Identification of potential biomarkers and inhibitors in SARS-CoV-2 infected macaques

Hanming Gu

Shanghai Jiao Tong University

Wei Wang

Shanghai Civil Aviation College

Gongsheng Yuan (✉ gsyuan14@gmail.com)

Fudan University

Research Article

Keywords: SARS-Cov-2, COVID-19, macaques, inhibitors

Posted Date: October 20th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-95211/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The COVID-19 pandemic caused by the infection with SARS-CoV-2 has overwhelmed many health systems globally. Our study is to identify differentially expressed genes (DEGs) and the associated biological pathways of COVID-19 to elucidate the potential pathogenesis and metabolism. The gene expression profile of the GSE155363 dataset was originally produced using the high-throughput Illumina HiSeq 4000 (Macaca mulatta). Kyoto Encyclopedia of Genes and Genomes pathway (KEGG) enrichment analyses were performed to discover their functional categories and biochemical pathways. The results suggested that four biological pathways: Fatty acid elongation, Biosynthesis of unsaturated fatty acids, Fatty acid metabolism, and Ribosome were mostly involved in the macaques with COVID-19. Thus, our study provides novel insights into the underlying pathogenesis of COVID-19.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a RNA coronavirus that causes coronavirus disease 2019 (COVID-19)¹. The common clinical symptoms of severe COVID-19 are fever, cough, pneumonia, and dyspnoea². The Chinese Center for Disease Control and Prevention reported that 80% of infected patients experienced mild diseases, 14% developed severe diseases and 5% developed critical diseases^{3,4}.

The factors that caused illness in COVID-19 patients are not fully understood and the development of diseases does not simply relate to viral load⁵⁻⁷. The aggressive inflammatory response to SARS-CoV-2 is considered as a major reason for disease severity and death in COVID-19 patients⁸⁻¹¹. The severe stage of COVID-19 is relevant to high levels of circulating cytokines and substantial inflammatory cell infiltration in the lungs^{8,12,13}. Given that the morbidity and mortality in COVID-19 is still rising worldwide, a better understanding of the mechanism of SARS-CoV-2 infection is necessary to the discovery of new therapeutic targets. Here, we analyzed the GEO data (GSE155363), provided by Price A and Rasmussen AL (Columbia University) to identify DEGs and the relevant biological process of COVID-19 in macaques. The KEGG pathway analysis, and protein-protein interaction were performed for finding those key gene nodes.

Methods

Data resources

Gene expression profile dataset GSE155363 was obtained from the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>). The data was produced by using an Illumina HiSeq 4000 (Macaca mulatta) (Columbia University, New York, NY10032). Our study analyzed eight positive samples (infection for 24 hours: GSM4700109, GSM4700110, GSM4700111, GSM4700112, GSM4700113, GSM4700114,

GSM4700115, GSM4700116) and eight negative control samples (GSM4700101, GSM4700102, GSM4700103, GSM4700104, GSM4700105, GSM4700106, GSM4700107, GSM4700108).

Data acquisition and preprocessing

The genes between COVID-19 macaques and healthy controls were conducted by R script. We used the Student's t-tests to identify DEGs with $P < 0.01$ and fold change ≥ 2 as being statistically significant.

KEGG and Reactome Pathway analysis

KEGG pathway was analyzed by the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (<http://david.ncifcrf.gov/>) online tools and Reactome Pathway Database (<https://reactome.org/>). $P < 0.05$ was considered statistically significant.

Module analysis

The Molecular Complex Detection (MCODE) tool was used to detect the connected regions in PPI networks. The functional and pathway enrichment analyses were performed using DAVID (<http://david.ncifcrf.gov/>), and $P < 0.05$ was used as the cutoff criterion.

Inhibitors prediction

The L1000fwd is used to discover the potential inhibitors in COVID-19 macaques. L1000fwd analyzes the similarity to rank inhibitors potentially able to reverse the gene signature (29420694). The adjusted p-value of 0.05 is considered as statistical significance.

Results

Identification of DEGs in COVID-19 macaques

To determine the key features of COVID-19 infection, the modular transcriptional signature of COVID-19 macaques was compared to that of the healthy controls. A total of 21 genes were identified to be differentially expressed in COVID-19 samples with the threshold of $P < 0.01$. The DEGs for COVID-19 and healthy controls were showed in Figure 1.

