure 1, red represents upregulation genes, green represents downregulated genes and gray represents no significantly difference genes. The immport website (https://www.immport.org/) provides all the confirmed immune gene sets. This immune gene set contains a total of 1811 immune-related genes. When we merged the differential gene data set and the immune-related data set, 29 differential immune-related genes were extracted from the DEGs (Figure 2).

Functional Enrichment Analysis of Immune-Related Genes in DEGs

Functional enrichment analysis includes Biological Process (BP), Cellular Component (CC) and Molecular Function (MF). Functional enrichment results showed that 29 immune-related genes were enriched to fifteen BP and one MF (Table II).

Table I. The top 20 upregulated and downregulated differential genes between MOCK and CoV groups.

	Symbol	logFC	AveExpr	t	*p*-value	B
UP	MX1	4.934150931	7.067095184	29.68419913	3.16E-06	5.230224663
	IFIT1	4.276905134	7.986125918	29.25855331	3.36E-06	5.190362911
	IFI6	4.276611378	9.13727315	33.56158562	1.85E-06	5.546804633
	IFI44	4.087984454	2.866980071	7.929755425	0.000922535	-0.095188279
	IFITM1	4.062213604	3.196868257	7.240068201	0.001342683	-0.522108799
	XAF1	3.961624147	1.980812073	10.23659076	0.000315949	1.098177328
	OAS2	3.769387759	3.559895612	18.7614896	2.33E-05	3.708635153
	ISG15	3.748128644	7.94783049	45.76739558	4.75E-07	6.174050423
	IFI44L	3.685597769	1.842798884	20.56056479	1.57E-05	4.051827246
	IFI27	3.362681164	5.083325474	9.164643808	0.000504037	0.583489096
	IRF7	3.170227481	7.725095617	20.43915333	1.61E-05	4.030172584
	OASL	2.530905477	5.030765605	7.633536744	0.001079922	-0.273928371
	BST2	2.444994352	2.77714752	4.115145408	0.012117336	-3.0630533
	EPSTI1	2.407254637	1.425849541	11.20633542	0.000215026	1.514317337
	HPCA	2.260092106	1.130046053	12.22476095	0.000148273	1.908260412
	DDX60	2.25794291	7.615367862	20.7537095	1.50E-05	4.08585633
	SERPINB3	2.249235472	2.064885097	4.61625227	0.007932017	-2.572613013
	CXCL10	2.154153739	1.299299092	7.941926764	0.000916686	-0.087986203
	IRF9	2.120830178	9.214124621	37.76674675	1.10E-06	5.813939374
	SNORD19	2.091806198	1.268125321	8.292064906	0.000766196	0.114590458
DOWN	EFCC1	-1.427663984	0.880498659	-2.906273809	0.039201298	-4.406429442
	P2RX3	-1.432198082	1.656366402	-2.943143449	0.037692514	-4.362153096
	APOBR	-1.434991526	2.381994992	-3.721482939	0.017324922	-3.475316282
	TRABD2B	-1.435647521	3.361274995	-6.674309045	0.001871157	-0.902340897
	GSDMC	-1.462419789	0.897876561	-5.407530891	0.004327262	-1.870250536
	ZAR1	-1.463076757	1.593192546	-3.555896569	0.020276565	-3.656045185
	NRN1L	-1.498288009	2.119242788	-4.219865092	0.011059456	-2.957445537
	ANKK1	-1.50009045	0.972267447	-4.039301263	0.012958636	-3.140595063
	C1QTNF9B-AS1	-1.503362322	1.228623522	-4.368484927	0.009739853	-2.810425504
	PCDHGB7	-1.503362322	1.228623522	-4.368484927	0.009739853	-2.810425504
	FAM65C	-1.524162072	1.51162843	-2.743267521	0.046753041	-4.604383371
	TDRD12	-1.544943501	0.939138417	-4.611367331	0.007963583	-2.577215391
	LINC00696	-1.554650344	0.777325172	-2.701418312	0.048949999	-4.655755137
	MIR1256	-1.554650344	0.777325172	-2.701418312	0.048949999	-4.655755137
	LOC100507557	-1.565104605	1.897930947	-4.260315529	0.010680353	-2.917100234
	SNORD18B	-1.578098387	1.219876277	-2.94568121	0.037591132	-4.359112532
	LOC158435	-1.625566341	1.511947754	-13.52393046	9.61E-05	2.356194083
	TREH	-1.634916169	1.672943845	-2.709663379	0.048508066	-4.645617022
	PRAP1	-1.861405924	1.944410773	-3.005910347	0.035274322	-4.287219036
	RASGEF1B	-1.878295157	1.369974662	-3.883968292	0.014906887	-3.302202152

