ISSN 2618-1584 Vol. 3, Issue 6: 1-5, DOI: 10.25236/FMSR.2021.030601

Differential gene expression in B lymphocyte immune process induced by COVID-19 infection

Kehan Lin[1,a], Hengxi Liu[2,b], Junying Li[3,c], Chenhao Shou[4,d], Yihan Xia[5,e], Ming Chu[6,f*], Xi Chen[7,g]

[1]*Nanjing Medical University, Nanjing, Jiangsu, China*
[2]*University of Washington--Seattle, Seattle, WA, United States*
[3]*The Barstow School-Ningbo Campus, Ningbo, Zhejiang, China*
[4]*Hangzhou Foreign Languages School, Hangzhou, Zhejiang, China*
[5]*Shanghai Pinghe Bilingual School, Shanghai, China*
[6,7]*Department of Immunology, School of Basic Medical Sciences, Peking University. NHC Key Laboratory of Medical Immunology (Peking University). Beijing, China.*
[a]*lilylin@njmu.edu.cn,* [b]*liuhengxi_us@outlook.com,* [c]*3117953878@qq.com,* [d]*2798452103@qq.com,* [e]*xiayihan@shphschool.com,* [f]*famous@pku.edu.cn*
**Corresponding Author: famous@pku.edu.cn*
These authors contributed equally to this work

Abstract: *The Sequencing reads of this study are derived from research shared by Aaron J. Wilk. We conducted a deep analysis of B lymphocyte data. A total of 2126 differentially expressed genes were listed out. Through GO enrichment analysis, 24 clusters related to B lymphocyte functioning were screened out, and genes with significant differential expression were mainly enriched in 10 KEGG signaling pathways related to immune regulatory functions. This suggests that there are differences in the expression of immune regulatory genes in the immune process of B lymphocytes during the Sars-Cov2 infection.*

Keywords: *COVID-19, B lymphocytes, expression difference*

1. Introduction

In December 2019, 6 cases of unexplained pneumonia were discovered in China Wuhan, which’s clinical manifestations were different from the SARS coronavirus in 2003. After investigation, the pathogen was confirmed as a new type of coronavirus, named COVID-19. Recently, the COVID-19 pandemic has become the most serious public health event and had caused severe consequences to society. Through analysis on the genetic sequencing, the COVID-19 virus belongs to the family of coronavirus, the genus of beta coronavirus. It is the third coronavirus found to infect mainly humans. Compared to SARS-CoV and MERS-CoV, there are significant differences in the genes of COVID-19.The disease is highly infectious and has a high mortality rate. The total number of peripheral blood white blood cells is normal or decreased in most cases, and lymphocyte counts can be seen. In the severe cases, peripheral blood lymphocytes decreases progressively (NHC, 2020). The change of B lymphocyte subsets is an important indicator which reflects the body's cellular immune function, and can indicate the severity of the disease, help to evaluate the changes of the disease and the prognosis of the disease.The abnormal gene expression plays an important role in the development, progression and immune process of the disease, but the mechanism of action in the immune process of B lymphocytes caused by the new coronavirus is still unclear. In this study, single-cell RNA sequencing was used to screen out the differentially expressed genes in peripheral blood mononuclear cells (PBMCs) between the COVID-19 patient groups and the healthy control groups. The B cell data was analyzed separately to trace the role of differential genes and reveal from the level of gene expression, the underlying immunological mechanism of B lymphocyte immunity caused by the new coronavirus.

ISSN 2618-1584 Vol. 3, Issue 6: 1-5, DOI: 10.25236/FMSR.2021.030601

2. Material and Methods

2.1 Data sources

In July 2020, Angela J. Rogers and Catherine A. Blish team published a research related to pathways in peripheral immune cells of COVID-19 in Nature Medicine. They collected blood and profiled peripheral blood mononuclear cells (PBMCs) from seven patients with COVID-19 (four of them had acute respiratory distress syndrome) and six healthy participants (Aaron J. Wilk， 2020).The sequencing reads of this research came from Aaron J. Wilk's data sharing, we deeply analyze the data of B lymphocyte.

