Age- and sex-transcriptome analysis provide evidences for the sex biases in the pathogenesis of COVID-19 and other respiratory infectious diseases

Moran Gershoni (gmoran@volcani.agri.gov.il)

Research

Keywords: age-related, sex-related, surfactant proteins, respiratory infectious diseases, COVID-19

Posted Date: July 17th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-43708/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Age- and sex-transcriptome analysis provide evidences for the sex biases in the pathogenesis of COVID-19 and other respiratory infectious diseases

Moran Gershoni*

1. Institute of Animal Sciences, Agricultural Research Organization, the Volcani Center, Rishon LeZion, Israel

*gmoran@volcani.agri.gov.il

Keywords

age-related, sex-related, surfactant proteins, respiratory infectious diseases, COVID-19
Abstract

Age and sex were shown to affect the prevalence and the manifestation of many respiratory infectious diseases. These can be attributed to age and sex related alterations in the immune system and in the lung functions. Since the outbreak of COVID-19, epidemiological studies consistently report that age and sex are major risk factors in both morbidity and mortality due to COVID-19. Thus, understanding age and sex dependent gene expression in the lung and in the immune system can provide mechanical evidences with respects to sex related higher risk of elderly to develop severe complications in respiratory infectious diseases. In this context, sex- and age- transcriptome analysis from hundreds of lung and blood samples, revealed significant downregulation of the lung surfactant and blood innate immune genes, that occur predominantly in elderly men. Depletion in lung surfactant leads to enhanced injury of alveolar epithelium and fibrotic destruction, and recruitment of the innate immune system is essential to control infection of new pathogens like SARS-CoV-2. Interestingly, surfactant proteins, which protect the lung from infection, are co-produced with the SARS-CoV-2 host receptor- ACE2, by the AT2 cells. Thus, infection by SARS-CoV-2 is expected to lead to decline in AT2 cells and a loss of surfactant proteins, especially in elderly men.
Background

Many infectious diseases, such as influenza and pneumonia, are more common in the elderly, and are associated with poor outcome. Partially, these can be attributed to age related alterations in the immune system, including modifications in several component of the innate immune system, like alterations in secretion and response to cytokines[1]. Sex differences were also described for the manifestations of many infectious diseases, including infections caused by viruses. These sex differences were shown to be associated with differential immune regulation[2,3]. A recent study in mice suggested that the stronger immune response of females might be the results of more activated innate immune pathways prior to infection[4]. In December 2019, a new lethal infectious respiratory disease emerged in Wuhan China[5,6] . This coronavirus disease 19 (COVID-19) has, so far, caused more than 10 million confirmed cases (28 June 2020[7]). Sequence analysis found a novel coronavirus, SARS-CoV-2, which is closely related to SARS-CoV, as the cause of the new lung disease[8]. Early epidemiologic analysis of more than 70 thousand cases in China indicates that about 80% of COVID-19 patients experience only mild symptoms, but some 5% developed severe pneumonia that in many cases leads to death. The overall case fatality rate (CFR) was estimated as 2.3. In this analysis three notable trends were observed[9]:

1) The male CFR was almost twice as high than of females (2.8 vs 1.7 respectively)

2) CFR sharply increases as a function of age, where CFR for patients below 50 is less than 0.4 and approximately 15 for patients above 80

3) Background health conditions associate with elevated CFR

Similar observations were then consistently reported from different countries. For instance, in South Korea, where an intensive policy of testing for outbreak monitoring was implemented[10], the local CDC reported on more than 1,250,000 tests and more than 12,700 confirmed cases, with an overall 2.22 CFR.
At that time about 58% of the confirmed cases were females but their CFR was reported to be 1.8, compared to the male CFR of 2.77. The differences in the age-related CFR were even higher than in China, with less than 0.2 for patients younger than 50, and 25 for patients older than 80 (28 June 2020[11]). Similar age and sex biases were also reported by the New York City health department[12]. UK data from ~4000 COVID-19 patients that were admitted to intensive care units since the outbreak showed an almost 3-fold more male than female cases in critical condition (72.5% and 27.5% respectively). This work also reported an age-related survival rate for patient in critical condition ranging from 75% for young adults to only 27% for patients older than 75. Such differences have not been observed in pneumonia caused by other viral infections[13]. A recent study that summarized data from 38 countries reported average male CFR to be 1.7-fold higher than average female CFR[14]. Finally, a large study from France, that included serological and clinical data, reported age and male bias for hospitalization, intensive care units (ICU) admittance, and mortality rate. This study also reported that although this sex bias occurs at most age intervals, the male to female mortality rate ratio increases with age, and was almost 3 fold for patients older than 80[15]. Molecular and clinical studies have reported that, of all confirmed COVID-19 cases, up to 20% will develop pneumonia, and of them some will develop acute respiratory distress syndrome (ARDS), primarily patients older than 65[16].