KEGG and Reactome Pathway analysis of DEGs in COVID-19 macaques

To further determine the roles of the DEGs from negative controls versus COVID-19 macaques' samples, we performed KEGG pathway and Reactome Pathway (RCTM) analysis as previously described¹⁴. KEGG and RCTM pathway analysis are used to define the functional meanings of genes and genomes. In our study, the top enriched biological KEGG pathways associated with COVID-19 included " Fatty acid elongation", " Biosynthesis of unsaturated fatty acids ", " Fatty acid metabolism " and "Ribosome" (Table 1). Moreover, the RCTM analysis showed the DGEs were involved in the synthesis of very long-chain fatty

acyl-CoAs from fatty acid metabolism group and diseases of DNA repair group (Figure 2 and Supplemental Table S1).

PPI network analysis of DEGs between healthy controls and COVID-19 macaques

To explore the relationship of DEGs at the protein level, we identified 14 significant functional proteins with String tool (<https://string-db.org/>) (Figure 3A), and created a PPI network by using the Cytoscape software (Figure 3B). We set the predefined criterion of combined score >0.7, a total of 7 interactions and 7 nodes were constructed to form the PPI network between healthy controls and COVID-19 macaques' samples (Figure 3B). Among these nodes, the top 5 hub genes with highest degree scores are shown in Table 2. We then created the top two significant modules of COVID-19 versus control samples to depict the functional annotation of the genes (Figure 3B).

By using the L1000FWD analysis, we identified top ten potential inhibitors with the highest scores (Figure 4 and Supplemental Table S2). Among them: angiogenesis-inhibitor (BRD-K08502430) is used as a PARP inhibitor; JAK3-INHIBITOR-VI (BRD-K04546108) is used for PARP and PI3K inhibitors; FGIN-1-27 (BRD-K09778810) is used as the dopamine receptor antagonist; HG-5-113-01 (BRD-K93747373) is used as a RAF inhibitor; WITHAFERIN-A (BRD-K88378636) is an NF- κ B pathway inhibitor; GDC-0941 (BRD-K52911425) and BRD-K79222491 (BRD-K79222491) are PI3K inhibitors; BRD-K55055802 (BRD-K55055802) is used as a cyclooxygenase inhibitor; LOFEPRAMINE (BRD-K82147103) is a dopamine receptor antagonist; DEXAMETHASONE (BRD-A69951442) is an anti-inflammatory and immunosuppressant. These inhibitors may be the potential therapies for patients with COVID-19.

Discussion

The most frequent symptoms of COVID-19 are fever, cough, fatigue and diarrhea¹⁵. Severe illness usually happens about 1 week after the onset of symptoms¹⁶. Thus, it is of the utmost importance to find the biomarkers during the early infection. We analyzed the GEO data by comparing the early infection samples (day 1) with the negative controls to find the DEGs and potential therapeutic targets. After infection on day 1, most of DEGs were upregulated in macaques. Among them, we identified 17 most changed genes including TRNAA-AGC, PQLC1, NEIL3, HACD1 and other unknown genes, which may be involved in the process of virus infection. Interestingly, we identified a large number of unknown genes that could be the novel targets for COVID-19 treatment. Thus, further studies will be focused on identifying the functions and activities of these unknown genes.

The KEGG pathways of DEGs after infection showed high relevance to the fat metabolism. Previously, our reports had shown that COVID-19 had a great impact on metabolic syndromes including the "Biosynthesis of unsaturated fatty acids", "fatty acid elongation", "Fatty acid metabolism", and "Metabolic pathways" in human patients^{14, 17}. Similarly, in macaques, the fat metabolism changed during the early stage of SARS-CoV-2 infection. The presence of one or more underlying health conditions including diabetes, chronic lung disease, and cardiovascular disease are risk factors for the enhanced

severity of the COVID-19 diseases¹⁸. In the U.S., 89% of adults who have diabetes are also overweight or have obesity¹⁸. Obesity disrupts breathing that leads to sleep apnea and disordered sleep/wake cycle, which further causes the circadian related diseases and promotes virus infection¹⁹⁻²⁵. Obesity also affects the mitochondrial structure and function²⁶, which further leads to physiological dysfunction and the related diseases^{27, 28}. Obesity was first recognized as an independent risk factor for the H1N1 pandemic in 2009^{29, 30}. In the report by Northwell Health, among the confirmed SARS-CoV-2 cases, hypertension, obesity, and diabetes were the most common comorbidities³¹.

