



SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes

Waradon Sungnak¹✉, Ni Huang¹, Christophe Bécavin², Marijn Berg^{3,4}, Rachel Queen⁵, Monika Litvinukova^{1,6}, Carlos Talavera-López¹, Henrike Maatz⁶, Daniel Reichart⁷, Fotios Sampaziotis^{8,9,10}, Kaylee B. Worlock¹¹, Masahiro Yoshida¹¹, Josephine L. Barnes¹¹ and HCA Lung Biological Network^{*✉}

We investigated SARS-CoV-2 potential tropism by surveying expression of viral entry-associated genes in single-cell RNA-sequencing data from multiple tissues from healthy human donors. We co-detected these transcripts in specific respiratory, corneal and intestinal epithelial cells, potentially explaining the high efficiency of SARS-CoV-2 transmission. These genes are co-expressed in nasal epithelial cells with genes involved in innate immunity, highlighting the cells' potential role in initial viral infection, spread and clearance. The study offers a useful resource for further lines of inquiry with valuable clinical samples from COVID-19 patients and we provide our data in a comprehensive, open and user-friendly fashion at www.covid19cellatlas.org.

The coronavirus disease 2019 (COVID-19) is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)¹. Detection of the virus was first reported in Wuhan², China and has since spread worldwide, emerging as a global pandemic³.

In symptomatic patients, nasal swabs have yielded higher viral loads than throat swabs⁴. The same distribution was observed in an asymptomatic patient⁴, implicating the nasal epithelium as a portal for initial infection and transmission. Cellular entry of coronaviruses depends on the binding of the spike (S) protein to a specific cellular receptor and subsequent S protein priming by cellular proteases. Similarly to SARS-CoV^{5,6}, SARS-CoV-2 employs ACE2 as a receptor for cellular entry. The binding affinity of the S protein and ACE2 was found to be a major determinant of SARS-CoV replication rate and disease severity^{4,7}. Viral entry also depends on TMPRSS2 protease activity and cathepsin B/L activity may be able to substitute for TMPRSS2⁷.

ACE2 and TMPRSS2 have been detected in both nasal and bronchial epithelium by immunohistochemistry⁸. Gene expression of ACE2 and TMPRSS2 has been reported to occur largely in alveolar epithelial type II cells^{9–11}, which are central to SARS-CoV pathogenesis, whereas a different study reported the absence of ACE2 in the upper airway¹². To clarify the expression patterns of ACE2 and TMPRSS2, we analyzed their expression and the expression of other genes potentially

associated with SARS-CoV-2 pathogenesis at cellular resolution, using single-cell RNA sequencing (scRNA-seq) datasets from healthy donors generated by the Human Cell Atlas consortium and other resources to inform and prioritize the use of precious, limited clinical material that is becoming available from COVID-19 patients.

We investigated gene expression of ACE2 in multiple scRNA-seq datasets from different tissues, including those of the respiratory tree, cornea, retina, esophagus, ileum, colon, heart, skeletal muscle, spleen, liver, placenta/decidua, kidney, testis, pancreas, prostate gland, brain, skin and fetal tissues. We note that studies may lack specific cell types due to their sparsity, the challenges associated with isolation or analysis methodology. Moreover, expression may be under-detected due to technical dropout effects. Thus, while positive (presence) results are highly reliable, absence should be interpreted with care.

ACE2 expression was generally low in all analyzed datasets. Consistently with independent studies^{10,11}, ACE2 was expressed in cells from multiple tissues, including airways, cornea, esophagus, ileum, colon, liver, gallbladder, heart, kidney and testis (Fig. 1a; first column). TMPRSS2 was highly expressed with a broader distribution (Fig. 1a; second column), suggesting that ACE2, rather than TMPRSS2, may be a limiting factor for viral entry at the initial infection stage. Cells from the respiratory tree, cornea, esophagus, ileum, colon, gallbladder and common bile duct expressed both genes in the same cell (Fig. 1a; third column). We also assessed ACE2 and TMPRSS2 expression in developmental datasets from fetal tissues, including liver, thymus, skin, bone marrow, yolk sac and lung, and found little to no expression of ACE2 in all but fetal liver and thymus (Fig. 1a) where there was no co-expression with TMPRSS2 (data not shown) except for a cluster of medullary thymic epithelial cells (Fig. 1a). ACE2 expression is noticeable in certain cell types in placenta/decidua without TMPRSS2 (Fig. 1a). Additional fetal data across relevant tissues and stages are needed to determine the generality of these findings.

To further characterize specific epithelial cell types expressing ACE2, we evaluated ACE2 expression within the lung and airway

¹Wellcome Sanger Institute, Cambridge, UK. ²Université Côte d'Azur, CNRS, IPMC, Sophia-Antipolis, France. ³Department of Pathology and Medical Biology, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands. ⁴Groningen Research Institute for Asthma and COPD, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands. ⁵Bioinformatics Core Facility, Newcastle University Biosciences Institute, Faculty of Medical Sciences, Newcastle University, Newcastle-upon-Tyne, UK. ⁶Cardiovascular and Metabolic Sciences, Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), Berlin, Germany. ⁷Department of Genetics, Harvard Medical School, Boston, MA, USA. ⁸Wellcome and MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK. ⁹Department of Medicine, Addenbrookes Hospital, Cambridge, UK. ¹⁰Cambridge Liver Unit, Cambridge University Hospitals, Cambridge, UK. ¹¹UCL Respiratory, Division of Medicine, University College London, London, UK. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: ws4@sanger.ac.uk; lung@humancellatlas.org

epithelium. We found that, despite a low level of expression overall, *ACE2* was expressed in multiple epithelial cell types across the airway, as well as in alveolar epithelial type II cells in the parenchyma, consistently with previous studies^{9–11}. Notably, nasal epithelial cells, including two previously described clusters of goblet cells and one cluster of ciliated cells, show the highest expression among all investigated cells in the respiratory tree (Fig. 1b). We confirmed enriched *ACE2* expression in nasal epithelial cells in an independent scRNA-seq study that includes nasal brushings and biopsies. The results were consistent; we found the highest expression of *ACE2* in nasal secretory cells (equivalent to the two goblet cell clusters in the previous dataset) and ciliated cells (Fig. 1b).

