

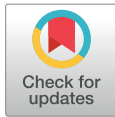


Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



ELSEVIER



BIOMEDICAL

Molecular Alterations Prompted by SARS-CoV-2 Infection: Induction of Hyaluronan, Glycosaminoglycan and Mucopolysaccharide Metabolism

Sergio Andonegui-Elguera,^{a,b,*} Keiko Taniguchi-Ponciano,^{a,*} Cesar Raul Gonzalez-Bonilla,^c Javier Torres,^d Hector Mayani,^e Luis Alonso Herrera,^{f,g} Eduardo Peña-Martínez,^a Gloria Silva-Román,^a Sandra Vela-Patiño,^a Aldo Ferreira-Hermosillo,^{a,h} Claudia Ramirez-Renteria,^{a,h} Roberto Carvente-Garcia,^a Carlos Mata-Lozano,^a Daniel Marrero-Rodríguez,^a and Moises Mercado^a

^aUnidad de Investigación Médica en Enfermedades Endocrinas, Hospital de Especialidades, Centro Medico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Ciudad de México, México

^bDepartamento de Desarrollo Tecnológico, Instituto Nacional de Medicina Genómica, Ciudad de México, México

^cTitular, Coordinación de Investigación en Salud, Centro Medico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Ciudad de México, México

^dUnidad de Investigación Médica en Enfermedades Infecciosas y Parasitarias, Hospital de Pediatría, Centro Medico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Ciudad de México, México

^eUnidad de Investigación Médica en Enfermedades Oncológicas, Hospital de Oncología, Centro Medico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Ciudad de México, México

^fUnidad de Investigación Biomédica en Cáncer, Instituto Nacional de Cancerología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Ciudad de México, México

^gDirección General, Instituto de Medicina Genómica, Ciudad de México, México

^hServicio de Endocrinología, Hospital de Especialidades, Centro Medico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Ciudad de México, México

Received for publication May 5, 2020; accepted June 13, 2020 (ARCMED_2020_627).

Background. The SARS-CoV-2 is the etiological agent causing COVID-19 which has infected more than 2 million people with more than 200000 deaths since its emergence in December 2019. In the majority of cases patients are either asymptomatic or show mild to moderate symptoms and signs of a common cold. A subset of patients, however, develop a severe atypical pneumonia, with the characteristic ground-glass appearance on chest x-ray and computerized tomography, which evolves into an acute respiratory distress syndrome, that requires mechanical ventilation and eventually results in multiple organ failure and death. The Molecular pathogenesis of COVID-19 is still unknown.

Aim of the study. In the present work we performed a stringent metanalysis from the publicly available RNAseq data from bronchoalveolar cells and peripheral blood mononuclear cells to elucidate molecular alterations and cellular deconvolution to identify immune cell profiles.

Results. Alterations in genes involved in hyaluronan, glycosaminoglycan and mucopolysaccharides metabolism were over-represented in bronchoalveolar cells infected by SARS-CoV-2, as well as potential lung infiltration with neutrophils, T CD4+ cell and macrophages. The blood mononuclear cells presented a proliferative state. Dramatic reduction of NK and T lymphocytes, whereas an exacerbated increase in monocytes.

Conclusions. In summary our results revealed molecular pathogenesis of the SARS-CoV-2 infection to bronchoalveolar cells inducing the hyaluronan and glycosaminoglycan metabolism that could shape partially the components of the ground-glass opacities observed in CT. And the potential immune response profile in COVID-19. © 2020 IMSS. Published by Elsevier Inc.

Key Words: Hyaluronan, Bronchoalveolar transcriptome, COVID-19, SARS-CoV-2.

*These authors contributed equally to this work.

Address reprint requests to: Moises Mercado, MD and Daniel Marrero-Rodríguez, PhD, CONACyT-Unidad de Investigación Médica en Enfermedades Endocrinas, Hospital de Especialidades, Centro Médico Nacional

Siglo XXI, Instituto Mexicano del Seguro Social, Av. Cuauhtémoc 330, Col. Doctores, Ciudad de México 06720, México; Phone: (+52) (55) 54401021; FAX: 5556276900x21913; E-mail: dan.mar57@gmail.com or moises.mercado@endocrinologia.org.mx

Introduction

Since its emerging in Wuhan, China in December 2019 the coronavirus disease (COVID-19) epidemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has progressed rapid to pandemic (1). As for April 27, 2020 more than 2 million people worldwide have been diagnosed with COVID-19 (2). It is associated with significant mortality due to the development of acute respiratory distress syndrome (ARDS) and eventually in multiple organ failure, particularly in certain high-risk populations such as the elderly, as well as patients with obesity, diabetes and hypertension. This complicated course is thought to be due to a so-called “cytokine storm” characterized by over production of TNF, IL6 and IL1 β , which together lead to vascular hyperpermeability, multiorgan failure and eventually death (3).

