Baseline pulmonary levels of CD8+ T cells and NK cells inversely correlate with expression of the SARS-CoV-2 entry receptor ACE2

Pascal H.G. Duijf [1, 2 *]

[1] Institute of Health and Biomedical Innovation, Queensland University of Technology (QUT), Faculty of Health, School of Biomedical Sciences, 37 Kent Street, Brisbane, QLD 4102, Australia.

[2] University of Queensland Diamantina Institute, The University of Queensland, Translational Research Institute, 37 Kent Street, Brisbane, QLD 4102, Australia.

[*] Correspondence: Pascal H.G. Duijf, email: pascal.duijf@qut.edu.au

Abstract

Coronavirus disease 19 (COVID-19) is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and currently has detrimental human health, community and economic impacts around the world. In SARS-CoV-2-infected lung, induction of immune cell-recruiting cytokines is initially poor. When induction does occur, this may in fact exacerbate infection. These observations suggest that baseline levels of leukocytes, already residing in the lung prior to infection, may be important for orchestrating an effective early immune response. Hence, we performed "*in silico* flow cytometry" on 1,927 human lung tissues to deconvolute the levels of seven leukocyte types involved in triggering an acute anti-viral cellular immune response. Baseline levels of CD8+ T cells, resting natural killer (NK) cells and activated NK cells are significantly lower in lung tissues with high expression of the SARS-CoV-2 host cell entry receptor ACE2. We observe this in univariate analyses, in multivariate analyses that include sex, age, race, body mass index and smoking history, and in two independent datasets. Elevated ACE2 expression increases sensitivity to coronavirus infection. Thus, our results suggest that a subgroup of individuals may be exceedingly susceptible to COVID-19 due to concomitant high pre-existing ACE2 expression and low baseline cytotoxic lymphocyte levels in the lung.

Keywords: COVID-19, SARS-CoV-2, ACE2, T cells, NK cells

Introduction

Coronaviruses are viruses belonging to the family of *Coronaviridae* [1]. They are large, single-stranded RNA viruses that often originate from bats and commonly infect mammals. While the majority of coronavirus infections cause mild symptoms, some can cause severe symptoms, such as pneumonia, respiratory failure and sepsis, which may lead to death [2, 3].

Coronavirus zoonosis constitutes a serious health risk for humans. Indeed, in recent history, transmissions of three types of coronaviruses to humans have led to varying numbers of deaths. The outbreak of the Severe Acute Respiratory Syndrome (SARS) epidemic, which is caused by the SARS coronavirus (SARS-CoV), originated in Guangdong, China in 2002 and led to nearly 800 deaths [4]. The Middle East Respiratory Syndrome coronavirus (MERS-CoV) outbreak, which emerged in Saudi Arabia in 2012, similarly caused about 800 deaths, although with over 8,000 cases, nearly four times as many cases were reported [4]. Finally, Coronavirus Disease 19 (COVID-19), caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), is currently causing a pandemic. On 1 May 2020, the World Health Organization reported over 3 million confirmed cases and over 220,000 patients to have succumbed to COVID-19 around the world [5]. However, the factual number of deaths is probably considerably higher [6]. In addition, this figure is still soaring, on 1 May 2020 at a rate exceeding 6,400 deaths per day [5].

To infect target cells, coronaviruses use their spike (S) glycoprotein to bind to receptor molecules on the host cell membrane. Angiotensin-converting enzyme 2 (ACE2) has been identified as the main SARS-CoV-2 entry receptor on human cells [7, 8], while the serine protease TMPRSS2, or potentially cathepsin B and L, are used for S-protein priming to facilitate host cell entry [7]. SARS-CoV-2 S-protein has a 10 to 20-fold higher affinity to human ACE2 than SARS-CoV S-protein [9]. Moreover, ACE2 expression proportionally increases the susceptibility to S protein-mediated coronavirus infection [10-12]. Hence, increased expression of ACE2 is thought to increase susceptibility to COVID-19 [13-15].

Epithelial cells of the respiratory tract, including the lung, are thought to be primary SARS-CoV-2 target cells [16-18]. These cells can sense viral infection via pattern recognition receptors (PRRs). PRRs, including Toll-like receptors and NOD-like receptors, recognize pathogen-associated molecular patterns (PAMPs) [19]. Upon PRR activation, a range of pro-inflammatory cytokines and chemokines are produced and released in order to activate the host's immune system. Interferons (IFNs), in particular type I and type III IFN, are among the principal cytokines to recruit immune cells [19, 20].

Six types of leukocytes have been implicated in detecting and responding to viral infections in the lung, a major site of SARS-CoV-2 infection, which also presents with severe COVID-19 symptoms. The cytotoxic activities of CD8+ T cells and NK cells can facilitate early control of viral infections by clearing infected cells and avoiding additional viral dissemination [21, 22]. Dendritic cells specialize in sensing infections, including by viruses, and inducing an immune response [23]. CD4+ T cells contribute to viral clearance by promoting production of cytokines and interactions between CD8+ T cells and dendritic cells [24]. M1 macrophages interact with pulmonary epithelial cells to fight viral infections in the lung [25]. Finally, neutrophils may contribute to clearance of viral infections through phagocytosis of virions and viral particles. However, their precise role is uncertain [26].