In addition, scRNA-seq data from an in vitro epithelial regeneration system from nasal epithelial cells corroborated the expression of *ACE2* in goblet/secretory cells and ciliated cells in air–liquid interface cultures (Extended Data Fig. 1). Notably, the differentiating cells in the air–liquid interface acquire progressively more *ACE2* (Extended Data Fig. 1). The results also suggest that this in vitro culture system may be biologically relevant for the study of SARS-CoV-2 pathogenesis.

It is worth noting that *TMPRSS2* was only expressed in a subset of *ACE2*⁺ cells (Extended Data Fig. 2), suggesting that the virus might use alternative pathways. It was previously shown that SARS-CoV-2 could enter *TMPRSS2*[−] cells using cathepsin B/L⁷. Indeed, other proteases were more promiscuously expressed than *TMPRSS2*, especially cathepsin B, which was expressed in more than 70–90% of *ACE2*⁺ cells (Extended Data Fig. 2). However, while *TMPRSS2* activity is documented to be important for viral transmission^{13,14}, the potential of cathepsin B/L or other proteases to functionally replace *TMPRSS2* has not been determined.

We next asked whether enriched expression of viral receptors and entry-associated molecules in the nasal region/upper airway might be relevant for viral transmissibility. Here, we assessed the expression of viral receptor genes that are used by other coronaviruses and influenza viruses in our datasets. We looked for *ANPEP*, used by HCoV-22944 (ref. 15) and *DPP4*, used by MERS-CoV45 (ref. 16), as well as enzymes *ST6GAL1* and *ST3GAL4*, which are important for the synthesis of $\alpha(2,6)$ -linked and $\alpha(2,3)$ -linked sialic acids recognized by influenza viruses¹⁷. Notably, their expression distribution coincided with viral transmissibility patterns based on a comparison to the basic reproduction number (R_0), which estimates the number of people who can become infected from a single infected person. The skewed distribution of the receptors/enzymes toward the upper airway is observed in viruses with higher R_0 /infectivity, including those of SARS-CoV/SARS-CoV-2 ($R_0 \sim 1.4$ – 5.0 (refs. 18–20)), influenza (mean $R_0 \sim 1.347$ (ref. 21)) and HCoV-229E (unidentified R_0 ; associated with common cold). This distribution is in distinct contrast with that of *DPP4*, the

receptor for MERS-CoV ($R_0 \sim 0.3$ – 0.8 (ref. 22)), a coronavirus with limited human-to-human transmission²³, in which expression skews toward lower airway/lung parenchyma (Fig. 2a). Therefore, our data highlight the possibility that viral transmissibility is dependent on the spatial distribution of receptor accessibility along the respiratory tract.

To gain more insight into the expression patterns of genes associated with *ACE2*, we performed Spearman's correlation analysis with Benjamini–Hochberg-adjusted *P* values to identify genes associated with *ACE2* across all cells within the lung epithelial cell datasets. While the correlation coefficients are relatively low (<0.12), likely due to low expression of *ACE2*, technical noise and dropout effects, the expression pattern of the top 50 *ACE2*-correlated genes across the respiratory tree is consistent with that of *ACE2*, with a skewed expression toward upper airway cells (Fig. 2b and Extended Data Fig. 3a,b). Notably, while some of the genes are associated with carbohydrate metabolism, possibly due to their role in goblet cell mucin synthesis, a number of genes associated with immune functions including innate and antiviral immune functions, are over-represented in the rank list, including *IDO1*, *IRAK3*, *NOS2*, *TNFSF10*, *OAS1* and *MX1* (Fig. 2b and Supplementary Table 1). Expression of these genes is highest in nasal goblet 2 cells (Fig. 2b), consistent with the phenotype previously described. Nonetheless, nasal goblet 1 and nasal ciliated 2 cells also significantly express these genes (Fig. 2b). Given their environmental exposure and high expression of receptor/receptor-associated enzymes (Fig. 2a), it is plausible that nasal epithelial cells are conditioned to express these immune-associated genes to reduce viral susceptibility.