The SARS-CoV-2 viral entry depends upon the binding of viral spike (S) protein to the host cell surface protein angiotensin-converting enzyme 2 (ACE2). Once internalized, the S protein undergoes proteolytic processing by transmembrane serine protease 2 (TMPRSS2), as well as by cathepsin L and B (CTSL and CTSB, respectively), resulting in cleavage at the S1/S2 site, and the ensuing S2-dependent fusion of viral and cellular membranes (4). Coronaviruses such as SARS-CoV and MERS-CoV have evolved strategies to dampen or delay IFN production, which usually trigger exacerbated inflammatory host response leading to severe lung pathology (5). Among these mechanisms is the disruption of the pathogen associated molecular patterns recognition by the viral nonstructural protein 15 (nsp15) (6). Despite the global treats of COVID-19, the host immune response against virus infection remains poorly understood (5). Therefore, in the present work we carried out comprehensive and stringent transcriptomic metanalysis of SARS-CoV-2 infected bronchoalveolar cells and peripheral blood mononuclear cells to unveil the molecular alterations caused by viral infection as well as deconvolution analysis to identify the immune cell profiles in COVID-19 patients.

Materials and Methods

Bronchoalveolar Lavage Fluid and Peripheral Blood Mononuclear Cells RNAseq Files Analysis

Paired end fastq files were downloaded from BIGD-GSA and SRI from NCBI. Accession numbers are CRA002390 and SRR10571724, SRR10571730, and SRR10571732. Data files consisted in 3 COVID-19 patients bronchoalveolar fluid lavage (BALF) and 3 bronchoalveolar fluid lavage controls, and also 3 COVID-19 peripheral blood mononuclear cells (PBMC) and 3 control peripheral blood mononuclear cells. Fastq files quality control of was performed using FastQC v. 0.11.9. Samples adapter were trimmed with the Cutadapt v tool. 2.10.

BALF and PBMC samples duplicates were spotted using Picard v. 2.22.4, the RNAseq reads alignments were done using STAR tools v. 2.7.3a, with the parameters previously described (Xiong Y, et al, 2020). The RNAseq reading count was carried out using the Rsubread package with the annotations of Homo sapiens GRCh.38.99 downloaded from Ensembl. Subsequent analyzes were performed using the stable Gene ID identifier with BioMart annotations: Ensembl Genes 100, Human genes (GRCh38.p13), the identifier was translated to HGNC symbol to facilitate biological interpretation. The BALF and PBMC counts were subjected to quality control using the NOISeq package, biases associated with length, GC content, RNA composition and batch effect, were identified and removed. BALF and PBMC samples were processed using full-quantile normalization for GC content and length, finally, the ARSyN function was used to eliminate batch effect. The normalization process was carried out using the NOISeq and EDASeq packages in R. Differentially expressed genes were calculated on the normalized counts using the noiseqbio function of the NOISeq package, using the predefined parameters and filtered by a probability greater than 0.99.

Deconvolution and Gene Ontology Analysis

The deconvolution analysis was performed using the online tool CIBERSORT (Cell Type Identification By Estimating Relative Samples Of RNA Transcripts), which solves the equation $m = f \times B$, where m is a mixture of mRNA, f is a vector that denotes the Cell fractions that make up the mixture and B is an array of gene expression profiles characteristic of each cell subtype. CIBERSORT is based on the application of v -SVR (nu-support vector regression) where limits of a hyperplane are determined that fits most of the possible points at a constant distance and where a regression is performed. Support vectors are genes selected from B . CIBERSORT was used for the present work with the default parameters, using the LM22 expression profile matrix (22 types of immune cells) and 100 permutations to calculate the p -value associated by Monte Carlo Sampling. The mRNA mixture used was a matrix constructed with the expression values of each sample obtained by bronchoalveolar lavage fluid and peripheral blood mononuclear cells RNAseq files analysis.

WebGestalt (<http://www.webgestalt.org>) was used for understanding the biological meaning behind the resulting list of genes, to obtain gene ontology and pathway information for significantly de-regulated genes in COVID-19 patients.

Results

SARS-CoV-2-infected Bronchoalveolar Molecular Alterations

In order to identify potential cellular and molecular alterations in lung tissue we analyzed top 100 genes up- and