SARS-CoV-2 is considerably more efficient in infection, replication and production of infectious virus particles in human lung tissue than SARS-CoV [17]. Strikingly, despite this, SARS-CoV-2 initially does not significantly induce type I, II or III IFNs or various other pro-inflammatory cytokines and chemokines in infected human lung cells and tissue [17, 27]. When this does occur, it may in fact promote further SARS-CoV-2 infection, as IFNs directly upregulate expression of the SARS-CoV-2 receptor ACE2 [28]. These observations suggest that the baseline levels of leukocytes, which already reside in the lung prior to infection, may be important in mounting a rapid immune response against SARS-CoV-2 infection.

Results

The efficiency of S protein-mediated coronavirus infection increases with elevated host lung cell-specific ACE2 expression levels, suggesting that this is a predictor of increased susceptibility to COVID-

19 [10-15]. In addition, the levels of leukocytes residing in the lung prior to infection may be particularly important for eliciting a rapid SARS-CoV-2-specific immune response. Thus, we investigated the relationship between ACE2 expression and the levels of seven leukocyte types implicated in anti-viral immune response in human lung tissue.

First, we used bulk RNAseq gene expression data from the 578 human lung tissues present in the Genotype-Tissue Expression (GTEx) database [29, 30]. Using an established "*in silico* flow cytometry" pipeline [31], we estimated the levels of CD8+ T cells, resting and activated NK cells, M1 macrophages, dendritic cells, CD4+ T cells and neutrophils in these tissues **(Fig. S1a-c)**. Next, we compared these to ACE2 expression levels in these lung tissues. This revealed that ACE2 expression is negatively correlated with the levels of CD8+ T cells, resting and activated NK cells and M1 macrophages ($p < 8\times10^{-6}$, Pearson correlations) **(Fig. 1a-d)**. However, there are no statistically significant correlations between ACE2 expression and the levels of CD4+ T cells, dendritic cells and neutrophils ($p > 0.05$) **(Fig. 1e-g)**. Thus, the levels of a majority of leukocytes involved in anti-viral immune responses are significantly reduced in lung tissues with elevated ACE2 expression levels.

It is possible that some of above observations are linked to phenotypic characteristics, such as sex, age, body mass index (BMI), race or smoking status. To test the robustness of our findings, we applied multivariable regression analysis that includes these five covariates **(Table S1)**, as well as the levels of the seven above leukocyte types or states. This showed that only 4 of the 12 variables significantly contribute to predicting ACE2 expression levels, specifically the levels of CD8+ T cells, resting NK cells, activated NK cells and M1 macrophages **(Table 1)**. Notably, none of the five added phenotypic covariates showed statistically significant contributions. Consistently, we found limited statistically significant correlations between these variables and ACE2 expression in univariate analyses, irrespective of whether they were analyzed as continuous data or binned into discrete ordinal categories **(Fig. S2a-j)**. Thus, the levels of four types of leukocytes that respond to viral infection in lung tissue are reduced with elevated ACE2 expression independently of phenotypic covariates.

Finally, we tested whether above observations could be validated in an independent cohort of individuals. For this, we used the, to our knowledge, largest publicly available lung tissue dataset. The Laval University, University of British-Columbia, Groningen University (LUG) dataset includes microarray gene expression data of 1,349 human lung tissues. Following determination of ACE2 expression levels and estimation of the levels of CD8+ T cells, resting NK cells, activated NK cells and M1 macrophages **(Fig. S1a-c)**, we found that three of the four also negatively correlated with ACE2 expression in this independent dataset ($p < 2\times10^{-8}$) **(Fig. 2a-d)**. With a correlation coefficient of $R = 0.096$, only M1 macrophages did not correlate with ACE2 expression in this dataset **(Fig. 2d)**.

Taken together, our observations indicate that three types of cytotoxic lymphocytes, specifically CD8+ T cells, resting NK cells and activated NK cells, are robustly and consistently reduced in lung tissue with elevated expression of the SARS-CoV-2 receptor ACE2.

Discussion

We investigated the baseline expression levels of the SARS-CoV-2 host cell entry receptor ACE2 and the levels of seven types of leukocytes in 1,927 human lung tissue samples. Although SARS-CoV-2 cellular tropism is broad [16-18], we focused on lung tissue. In addition to epithelial cells elsewhere in the respiratory tract, alveolar epithelial cells are thought to be a primary SARS-CoV-2 entry point [16, 28]. Consistently, the SARS-CoV-2 receptor ACE2 is expressed in these cells at the mRNA and protein levels [28, 32-34]. Moreover, in severely affected COVID-19 patients, the lungs are among the few organs that present with the most life-threatening symptoms. "Cytokine storm"-induced acute respiratory distress syndrome (ARDS), widespread alveolar damage, pneumonia, and progressive respiratory failure have been observed [35, 36]. These indications frequently require admission to intensive care units (ICUs) and mechanical ventilation and may ultimately be fatal.