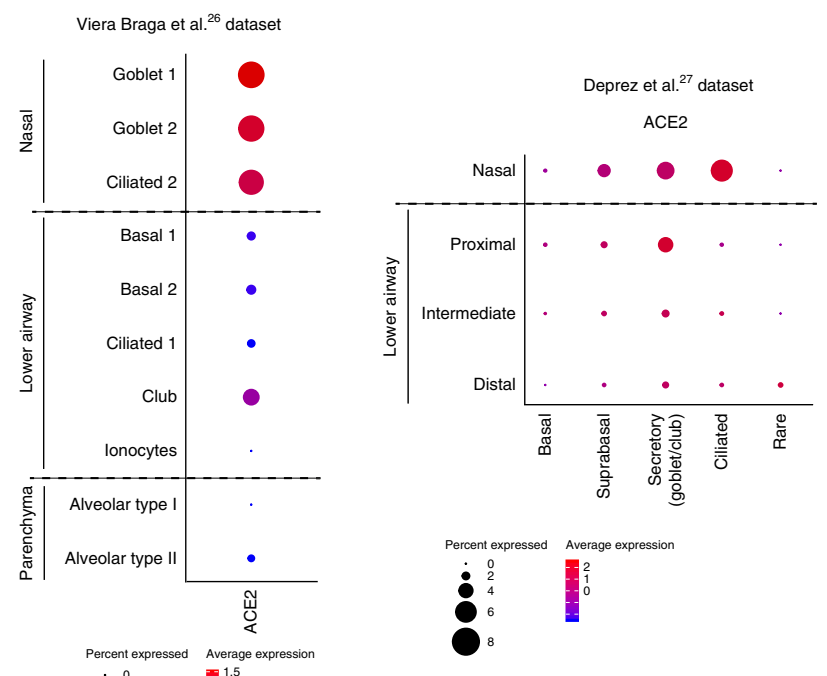
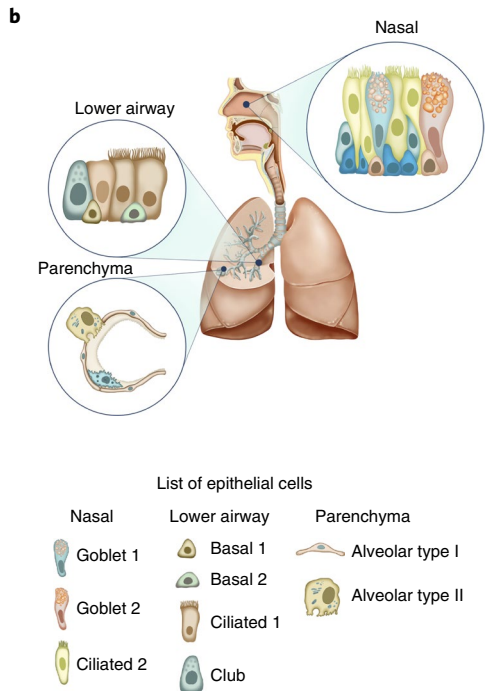
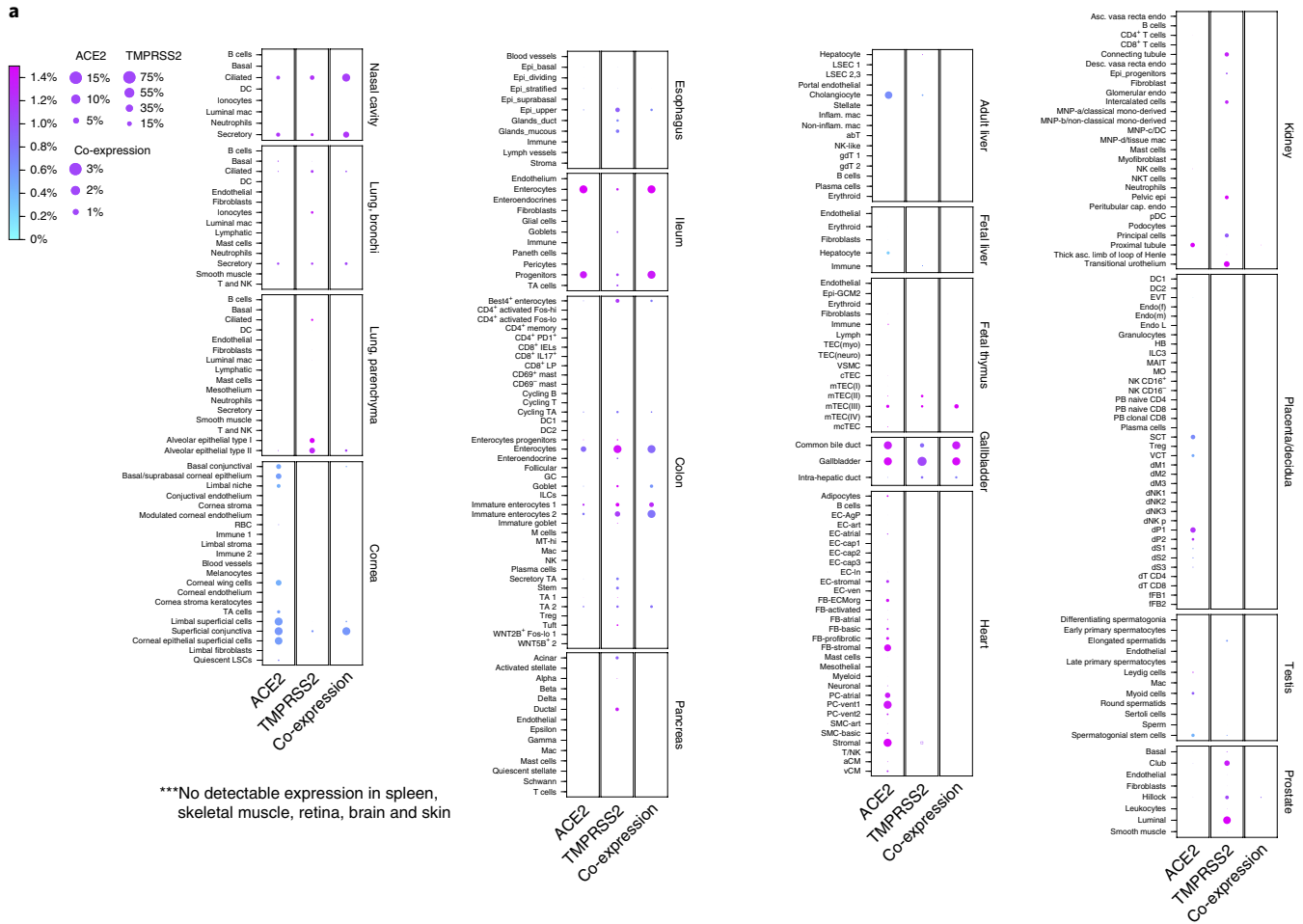
In this study, we explored multiple scRNA-seq datasets generated within the Human Cell Atlas (HCA) consortium and other resources and found that the SARS-CoV-2 entry receptor *ACE2* and viral entry-associated protease *TMPRSS2* are highly expressed in nasal goblet and ciliated cells. This finding implicates these cells as loci of original infection and possible reservoirs for dissemination within and between individuals. Co-expression in other barrier surface tissues could also suggest further investigation into alternative transmission routes. For example, the co-expression in esophagus, ileum and colon could explain viral fecal shedding observed clinically²⁴, with implications for potential fecal–oral transmission, whereas the co-expression in superficial conjunctival cells could explain an ocular phenotype observed in a small portion of COVID-19 patients²⁵ with the potential of spread through the nasolacrimal duct.