2.2 Data analysis

The R package Seurat was used for data scaling, transformation, clustering, dimensionality reduction, differential expression analysis and most visualization (Butler A, 2018).First of all, differential expressed genes (DE genes) were calculated by comparing gene expression of COVID-19 patients and healthy donors. Heatmap of DE genes were drawn using R package. For each COVID-19 sample, it was colored by average log(fold-change). Colors of the heatmap are used to indicate the amount of expressing genes, red for up-regulated genes and blue for down-regulated genes.Additionally, we perform enrichment analysis of DE genes, including identifying relevant biological functions with GO function and finding gene enrichment pathway using KEGG analysis. This contributes to further screening and filtering of significant genes in the process of B lymphocyte immunity.

3. Results

3.1 Identification of differentially expressed genes

The identification of differentially expressed genes (DEGs) has conditions that the logarithm of the fold change was one and the adjusted P-value was 0.05. Among a total of 2126 differentially expressed genes, there are 1293 up-regulated genes and 832 down-regulated genes. Figure 1 displays a volcano plot of differential expressed genes.(Figure 1)Fig 1. Volcano plot for differentially expressed genes for Covid-19 patients with healthy people. The horizontal axis represents the fold change of DEGs. The vertical axis represents the significance of DEGs. The blue dots represent differentially down-regulated genes; the red dots represent differentially up-regulated genes.

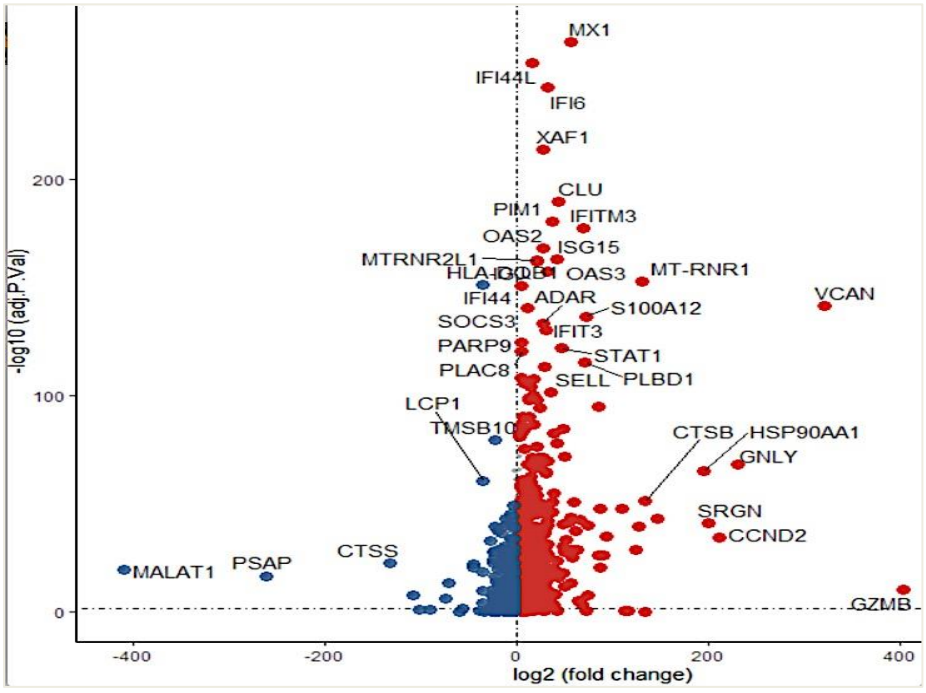

Figure 1. Volcano plot for differentially expressed genes for Covid-19 patients with healthy people

ISSN 2618-1584 Vol. 3, Issue 6: 1-5, DOI: 10.25236/FMSR.2021.030601

3.2 GO Enrichment Analysis

GO enrichment analysis was performed on the differentially expressed genes in the COVID-19 patient group and the healthy control group, and a total of 7263 GO items were obtained. See Figure 2 for THE GO enrichment analysis diagram.Among these items,5906 Biological Process(BP) items mainly involve neutrophil degranulation, neutrophil activation involved in immune response, mRNA catabolic process, RNA catabolic process, translation initiation,cotranslational protein targeting to membrane,viral gene expression and SRP-dependent contranslational protein targeting to membrane;624 Cellular Component(CC) items mainly involves cytosolic ribosome, focal adhesion, cell-substrate junction, secretory granule lumen, cytoplasmic vesicle lumen, vesicle lumen, cytosolic large ribosomal subunit, ribosomal subunit;733 Molecular Function(MF) items mainly involves cadherin binding, structural constituent of ribosome, ubiquitin-like protein ligase binding, ubiquitin protein ligase binding, translation factor activity RNA binding, tranlation regulator activity, translation initiation factor activity, double stranded RNA binding.Figure.2 The GO enrichment analysis diagram. The abscissa is the enrichment multiple, the ordinate is the item, and the color represents the PADJ value