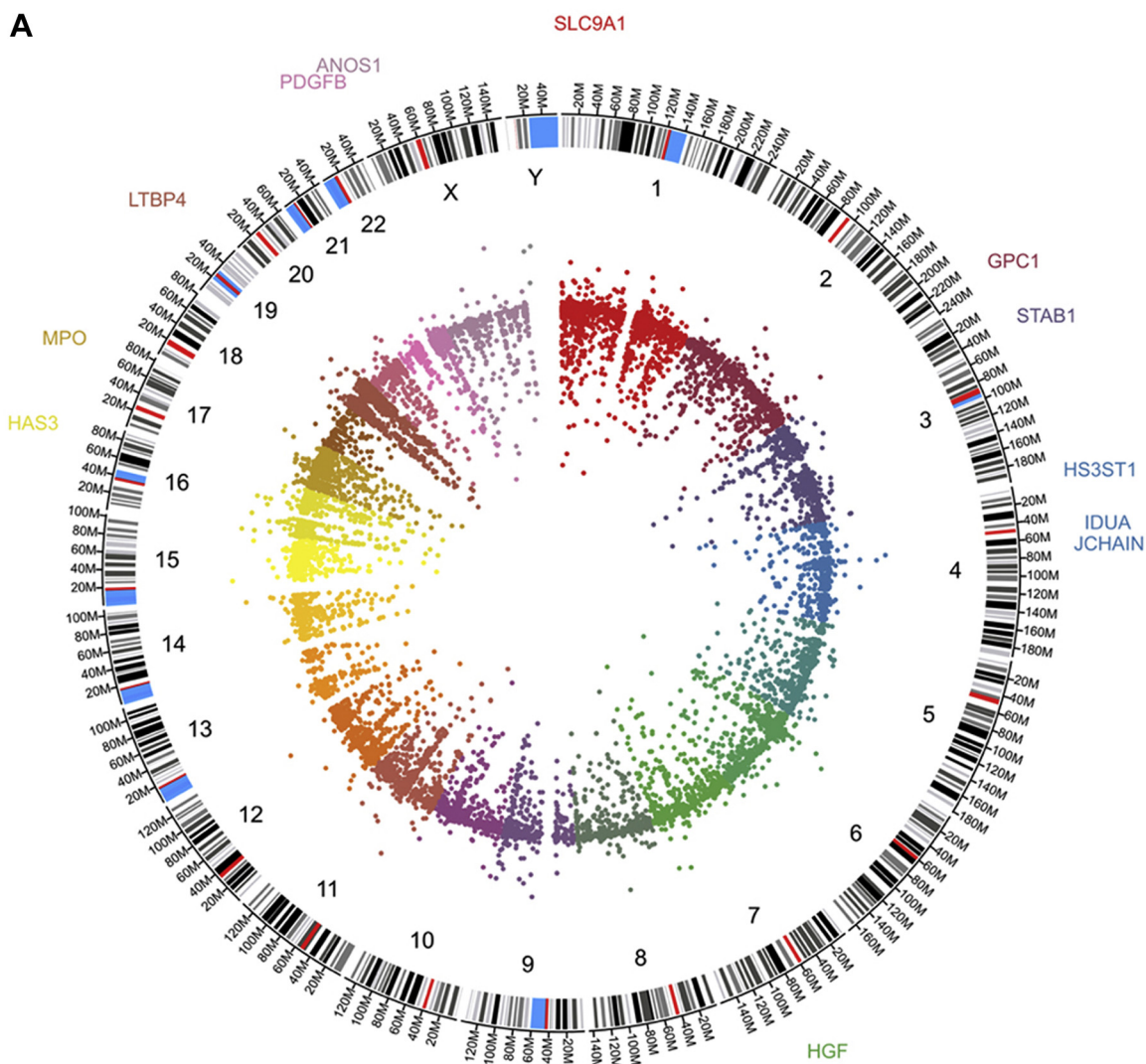


Figure 1. Panel A depicts a circos plot from the bronchoalveolar lavage fluid cells transcriptome from COVID-19 patients. The inner ring is a scatter plot representing each transcript identified, followed by chromosome number and size. The outer ring presents genes identified participating in hyaluronan, glycosaminoglycan, and mucopolysaccharide metabolisms. Panels B and C show Gene Ontology (GO) terms for biological processes and molecular functions, respectively, for the top 100 up-regulated genes. Panels D and E show GO terms for biological processes and molecular functions, respectively, for the top 100 down-regulated genes.

down-regulated. The gene ontology (GO) at the biological process level of the 100 most up-regulated genes showed participation in hyaluronan metabolic process (HGF, PDGFB, and SLC9A1), glycosaminoglycan metabolic process (HS3ST1, IDUA, and GPC1), aminoglycan metabolic process (HS3ST1, PDGFB, and GPC1), and mucopolysaccharide metabolic process (HGF, IDUA, and SLC9A1) among others. The GO at the molecular function level showed again participation in glycosaminoglycan binding (MPO, JCHAIN, and ANOS1) and chemoattractant activity (HGF and PDGFB). Very interestingly, not only the up-regulated genes showed participation in COVID-19 relevant events. The gene ontology (GO) at the biological process level of the 100 most down-regulated genes showed participation

primarily in chemokine events such as chemokine-mediated signaling pathway, response to chemokine, (CCL17, CXCL9, and SLIT3), and antimicrobial humoral response (AZU1, PPBP, and ITLN1). The GO at the molecular function level showed involvement in heparin binding (PTCH1, RSPO3, and SAA1) and G protein-coupled receptor binding (CRH, CXCL9, and GNAS14) (Figure 1).

Further characterization of the molecular alterations was followed. Once we identify the transcriptome profile of the SARS-CoV-2 infected bronchoalveolar cells, we carried out cellular deconvolution to identify potential infiltrating immune cells. We observed more presence of Neutrophils, T CD4+ lymphocytes, and macrophages in COVID-19 patients BALF compared to control (Figure 2).