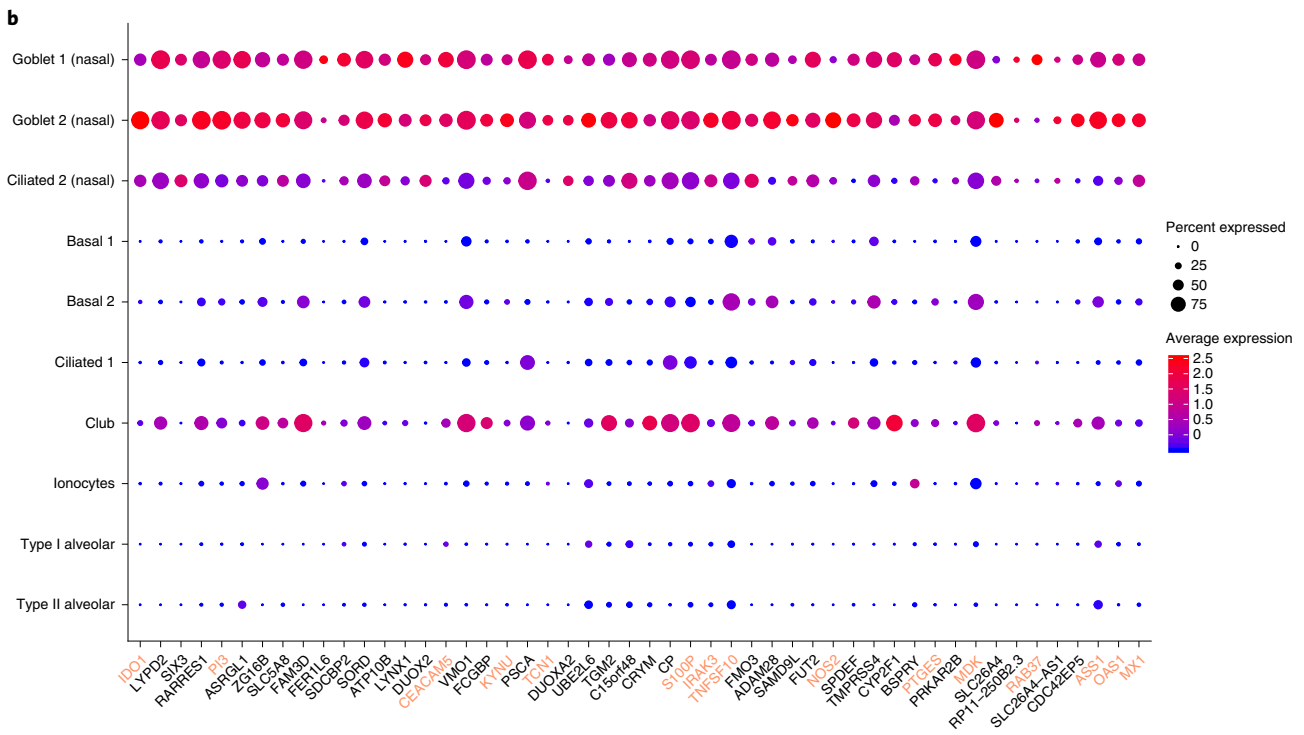
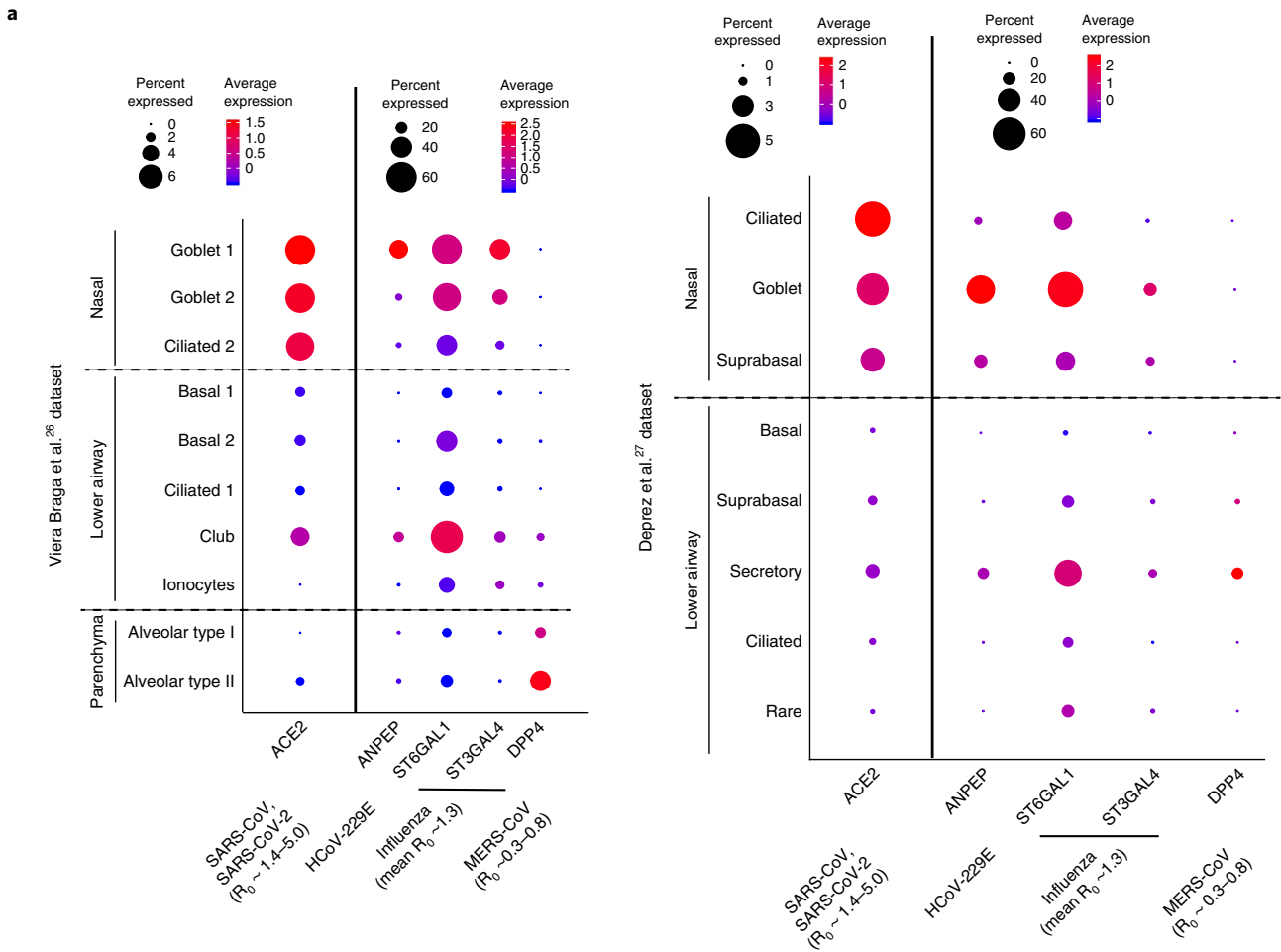
The results confirmed the expression of *ACE2* in multiple tissues shown in previous studies^{10,11} with added information on tissues not previously investigated, including nasal epithelium and cornea and its co-expression with *TMPRSS2*. We