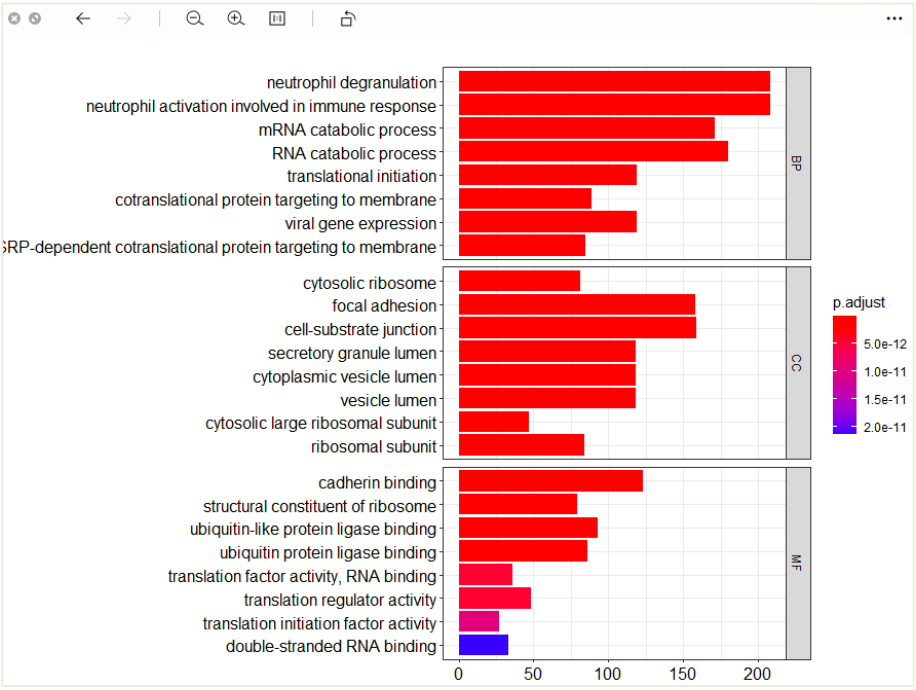

Figure.2 The GO enrichment analysis diagram.

3.3 KEGG enrichment analysis

KEGG enrichment analysis of the differentially expressed genes revealed that the Coronavirus disease (COVID-19) is the most significant pathway associated with the up-regulated genes. Spliceosome is the most significant pathway associated with down-regulated genes. Most of the other differentially expressed genes enriched in pathways of Amyotrophic lateral sclerosis, Parkinson's disease, Prion disease, Epstein-Barr virus infection, Ribosome, Protein processing in the endoplasmic reticulum, Leishmaniasis, and Proteasome. Figure 3 displays a KEGG enrichment analysis of differentially expressed genes between Covid-19 patients and healthy people.(Figure 3)Fig. 3 The KEGG enrichment analysis of differentially expressed genes between Covid-19 patients and healthy people. The horizontal axis represents the significance of enrichment. The vertical axis represents the KEGG pathways. The size of dots represents the number of differentially expressed genes in the KEGG enrichment analysis, and the color of dots represents the adjusted p-value. Gene ratio: the numerator is the number of different genes enriched in the GO term and the denominator is the total number of genes used in KEGG enrichment analysis.