The complete genome sequence of SARS-COV-2 was made public in early January 2020 (http://virological.org/). Initial sequence comparisons revealed that SARS-CoV-2 is approximately 79% identical to SARS-CoV at the nucleotide level[17]. Structural studies and biochemical experiments show SARS-CoV-2 virions to be optimized for binding to the human ACE2 receptor ([17]and references within). An additional study found that, like SARS-CoV, SARS-CoV-2 spread also depends on the proteolytic activity of human TMPRSS2 protease[18]. Thus, tissues expressing both genes are prone to infection and virus replication, as demonstrated in the respiratory and digestive systems[19].
Taken together, the above evidences suggesting that age-related sexual differences that exist prior to infection in the lung and in the immune system might explain age- and sex- biased pathogenesis observed in several infectious disease and specifically in COVID-19. Therefore, the hypothesis of this study is that age and sex related biological factors, specifically in the lung and the immune system, are the cause for the high susceptibility of elderly and male patients who contract SARS-Cov-2, to develop ARDS and a pathological inflammatory condition that leads to a higher mortality rate. Identification of genes that have differential expression (DE), specifically between elderly males and elderly females, may provide evidences for the age and sex biased pathogenesis of COVID-19 and other respiratory infectious diseases. To test this hypothesis, an age and sex differential gene expression analysis from hundreds of publicly available lung and blood samples, followed by comprehensive bioinformatics functional annotation and disease association, was carried out. These analyses revealed a significant age-related decline in the expression of the lung surfactant proteins (SFTPs) in men. In addition, downregulation of dozens of genes associated with the innate immune system in the blood, including the Toll-Like Receptor 4 (TLR4) signaling pathway, was found. All these differences were found to be predominant in older men as compared to older women. Deficiency in surfactant proteins genes leads to idiopathic pulmonary fibrosis and incomplete repair of injured alveolar epithelium, mainly in older adults[20,21]. The recruitment of the innate immune system may be a determinant for efficient protection of the host, in the case of SARS-CoV-2 infection, since this virus is newly introduced to the immune system and there is no pre-existing adaptive immune memory. Finally, in attempting to provide genetic markers that can pinpoint risk factors associated with the identified age- and sex- related genes, a screening of public databases identified eQTLs and GWAS QTLs that associated with the expression and malfunction of these genes. Such genetic markers should be further evaluated to assess their clinical significance in COVID-19, specifically in elderly.

Material and methods
Data acquisition. RNA-seq data, samples and donor annotations, genes annotations, and eQTLs information were obtained from the GTEx project version 8[22]. The population variation data and the minor alleles frequencies (MAFs) were obtained from the gnomAD database[23] (N=160,000).

Analysis of sex and age differential transcriptome. Age and sex differential expression was calculated as previously described[24]. First, all the samples were annotated according to the tissue and the donor’s age and sex. Then, all the samples from the Lungs (N_all = 578; N_men=395 N_women=183) and Blood (N_all= 755; N_men=501 N_women=183) were further obtained from the full dataset, and divided into four groups according to age (older or younger than 60) and sex. Only protein coding genes were included in the analysis. The NOISeqBIO[25,26] algorithm was then used to compare Transcripts Per Million (TPM) expression values between the following comparisons:

| Men younger than 60 | Men older than 60 |
|---------------------|-------------------|
| Women younger than 60 | Women older than 60 |
| Women older than 60 | Men older than 60 |

A NOISeqBIO probability cutoff of 0.95 was used to identify genes with significant differential expression, as this cutoff value was shown to correct for multiple testing in similar datasets[27]. Differentially expressed genes were annotated and enrichment analysis was performed using the GeneAnalytics server, which can identify gene enrichment for several terms and data sources, including diseases, biological pathways, GO terms, and tissue expression[28]. To test Co-localization of suspected eQTLs, genomic variants which were reported to be associated with diseases and traits relevant to COVID-19 (e.g., pulmonary diseases) were obtained from the GWAS atlas[29].

Results

The working hypothesis of this study is that significant age-related differences in genetic expression between men and women in the respiratory and the immune systems can reveal contributors to the age
and sex biases pathogenesis of COVID-19 and other respiratory infections. The immune system can be sufficiently monitored by profiling blood transcriptome[30] and the lung represent the most affected tissue in ARDS[16]. Thus, to test the working hypothesis, the following analyses were performed:

Differential expression in blood and lung between: Women older than 60 versus women younger than 60; men older than 60 versus men younger than 60; men older than 60 versus women older than 60.

The results of these analyses are shown in supplementary tables 1-6, and the number of differentially expressed (DE) genes identified in these analyses within and between the tested conditions are summarized in figure 1. Overall, about 3000 and 2000 age related DE (ADE) genes were identified in the blood and lungs, respectively. The vast majority of these genes showed age-related differential expression in men. 212/3000 of the blood genes and 20/2000 of the lung genes that have significant ADE also showed significant DE between older men to older women (figures 1a and 1b). Such genes are the main candidates to contribute to the age- and sex-related manifestation that have been observed in COVID-19. In order to assess the possible role of these ADE genes in ARDS and COVID-19, a functional analysis was carried out and the enrichment of these ADE genes to biological pathways, diseases, phenotypes and GO terms was computed using the GeneAnalytics tool[28]. The results of these analyses are presented in supplementary tables 7-12 and summarized in tables 1 and 2. Genes that undergo significant downregulation in the blood of old men compared to old women are mainly associated with the innate immune system response for viral infection (supplementary tables 7-9, table 1 and figure 2).