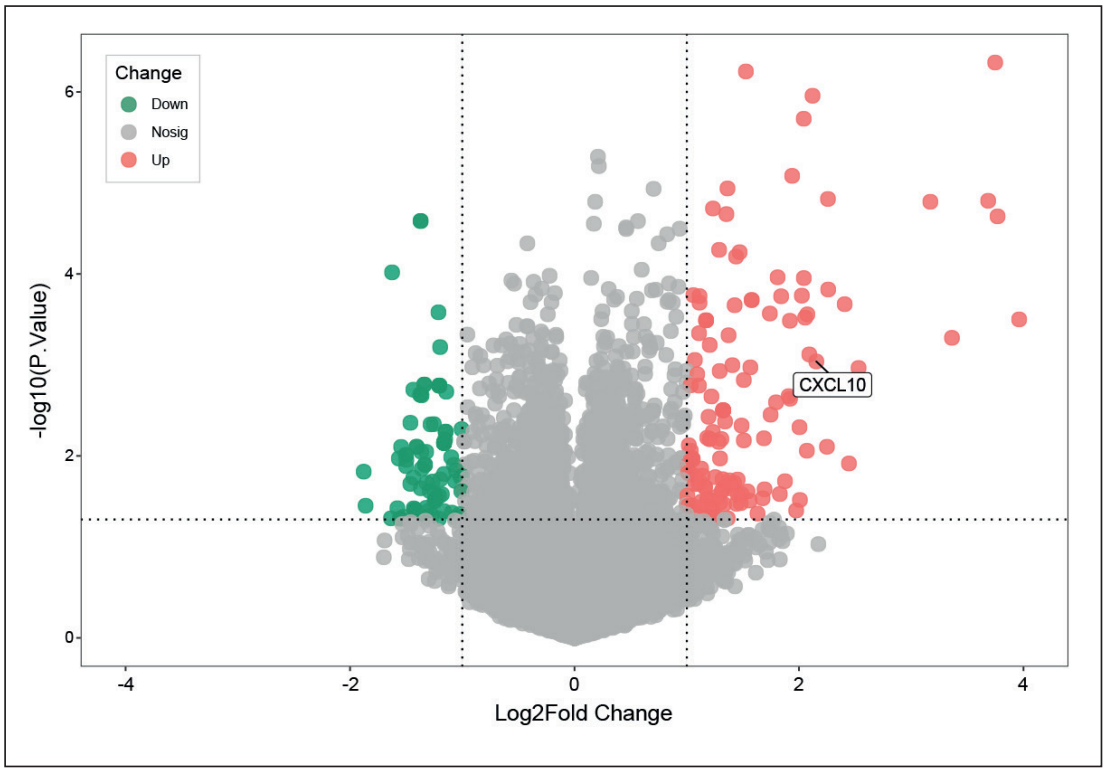

Figure 1. Volcano of differential genes, red represents upregulated genes, green represents downregulated genes.