Early after infection, rapid activation of the innate immune system is of paramount importance for the clearance of virus infections. Infected cells typically do so through release of pro-inflammatory cytokines and chemokines, in particular type I and III interferons [19, 20]. Notably, however, several studies have highlighted multiple complexities related specifically to SARS-CoV2 and innate immune

system activation at early stages. First, unlike SARS-CoV, SARS-CoV-2-infected lung tissue initially fails to induce a range of immune cell-recruiting molecules, including several interferons [17, 27], suggesting that leukocytes are ineffectively recruited to infected lung shortly after infection. Second, the host cell entry receptor ACE2 has been identified as an interferon target gene [28]. Thus, even when interferons are upregulated in order to recruit immune cells, concomitant upregulation of ACE2 expression may in fact exacerbate SARS-CoV-2 infection [28].

These findings suggest that the levels of immune cells that already reside in the lung prior to infection are more critical for dampening SARS-CoV-2 infection at early stages than they are for fighting infections of other viruses. Consistently, cytotoxic lymphocytes are markedly reduced in both numbers and functionality in COVID-19 patients [37]. Notably, cytotoxic lymphocytes, including CD8+ T cells and NK cells, are key early responders to virus infections and these are the cells whose baseline levels we here identify as significantly reduced in lung tissue with elevated ACE2 expression. Therefore, our results suggest that individuals with increased baseline susceptibility to SARS-CoV-2 infection in the lungs may also be less well equipped from the outset to mount a rapid anti-viral cellular immune response.

Several observations indicate that these cytotoxic lymphocytes are critically important for effective control of SARS-CoV-2 infection. Recent studies showed that CD8+ T cells in peripheral blood are considerably reduced and functionally exhausted in COVID-19 patients, in particular in elderly patients and in severely affected patients that require ICU admission [37-39]. Reduced CD8+ T cell counts also predict poor COVID-19 patient survival [38].

NK cells are the first immune cells to respond to viral infections. They can recognize and eliminate virus-infected cells following cytokine stimulation or antibody-mediated activation [22]. However, they can also do so when infected cells express so-called 'induced-self' ligands or viral antigens [22, 40]. This is probably exceedingly important during the initial stages of COVID-19 pathogenesis, as SARS-CoV-2-infected lung fails to significantly induce pro-inflammatory cytokines and chemokines early after infection [17, 27]. In line with this, at later stages during COVID-19

progression, the total numbers of NK cells in peripheral blood are substantially lower compared to healthy individuals [37]. Moreover, they are functionally exhausted [37].

We found that the five phenotypic parameters, sex, age, BMI, race and smoking history did not statistically significantly contribute to variation in ACE2 expression in human lung tissue, neither in univariate nor in multivariate analyses. This is consistent with some studies but inconsistent with others [36, 41-43]. These paradoxical observations may be partially explained by varying gender, age and race distributions within each study cohort.

Further research will be required to elucidate the precise mechanisms of SARS-CoV-2-induced activation of the innate immune system early after infection. However, the link that we identified between high baseline ACE2 expression and reduced cytotoxic lymphocyte levels in human lung tissue prior SARS-CoV-2 infection is striking. It suggests that increased susceptibility to SARS-CoV-2 infection in the lungs may be accompanied by a poorer ability to mount a rapid innate immune response at early stages. This may predict long-term outcome of individuals infected with COVID-19, given that CD8+ T cell and NK cell counts in peripheral blood are low in COVID-19 patients [37-39]. Finally, it may contribute to the substantial variation in COVID-19 clinical presentation, ranging from asymptomatic to severe respiratory and other symptoms.

Methods

Discovery dataset and processing

Gene expression data and corresponding phenotype data from human lung tissues (n=578) were obtained from the Genotype-Tissue Expression (GTEx) Portal (https://gtexportal.org), managed by the National Institutes of Health (NIH). Gene expression data were publicly available. Access to phenotype data required authorization. The GTEx protocol was previously described [29, 30]. Briefly, total RNA was extracted from tissue. Following mRNA isolation, cDNA synthesis and library preparation, samples were subjected to HiSeq2000 or HiSeq2500 Illumina TrueSeq RNA sequencing. Gene

expression levels were obtained using RNA-SeQC v1.1.9 [44] and expressed in transcripts per million (TPM). Log-transformed expression levels were reported, unless otherwise indicated.

Validation dataset and processing

For validation purposes, the Laval University, University of British-Columbia, Groningen University (LUG) lung tissue dataset (n=1,349) was used. This dataset was accessed via Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo), accession number GSE23546, and was previously described [45]. Briefly, total RNA from human lung tissue samples was isolated, quantified, quality-checked and used to generate cDNA, which was amplified and hybridized to Affymetrix gene expression arrays. Arrays were scanned and probe-level gene expression values were normalized using robust multichip average (RMA). These normalized values were obtained from GEO. To collapse probe-level expression data to single expression levels per gene, for each gene the probe with the highest median absolute deviation (MAD) was used. The MAD for each probe p was calculated using equation (1).

$$\mathrm{MAD}(p) = M(|p_i - M(p)|) \quad (1)$$

Herein, M is the median, p_i denotes probe p's expression level in sample i, and $M(p)$ represents the median signal of probe p.