Among these genes are IFITM1 and IFITM2 which are part of the IFN Alpha/beta Signaling pathway (table 1 and figure 2) and were previously shown to be important for the resistance against RNA viruses, including SARS-CoV[31]. Another innate immune response pathway that was found to be downregulated is the Tol-Like-Receptor 4 (TLR4) signaling pathway (table 1, figure 2). Results from a previous study showed that SARS-CoV membrane protein may function as a cytosolic pathogen-associated molecular pattern (PAMP) which stimulates IFN-β production by activating TLR-related TRAF3-independent
signaling cascade[32]. In addition, several genes like SERPINB1 that are part of the innate immune response are also associated with lung diseases (table 1, figure 2). Interestingly, SERPINB1 was reported to have a protective immunomodulatory activity that prevents lung epithelial tissue injury[33]. Genes that undergo significant age-related downregulation in male lungs are specifically associated with pulmonary surfactant metabolism dysfunction and pulmonary fibrosis (supplementary tables 10-12, table 2, figure 3). The surfactant genes SFTPB and SFTPC have significant additional downregulation in old men compared to old women (supplementary tables 4-6, table 2 and figure 3). Functional deficiency of these genes leads to fatal neonatal respiratory distress and pulmonary alveolar proteinosis[34]. Age related downregulation of lung surfactant could be the results of reduced expression, but could also be due to alteration in the tissue cell’s composition. SFTPs genes were shown to be co-expressed with ACE2, the SARS-CoV-2 host receptor, by the alveolar type II (AT2) cells. Thus, changes in AT2 cells distribution in the lung tissue is expected to be reflected by ACE2 age and sex related expression like the SFTPs genes. However, ACE2 expression analysis revealed no such age or sex related differences in the lungs (supplementary tables 4-6 and figure 3). To this point, the DE analyses have pointed to the association of two systems with the high age and sex related susceptibility to develop severe complications in respiratory infectious diseases including in COVID-19: the innate immune response and the lung surfactant metabolism. Thus, one can ask if the likelihood of having downregulation in these pathways is independent or tends to be correlated. To answer this, blood and lung samples from the same donors were obtained (n= 154), and the Pearson correlation co-efficiency values were calculated between the blood innate immune system ADE genes (n=82) and the lung SFTPs genes expression. This analysis revealed that the expressions of the lung SFTPs genes are highly correlated with each other (r > 0.9) and, more importantly, the expression of the SFTPs genes tend to be correlated with most blood innate immune genes. Specifically, a significant high correlation between SFTPB, TLR5 and SERPINB1 (r = 0.76; p < 0.00001, figure 4) was found. These results suggest that patients with low expression of the
blood innate immune ADE genes (that can be detected from blood samples) also tend to have reduced
SFTPs expression. Finally, screening for genetic variants that associated with SFTPs gene dysfunction and
expression can be useful for further evaluation of the clinical significance of these lung genes in COVID-
19 and other respiratory diseases. To this end, the GTEx eQTL database was screened to identify lung
eQTLs that significantly alter the expression of the SFTPs genes. In addition, the GWAS atlas[29] was
screened to identify variants associated with relevant diseases and phenotypes (e.g. pulmonary diseases
and pulmonary functions) and are co-localized to the same eQTLs loci (+/- 0.5Mb). Such co-localized
variants are prime candidates to undergo further evaluation and assessment of clinical significance. The
results of these screens are summarized in table 3.

Discussion

Age related physiological changes are involved in impairing respiratory infections and its poor outcomes
in the elderly. This high vulnerability can be attributed to modified immune response and to other
alterations also occur in the lung as a consequence of the ageing process, which gradually lead to a
decline in lung function[35]. Sex related biological differences were also shown to affect the prevalence
and the course of many diseases. Many of these differences can be relate to differential gene
expression[36]. previous study found that in many cases, this sex differential expression is age
dependent[24]. Previously, a meta-analysis that preformed on more than 80 studies reported that males
are more commonly affected with infections in the lower respiratory tract, and that the course of most
respiratory diseases is more severe in males than in females, leading to higher mortality in males,
especially in community-acquired pneumonia[37].

One of the open questions since the outbreak of COVID-19 regards the causes for the age and sex biases
in the morbidity and mortality rates due to SARS-CoV-2 infection. Clinical, epidemiological, and
pathological evidences from COVID-19 studies demonstrate the involvement of the immune and the
respiratory systems, in an age and sex dependent manner, in the pathophysiology of the disease[15,38]. The increased sex- and age- related susceptibility is likely a result of naturally occurring differences between the sexes that express upon the properties of the SARS-CoV-2 infection. Sex related gene expression in the immune system was previously shown to affect the response to a pathogen prior to infection[4]. In this context, the age and sex differential transcriptome of lungs and blood was analyzed from hundreds of samples. These analyses revealed hundreds of ADE genes, mainly in men. The differences in the amount of ADE genes in these tissues between men and women could be explained by the differences in the sample sizes, which are much larger for men (for lung $N_{men}=395$; $N_{women}=183$, and for blood $N_{men}=501$; $N_{women}=183$). However, comparing the fold changes in gene expression between old to young, out of 19,179 protein coding genes that were analyzed, 745 and 225 genes have an average of 1.5-fold change or greater, in the lung of men versus women, respectively. This suggests that while the differences in the number of identified ADE genes between men and women was likely affected by the sample sizes, it should also be attributed to the biological differences between men and women in the aging process. Functional analysis of the ADE genes within and between men’s and women’s blood tissue revealed that the most affected age-related pathway in men is the innate immune system (table 1). Most of the blood ADE genes that associated with the innate immune system have a significantly reduced expression, which suggests a male age-related down regulation of the innate immune system. Functional analysis of ADE genes between elderly men and elderly women found the enrichment of a set of 82 genes of the innate immune system, which all had significantly reduced expression in elderly men as compared to elderly women. These findings suggest that women’s innate immune system is less affected by age. Some of the DE genes between elderly men to elderly women are specifically associated with the TLR4 signaling pathway (e.g., TLR2, TLR4 and TLR5, table 1 and figure 2). The innate immune response is the first mechanism of defense against infections. TLRs are important for the development and activation of innate immunity. The role of TLR4 is typically part of the response to the LPS of gram-
negative bacteria[39]. However, a recent study suggested that cell surface TLR4 is most likely to be involved in recognizing SARS-CoV-2 and to inducing inflammatory responses, and thus plays a crucial role in the initial virus-induced inflammatory consequences associated with COVID-19[40]. Another group of blood downregulated DE genes between elderly women to elderly men, are directly associated with lung diseases. Among these genes are SERPBINA1 and SERPBINB1. These genes have a serine protease elastase activity, and deficiency of serpinB1 has been shown to decrease the survival rate and increases morbidity associated with murine pulmonary influenza. This was shown to be likely due to enhanced injury of lung epithelial and the failing to downregulate pro-inflammatory cytokine production, similar to the pathological finding in COVID-19 and other respiratory infectious diseases[41]. Mouse models found that the regulation of pulmonary innate immunity by SERPINB1 is an essential process in the host response to infection, accomplished by controlling the recruitment of neutrophils to the infected lung. SERPINB1 deficiency leads to overproduction of pro-inflammatory cytokines that are associated with fatal outcome of human influenza[41,42]. Mutations in the SERPINA1 gene are known to cause Alpha-1 antitrypsin (AAT) deficiency. With insufficient functional AAT, neutrophil elastase destroys alveoli and typically causes age-related lung disease[43–45].