Fig. 1 | Expression of *ACE2* and *TMPRSS2* across different tissues and its enrichment in nasal epithelial cells. **a**, RNA expression of SARS-CoV-2 entry receptor *ACE2* (first column), entry-associated protease *TMPRSS2* (second column) and their co-expression (third column) from multiple scRNA-seq datasets across different tissues. DC, dendritic cells; mac, macrophages; RBC, red blood cells; TA, transit-amplifying cells; LSC, limbal stem cells; Epi, epithelial cells; IEL, intraepithelial lymphocytes; ILC, innate lymphoid cells; GC, germinal center B cells; MT, mitochondria; LSEC, liver sinusoidal endothelial cells; VSMC, vesicular smooth muscle cells; cTEC, cortical thymic epithelial cells; mTEC, medullary thymic epithelial cells; mcTEC, medullary/cortical thymic epithelial cells; EC, endothelial cells; FB, fibroblasts; SMC, smooth muscle cells; aCM, atrial cardiomyocytes; vCM, ventricular cardiomyocytes; MNP, mononuclear phagocytes; EVT, extravillous trophoblast cells; HB, Hofbauer cells; MAIT, mucosal-associated invariant T cells; MO, monocytes; SCT, syncytiotrophoblast cells; VCT, villous cytotrophoblast cells; dM, decidual macrophages; dP, decidual perivascular cells; dS, decidual stromal cells. Raw expression values were normalized, log transformed and summarized by published cell clustering where available or reproduced clustering annotated using marker genes and cell type nomenclature from the respective studies. The size of the dots indicates the proportion of cells in the respective cell type having greater-than-zero expression of *ACE2* (first column), *TMPRSS2* (second column) or both (third column), while the color indicates the mean expression of *ACE2* (first and third columns) or *TMPRSS2* (second column). **b**, Schematic illustration depicts major anatomical regions in the human respiratory tree demonstrated in this study: nasal, lower airway and lung parenchyma (left). Expression of *ACE2* is from airway epithelial cell datasets: Vieira Braga et al.²⁶ (middle) and Deprez et al.²⁷ (right). Datasets were retrieved from existing sources and cell clustering and nomenclature were retained based on the respective studies. For gene expression results in the dot plots, the dot size represents the proportion of cells within the respective cell type expressing the gene and the dot color represents the average gene expression level within the particular cell type.

clearly detected nasal *ACE2* mRNA expression, for which protein confirmation is needed to resolve conflicting results in literature^{8,12}. Our findings may have important implications for

understanding viral transmissibility, considering that the primary viral transmission is through infectious droplets. Moreover, as SARS-CoV-2 is an enveloped virus, its release does not require cell





lysis. Thus, the virus might exploit existing secretory pathways in nasal goblet cells sustained at a presymptomatic stage. These discoveries could have translational implications. For example, given

that nasal carriage is likely to be a key feature of transmission, drugs/vaccines administered intranasally could be highly effective in limiting spread.

Fig. 2 | Respiratory expression of viral receptor/entry-associated genes and implications for viral transmissibility and genes associated with ACE2 expression. **a**, Expression of ACE2 (an entry receptor for SARS-CoV and SARS-CoV-2), ANPEP (an entry receptor for HCoV-229E), ST6GAL1/ST3GAL4 (enzymes important for synthesis of influenza entry receptors) and DPP4 (an entry receptor for MERS-CoV) from the airway epithelial datasets: Vieira Braga et al.²⁶ (left) and Deprez et al.²⁷ (right). The basic reproductive number (RO) for respective viruses, if available, is shown. **b**, Respiratory epithelial expression of the top 50 genes correlated with ACE2 expression based on Spearman's correlation analysis (with Benjamini-Hochberg-adjusted *P* values) performed on all cells within the Vieira Braga et al.²⁶ airway epithelial dataset. The colored gene names represent genes that are immune-associated (GO:0002376, immune system process or GO:0002526, acute inflammatory response). For gene expression results in the dot plots, the dot size represents the proportion of cells within the respective cell type expressing the gene and the color represents the average gene expression level within the particular cell type.

This collaborative effort by HCA Biological Network (the lung) illustrates the opportunities from integrative analyses of HCA data, with future examples of consortium work expected soon.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgments, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41591-020-0868-6>.