***In silico* cytometry**

The levels of seven types of leukocytes involved in anti-viral cellular immune response, specifically CD8+ T cells, resting NK cells, activated NK cells, M1 macrophages, CD4+ T cells, dendritic cells and neutrophils, were estimated in the discovery and validation lung tissue samples using a previously described approach [31]. Specifically, the following workflow was used. First, only non-log-transformed expression values were used. Thus, where required, expression values for all samples in the discovery and validation datasets were reverse-$\log_2$-transformed using equation 2.

$$c = 2^{c_l} - 1 \quad (2)$$

Herein, c denotes the calculated non-$\log_2$-transformed expression counts and c_l denotes the previously reported $\log_2$-transformed expression counts. Next, to compensate for potential technical differences

between signatures and bulk sample gene expression values due to inter-platform variation, bulk-mode batch correction was applied. To ensure robustness, deconvolution was statistically analyzed using 100 permutations. Pearson correlation coefficients R, root mean squared errors (RMSA) and p values are reported on a per-sample level in Fig. S1a-c.

Univariate statistical analyses

Log_2-transformed expression levels of ACE2 in lung tissue samples were compared to the estimated levels of seven leukocyte types or states. Pearson correlation analyses were performed to determine Pearson correlation coefficients R and p values. P values were adjusted at a false discovery rate of 0.05 to yield q values, as previously described [46]. Straight lines represent the minimized sum of squares of deviations of the data points with 95% confidence intervals shown. Continuous phenotypic covariates were analyzed in the same way and, additionally, as discrete ordinal categories after binning. Discrete and binned phenotype data were statistically evaluated using Mann-Whitney U tests. All analyses were performed in the *R* computing environment (*R* Project for Statistical Computing, Vienna, Austria).

Multivariate regression analysis

Multivariate analysis was performed using standard ordinary least squares regression, summarized in equation 3.

$$Y = \beta_0 + \sum_{k=1}^{n} (\beta_k X_k) + \varepsilon \quad (3)$$

Herein, β_0 denotes the intercept, while β_k represents the slope of each variable X_k in a model with n variables and ε denotes the random error component. This analysis was performed using *R*.

Acknowledgements

I thank Dr. Marianna Datseris, M.D. for critically reading the manuscript. This study used publicly available and controlled data from the Genotype-Tissue Expression (GTEx) Project. GTEx was supported by the Common Fund of the Office of the Director of the National Institutes of Health

(commonfund.nih.gov/GTEx). Additional funds were provided by the NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. Donors were enrolled at Biospecimen Source Sites funded by NCI\Leidos Biomedical Research, Inc. subcontracts to the National Disease Research Interchange (10XS170), Roswell Park Cancer Institute (10XS171), and Science Care, Inc. (X10S172). The Laboratory, Data Analysis, and Coordinating Center (LDACC) was funded through a contract (HHSN268201000029C) to the The Broad Institute, Inc. Biorepository operations were funded through a Leidos Biomedical Research, Inc. subcontract to Van Andel Research Institute (10ST1035). Additional data repository and project management were provided by Leidos Biomedical Research, Inc. (HHSN261200800001E). The Brain Bank was supported supplements to University of Miami grant DA006227. Statistical Methods development grants were made to the University of Geneva (MH090941 & MH101814), the University of Chicago (MH090951, MH090937, MH101825, & MH101820), the University of North Carolina - Chapel Hill (MH090936), North Carolina State University (MH101819), Harvard University (MH090948), Stanford University (MH101782), Washington University (MH101810), and to the University of Pennsylvania (MH101822). The datasets used for the analyses described in this study were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000424.v8.p2.

References

1. Coronaviridae Study Group of the International Committee on Taxonomy of, V., *The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2.* Nat Microbiol, 2020. **5**(4): p. 536-544.
2. Liu, Y., et al., *Viral dynamics in mild and severe cases of COVID-19.* Lancet Infect Dis, 2020.
3. Gattinoni, L., et al., *COVID-19 pneumonia: different respiratory treatments for different phenotypes?* Intensive Care Med, 2020.
4. Xie, M. and Q. Chen, *Insight into 2019 novel coronavirus - An updated interim review and lessons from SARS-CoV and MERS-CoV.* Int J Infect Dis, 2020. **94**: p. 119-124.
5. *Coronavirus disease 2019 (COVID-19), Situation Report 102.* World Health Organization, 2020(1 May, 2020).