Functional analysis of the lung ADE genes within and between men and women reveals that the most affected age-related pathways differing between men and women are the surfactant metabolism and pulmonary fibrosis (table 2), including all the SFTPs genes. The surfactant proteins, encoded by the SFTPs genes, are expressed and produced by the AT2 cells and are essential to maintaining lung homeostasis. Deficiency in these genes has been found to cause idiopathic pulmonary fibrosis and incomplete repair of injured alveolar epithelium, mainly in older adults[20,21]. Interestingly, SARS-CoV-2 host receptor, ACE2, was also found to be expressed by AT2 cells[46]. Investigation of histopathology sections from COVID-19 patients’ discovered SARS-CoV-2 particles in AT2 epithelia, which was confirmed by RT-PCR[47]. Thus, the lung AT2 cells which are likely the SARS-CoV-2’s target are also the producers
of the surfactant proteins. Depletion of AT2 cells due to SARS-CoV-2 infection, or other pathogens that targeting the AT2 CELLS, is therefore expected to cause surfactant deficit, which has previously been shown to be associated with incomplete repair of injured alveolar epithelium and fibrotic elimination[20]. Similar to surfactant deficit, the main histological finding from autopsies taken from lungs of COVID-19 patients includes injury to the alveolar epithelial cells and fibrosis[48]. Altogether, these results show that the basal expression levels of the SFTPs genes in men older than 60 tend to be very low (figure 3, Supp. Table 4-6) as compared to younger men and women in all ages. Thus, older men are expected to be much more vulnerable to respiratory infection, specifically by pathogen that infect the lung AT2 cells, like in SARS-CoV-2 infection, due to the additional depletion of AT2 cell and loss of surfactant proteins.

**Conclusions**

Altogether, the findings of this study provide evidence for the high prevalence and poor outcome of elderly, specifically of males, to respiratory infectious diseases, and can illuminate a possible sex biased mechanism in COVID-19 pathogenesis: In the absence of adaptive immune memory, downregulation of the innate immune system in the elderly, especially in old men, suppresses their initial response to SARS-CoV-2 infection, likely through genes such as the TLRs. This, in turn, reduces virus clearance and increases the susceptibility of old men to developing pulmonary infection. The SARS-CoV-2 target host cells in the lungs are the AT2 cells, the producers of the surfactant proteins. The initial expression levels of the surfactant genes tend to be much lower in old men. Thus, additional decline of AT2 cells due to viral infection will abolish surfactant proteins and thus promote massive injury of the alveolar epithelium and fibrosis. This pathological process enhances hyper-inflammation that should be modulated by additional components of the innate immune system, like the SERPINs genes. However, the expression of these genes is the lowest in elderly men. Since the expression of the innate immune
ADE genes and the SFTPs genes tend to be correlated, it is likely that subgroup of elderly male COVID-19 patients, will have an overall poor response to SARS-CoV-2 infection in both systems.

**Ethics approval and consent to participate**

Not applicable

**Consent for publication**

Not applicable

**Availability of data and materials**

All the data is publicly available. The GTEx Analysis V8 RNA-seq and eQTL data are available on the GTEx portal (http://www.gtexportal.org/home/datasets).

**Competing interests**

The author declares no competing interest

**Funding**

This work was partially supported by a grant from the Israeli Science Foundation (ISF 946/19 to M.G.). The grant foundations had no competing interest.

**Authors’ contributions**

M.G. designed the study, performed the analyses, performed data mining and wrote the paper

**Acknowledgements**

I want to thank Shmuel Pietrokovski for helpful discussion and for pointing me to relevant literature. I want to thank Aviva Isaacs for critical reading of the manuscript.

**References**
1. Gavazzi G, Krause KH. Ageing and infection. Lancet Infect. Dis. Lancet Publishing Group; 2002. p. 659–66.