Received: 10 March 2020; Accepted: 31 March 2020;
Published online: 23 April 2020

References

- World Health Organization. Naming the coronavirus disease (COVID-19) and the virus that causes it. [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-\(covid-2019\)-and-the-virus-that-causes-it](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it) (2020).
- Chen, N. et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* **395**, 507–513 (2020).
- Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
- Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
- Li, W. et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450–454 (2003).
- Matsuyama, S. et al. Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. *J. Virol.* **84**, 12658–12664 (2010).
- Hoffmann, M. et al. SARS-CoV-2 entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280.e8 (2020).
- 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* **7**, e35876 (2012).
- Zhao, Y. et al. Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCoV. Preprint at <https://www.biorxiv.org/content/10.1101/2020.01.26.919985v1> (2020).
- 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.* <https://doi.org/10.1007/s11684-020-0754-0> (2020).
- Qi, F., Qian, S., Zhang, S. & Zhang, Z. Single cell RNA sequencing of 13 human tissues identify cell types and receptors of human coronaviruses. *Biochem. Biophys. Res. Commun.* <https://doi.org/10.1016/j.bbrc.2020.03.044> (2020).
- Hamming, I. et al. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J. Pathol.* **203**, 631–637 (2004).
- Zhou, Y. et al. Protease inhibitors targeting coronavirus and filovirus entry. *Antiviral Res.* **116**, 76–84 (2015).
- Iwata-Yoshikawa, N. et al. TMPRSS2 contributes to virus spread and immunopathology in the airways of murine models after coronavirus infection. *J. Virol.* **93**, e01815-18 (2019).
- Yeager, C. L. et al. Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* **357**, 420–422 (1992).
- Raj, V. S. et al. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* **495**, 251–254 (2013).
- Broszeit, F. et al. N-glycolylneuraminic acid as a receptor for influenza A viruses. *Cell Rep.* **27**, 3284–3294 (2019).
- Li, Q. et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).
- Wallinga, J. & Teunis, P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am. J. Epidemiol.* **160**, 509–516 (2004).
- Riou, J. & Althaus, C. L. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Euro Surveill.* **25**, 2000058 (2020).
- Coburn, B. J., Wagner, B. G. & Blower, S. Modeling influenza epidemics and pandemics: insights into the future of swine flu (H1N1). *BMC Med.* **7**, 30 (2009).
- Kucharski, A. J. & Althaus, C. L. The role of superspreading in Middle East respiratory syndrome coronavirus (MERS-CoV) transmission. *Euro Surveill.* **20**, 14–18 (2015).
- Killerby, M. E. et al. Middle East Respiratory Syndrome coronavirus transmission. *Emerg. Infect. Dis.* **26**, 191–198 (2020).
- Xu, Y. et al. Characteristics of pediatric SARS-CoV-2 infection and potential evidence for persistent fecal viral shedding. *Nat. Med.* **26**, 502–505 (2020).
- Guan, W.-J. et al. Clinical characteristics of coronavirus disease 2019 in China. *N. Engl. J. Med.* <https://doi.org/10.1056/NEJMoa2002032> (2020).
- Vieira Braga, F. A. et al. A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat. Med.* **25**, 1153–1163 (2019).
- Deprez, M. et al. A single-cell atlas of the human healthy airways. Preprint at <https://www.biorxiv.org/content/10.1101/2019.12.21.884759v1> (2019).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020

HCA Lung Biological Network

Nicholas E. Banovich¹², Pascal Barbry¹³, Alvis Brazma¹⁴, Joseph Collin¹⁵, Tushar J. Desai¹⁶, Thu Elizabeth Duong^{17,18}, Oliver Eickelberg¹⁹, Christine Falk^{20,21}, Michael Farzan²², Ian Glass²³, Ravindra K. Gupta²⁴, Muzlifah Haniffa^{15,25,26}, Peter Horvath^{27,28}, Norbert Hubner^{6,29,30,31}, Deborah Hung^{7,32,33}, Naftali Kaminski³⁴, Mark Krasnow³⁵, Jonathan A. Kropski^{36,37,38},

Malte Kuhnemund³⁹, Majlinda Lako¹⁵, Haeock Lee⁴⁰, Sylvie Leroy^{13,41}, Sten Linnarson⁴², Joakim Lundeberg⁴³, Kerstin B. Meyer²⁵, Zhichao Miao²⁵, Alexander V. Misharin⁴⁴, Martijn C. Nawijn⁴⁵, Marko Z. Nikolic⁴⁶, Michela Noseda^{47,48}, Jose Ordovas-Montanes^{49,50,51,52}, Gavin Y. Oudit⁵³, Dana Pe'er⁵⁴, Joseph Powell^{55,56}, Steve Quake⁵⁷, Jay Rajagopal^{52,58}, Purushothama Rao Tata⁵⁹, Emma L. Rawlins⁶⁰, Aviv Regev^{51,61,62,63}, Paul A. Reyfman⁶⁴, Orit Rozenblatt-Rosen^{51,61}, Kourosh Saeb-Parsy⁶⁵, Christos Samakovlis^{43,66}, Herbert B. Schiller⁶⁷, Joachim L. Schultze^{68,69,70}, Max A. Seibold⁷¹, Christine E. Seidman^{7,62,72}, Jonathan G. Seidman⁷, Alex K. Shalek^{51,73,74}, Douglas Shepherd⁷⁵, Jason Spence^{76,77,78}, Avi Spira^{79,80}, Xin Sun⁸¹, Sarah A. Teichmann^{25,82}, Fabian J. Theis⁸³, Alexander M. Tsankov⁸⁴, Ludovic Vallier^{85,86}, Maarten van den Berge⁸⁷, Jeffrey Whitsett⁸⁸, Ramnik Xavier^{32,51}, Yan Xu⁸⁹, Laure-Emmanuelle Zaragosi¹³, Darin Zerti^{15,90}, Hongbo Zhang⁹¹, Kun Zhang⁹², Mauricio Rojas⁹³ and Francisco Figueiredo⁹⁴