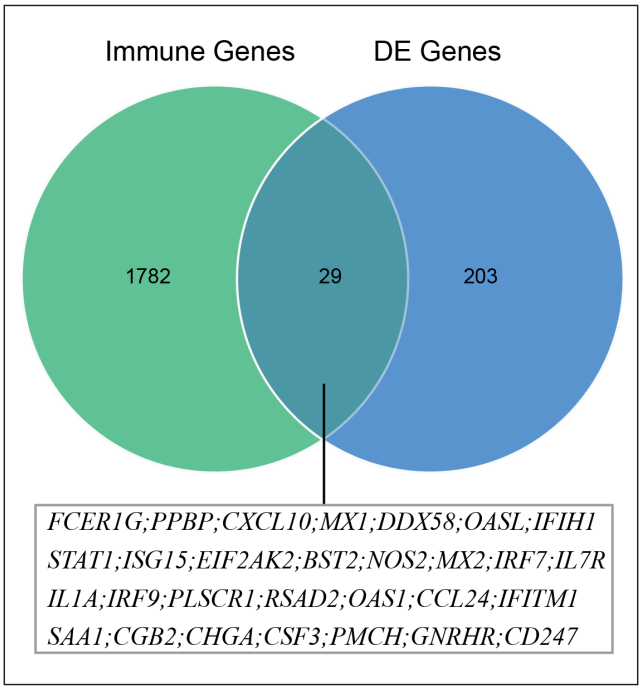

Figure 2. Screening of 29 immune-related genes, green represents immune-related gene set, blue represents differential gene set.

Table II. Enrichment of immune-related genes function in DEGs.

Category	Term ID	Description	LogP_val	Ratio
[a]BP	GO: 0098542	Defense response to other organism	-25.381	20/596
BP	GO: 0019221	Cytokine-mediated signaling pathway	-22.864	20/796
BP	GO: 0034341	Response to interferon-gamma	-11.839	9/202
BP	GO: 0035455	Response to interferon-alpha	-8.005	4/21
BP	GO: 1903555	Regulation of tumor necrosis factor superfamily cytokine production	-10.595	4/172
BP	GO: 1990869	Cellular response to chemokine	-7.166	3/97
BP	GO: 0031623	Receptor internalization	-3.466	3/114
BP	GO: 0002705	Positive regulation of leukocyte mediated immunity	-6.771	3/134
BP	GO: 0002262	Myeloid cell homeostasis	-4.681	3/148
BP	GO: 0090066	Regulation of anatomical structure size	-2.498	4/515
BP	GO: 0003013	Circulatory system process	-3.718	4/554
BP	GO: 0007249	I-kappaB kinase/NF-kappaB signaling	-7.288	3/279
BP	GO: 0009611	Response to wounding	-4.614	4/691
BP	GO: 0002366	Leukocyte activation involved in immune response	-5.79	4/711
bMF	GO: 0003725	Double-stranded RNA binding	-7.549	5/75

[a]Represents biological processes; [b]Represents molecular functions.

The four most significant functions of enrichment are: Defense response to other organism, Cytokine-mediated signaling pathway, Response to interferon-gamma, and Response to interferon-alpha (Figure 3).

PPI Analysis and Clustering of Immune-Related Genes in DEGs

In order to clarify the potential interaction between 29 immune-related genes, we used string database for PPI analysis. The results show that the