2. Klein SL, Jedlicka A, Pekosz A. The Xs and Y of immune responses to viral vaccines [Internet]. Lancet Infect. Dis. 2010 [cited 2020 Jul 7]. p. 338–49. Available from: https://www.sciencedirect.com/science/article/pii/S1473309910700499

3. Fish EN. The X-files in immunity: Sex-based differences predispose immune responses [Internet]. Nat. Rev. Immunol. 2008 [cited 2020 Jul 7]. p. 737–44. Available from: https://www.nature.com/articles/nri2394

4. Gal-Oz ST, Maier B, Yoshida H, Seddu K, Elbaz N, Czysz C, et al. ImmGen report: sexual dimorphism in the immune system transcriptome. Nat Commun [Internet]. Nature Publishing Group; 2019 [cited 2020 Jul 7];10:1–14. Available from: https://doi.org/10.1038/s41467-019-12348-6

5. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet [Internet]. 2020 [cited 2020 Apr 19];395:497–506. Available from: https://www.sciencedirect.com/science/article/pii/S0140673620301835

6. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern [Internet]. Lancet. 2020 [cited 2020 Apr 19]. p. 470–3. Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30185-9/fulltext

7. COVID-19 situation reports [Internet]. [cited 2020 Apr 19]. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

8. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. Massachusetts Medical Society; 2020;382:727–33.

9. China CDC. METHOD Study Design  •. cdn.onb.it [Internet]. 2020 [cited 2020 Apr 19];5. Available from:
10. Al-Rousan N, Al-Najjar H. Data Analysis of Coronavirus CoVID-19 Epidemic in South Korea Based on Recovered and Death Cases. J Med Virol [Internet]. 2020 [cited 2020 Apr 19]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/32270521

11. List | Press Release | News Room : KCDC [Internet]. [cited 2020 Apr 19]. Available from: https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030

12. COVID-19: Data - NYC Health [Internet]. [cited 2020 Apr 19]. Available from: https://www1.nyc.gov/site/doh/covid/covid-19-data.page#download

13. ICNARC – Reports [Internet]. [cited 2020 Apr 19]. Available from: https://www.icnarc.org/Our-Audit/Audits/Cmp/Reports

14. Scully EP, Haverfield J, Ursin RL, Tannenbaum C, Klein SL. Considering how biological sex impacts immune responses and COVID-19 outcomes. Nat. Rev. Immunol. Nature Research; 2020.

15. Salje H, Tran Kiern C, Lefrancq N, Courtejoie N, Bosetti P, Paireau J, et al. Estimating the burden of SARS-CoV-2 in France. Science (80- ). American Association for the Advancement of Science (AAAS); 2020;eabc3517.

16. Mason RJ. Pathogenesis of COVID-19 from a cell biology perspective [Internet]. Eur. Respir. J. European Respiratory Society; 2020 [cited 2020 Jun 23]. p. 2000607. Available from: http://erj.ersjournals.com/lookup/doi/10.1183/13993003.00607-2020

17. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. Nat Med. Springer Science and Business Media LLC; 2020;26:450–2.

18. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell
Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell. Cell Press; 2020;181:271-280.e8.

19. Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, et al. Virological assessment of hospitalized patients with COVID-19. Nature. Springer Science and Business Media LLC; 2020;1–10.

20. Li Y, Wu Q, Sun X, Shen J, Chen H. Organoids as a Powerful Model for Respiratory Diseases [Internet]. Stem Cells Int. 2020 [cited 2020 Jun 23]. Available from: https://www.hindawi.com/journals/sci/2020/5847876/

21. Wang Y, Kuan PJ, Xing C, Cronkhite JT, Torres F, Rosenblatt RL, et al. Genetic Defects in Surfactant Protein A2 Are Associated with Pulmonary Fibrosis and Lung Cancer. Am J Hum Genet. 2009;84:52–9.

22. Melé M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, et al. The human transcriptome across tissues and individuals. Science (80-. ). American Association for the Advancement of Science; 2015;348:660–5.

23. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. bioRxiv [Internet]. Cold Spring Harbor Laboratory; 2019 [cited 2020 Apr 19];531210. Available from: https://www.biorxiv.org/content/10.1101/531210v3

24. Gershoni M, Pietrokovski S. The landscape of sex-differential transcriptome and its consequent selection in human adults. BMC Biol. 2017;15.

25. Tarazona S, García-Alcalde F, Dopazo J, Ferrer A, Conesa A. Differential expression in RNA-Seq: A matter of depth. Genome Res. Cold Spring Harbor Laboratory Press; 2011;21:2213–23.

26. Tarazona S, Furió P, Ferrer A, Conesa A. NOISeq: An R package for differential expression in RNA-Seq using biological replicates. 2013.
27. Tarazona S, Furió-Tari P, Turrà D, Di Pietro A, Nueda MJ, Ferrer A, et al. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. Nucleic Acids Res [Internet]. 2015 [cited 2020 Apr 19];43. Available from: https://academic.oup.com/nar/article-abstract/43/21/e140/2468096

28. Fuchs SBA, Lieder I, Stelzer G, Mazor Y, Buzhor E, Kaplan S, et al. GeneAnalytics: An Integrative Gene Set Analysis Tool for Next Generation Sequencing, RNAseq and Microarray Data. Omi A J Integr Biol. Mary Ann Liebert Inc.; 2016;20:139–51.

29. Watanabe K, Stringer S, Frei O, Umčević Mirkov M, de Leeuw C, Polderman TJC, et al. A global overview of pleiotropy and genetic architecture in complex traits. Nat Genet. Nature Publishing Group; 2019;51:1339–48.

30. Chaussabel D, Pascual V, Banchereau J. Assessing the human immune system through blood transcriptomics [Internet]. BMC Biol. BioMed Central; 2010 [cited 2020 Jun 28]. p. 1–14. Available from: https://link.springer.com/articles/10.1186/1741-7007-8-84

31. Huang IC, Bailey CC, Weyer JL, Radoshitzky SR, Becker MM, Chiang JJ, et al. Distinct patterns of IFITM-mediated restriction of filoviruses, SARS coronavirus, and influenza A virus. PLoS Pathog. 2011;7.