6. Li, R., et al., *Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2).* Science, 2020. **368**(6490): p. 489-493.
7. Hoffmann, M., et al., *SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor.* Cell, 2020. **181**(2): p. 271-280 e8.
8. Zhou, P., et al., *A pneumonia outbreak associated with a new coronavirus of probable bat origin.* Nature, 2020. **579**(7798): p. 270-273.
9. Wrapp, D., et al., *Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation.* Science, 2020. **367**(6483): p. 1260-1263.
10. Hofmann, H., et al., *Susceptibility to SARS coronavirus S protein-driven infection correlates with expression of angiotensin converting enzyme 2 and infection can be blocked by soluble receptor.* Biochem Biophys Res Commun, 2004. **319**(4): p. 1216-21.
11. Mossel, E.C., et al., *Exogenous ACE2 expression allows refractory cell lines to support severe acute respiratory syndrome coronavirus replication.* J Virol, 2005. **79**(6): p. 3846-50.
12. Li, W., et al., *Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus.* Nature, 2003. **426**(6965): p. 450-4.
13. Fang, L., G. Karakiulakis, and M. Roth, *Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection?* Lancet Respir Med, 2020. **8**(4): p. e21.
14. Guo, J., et al., *Coronavirus Disease 2019 (COVID-19) and Cardiovascular Disease: A Viewpoint on the Potential Influence of Angiotensin-Converting Enzyme Inhibitors/Angiotensin Receptor Blockers on Onset and Severity of Severe Acute Respiratory Syndrome Coronavirus 2 Infection.* J Am Heart Assoc, 2020. **9**(7): p. e016219.
15. Wang, P.H. and Y. Cheng, *Increasing Host Cellular Receptor—Angiotensin-Converting Enzyme 2 (ACE2) Expression by Coronavirus may Facilitate 2019-nCoV Infection.* bioRxiv, 2020. **Preprint**: p. doi: 10.1101/2020.02.24.963348.
16. Sungnak, W., et al., *SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes.* Nat Med, 2020.
17. Chu, H., et al., *Comparative replication and immune activation profiles of SARS-CoV-2 and SARS-CoV in human lungs: an ex vivo study with implications for the pathogenesis of COVID-19.* Clin Infect Dis, 2020.
18. Chu, H., et al., *Comparative tropism, replication kinetics, and cell damage profiling of SARS-CoV-2 and SARS-CoV with implications for clinical manifestations, transmissibility, and laboratory studies of COVID-19: an observational study.* Lancet Microbe, 2020. **April 21, 2020**: p. 1-10.
19. Newton, A.H., A. Cardani, and T.J. Braciale, *The host immune response in respiratory virus infection: balancing virus clearance and immunopathology.* Semin Immunopathol, 2016. **38**(4): p. 471-82.

20. Crouse, J., U. Kalinke, and A. Oxenius, *Regulation of antiviral T cell responses by type I interferons.* Nat Rev Immunol, 2015. **15**(4): p. 231-42.
21. Schmidt, M.E. and S.M. Varga, *The CD8 T Cell Response to Respiratory Virus Infections.* Front Immunol, 2018. **9**: p. 678.
22. Hammer, Q., T. Ruckert, and C. Romagnani, *Natural killer cell specificity for viral infections.* Nat Immunol, 2018. **19**(8): p. 800-808.
23. Cook, P.C. and A.S. MacDonald, *Dendritic cells in lung immunopathology.* Semin Immunopathol, 2016. **38**(4): p. 449-60.
24. Sant, A.J. and A. McMichael, *Revealing the role of CD4(+) T cells in viral immunity.* J Exp Med, 2012. **209**(8): p. 1391-5.
25. Harper, R.W., *Partners in Crime: Epithelial Priming of Macrophages during Viral Infection.* Am J Respir Cell Mol Biol, 2017. **57**(2): p. 145-146.
26. Galani, I.E. and E. Andreakos, *Neutrophils in viral infections: Current concepts and caveats.* J Leukoc Biol, 2015. **98**(4): p. 557-64.
27. Blanco-Melo, D., et al., *SARS-CoV-2 launches a unique transcriptional signature from in vitro, ex vivo, and in vivo systems.* bioRxiv, 2020. **Preprint**: p. 24 March 2020.
28. Ziegler, C.G.K., et al., *SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues.* Cell, 2020. **In Press**.
29. Consortium, G.T., *The Genotype-Tissue Expression (GTEx) project.* Nat Genet, 2013. **45**(6): p. 580-5.
30. Consortium, G.T., et al., *Genetic effects on gene expression across human tissues.* Nature, 2017. **550**(7675): p. 204-213.
31. Newman, A.M., et al., *Robust enumeration of cell subsets from tissue expression profiles.* Nat Methods, 2015. **12**(5): p. 453-7.
32. Bertram, S., et al., *Influenza and SARS-coronavirus activating proteases TMPRSS2 and HAT are expressed at multiple sites in human respiratory and gastrointestinal tracts.* PLoS One, 2012. **7**(4): p. e35876.
33. Hamming, I., et al., *Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis.* J Pathol, 2004. **203**(2): p. 631-7.
34. Zou, X., et al., *Single-cell RNA-seq data analysis on the receptor ACE2 expression reveals the potential risk of different human organs vulnerable to 2019-nCoV infection.* Front Med, 2020.
35. Huang, C., et al., *Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China.* Lancet, 2020. **395**(10223): p. 497-506.
36. Chan, J.F., et al., *A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster.* Lancet, 2020. **395**(10223): p. 514-523.