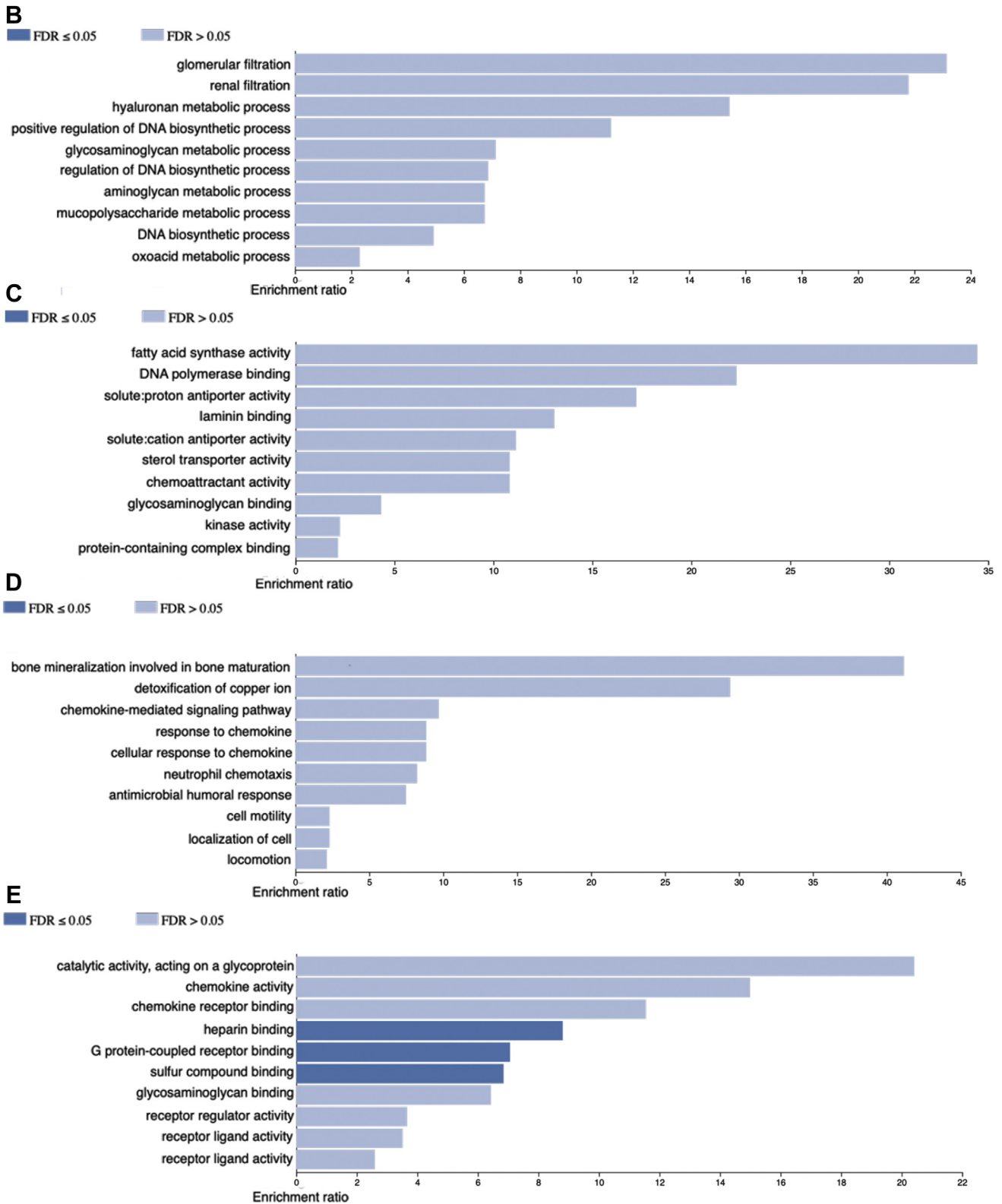


Figure 1. (Continued)

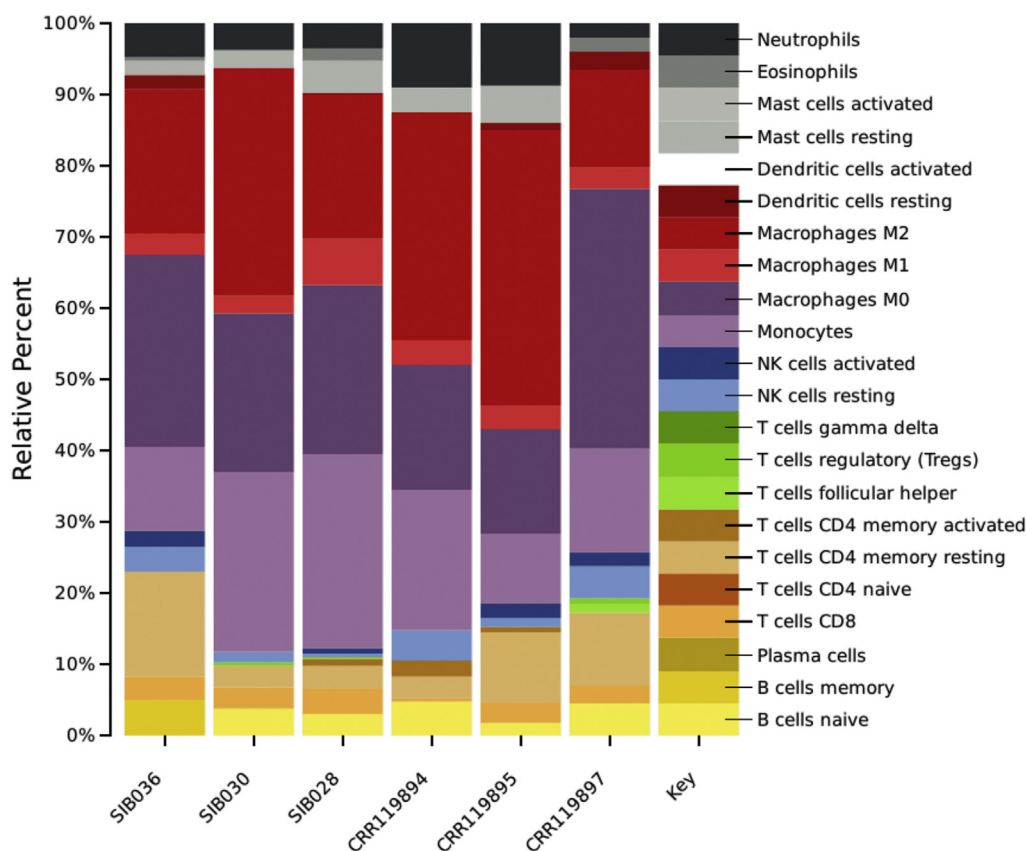


Figure 2. Deconvolution results by CIBERSORT algorithm show probability percentage of presence of Neutrophils, NK cells, T CD4+ cell and macrophages in COVID-19 patients BALF. SIB036, 030 and 028 correspond to BALF control and CRR119894, –119895 and –119897 correspond to COVID-19 BALF.