32. Wang Y, Liu L. The membrane protein of severe acute respiratory syndrome coronavirus functions as a novel cytosolic pathogen-associated molecular pattern to promote beta interferon induction via a toll-like-receptor-related TRAF3-independent mechanism. MBio [Internet]. American Society for Microbiology; 2016 [cited 2020 Jun 23];7. Available from: https://mbio.asm.org/content/7/1/e01872-15

33. Zhao P, Hou L, Farley K, Sundrud MS, Remold-O’Donnell E. SerpinB1 regulates homeostatic expansion of IL-17 + γδ and CD4 + Th17 cells. J Leukoc Biol [Internet]. 2014 [cited 2020 Jun 23];95:521–30. Available from: https://jlb.onlinelibrary.wiley.com/doi/abs/10.1189/jlb.0613331

34. Wallot M, Wagenvoort C, Demello D, Müller KM, Floros J, Roll C. Congenital alveolar proteinosis
caused by a novel mutation of the surfactant protein B gene and misalignment of lung vessels in consanguineous kindred infants. Eur J Pediatr. 1999;158:513–8.

35. Haq K, McElhaney JE. Ageing and respiratory infections: The airway of ageing. Immunol Lett. Elsevier; 2014;162:323–8.

36. Ober C, Loisel DA, Gilad Y. Sex-specific genetic architecture of human disease [Internet]. Nat. Rev. Genet. Nature Publishing Group; 2008 [cited 2020 Jul 8]. p. 911–22. Available from: www.nature.com/reviews/genetics

37. Falagas ME, Mourtzoukou EG, Vardakas KZ. Sex differences in the incidence and severity of respiratory tract infections. Respir. Med. W.B. Saunders; 2007. p. 1845–63.

38. Rivellese F, Prediletto E. ACE2 at the centre of COVID-19 from paucisymptomatic infections to severe pneumonia. Autoimmun. Rev. Elsevier B.V.; 2020.

39. Piras V, Selvarajoo K. Beyond MyD88 and TRIF pathways in Toll-like receptor signaling. Front Immunol [Internet]. Frontiers Research Foundation; 2014 [cited 2020 Jun 24];5. Available from: www.frontiersin.org

40. Choudhury A, Mukherjee S. In silico studies on the comparative characterization of the interactions of SARS-CoV-2 spike glycoprotein with ACE-2 receptor homologs and human TLRs. J Med Virol. John Wiley and Sons Inc.; 2020;

41. Gong D, Farley K, White M, Hartshorn KL, Benarafa C, Remold-O'Donnell E. Critical role of SerpinB1 in regulating inflammatory responses in pulmonary influenza infection. J Infect Dis [Internet]. 2011 [cited 2020 Jun 24];204:592–600. Available from: https://academic.oup.com/jid/article-abstract/204/4/592/2192600

42. Benarafa C, LeCuyer TE, Baumann M, Stolley JM, Cremona TP, Remold-O'Donnell E. SerpinB1
protects the mature neutrophil reserve in the bone marrow. J Leukoc Biol. Wiley-Blackwell; 2011;90:21–9.

43. Hersh CP, Campbell EJ, Scott LR, Raby BA. Alpha-1 antitrypsin deficiency as an incidental finding in clinical genetic testing. Am. J. Respir. Crit. Care Med. American Thoracic Society; 2019. p. 246–8.

44. Tasch JJ, McLaughlan AT, Nasir AA. A novel approach to screening for alpha-1 antitrypsin deficiency: Inpatient testing at a teaching institution. Chronic Obstr Pulm Dis. COPD Foundation; 2018;5:106–10.

45. Gramegna A, Aliberti S, Confalonieri M, Corsico A, Richeldi L, Vancheri C, et al. Alpha-1 antitrypsin deficiency as a common treatable mechanism in chronic respiratory disorders and for conditions different from pulmonary emphysema? A commentary on the new European Respiratory Society statement 11 Medical and Health Sciences 1102 Cardi. Multidiscip. Respir. Med. BioMed Central Ltd.; 2018.

46. Ni L, Zhou L, Zhou M, Zhao J, Wang DW. Combination of western medicine and Chinese traditional patent medicine in treating a family case of COVID-19. Front Med [Internet]. 2020 [cited 2020 Jun 24];14:210–4. Available from: https://doi.org/10.1007/s11684-020-0757-x

47. Yao XH, Li TY, He ZC, Ping YF, Liu HW, Yu SC, et al. [A pathological report of three COVID-19 cases by minimally invasive autopsies]. Zhonghua bing li xue za zhi = Chinese J Pathol [Internet]. 2020 [cited 2020 Jun 24];49:E009. Available from: http://www.ncbi.nlm.nih.gov/pubmed/32172546

48. Polak SB, Van Gool IC, Cohen D, von der Thüsen JH, van Paassen J. A systematic review of pathological findings in COVID-19: a pathophysiologica timeline and possible mechanisms of disease progression. Mod Pathol [Internet]. Nature Publishing Group; 2020 [cited 2020 Jun 24];1–11. Available from: http://www.nature.com/articles/s41379-020-0603-3
Table 1. Summary of pathways and diseases associated with blood ADE genes