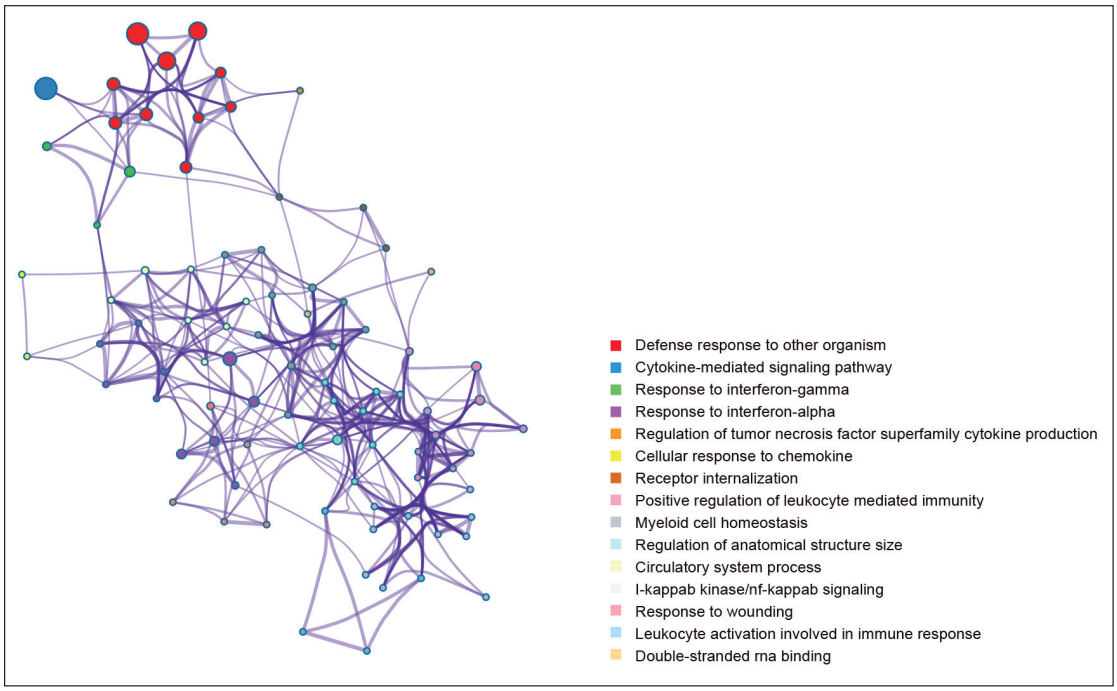

Figure 3. Functional enrichment analysis of 29 immune-related genes.

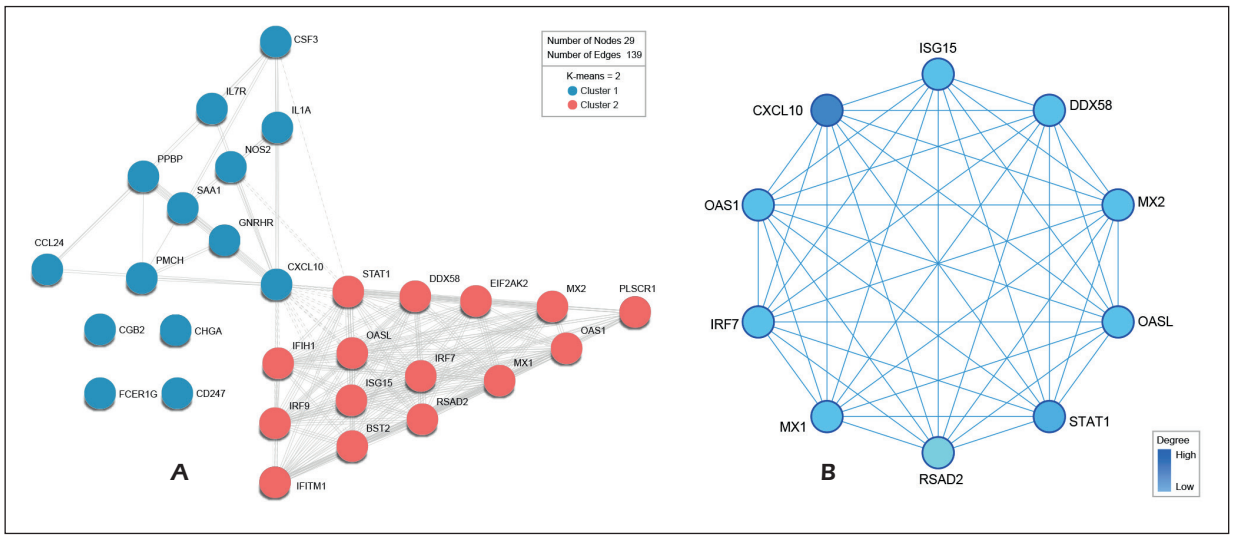

Figure 4. The protein interaction network of 29 immune-related genes (**A**) and the 10 core genes in the net-work (**B**).

protein interaction network has a total of 29 nodes and 139 edges. K-means clustering results show that the protein-protein interaction (PPI) network is clustered into two subgroups (Figure 4A).