Transcriptome Profiles from COVID-19 Patients' Peripheral Blood Mononuclear Cells

The PBMC transcriptomic profiles from COVID-19 differs from the control group (Figure 3). The top 100 up-regulated genes GO at biological process level showed involvement in G1/S transition of mitotic cell cycle (CDC6, MCM10 and RRM2) and regulation of protein processing (BIRC7, SERPINE1 and C1QA) among others. WNT-protein binding (FZD5 and MERTK), RAS GTPase binding (NRG1, SRGAP1 and IQGAP3), was observed at the molecular function terms of the GO. Consistently with the observed in the up-regulated genes, the down-regulated genes showed involvement in negative regulation of cell proliferation (CDKN1C, SFN and RUNX3) and cell maturation (SOX8, RHEX and GATA2) at GO biological process. In the GO molecular function terms cytokine receptor binding (CXCL1, EFNA5 and RHEX) among others (Figure 3).

Very interestingly the cellular deconvolution analysis showed dramatic reduction of NK and T cells, whereas an exacerbated increase in monocytes (Figure 4).

Discussion

The present study constitutes a comprehensive meta-analysis of the available RNA sequencing data from pulmonary and blood mononuclear cells from patients with COVID-19 infection. In order to be able to design effective therapeutic strategies against SARS-CoV-2 infection, it is crucial to understand the molecular basis of its interaction with host cells. The infection of SARS-CoV-2 can cause severe pulmonary disease and complications with severe mortalities, currently there is no effective drug (5).

The acute respiratory distress syndrome that complicates a substantial number of patients with COVID-19 infection is characterized by a generalized endothelial dysfunction with a substantial fluid leak into the interstitial space thereby reducing lung compliance and thus, effective ventilation (7,8). This is the basis for the so-called ground-glass appearance on chest CT (9) and correlates with recent autopsy findings showing filling of the lungs with a clear fluid or gel-like material, much resembling what occur in wet drowning (10). Although the fluid filling components of the lungs are not yet defined,