| GA Score | SuperPath/ Disease Name | Matched Genes |
|----------|-------------------------|---------------|
|          | Blood up regulated in older men |               |
| 21.08    | Innate Immune System     | CEACAM8, JUN, HSP90AB1, FBXO7, DUSP2, EPX, DUSP1, GZMM, HSPA1A, CCL4, MPO, ELANE, LTF, CCL3, PLAUR, SLC2A3, HSPA5 |
| 9.14     | Rheumatoid Arthritis     | HSPA5, DUSP1, DDIT4, CCL3, CXCR4, ELANE, BAG3, CYTIP, CCL4, MPO, JUN, PER1, NKG7, PLIN2, JUND |
| 18.36    | Chemokine Superfamily    | CCL4L2, CXCR4, CCL4, CCL3, XCL2 |
|          | Pathway: Ligand-Receptor |               |
|          | Blood down regulated in older men |               |
| 166.10   | Innate Immune System     | PAK2, METTL7A, CD55, ATP6V1D, ATP6V1H, CD53, CDC42, CD93, PDE1B, NLRC4, CD79B, ARPC1B, CAPZA2, CASP4, CD177, MVP, NFKB1, CD300LF, MYD88, MANBA, MAPK3, CD33, CD1D, CD47, CD59, CDA, LY86, MAP2K3, NCF2, CD58, CKAP4, NFAM1, CD46, CEBPD, BAD, |
|                         | Gene Names                                      |
|-------------------------|------------------------------------------------|
| Class I MHC Mediated Antigen Processing and Presentation | BCL10, LTB, CCR1, PDKX, PIAS1, PECAM1, MMP25, FBXW2, LYZ, MEFV, FGR [and sup. Table 8] |
| Antigen Processing-Cross Presentation | CTSB, CUL3, ICAM3, CLTC, CD36, FBXW5, IRS2, LAIR1, IFI30, IFITM1, GRB2, LILRB4, GSK3A, FBXL3, FBXO40, FBXO9, ITGβ2, FBXO30, FGF8, AGO1, FCGR1A, CD274, FBXL5, ANAPC13, AP2S1, CSK, KLHL2, ACTR1A, ANAPC5, KLHL21, CTSA, CUL2, FCGR1B, LILRB3, LILRA6, CD14, CD300LD, CTSK, AP1M2, KBTBD7, AGO4, CTSC, ASB7, LILRA5 |
| Toll-like Receptor Signaling Pathway | CD36, FCGR1A, FCGR1B, CD14, CYBA, B2M, BTK, CTSS, MYD88, NCF2, LY96, NCF4, NCF1, PSMA6, PSMD13, PSMB7, PSMC2, TIRAP, TLR4, PSMB1, PSMD6, TLR1, PSMA1, PSMC1, TLR2 [and sup. Table 8] |
| Pathway                                      | p-value | Gene Symbols                                                                 |
|----------------------------------------------|---------|-------------------------------------------------------------------------------|
| Phagosome                                    | 24.02   | CLEC7A, ATP6V1E1, CD36, ACTG1, ACTB, C1R, Fcgr3B, Itgb2, Fcgr1A, Hgs, Cd14,   |
|                                              |         | ATP6V0B, DynC1I2, Cyba, DynC1l1,                                            |
|                                              |         | ATP6V0E1, ATP6V1E2, Fcgr2A, ATP6V0D1,                                         |
|                                              |         | ATP6V1B2, ATP6V1G1 [and sup. Table 8]                                        |
| Blood Down regulated in older men compared  | 85.93   | Cxcr1, Ifi16, Gbp2, Clec4e, Cd63, Il1r2,                                   |
| to older women                               |         | Gmfg, Ifitm1, Cystm1, Cap1, Fcgr3B,                                        |
|                                              |         | Irak3, Gyg1, Il17ra, Il18r1, Il18rap, Alox5,                                |
|                                              |         | Arg1, Hck, Cxcr2, Fpr1, Lirlb3, Adgrg3,                                   |
|                                              |         | Fcer1g, Lirla5, Apbb1ip, Ifngr1, Bcl2l1,                                   |
|                                              |         | Il13ra1, Gca, Ifitm2, Atp6v0e1, Krt1,                                      |
|                                              |         | Fcgr2a, Dok3 [and sup. Table 9]                                             |
| Innate Immune System                         | 14.49   | Irak3, Tlr4, Rps6ka1, S100a8, S100a9,                                       |
|                                              |         | Ube2d1, S100a12, Tlr5, Tlr2                                                |
| Activated TLR4 Signaling                     | 9.69    | Anxa3, Arg1, Casp5, Cd177, Clec4e, Ifitm1,                                 |
|                                              |         | Ifitm2, Il18r1, Mak, Olah, Olfm4, Rps4x,                                  |
|                                              |         | Tlr2, Tmem119, Zdhhc19                                                     |
| Influenza                                    | 9.69    | Anxa1, Arg1, Cxcr1, Cxcr2, Ifngr1,                                        |
|                                              |         | Ifitm1, Cxcr1, Cxcr2, Ifngr1,                                              |
|                                              |         | Ifitm2, Il18r1, Mak, Olah, Olfm4, Rps4x,                                  |
|                                              |         | Tlr2, Tmem119, Zdhhc19                                                     |
| Lung Disease                                 | 12.63   | Anxa1, Arg1, Cxcr1, Cxcr2, Ifngr1,                                        |
|                                              |         | S100a8, S100a9, Serpina1, Serpinb1,                                        |
|                                              |         | Slc11a1, Tlr2, Tlr4, Tlr5, Txn                                               |