37. Zheng, M., et al., *Functional exhaustion of antiviral lymphocytes in COVID-19 patients.* Cell Mol Immunol, 2020.
38. Diao, B., et al., *Reduction and Functional Exhaustion of T Cells in Patients with Coronavirus Disease 2019 (COVID-19).* medRxiv preprint, 2020. **doi: 10.1101/2020.02.18.20024364**.
39. Liu, J., et al., *Longitudinal characteristics of lymphocyte responses and cytokine profiles in the peripheral blood of SARS-CoV-2 infected patients.* EBioMedicine, 2020. **55**: p. 102763.
40. Smith, H.R., et al., *Recognition of a virus-encoded ligand by a natural killer cell activation receptor.* Proc Natl Acad Sci U S A, 2002. **99**(13): p. 8826-31.
41. Lloyd-Sherlock, P., et al., *Bearing the brunt of covid-19: older people in low and middle income countries.* BMJ, 2020. **368**: p. m1052.
42. Yang, X., et al., *Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study.* Lancet Respir Med, 2020.
43. Smith, J.C. and J.M. Sheltzer, *Cigarette smoke triggers the expansion of a subpopulation of respiratory epithelial cells that express the SARS-CoV-2 receptor ACE2.* bioRxiv, 2020. **Preprint**: p. 31 March 2020.
44. DeLuca, D.S., et al., *RNA-SeQC: RNA-seq metrics for quality control and process optimization.* Bioinformatics, 2012. **28**(11): p. 1530-2.
45. Obeidat, M., et al., *Molecular mechanisms underlying variations in lung function: a systems genetics analysis.* Lancet Respir Med, 2015. **3**(10): p. 782-95.
46. Benjamini, Y. and Y. Hochberg, *Controlling the False Discovery Rate: A practical and powerful approach to multiple testing.* J R Stat Soc Ser B-Stat Methodol, 1995. **57**(1): p. 289-300.

Figure 1

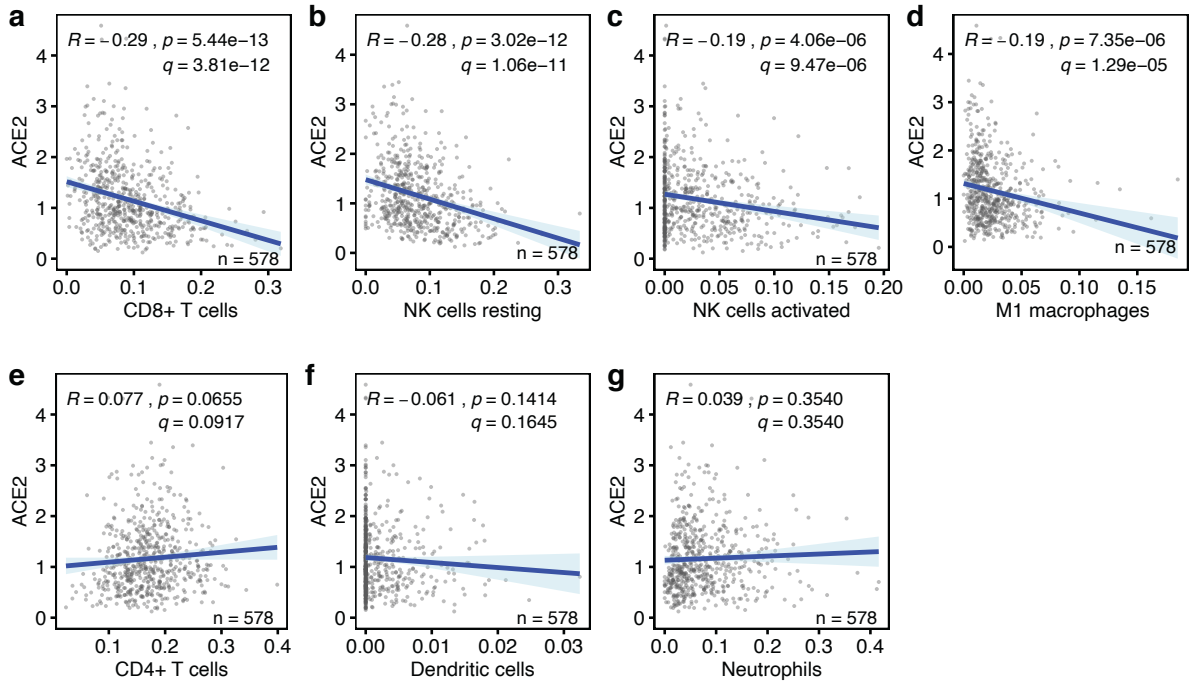

Figure 1. Baseline levels of CD8+ T cells, resting NK cells, activated NK cells and M1 macrophages are reduced in human lung tissue with elevated ACE2 expression.

The correlations between the level of SARS-CoV-2 host cell receptor ACE2 expression (x-axis) and the levels of indicated immune cell types (y-axes) in human lung tissue are shown. Data are from the GTEx dataset (n=578). Regression lines and 95% confidence intervals are shown. *R* and *p* values: Pearson correlations; *q* values: Benjamini-Hochberg-adjusted *p* values using a false discovery rate of 0.05.