Identification and Verification of Core Immune-Related Genes

Cytohubba is often used to screen the core genes of the network. We use the degree method to identify the core modules of the PPI network. A total of 10 core genes were identified, which are CXCL10, ISG15, DDX58, MX2, OASL, STAT1, RSAD2, MX1, IRF7, OAS1 (Figure 4B). This ten genes were used for subsequent analysis. In view of results of k-means cluster analysis, we can see that CXCL10 also serves as the core gene linking the two subgroups, so we finally determined CXCL10 as the core gene of our research. To further verify its expression stability, we used a GSE5972 data set containing 10 normal samples and 55 SARS samples for verification. The results showed that compared with normal samples, patients with SARS infection had high CXCL10 expression, and there was statistical significance between the groups (p<0.001) (Figure 5).

Analysis of ssGSEA Based on Core Gene CXCL10

In order to identify immune-related signature related to CXCL10, we used ssGSEA to analyze the GSE147507 data set, and stratified risk according to the median expression of CXCL10. As shown in Figure 6, immune-related signature related to CD4 and CD8 were significantly enriched in high-risk patients, while some immune-related signature related to Monocytes and Dendritic Cells (DC) were significantly enriched in low-risk groups. Volcano of GSVA analysis for CXCL10 is showed in **Figure S2**.

Discussion

In this study, we first revealed the expression changes of immune-related genes in the body after COVID-19 infection, and tried to identify the key chemokine CXCL10 that causes the cytokine storm.

Differential gene analysis, as a mature and reliable bioinformatics algorithm, has made gratifying achievements in tapping the potential biomarker of tumor treatment. Liu et al[10] successfully analyzed the breast cancer-related biomarker by analyzing the difference between breast cancer and para-cancer samples in the TCGA database, and constructed the signature of breast cancer prognosis. In our research, we also used this algorithm to conduct a differential analysis of the gene set expression, and obtained 29 immune-related differential genes. The external verification set suggest that core gene CXCL10 is indeed highly expressed in infected patients. This conclusion further confirms that the biomarker gene we are looking for has certain credibility and application value.

The cytokine storm was first proposed in 1993 as the pathogenesis of graft-versus-host disease (GVHD)[11]. Virus invasion leads to an imbalance of the body's immune system, which in turn induces an inflammatory storm, leading to multifunctional organ failure.