Table 2. Summary of pathways and diseases associated with Lung ADE genes
| GA top Scores | SuperPath/ Disease Name | Matched Genes |
|---------------|--------------------------|----------------|
| **Lung up regulated in older men** | | |
| 10.39 | Defensins | 3 | DEFA5, ITLN1, CAMP |
| 5.31 | Diabetes Mellitus, Noninsulin-Dependent | 7 | APOA2, ESM1, ITLN1, GH1, SCT, REG1B, REG1A |
| 4.78 | Crohn's Disease | 5 | HCAR3, REG1A, REG1B, IFNG, DEFA5 |
| **Lung down regulated in older men** | | |
| 58.50 | Surfactant Metabolism | 11 | ABCA3, DMBT1, CTH, PGA4, NAPSA, SFTPB, SFTPD, SFTPC, SLC34A2, SFTPA2, SFTPA1 |
| 31.27 | Surfactant Metabolism Dysfunction 4 (SMDP4) | 5 | SFTPB, SFTPD, SFTPC, SFTPA2, SFTPA1 |
| | Defective CSF2RA Causes Pulmonary Surfactant Metabolism Dysfunction 4 (SMDP4) | | ABCA3, ALDH3A1, FMOS5, CRTAC1, BPIFB1, LAMP3, PEBP4, MFSD2A, PLA2G4F, PLA2G3, SFTPA1, SFTPA2, SFTPC, SFTPD, ZNF385B |
| 13.50 | Pulmonary Fibrosis | 15 | ABCA3, ALDH3A1, FMOS5, CRTAC1, BPIFB1, LAMP3, PEBP4, MFSD2A, PLA2G4F, PLA2G3, SFTPA1, SFTPA2, SFTPC, SFTPD, ZNF385B |
| 9.02 | Lung Disease | 10 | SFTPA1, SFTPA2, SFTPB, SFTPC, SFTPD, SLPI, NPPB, CFTR, ABCA3, MUC1 |
| 8.78 | Respiratory Failure | 9 | CFTR, ABCA3, NPPB, REN, FGA, NPPA, SFTPC, SFTPD, SFTPB |
### Table 3. Summary of genetic variants associated with Lung SFTPs genes expression and function.

| Gene      | Top eQTLs hits                      | Top GWAS hits (phenotype)                                      |
|-----------|-------------------------------------|---------------------------------------------------------------|
| SFTPB     | rs4832183; rs4832182                 | -                                                             |
| SFTPC     | -                                   | rs8192327 (FEV1/FVC)                                          |
| SFTPA1    | rs1650158                           | rs2559728 (FEV1/FVC)                                          |
| SFTPA2    | rs1650158; rs35565917               | rs2559728 (FEV1/FVC)                                          |
| SFTPD     | rs149407882                          | rs1923539-A; rs1923539-A; rs3923564-G (Chronic obstructive pulmonary disease) |

**Lung Down regulated in older men compared to older women**

- Defective CSF2RA Causes Pulmonary Surfactant Metabolism Dysfunction 4 (SMDP4)
  - SFTPB, SFTPC

- Respiratory Distress Syndrome in Premature Infants
  - SFTPB, SFTPC
Figure 1. Numbers of identified ADE genes in all analyses. Schematic representation of the DE genes found between the tested conditions in the lung (a) and in the blood (b).

Figure 2. Expression of blood ADE genes. Distribution of gene expression levels of blood ADE genes, in women older than 60 (pink), women younger than 60 (red), men older than 60 (light blue), and men younger than 60 (blue). The TPM values are presented on the Y-axis.

Figure 3. Expression of lung ADE genes. Distribution of gene expression levels of lung ADE genes, in women older than 60 (pink), women younger than 60 (red), men older than 60 (light blue), and men younger than 60 (blue). The TPM values are presented on the Y-axis.

Figure 4. Surfactants genes expression tend to be correlate with the ADE genes of the innate immune system. The Correlation matrix as a heatmap, between the lung SFTPs genes (X-axis) and the DE innate immune genes between elderly men to elderly women (Y-axis). The matrix is sorted according to the highest correlation scores between the innate immune ADE genes to lung SFTPB.
Figures

Figure 1

Numbers of identified ADE genes in all analyses. Schematic representation of the DE genes found between the tested conditions in the lung (a) and in the blood (b).

Figure 2
Expression of blood ADE genes. Distribution of gene expression levels of blood ADE genes, in women older than 60 (pink), women younger than 60 (red), men older than 60 (light blue), and men younger than 60 (blue). The TPM values are presented on the Y-axis.

Figure 3

Expression of lung ADE genes. Distribution of gene expression levels of lung ADE genes, in women older than 60 (pink), women younger than 60 (red), men older than 60 (light blue), and men younger than 60 (blue). The TPM values are presented on the Y-axis.
Figure 4

Surfactants genes expression tend to be correlate with the ADE genes of the innate immune system. The Correlation matrix as a heatmap, between the lung SFTPs genes (X- axis) and the DE innate immune genes between elderly men to elderly women (Y- axis). The matrix is sorted according to the highest correlation scores between the innate immune ADE genes to lung SFTPB.

Supplementary Files

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

- Supp.table7.GeneAnalyticsBloodmenUpregulated.xlsx
- Supp.table1.Blood.Male.ADE.csv
- Supp.table8.GeneAnalyticsBloodmenDownregulated.xlsx
- Supp.table3.Blood.oldmen.oldwomen.DE.csv
- Supp.table9.GeneAnalyticsBlood.oldmen.oldwomen.xlsx
- Supp.table10.GeneAnalyticsLung.menupregulated.xlsx
- Supp.table2.Blood.Female.ADE.csv
- Supp.table11.GeneAnalyticsLung.mendownregulated.xlsx
- Supp.table4.Lung.Male.ADE.csv
- Supp.table12.GeneAnalyticsLung.oldmen.oldwomen.xlsx
• Supp.table6.Lung.oldmen.oldwomen.DE.csv
• Supp.table5.Lung.Female.ADE.csv