¹²Translational Genomics Research Institute, Phoenix, AZ, USA. ¹³Université Côte d'Azur, CHU Nice, CNRS, IPMC, Sophia-Antipolis, France. ¹⁴European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Cambridge, UK. ¹⁵Biosciences Institute, Faculty of Medical Sciences, Newcastle University, Newcastle-upon-Tyne, UK. ¹⁶Department of Medicine and Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA, USA. ¹⁷Department of Pediatrics Division of Respiratory Medicine, University of California San Diego, San Diego, CA, USA. ¹⁸Rady Children's Hospital San Diego, San Diego, CA, USA. ¹⁹Division of Pulmonary Sciences and Critical Care Medicine, Department of Medicine, University of Colorado, Aurora, CO, USA. ²⁰Institute of Transplant Immunology, Hannover Medical School, Hannover, Germany. ²¹German Center for Infectious Diseases DZIF, Brunswick, Germany. ²²Department of Immunology and Microbiology, The Scripps Research Institute, Jupiter, FL, USA. ²³Department of Pediatrics, Genetic Medicine, University of Washington, Seattle, WA, USA. ²⁴Cambridge Institute for Therapeutic Immunology and Infectious Diseases, Jeffrey Cheah Biomedical Centre, University of Cambridge, Cambridge, UK. ²⁵Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge, UK. ²⁶Department of Dermatology and NIHR Newcastle Biomedical Research Centre, Newcastle Hospitals NHS Foundation Trust, Newcastle-upon-Tyne, UK. ²⁷Synthetic and Systems Biology Unit, Biological Research Centre (BRC), Szeged, Hungary. ²⁸Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland. ²⁹DZHK (German Centre for Cardiovascular Research), Partner Site Berlin, Berlin, Germany. ³⁰Berlin Institute of Health (BIH), Berlin, Germany. ³¹Charité-Universitätsmedizin, Berlin, Germany. ³²Department of Molecular Biology, Massachusetts General Hospital, Boston, MA, USA. ³³Infectious Disease and Microbiome Program, Broad Institute of MIT and Harvard, Cambridge, MA, USA. ³⁴Pulmonary, Critical Care and Sleep Medicine, Yale University School of Medicine, New Haven, CT, USA. ³⁵Department of Biochemistry and Wall Center for Pulmonary Vascular Disease, Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA, USA. ³⁶Division of Allergy, Pulmonary and Critical Care Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. ³⁷Department of Veterans Affairs Medical Center, Nashville, TN, USA. ³⁸Department of Cell and Developmental Biology, Vanderbilt University, Nashville, TN, USA. ³⁹Cartana AB, Stockholm, Sweden. ⁴⁰Department of Biomedicine and Health Sciences, The Catholic University of Korea, Seoul, Korea. ⁴¹Institut de Pharmacologie Moléculaire et Cellulaire, Sophia-Antipolis, France. ⁴²Division of Molecular Neurobiology, Department of Medical Biochemistry and Biophysics, Karolinska Institute, Solna, Sweden. ⁴³SciLifeLab, Stockholm University/KTH Royal Institute of Technology, Solna, Sweden. ⁴⁴Division of Pulmonary and Critical Care Medicine, Northwestern University, Chicago, IL, USA. ⁴⁵Department of Pathology and Medical Biology, University of Groningen, GRIAC Research Institute, University Medical Center Groningen, Groningen, the Netherlands. ⁴⁶UCL Respiratory, Division of Medicine, University College London, London, UK. ⁴⁷National Heart and Lung Institute, Imperial College London, London, UK. ⁴⁸British Heart Foundation Centre for Research Excellence and Centre for Regenerative Medicine, Imperial College London, London, UK. ⁴⁹Division of Gastroenterology, Boston Children's Hospital, Boston, MA, USA. ⁵⁰Program in Immunology, Harvard Medical School, Boston, MA, USA. ⁵¹Broad Institute of MIT and Harvard, Cambridge, MA, USA. ⁵²Harvard Stem Cell Institute, Cambridge, MA, USA. ⁵³Division of Cardiology, University of Alberta, Edmonton, Alberta, Canada. ⁵⁴Computational and Systems Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA. ⁵⁵Garvan-Weizmann Centre for Cellular Genomics, Garvan Institute of Medical Research, Sydney, New South Wales, Australia. ⁵⁶UNSW Cellular Genomics Futures Institute, University of New South Wales, Sydney, New South Wales, Australia. ⁵⁷Depts of Bioengineering and Applied Physics, Stanford University and the Chan Zuckerberg Biohub, Stanford University, Stanford, CA, USA. ⁵⁸Center for Regenerative Medicine, Massachusetts General Hospital, Boston, Boston, MA, USA. ⁵⁹Department of Cell Biology, Regeneration Next Initiative, Duke University School of Medicine, Durham, NC, USA. ⁶⁰Wellcome Trust/CRUK Gurdon Institute and Department Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK. ⁶¹Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA. ⁶²Howard Hughes Medical Institute, Chevy Chase, MD, USA. ⁶³Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. ⁶⁴Division of Pulmonary and Critical Care Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. ⁶⁵Department of Surgery, University of Cambridge and NIHR Cambridge Biomedical Research Centre, Cambridge, UK. ⁶⁶Cardiopulmonary Institute, Justus Liebig University, Giessen, Germany. ⁶⁷Comprehensive Pneumology Center (CPC)/Institute of Lung Biology and Disease (ILBD), Helmholtz Zentrum München, Member of the German Center for Lung Research (DZL), Munich, Germany. ⁶⁸Department for Genomics & Immunoregulation, LIMES-Institute, University of Bonn, Bonn, Germany. ⁶⁹PRECISE Platform for Single Cell Genomics & Epigenomics, German Center for Neurodegenerative Diseases and University of Bonn, Bonn, Germany. ⁷⁰Department of Pediatrics, Center for Genes, Environment, and Health, Denver, CO, USA. ⁷¹Department of Pediatrics, Center for Genes, Environment and Health, National Jewish Health, Denver, CO, USA. ⁷²Cardiovascular Division, Brigham & Women's Hospital, Boston, MA, USA. ⁷³Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, USA. ⁷⁴Institute for Medical Engineering and Science (IMES), Koch Institute for Integrative Cancer Research and Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA. ⁷⁵Center for Biological Physics and Department of Physics, Arizona State University, Tempe, AZ, USA. ⁷⁶Gastroenterology, Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI, USA. ⁷⁷Department of Cell and Developmental Biology, University of Michigan Medical School, Ann Arbor, MI, USA. ⁷⁸Department of Biomedical Engineering, University of Michigan College of Engineering, Ann Arbor, MI, USA. ⁷⁹Department of Medicine, Boston University School of Medicine, Boston, MA, USA. ⁸⁰Johnson & Johnson Innovation, Cambridge, MA, USA. ⁸¹Department of Pediatrics and Department of Biological Sciences, University of California San Diego, San Diego, CA, USA. ⁸²Theory of Condensed Matter Group, Cavendish Laboratory/Department of Physics,