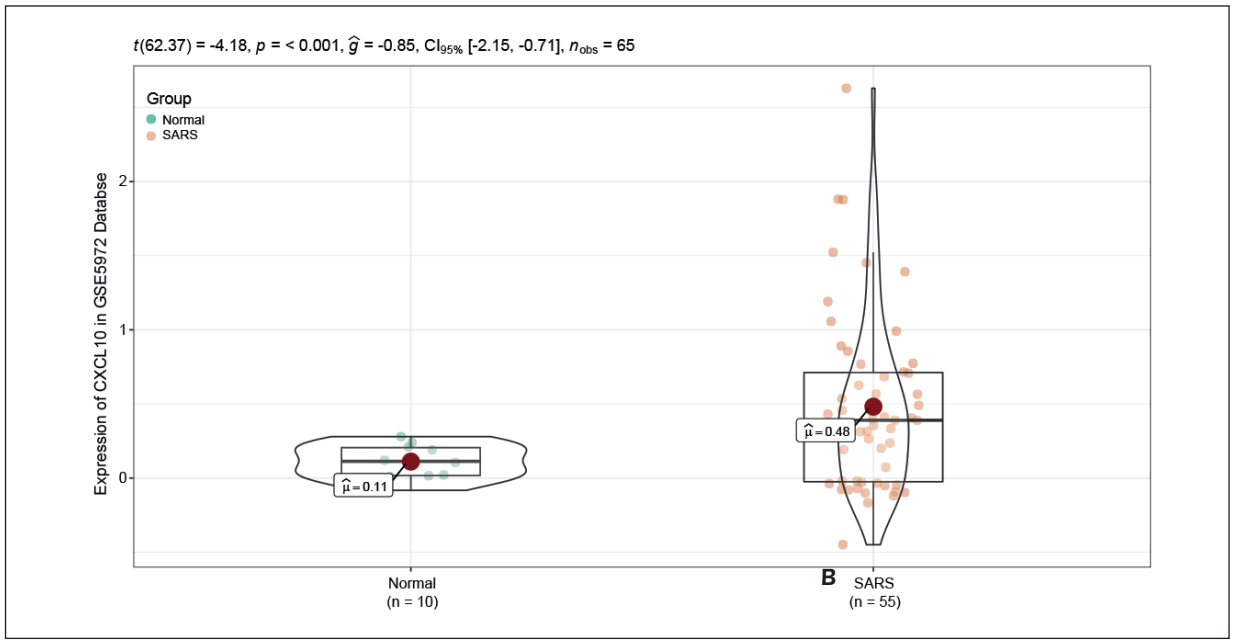

Figure 5. Verification of CXCL10 expression in SARS group and normal group, green represents normal group, pink represents SARS group.

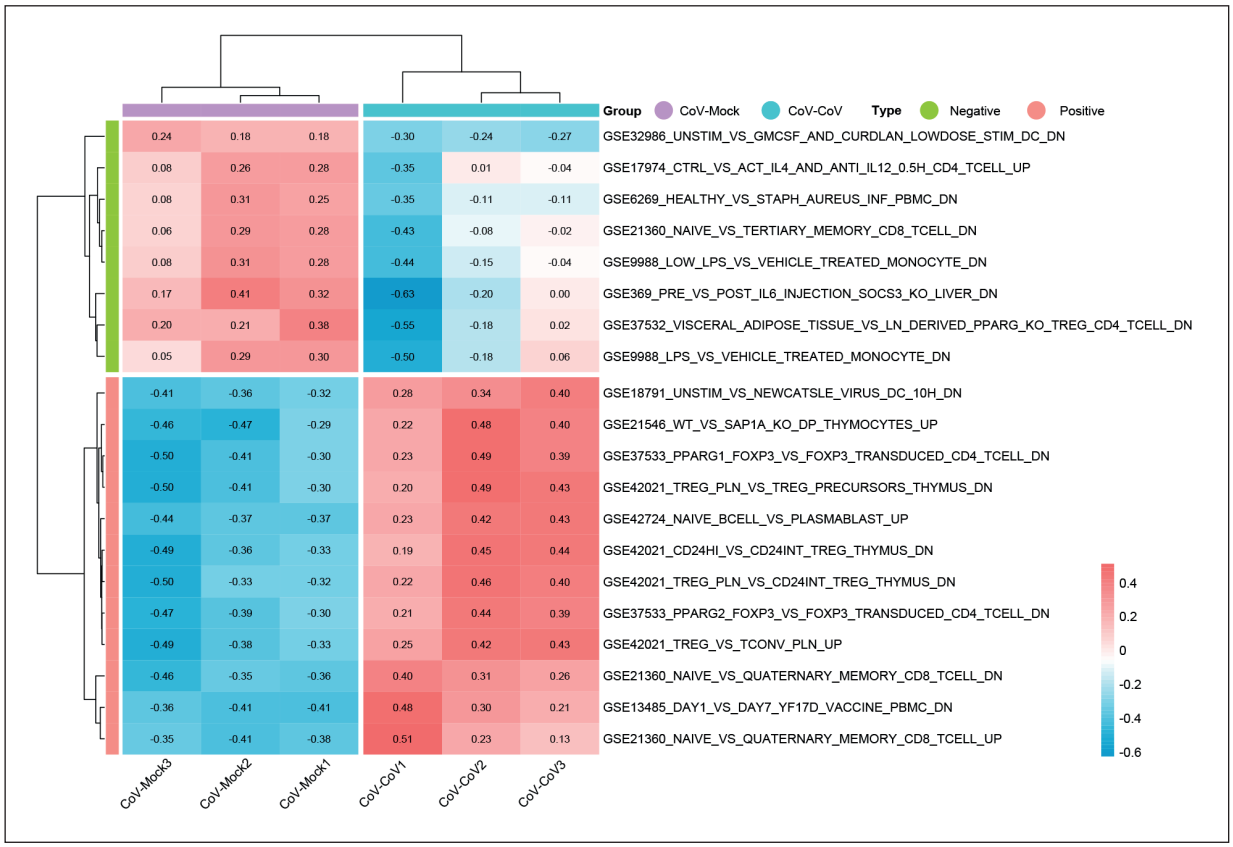

Figure 6. Heatmap of GSVA analysis for CXCL10, Red means high enrichment score, blue means low en-richment score.