ISSN 2618-1584 Vol. 3, Issue 6: 1-5, DOI: 10.25236/FMSR.2021.030601

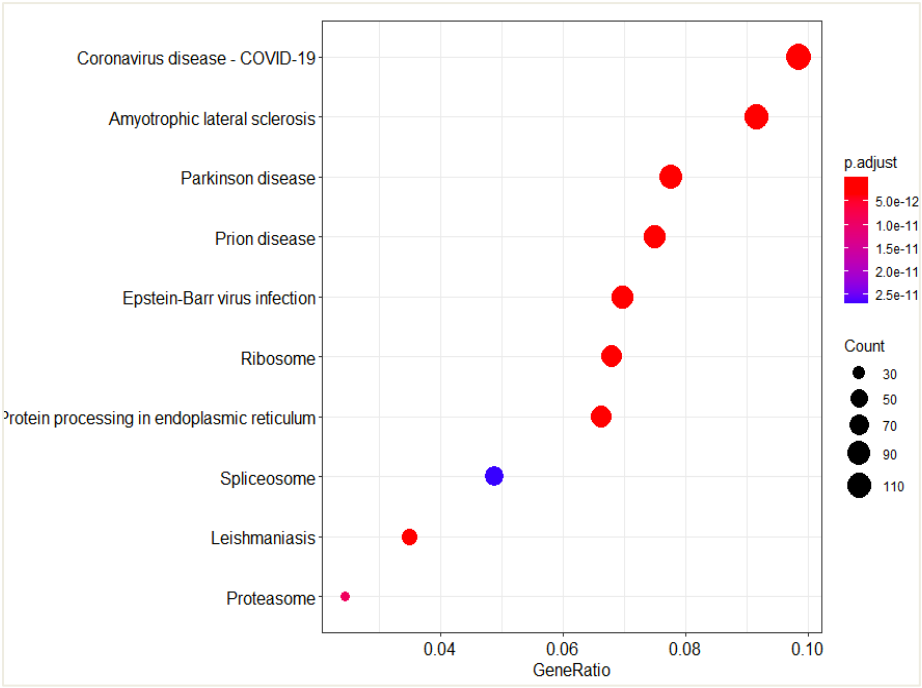

Figure. 3 The KEGG enrichment analysis of differentially expressed genes between Covid-19 patients and healthy people.

4. Discussion

In some of the covid-19 infected patients, the progression of lung infection was fierce and can cause the acute respiratory syndrome. Treatments were not pronounced. A decrease in peripheral blood lymphocyte count appeared in a majority of infected patients. In severe cases, a progressive dropping in peripheral blood lymphocyte count was observed (Guan WJ, 2020).

In this study, the clustering result of differential gene expression revealed a major relationship between differential gene expression and B cell differentiation and immune response B cell activation. Most of the genes in the related clusters were up-regulated. This finding suggests an association between gene up-regulation in B cell-related clusters and immune dysfunction caused by Sars-Cov2. KEGG enrichment showed that the differential gene mainly act on neuro-related pathways, suggesting that the immune response may induce epilepsy.

This article explored the immunological mechanism of B lymphocytes caused by Sars-Cov2 infection at the genetic level, and screened out differentially expressed genes related to immune regulation, together with their signaling pathway. However, there is still a lack of effective and solid evidence. More animal or cellular experiments and clinical researches with large scales are needed.

References

[1] Berthelot JM;Lioté F;Maugars Y;Sibilia J. Lymphocyte Changes in Severe COVID-19: Delayed Over-Activation of STING, Frontiers in immunology,2020, 11,607069.
[2] Meshcheryakova A;Pietschmann P;Zimmermann P;Rogozin IB;Mechtcheriakova D.AID and APOBECs as Multifaceted Intrinsic Virus-Restricting Factors: Emerging Concepts in the Light of COVID-19, Frontiers in immunology,2021,12,690416.
[3] You M;Chen L;Zhang D;Zhao P;Chen Z;Qin EQ;Gao Y;Davis MM;Yang P.Single-cell epigenomic landscape of peripheral immune cells reveals establishment of trained immunity in individuals convalescing from COVID-19,Nature cell biology,2021,23,620-630.
[4] Cheema PS;Nandi D;Nag A.Exploring the therapeutic potential of forkhead box O for outfoxing COVID-19,Open biology,2021,11,210069.
[5] Zhao J;Jaffe A;Li H;Lindenbaum O;Sefik E;Jackson R;Cheng X;Flavell RA;Kluger Y.Detection of differentially abundant cell subpopulations in scRNA-seq data,Proceedings of the National Academy of Sciences of the United States of America,2021,118,22.

ISSN 2618-1584 Vol. 3, Issue 6: 1-5, DOI: 10.25236/FMSR.2021.030601

[6] Liang C;Bencurova E;Psota E;Neurgaonkar P;Prelog M;Scheller C;Dandekar T. Population-Predicted MHC Class II Epitope Presentation of SARS-CoV-2 Structural Proteins Correlates to the Case Fatality Rates of COVID-19 in Different Countries,International journal of molecular sciences,2021,22.