Figure 2

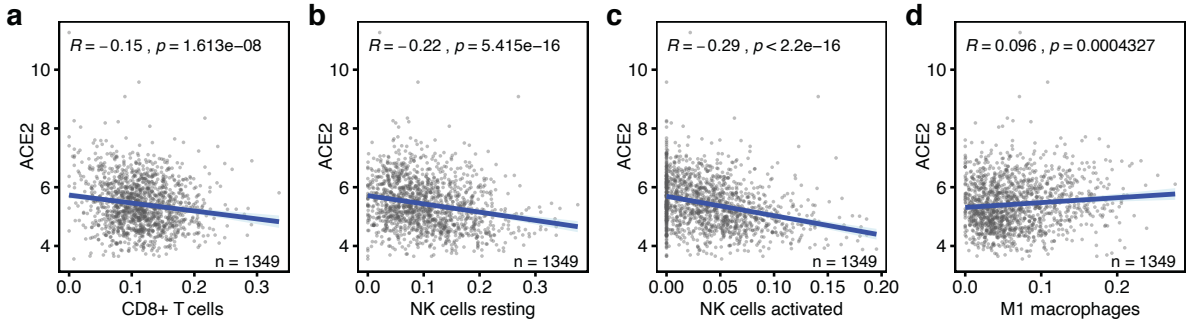

Figure 2. Baseline levels of CD8+ T cells, resting NK cells and activated NK cells are reduced in human lung tissue with elevated ACE2 expression in an independent dataset.

Correlations between ACE2 expression (x-axis) and the levels indicated immune cell types (y-axes) in human lung tissue are shown as in Figure 1. Data are from the LUG dataset (n=1,349). *R* and *p* values: Pearson correlations.

Table 1. Multivariate analysis of ACE2 expression in human lung tissue

Variable	β value	S.E.	t value	p value	
CD8+ T cells	-2.719	0.584	-4.656	**4.02e-06**	****
NK cells resting	-3.128	0.583	-5.366	**1.18e-07**	****
NK cells activated	-2.281	0.704	-3.241	**0.0013**	**
M1 macrophages	-4.963	1.339	-3.707	**0.0002**	***
CD4+ T cells	0.453	0.513	0.882	0.3781	
Dendritic cells	-12.793	6.688	-1.913	0.0563	
Neutrophils	-0.398	0.469	-0.848	0.3971	
Sex = male (vs. female)	0.065	0.056	1.16	0.2466	
Age	0.004	0.002	1.799	0.0726	
BMI	-0.004	0.006	-0.556	0.5787	
Race = Caucasian (vs. non-Caucasian)	-0.096	0.073	-1.313	0.1899	
Smoking status = yes (vs. no)	-0.025	0.058	-0.433	0.6650	

Abbreviations: S.E., Standard Error.

Supplementary Information

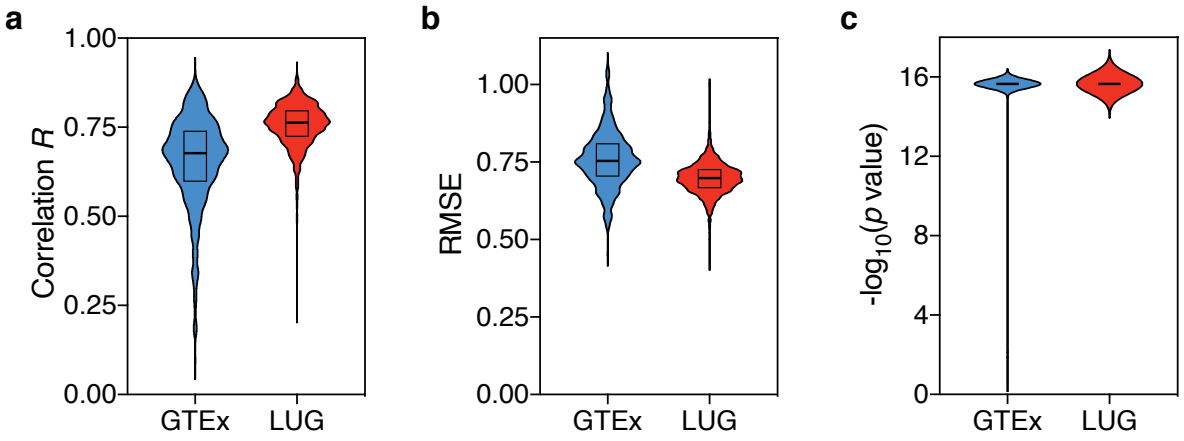

Figure S1. *In silico* cytometry statistics.

Per-sample statistics of *in silico* cytometry on human lung tissues are shown for the GTEx (n=578) and LUG (n=1,349) datasets. **(a)** Pearson correlation coefficients *R*. **(b)** Root mean square errors (RMSE). **(c)** *p* values. *P* values < 2.2×10^{-16} were processed as equal to 2.2×10^{-16}.