It is the main cause of death (SARS-CoV) with similar clinical manifestations and pathological characteristics as COVID-19[12]. At that time, glucocorticoid shock therapy for cytokine storms played an important role in disease control[13]. Huang et al[7] found that plasma inflammatory factors IL-1β, IFN-γ, IFN-γ inducible protein 10 (IP-10) and MCP-1 in patients with 2019-nCoV infection were significantly increased, which may lead to activated Th1 response. Further studies showed that ICU patients have higher concentrations of G-CSF, IP-10, MCP-1, macrophage inflammatory protein 1α (MIP1α) and TNF-α than non-ICU patients, and also pointed out that cytokine storm and the severity of disease in patients with 2019-nCoV infection is related. With regard to the suppression of inflammatory factor storms, existing research is limited to the suppression of individual cytokines, and there is no effective control from the overall level. The use of glucocorticoids by COVID-19 as a classic drug to suppress the storm of inflammatory factors remains controversial as to whether it is used in patients with severe COV-19[14]. Hydroxychloroquine has become one of the important drugs for the treatment of COVID-19 by inhibiting systemic inflammatory response[15]. The current treatment for cytokine storms, although delaying the progression of the disease to a certain extent, is not ideal. Therefore, how to find the critical genes related to COVID-19 from the complex cytokine storm is the key to inhibit the development of the disease.

Through the GEO database, we found that CXCL10 may be the core gene that suppresses the COVID-19 inflammatory factor storm. Recombinant Rat IP-10/CXCL10 is a chemokine superfamily CXC that was selected from the cDNA gene library by hybridization after stimulating lymphocyte U937 strain with IFN-γ in 1983 by Kumar et al[16]. The ELR-subfamily, also known as CXC chemokine ligand 10 (CXCL10), is a type of protein that can be induced by interferon to chemoattract lymphocytes. It has been confirmed that CXCL10 has many biological functions, such as chemotaxis monocytes, activated T cells and NK cells for expansion[17-19]. This conclusion suggests that it has potential application value in anti-infection and treatment of immune diseases. Our research also shows that CXCL10 has important value in the immunotherapy of the new crown. Because after cluster analysis of the differential immune gene protein interaction network, it was found that CXCL10 is not the only core gene in PPI but also the hinge gene linking the two clustering modules. Based on the above conclusions, we further speculate that CXCL10 and its receptor antagonists may be an important target for the treatment of inflammatory factor storms.

Gene Set Variation Analysis, called Gene Set Variation Analysis, is used to study the differences of the gene sets of interest between different samples, so as to explain the causes of phenotypic differences from the perspective of bioinformatics. Recognizing the mechanism of immune inflammation in neo-pneumonia is an important prerequisite for treatment. CXCL10 is an important pivotal gene. We speculate that this gene is likely to play an important role in the immune development of patients with new coronary pneumonia. The results of CXCL10 gene ssGSEA suggest that CD4 and CD8 are highly enriched in samples with high expression of CXCL10, while monocytes and Dendritic Cells (DC) in samples with low expression of CXCL10 are more enriched in dendritic cells. This result can be used to understand the pathological process and the molecular mechanism of disease progression provides a certain reference.

This report provides a new direction on how to suppress the cytokine storm and curb the development of COVID-19 disease. Starting from this point, new drug development and treatment will benefit patients' clinical benefit and epidemic control. Although this is the first study on the genetic level of the COVID-19 cytokine storm, this article still has the following limitations. Only three biological replicates of the new crown sample analysis were performed, and later, a larger sample size was required for more in-depth analysis. In addition, because the GEO database lacks sequencing data of other new crown samples, we used SARS samples when verifying the expression of CXCL10. This will inevitably lead to a certain bias in the verification results, and we will verify it again after the database update later.

Conclusions

The expression of CXCL10 increases significantly after COVID-19 infection in the body, which may be the key chemokine that causes the cytokine storm in the body.