Obesity appears to predispose COVID-19 patients to increase the severity of the disease³². Obesity-caused cellular dysfunction triggers a various range of signaling pathways, including activation of NF- κ B, IRE-1, mTOR, ERKs, and PKR³³. These pathways collaborate to produce two key metabolic effects. First, each pathway converges upon and inhibits insulin signaling pathways. Second, these signals are involved in two main inflammatory pathways³⁴. The transcription factor NF- κ B enhanced immunity by regulating the expression of genes involved in inflammation³⁵. A number of reports have demonstrated a critical role of the NF- κ B signaling pathway in the liver, bone, cartilage, and central nervous system in the development of inflammation-associated diseases³⁶⁻³⁸. High metabolic demands can induce cell death, which forms a pro-inflammatory signal to NF- κ B activation³³. Moreover, activation of innate immune receptors through MyD88 could activate NF- κ B in response to an excess of free fatty acids in a high fat diet^{33, 39, 40}. Thus, NF- κ B signaling is a critical pathway that contributes to the pathology of metabolic disorders.

In conclusion, the results suggested that several biological pathways such as "Fatty acid elongation", "Biosynthesis of unsaturated fatty acids", "Fatty acid metabolism" and "Ribosome" are commonly involved in the development of COVID-19. This study thus discovered novel biomarkers and inhibitors of COVID-19 in macaques, which may provide better targets for therapeutic intervention in the future.

Declarations

Conflict of Interest statement

The authors declare that there are no conflicts of interest.

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

Acknowledgment

We thank the anonymous referees for their useful suggestions.

References

- [1] Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR: Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *Int J Antimicrob Agents* 2020, 55:105924.
- [2] Lima C: Information about the new coronavirus disease (COVID-19). *Radiol Bras* 2020, 53:V-VI.
- [3] Merad M, Martin JC: Pathological inflammation in patients with COVID-19: a key role for monocytes and macrophages. *Nat Rev Immunol* 2020, 20:355-62.
- [4] George PM, Wells AU, Jenkins RG: Pulmonary fibrosis and COVID-19: the potential role for antifibrotic therapy. *Lancet Respir Med* 2020, 8:807-15.
- [5] Kirkcaldy RD, King BA, Brooks JT: COVID-19 and Postinfection Immunity: Limited Evidence, Many Remaining Questions. *JAMA* 2020, 323:2245-6.
- [6] Yi Y, Lagniton PNP, Ye S, Li E, Xu RH: COVID-19: what has been learned and to be learned about the novel coronavirus disease. *Int J Biol Sci* 2020, 16:1753-66.
- [7] Zhang R, Li Y, Zhang AL, Wang Y, Molina MJ: Identifying airborne transmission as the dominant route for the spread of COVID-19. *Proc Natl Acad Sci U S A* 2020, 117:14857-63.
- [8] Costela-Ruiz VJ, Illescas-Montes R, Puerta-Puerta JM, Ruiz C, Melguizo-Rodriguez L: SARS-CoV-2 infection: The role of cytokines in COVID-19 disease. *Cytokine Growth Factor Rev* 2020, 54:62-75.
- [9] Del Valle DM, Kim-Schulze S, Huang HH, Beckmann ND, Nirenberg S, Wang B, Lavin Y, Swartz TH, Madduri D, Stock A, Marron TU, Xie H, Patel M, Tuballes K, Van Oekelen O, Rahman A, Kovatch P, Aberg JA, Schadt E, Jagannath S, Mazumdar M, Charney AW, Firpo-Betancourt A, Mendu DR, Jhang J, Reich D, Sigel K, Cordon-Cardo C, Feldmann M, Parekh S, Merad M, Gnajatic S: An inflammatory cytokine signature predicts COVID-19 severity and survival. *Nat Med* 2020.
- [10] Gu H, Yuan G: Identification of potential key genes for SARS-CoV-2 infected human bronchial organoids based on bioinformatics analysis. *bioRxiv* 2020:2020.08.18.256735.
- [11] Yuen KS, Ye ZW, Fung SY, Chan CP, Jin DY: SARS-CoV-2 and COVID-19: The most important research questions. *Cell Biosci* 2020, 10:40.
- [12] Li S, Jiang L, Li X, Lin F, Wang Y, Li B, Jiang T, An W, Liu S, Liu H, Xu P, Zhao L, Zhang L, Mu J, Wang H, Kang J, Li Y, Huang L, Zhu C, Zhao S, Lu J, Ji J, Zhao J: Clinical and pathological investigation of patients with severe COVID-19. *JCI Insight* 2020, 5.
- [13] Li H, Liu L, Zhang D, Xu J, Dai H, Tang N, Su X, Cao B: SARS-CoV-2 and viral sepsis: observations and hypotheses. *Lancet* 2020, 395:1517-20.