A

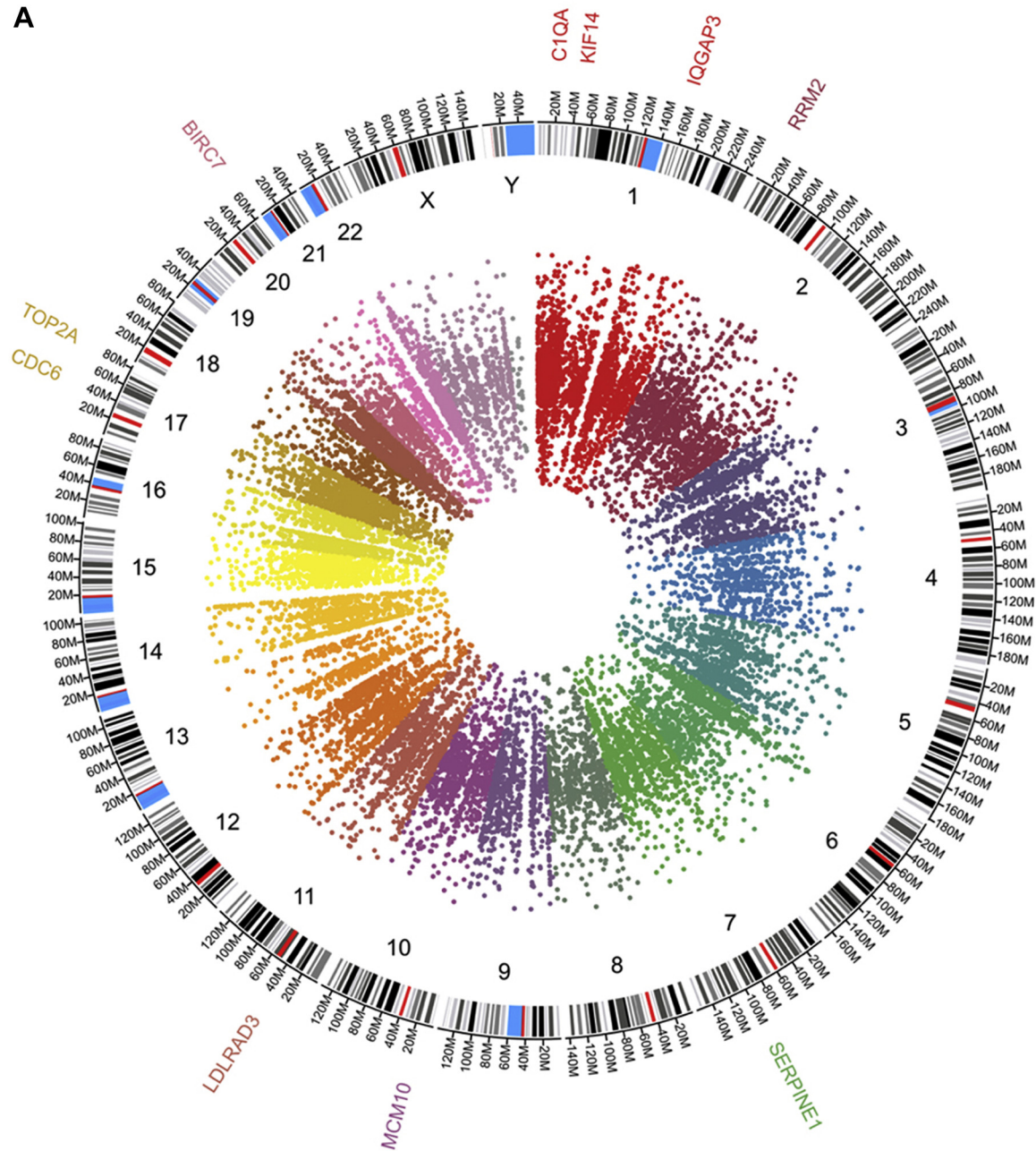


Figure 3. Panel A depict circo plot from the peripheral blood mononuclear cells transcriptome from COVID-19 patients. Inner ring is the scatter plot representing each transcript identified followed by chromosome number and size, the outer ring present genes identified participating in G1/S transition of mitotic cell cycle and protein processing. B and C panels show Gene Ontology at biological process and molecular function terms, respectively, from top 100 up-regulated genes. D and E panels show Gene Ontology at biological process and molecular function terms, respectively, from top 100 down-regulated genes.

hyaluronan accumulation has been documented in patients with ARDS, and the potential role of polysaccharides for the interstitial edema of the lung have not received much attention (11). Hyaluronan is glycosaminoglycan component of the extracellular matrix found in lung, and can absorb water 1000 times its molecular weight, and the alteration in hyaluronan synthesis and accumulation in ARDS is established (12). Our molecular findings in

bronchoalveolar cells regarding the upregulation of genes encoding proteins involved in hyaluronan, glycosaminoglycan, aminoglycan and mucopolysaccharide metabolism support the notion that these complex polysaccharides play indeed a role in the genesis of this severe form of interstitial pulmonary edema. Thus, inhibition of hyaluronan synthesis could represent a useful measure in the management of these patients (13).