Acknowledgements

The authors would like to thank the GEO website for providing original data in the present study.

Availability of Data and Materials

Original data could be obtained from the GEO website (https://www.ncbi.nlm.nih.gov/geo/).

Authors' contributions

ZN and ZY contributed to the data analysis and wrote the main manuscript. WX contributed to the process of revising the manuscript and designed the experiments, reviewed drafts of the paper and approved the final draft. All authors read and approved the manuscript and agree to be accountable for all aspects of the research in ensuring that the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflict of Interests

The authors declare that they have no conflict of interests.

References

1) Velavan TP, Meyer CG. The COVID-19 epidemic. Trop Med Int Health 2020; 25: 278-280.
2) Weston S, Frieman MB. COVID-19: knowns, unknowns, and questions. mSphere 2020; 5.
3) Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. J Autoimmun 2020; 109: 102433.
4) Chousterman BG, Swirski FK, Weber GF. Cytokine storm and sepsis disease pathogenesis. Semin Immunopathol 2017; 39: 517-528.
5) Tisoncik JR, Korth MJ, Simmons CP, Farrar J, Martin TR, Katze MG. Into the eye of the cytokine storm. Microbiol Mol Biol Rev 2012; 76: 16-32.
6) Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, Yuan ML, Zhang YL, Dai FH, Liu Y, Wang QM, Zheng JJ, Xu L, Holmes EC, Zhang YZ. A new coronavirus associated with human respiratory disease in China. Nature 2020; 579: 265-269.
7) Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020; 395: 497-506.
8) Li G, Fan Y, Lai Y, Han T, Li Z, Zhou P, Pan P, Wang W, Hu D, Liu X, Zhang Q, Wu J. Coronavirus infections and immune responses. J Med Virol 2020; 92: 424-432.
9) Pedersen SF, Ho YC. SARS-CoV-2: a storm is raging. J Clin Invest 2020.
10) Liu J, Li P, Wang R, Li J, Zhang M, Song Z, Liu P. High expression of DLG3 is associated with decreased survival from breast cancer. Clin Exp Pharmacol Physiol 2019; 46: 937-943.
11) Ferrara JL. Cytokine dysregulation as a mechanism of graft versus host disease. Curr Opin Immunol 1993; 5: 794-799.
12) Zhang J, Lin G, Zeng J, Lin J, Tian J, Li G. Challenges of SARS-CoV-2 and lessons learnt from SARS in Guangdong Province, China. J Clin Virol 2020; 126: 104341.
13) Chen RC, Tang XP, Tan SY, Liang BL, Wan ZY, Fang JQ, Zhong N. Treatment of severe acute respiratory syndrome with glucosteroids: the Guangzhou experience. Chest 2006; 129: 1441-1452.
14) Li T, Lu H, Zhang W. Clinical observation and management of COVID-19 patients. Emerg Microbes Infect 2020; 9: 687-690.
15) Gautret P, Lagier JC, Parola P, Hoang VT, Meddeb L, Mailhe M, Doudier B, Courjon J, Giordanengo V, Vieira VE, Dupont HT, Honoré S, Colson P, Chabrière E, La Scola B, Rolain JM, Brouqui P, Raoult D. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. Int J Antimicrob Agents 2020: 105949.
16) Kumar V, Lust J, Gifaldi A, Bennett M, Sonnenfeld G. Lack of correlation between mycoplasma induced IFN-gamma production in vitro and natural killer cell activity against FLD-3 cells. Immunobiology 1983; 165: 445-458.
17) Karin N, Razon H. Chemokines beyond chemo-attraction: CXCL10 and its significant role in cancer and autoimmunity. Cytokine 2018; 109: 24-28.
18) Karin N. Chemokines and cancer: new immune checkpoints for cancer therapy. Curr Opin Immunol 2018; 51: 140-145.
19) Koper OM, KamiÐska J, Sawicki K, Kemona H. CXCL9, CXCL10, CXCL11, and their receptor (CXCR3) in neuroinflammation and neurodegeneration. Adv Clin Exp Med 2018; 27: 849-856.