- [14] Gu H, Yuan G: Identification of key genes in SARS-CoV-2 patients on bioinformatics analysis. *bioRxiv* 2020:2020.08.09.243444.
- [15] Burke RM, Killerby ME, Newton S, Ashworth CE, Berns AL, Brennan S, Bressler JM, Bye E, Crawford R, Harduar Morano L, Lewis NM, Markus TM, Read JS, Rissman T, Taylor J, Tate JE, Midgley CM, Case Investigation Form Working G: Symptom Profiles of a Convenience Sample of Patients with COVID-19 - United States, January-April 2020. *MMWR Morb Mortal Wkly Rep* 2020, 69:904-8.
- [16] He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, Lau YC, Wong JY, Guan Y, Tan X, Mo X, Chen Y, Liao B, Chen W, Hu F, Zhang Q, Zhong M, Wu Y, Zhao L, Zhang F, Cowling BJ, Li F, Leung GM: Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med* 2020, 26:672-5.
- [17] Gu H, Yuan G: Identification of potential biomarkers and inhibitors for SARS-CoV-2 infection. *medRxiv* 2020:2020.09.15.20195487.
- [18] Rebello CJ, Kirwan JP, Greenway FL: Obesity, the most common comorbidity in SARS-CoV-2: is leptin the link? *Int J Obes* 2020, 44:1810-7.
- [19] Slater G, Steier J: Excessive daytime sleepiness in sleep disorders. *J Thorac Dis* 2012, 4:608-16.
- [20] Huttunen R, Syrjanen J: Obesity and the risk and outcome of infection. *Int J Obes* 2013, 37:333-40.
- [21] Yuan G, Hua B, Yang Y, Xu L, Cai T, Sun N, Yan Z, Lu C, Qian R: The Circadian Gene Clock Regulates Bone Formation Via PDIA3. *J Bone Miner Res* 2017, 32:861-71.
- [22] Yuan G, Hua B, Cai T, Xu L, Li E, Huang Y, Sun N, Yan Z, Lu C, Qian R: Clock mediates liver senescence by controlling ER stress. *Aging* 2017, 9:2647-65.
- [23] Zhu Z, Hua B, Xu L, Yuan G, Li E, Li X, Sun N, Yan Z, Lu C, Qian R: CLOCK promotes 3T3-L1 cell proliferation via Wnt signaling. *IUBMB Life* 2016, 68:557-68.
- [24] Zhu Z, Hua B, Shang Z, Yuan G, Xu L, Li E, Li X, Sun N, Yan Z, Qian R, Lu C: Altered Clock and Lipid Metabolism-Related Genes in Atherosclerotic Mice Kept with Abnormal Lighting Condition. *Biomed Res Int* 2016, 2016:5438589.
- [25] Zhu Z, Xu L, Cai T, Yuan G, Sun N, Lu C, Qian R: Clock represses preadipocytes adipogenesis via GILZ. *J Cell Physiol* 2018, 233:6028-40.
- [26] Breininger SP, Malcomson FC, Afshar S, Turnbull DM, Greaves L, Mathers JC: Effects of obesity and weight loss on mitochondrial structure and function and implications for colorectal cancer risk. *Proc Nutr Soc* 2019, 78:426-37.
- [27] Cai T, Hua B, Luo D, Xu L, Cheng Q, Yuan G, Yan Z, Sun N, Hua L, Lu C: The circadian protein CLOCK regulates cell metabolism via the mitochondrial carrier SLC25A10. *Biochim Biophys Acta Mol Cell Res*

2019, 1866:1310-21.

[28] Yuan G, Yang S, Liu M, Yang S: RGS12 is required for the maintenance of mitochondrial function during skeletal development. *Cell Discov* 2020, 6:59.