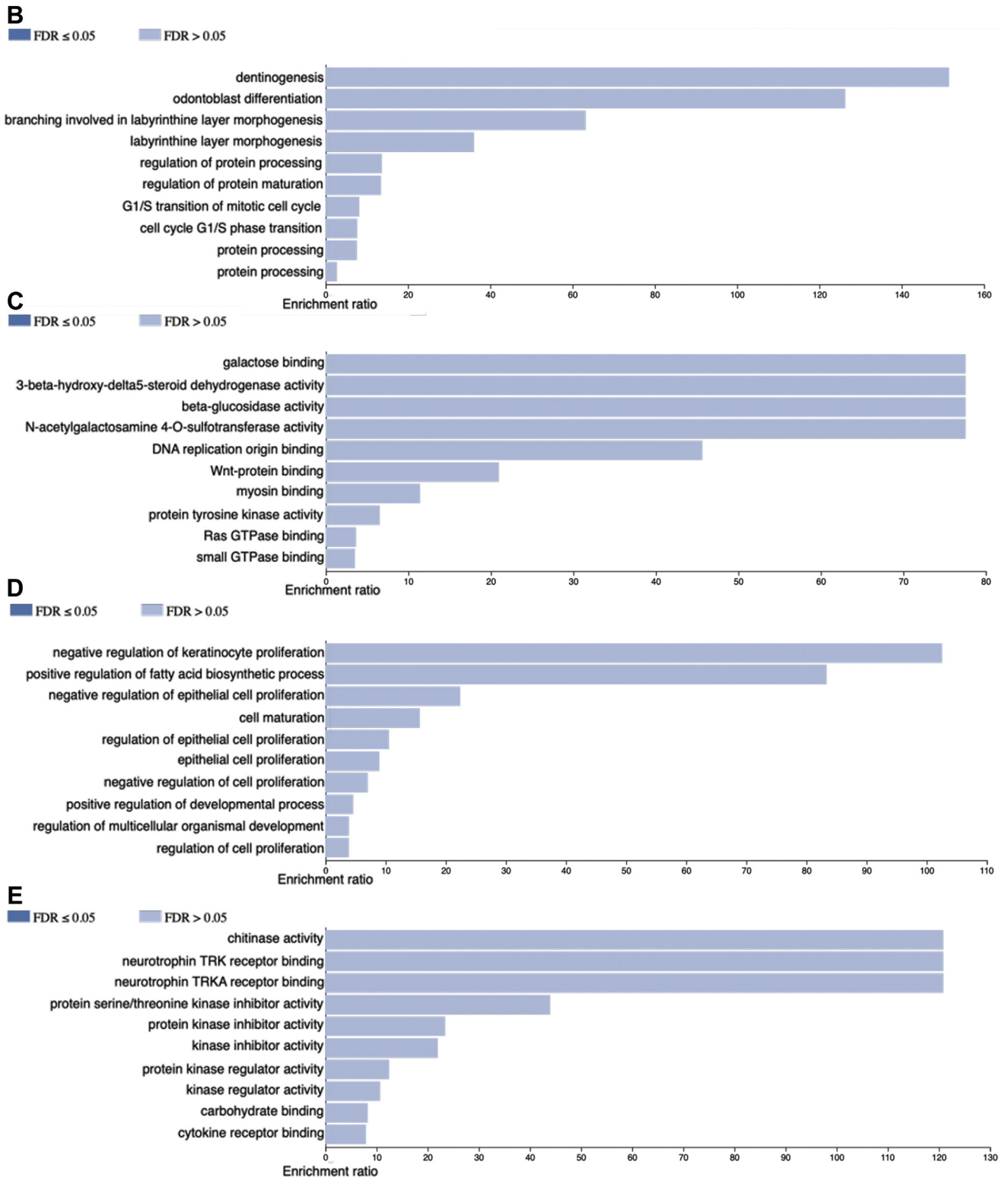


Figure 3. (Continued)

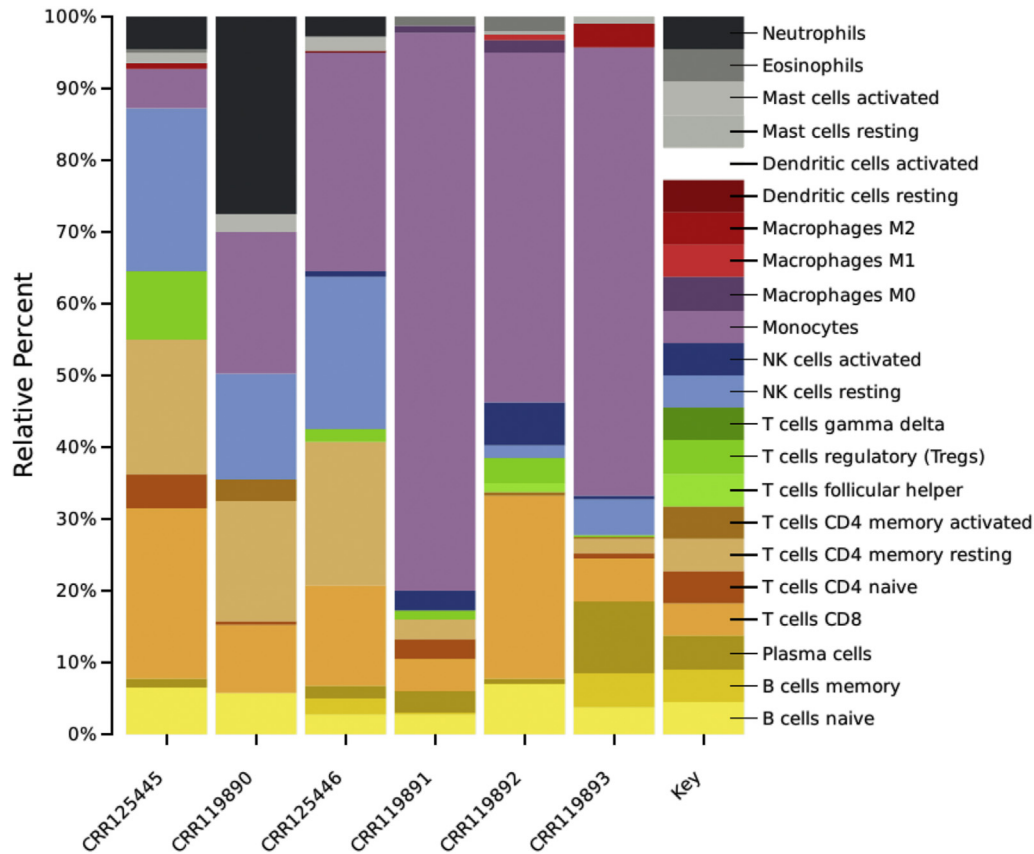


Figure 4. Deconvolution results by CIBERSORT algorithm show probability percentage of reduction of neutrophils, NK and T cells, whereas an exacerbated increase in monocytes in COVID-19 patients PBMC. CRR125445, 119890 and 125446 correspond to control PBMC whereas CRR119891, 119892 and 119893 correspond to COVID-19 PBMC.

Several inflammatory cytokines produced in COVID-19 could be strong inducers of hyaluronan metabolism (10) therefore the relevance to identify immune cells infiltrating lung tissue. Using molecular deconvolution analysis, we identified the presence of neutrophils, T CD4+ lymphocytes and macrophages infiltrating the lungs of COVID-19 patients, consistently our findings with the previously reported (14). Increased infiltrating macrophages are identified in autopsies and are thought to be responsible for fueling and perpetuating the inflammatory process (14). Neutrophils recruited to the site of infection can induce a prominent proinflammatory response directly by recognizing and clearing pathogens or by recruiting other immune cells (15). In summary, potentially the ARDS could be prompted by SARS-CoV-2 and fueled by host immune response.

Our in-silico findings revealed that peripheral mononuclear blood cells are in a proliferative state, perhaps, these proliferation events could represent clonal expansion of the remaining immune cells specific to viral infection. Interestingly, the molecular deconvolution analysis showed an evident shift of the immune profiles between COVID-19 and controls showing dramatic reduction of NK and T lymphocytes, whereas an exacerbated increase in monocytes.