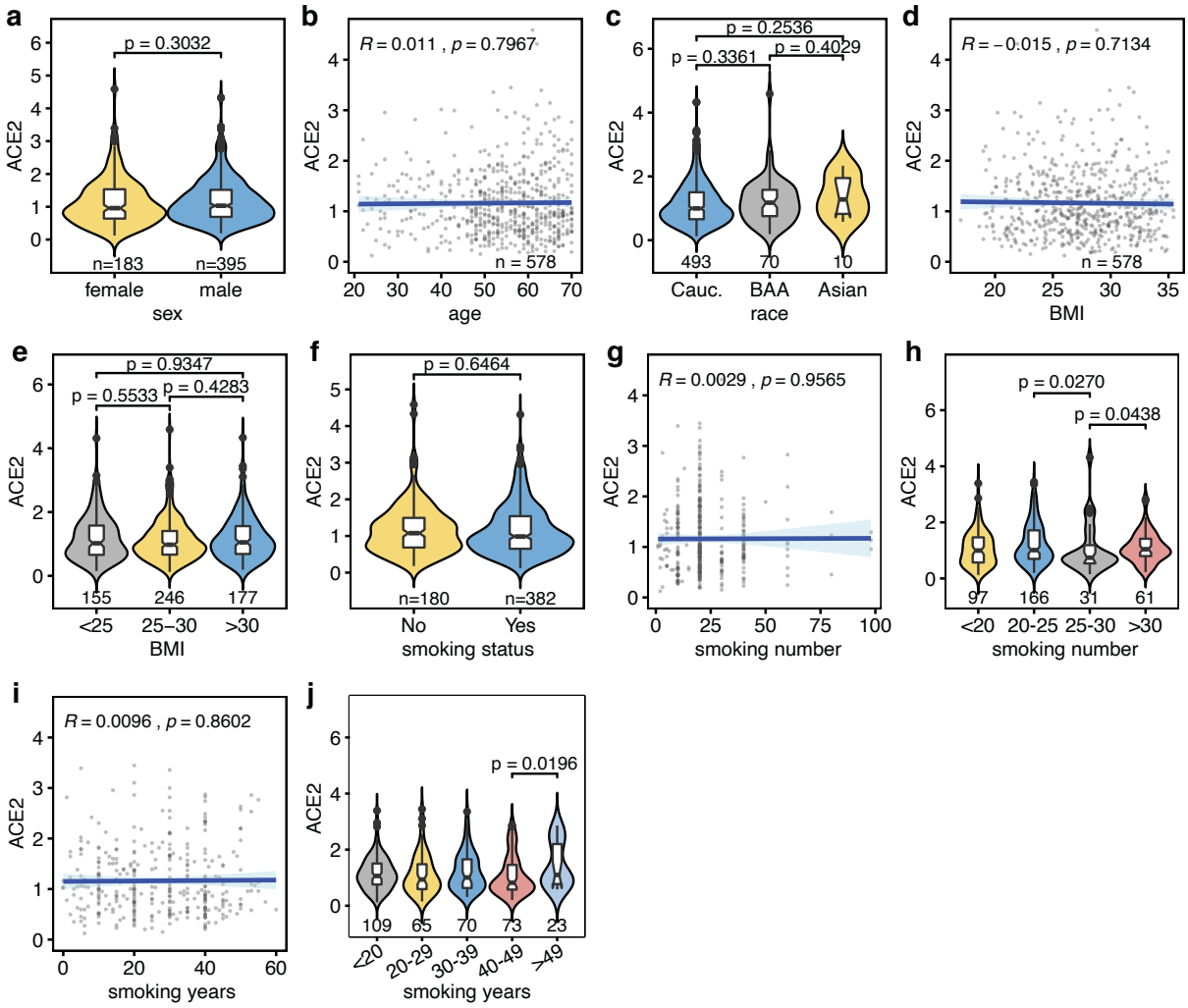

Figure S2. Univariate analyses of phenotypic covariates included in multivariate analysis. Univariate analyses of five covariates that were included in multivariate analysis of ACE2 expression in human lung tissue using the GTEx dataset (Table 1). **(a)** Sex. **(b)** Age, **(c)** Race. **(d, e)** Body mass index. **(f-j)** Smoking behavior, referring to smoking status (smoker/non-smoker) (f), number of units smoked during the smoke period (g, h) and number of years smoked (i, j). *P* values in categorical analyses: Mann-Whitney *U* tests. *R* and *p* values in continuous analyses: Pearson correlations. *P* values in panels h and j are only shown if $p < 0.05$. Samples sizes are shown on the x-axes.

Table S1. Phenotypic characteristics of GTEx lung tissue donors

Feature		Number of individuals (%)	Median (range)
Sex			
	Female	183 (31.7)	
	Male	395 (68.3)	
Age			
	<40	73 (12.6)	56 (21 - 70)
	40-49	93 (16.1)	
	50-59	200 (34.6)	
	60-69	190 (32.9)	
	>69	22 (3.8)	
Body mass index			
	<25	155 (26.8)	27.7 (17.0 - 35.4)
	25-30	246 (42.6)	
	>30	177 (30.6)	
Race			
	Caucasian	493 (85.9)	
	Black/African American	70 (12.2)	
	Asian	10 (1.7)	
	American Indian/Alaska Native	1 (0.2)	
Smoking status			
	Non-smoker	180 (32.0)	
	Smoker	382 (68.0)	