[29] Milner JJ, Beck MA: The impact of obesity on the immune response to infection. *Proc Nutr Soc* 2012, 71:298-306.

[30] Huttunen R, Syrjanen J: Obesity and the outcome of infection. *Lancet Infect Dis* 2010, 10:442-3.

[31] Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, the Northwell C-RC, Barnaby DP, Becker LB, Chelico JD, Cohen SL, Cookingham J, Coppa K, Diefenbach MA, Dominello AJ, Duer-Hefe J, Falzon L, Gitlin J, Hajizadeh N, Harvin TG, Hirschwerk DA, Kim EJ, Kozel ZM, Marrast LM, Mogavero JN, Osorio GA, Qiu M, Zanos TP: Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *JAMA* 2020, 323:2052-9.

[32] Sattar N, McInnes IB, McMurray JJV: Obesity Is a Risk Factor for Severe COVID-19 Infection: Multiple Potential Mechanisms. *Circulation* 2020, 142:4-6.

[33] Baker RG, Hayden MS, Ghosh S: NF-kappaB, inflammation, and metabolic disease. *Cell Metab* 2011, 13:11-22.

[34] Odegaard JI, Chawla A: Pleiotropic actions of insulin resistance and inflammation in metabolic homeostasis. *Science* 2013, 339:172-7.

[35] Lawrence T: The nuclear factor NF-kappaB pathway in inflammation. *Cold Spring Harb Perspect Biol* 2009, 1:a001651.

[36] Choi MC, Jo J, Park J, Kang HK, Park Y: NF-kappaB Signaling Pathways in Osteoarthritic Cartilage Destruction. *Cells* 2019, 8.

[37] Yuan G, Xu L, Cai T, Hua B, Sun N, Yan Z, Lu C, Qian R: Clock mutant promotes osteoarthritis by inhibiting the acetylation of NFkappaB. *Osteoarthritis Cartilage* 2019, 27:922-31.

[38] Yuan G, Yang S, Ng A, Fu C, Oursler MJ, Xing L, Yang S: RGS12 Is a Novel Critical NF-kappaB Activator in Inflammatory Arthritis. *iScience* 2020, 23:101172.

[39] Rogero MM, Calder PC: Obesity, Inflammation, Toll-Like Receptor 4 and Fatty Acids. *Nutrients* 2018, 10.

[40] Shi H, Kokoeva MV, Inouye K, Tzameli I, Yin H, Flier JS: TLR4 links innate immunity and fatty acid-induced insulin resistance. *J Clin Invest* 2006, 116:3015-25.

Tables

Due to technical limitations, full-text HTML conversion of the tables could not be completed. However, they can be downloaded and accessed in the Supplementary Files.

Figures

Figure 1

Heatmap of the gene expression patterns (healthy controls vs macaques with COVID19 infection). (A) The color key represents differential gene expression magnitude (log2FC). (B) The table showed the significant changed genes in COVID-19 macaques from (A).

Figure 2

The RCTM map showed the main reactive pathways. Note that the metabolism and DNA repair was mainly involved in COVID-19.

Figure 3

The String proteins (A) and top modules (B) from the protein-protein interaction network between COVID-19 macaques and controls.

B

Inhibitors	p-value	Z-score	combined score
angiogenesis-inhibitor	1.73E-03	-1.77	4.88
JAK3-inhibitor-VI	7.68E-02	-1.55	1.73
FGIN-1-27	6.67E-02	-1.65	1.94
HG-5-113-01	6.20E-02	-1.87	2.26
withaferin-a	6.23E-02	-1.85	2.23
GDC-0941	6.80E-02	-1.73	2.02
BRD-K79222491	6.64E-02	-1.69	1.99
BRD-K55055802	6.79E-02	-1.7	1.98
lofepramine	7.34E-02	-1.67	1.89
dexamethasone	7.29E-02	-1.59	1.81

Figure 4

Inhibitor prediction against SARS-CoV-2 infection by L1000FDW. (A) Input genes are represented by the significantly changed genes obtained from the analysis of the GSE155363 dataset. Dots are based on the Mode of Action (MOA) of the respective inhibitor. (B) The top ten inhibitors with a high significance p-value and combined score were selected.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Tables.docx](#)
- [SupplementalTableS1RCTMpathways.csv](#)
- [SupplementalTableS2inhibitors.csv](#)