Changes in myeloid, NK and B cells in the COVID-19 patients has been reported (16). Important lymphopenia is also observed, indicating an impairment of the immune system during the SARS-CoV-2 infection (17). During viral infections it is generally accepted that host immune responses determine both protection against viral infections and the pathogenesis of respiratory injury (16). Coordinated immune response, innate and adaptive, may lead to rapid control of virus, whereas failed immune response may lead to viral spreading, cytokine storm and eventually death (18).

An additional interesting observation was the potential alteration in mast cells in both BALF and PBMC data. Mast cells could be found in submucosa of the respiratory tract and the nasal cavity. Mast cells secrete L-6, IL-1, IL-31, IL-33 and TNF, and chemokines CC5, CCL2, MCP-1 and CXCL8 which attract white blood cells to the inflammatory sites. Experiments with coronavirus in mice have shown that this pathogenic RNA microorganism can activate the cells of both innate and adaptive immune systems, including MCs (19,20).

In conclusion the present data shows several molecular events that are potentially involved in the pathogenesis of SARS-CoV-2 infection and its pulmonary and systemic

complications. Specifically, our finding that genes encoding enzymes involved in hyaluronan and glycosaminoglycan metabolism are upregulated in pulmonary cells obtained by BAL of these patients could very well be the basis of the severe interstitial pneumonitis that characterizes this disease. Whether if our transcriptomic findings in PBMC indicating a proliferative state of immune cells reflect a clonal expansion resulting from recognition of specific viral antigens or simply a compensatory mechanism to replenish the number of immune cells, remains to be established.

Competing Interest

The authors declare no conflict of interest.

Acknowledgments

DMR is a recipient of the National Council for Science and Technology “Catedra CONACyT” program. This work was partially supported by grants 289499 from Fondos Sectoriales del Consejo Nacional de Ciencia y Tecnologia Mexico and R-2015-785-015 from Instituto Mexicano del Seguro Social (MM). The authors particularly thanks to the Instituto Mexicano del Seguro Social for the support to the formation of the Scientific Coalition for COVID-19 research.

References

1. Bi Q, Wu Y, Mei Y, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect Dis* 2020;20:30287–30295.
2. Nathan N, Prevost B, Corvol H. Atypical presentation of COVID-19 in young infants. *Lancet* 2020;10235:30980–30986.
3. Jose R, Manuel A. COVID-19 cytokine storm: the interplay between inflammation and coagulation. *Lancet Respir Med* 2020;8:e46–e47.
4. Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 2020;181:1–10.
5. Xiong Y, Liu Y, Cao L, et al. Transcriptomic characteristics of bronchoalveolar lavage fluid and peripheral blood mononuclear cells in COVID-19 patients. *Emerg Microbes Infect* 2020;9:761–770.
6. Hackbart M, Denga X, Bakera S. Coronavirus endoribonuclease targets viral polyuridine sequences to evade activating host sensors. *Proc Natl Acad Sci U S A* 2020;117:8094–8103.
7. Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020;395:507–513.
8. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395:497–506.
9. Lee E, Ng M, Khong P. COVID-19 pneumonia: what has CT taught us? *Lancet Infect Dis* 2020;4:384–385.
10. Shi Y, Wang Y, Shao C, et al. COVID-19 infection: the perspectives on immune responses. *Cell Death Differ* 2020;27:1451–1454.
11. Hällgren R, Samuelsson T, Laurent T, Modig J. Accumulation of hyaluronan (hyaluronic acid) in the lung in adult respiratory distress syndrome. *Am Rev Respir Dis* 1989;139:682–687.
12. Bell T, Brand O, Morgan D, et al. Defective lung function following influenza virus is due to prolonged, reversible hyaluronan synthesis. *Matrix Biol* 2019;80:14–28.
13. Collum S, Chen N, Hernandez A, et al. Inhibition of hyaluronan synthesis attenuates pulmonary hypertension associated with lung fibrosis. *Br J Pharmacol* 2017;174:3284–3301.
14. Liao M, Liu Y, Yuan J, et al. The landscape of lung bronchoalveolar immune cells in COVID-19 revealed by single-cell RNA sequencing. Preprint at medRxiv, 2020. <https://doi.org/10.1101/2020.02.23.20026690>. Accessed May 4, 2020.
15. Naumenko V, Turk M, Jenne C, et al. Neutrophils in viral infection. *Cell Tissue Res* 2018;371:505–516.
16. Wen W, Su W, Tang H, et al. Immune cell profiling of COVID-19 patients in the recovery stage by single-cell sequencing. Preprint at medRxiv, 2020. <https://doi.org/10.1101/2020.03.23.20039362>. Accessed May 4, 2020.
17. Wang F, Nie J, Wang H, et al. Characteristics of peripheral lymphocyte subset alteration in COVID-19 pneumonia. *J Infect Dis* 2020;221:1762–1769.
18. Wang Z, Wan Y, Qiu C, et al. Recovery from severe H7N9 disease is associated with diverse response mechanisms dominated by CD8⁺ T cells. *Nat Commun* 2015;6:6833.
19. Kritas S, Ronconi G, Caraffa A, et al. Mast cells contribute to coronavirus-induced inflammation: new anti-inflammatory strategy. *J Biol Regul Homeost Agents* 2020;34:1–14.
20. Theoharides T. COVID-19, pulmonary mast cells, cytokine storms, and beneficial actions of luteolin. *Biofactors* 2020;46:306–308.