University of Cambridge, Cambridge, UK. ⁸³Institute of Computational Biology, Helmholtz Zentrum München and Departments of Mathematics and Life Sciences, Technical University Munich, Munich, Germany. ⁸⁴Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ⁸⁵Wellcome and MRC Cambridge Stem Cell Institute, Jeffrey Cheah Biomedical Centre, Cambridge, UK. ⁸⁶Department of Surgery, Cambridge Biomedical Campus, Cambridge, UK. ⁸⁷Department of Pulmonary Diseases and Tuberculosis, University of Groningen, GRIAC Research Institute, University Medical Center Groningen, Groningen, Netherlands. ⁸⁸Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. ⁸⁹Divisions of Pulmonary Biology and Biomedical Informatics, Perinatal Institute, Cincinnati Children's Hospital Medical Center, University of Cincinnati College of Medicine, Cincinnati, OH, USA. ⁹⁰Microscopy Centre and Department of Applied Clinical Sciences and Biotechnology, University of L'Aquila, via Vetoio, Coppito, Italy. ⁹¹Key Laboratory for Stem Cells and Tissue Engineering, Ministry of Education and Department of Histology and Embryology of Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou, China. ⁹²Department of Bioengineering, University of California San Diego, La Jolla, CA, USA. ⁹³Division of Pulmonary, Allergy and Critical Care Medicine, University of Pittsburgh, Pittsburgh, PA, USA. ⁹⁴Department of Ophthalmology and Faculty of Medical Sciences, Newcastle NHS Foundation Trust, Newcastle upon Tyne, UK.

Methods

Datasets were retrieved from published and unpublished datasets in multiple human tissues, including airways^{26,27}, cornea (personal communication; Lako laboratory, Newcastle, UK), skeletal muscle (personal communication, Teichmann laboratory, Wellcome Sanger Institute and Zhang laboratory, Sun-Yat-Sen University, Guangzhou, China), ileum²⁸, colon²⁹, pancreas³⁰, liver³¹, gallbladder (personal communication; Vallier laboratory, University of Cambridge, UK), heart (Teichmann laboratory, Hubner laboratory/Berlin, Seidmanns/Harvard and Nosedal laboratory/Imperial College London, UK), kidney³², placenta/decidua³³, testis³⁴, prostate gland³⁵, brain³⁶, skin³⁷, retina³⁸, spleen³⁹, esophagus³⁹ and fetal tissues^{40,41}. Raw expression values were normalized and log transformed. We retained cell clustering based on the original studies when available.

For each dataset where per-cell annotation was not available, we re-processed the data from a raw or normalized (whichever was deposited alongside the original publication) quantification matrix. The standard scanpy (v.1.4.3) clustering procedure was followed. When batch information was available, the harmony package was used to correct batch effects in the principal component space and the corrected principal components were used for computing nearest-neighbor graphs. To re-annotate the cells, multiple clusterings of different resolutions were generated among which the one best matching published clustering was picked and manual annotation was undertaken using marker genes described in the original publication. Full details can be found in analysis notebooks available at github.com/Teichlab/covid19_MS1.

Illustration of the results was generated using scanpy and Seurat (v.3.1). For correlation analysis with *ACE2*, we performed Spearman's correlation with statistical tests using the R Hmisc package (v.4.3-1) and *P* values were adjusted with the Benjamini-Hochberg method with the R stats package (v.3.6.1) on the Vieira Braga et al.²⁶ airway epithelial dataset and the Deprez et al.³⁷ airway dataset. We also tested multiple additional approaches, including Kendall's correlation, data transformation by *sctransform* function in the Seurat package and data imputation by the Markov affinity-based graph imputation of cells algorithm to compare correlation results. While imputation significantly improved correlations, the top genes correlated with *ACE2* are largely the same as the analysis performed on un-imputed data. With the uncertainty of the extent that imputation artificially distorted the data, we reported results with no imputation, even though correlations were low. The correlation coefficients for all genes are included as Supplementary Data 1. The top 50 genes in each dataset were characterized based on gene ontology classes from the Gene Ontology database and associated pathways in PathCards were from the Pathway Unification database.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The published datasets can be found as follows: pulmonary airways (European Genome-phenome Archive: EGAS00001001755, EGAS00001002649; EGAS00001004082; [lungcellatlas.org](https://www.lungcellatlas.org) and www.genomique.eu/cellbrowser/HCA), ileum (NCBI: GSE134809), colon (Single Cell Portal: SCP259; singlecell.broadinstitute.org/single_cell), pancreas (NCBI: GSE84133), liver (NCBI: GSE115469), kidney (www.kidneycellatlas.org), placenta/decidua (EBI Array Express: E-MTAB-6701; maternal-fetal-interface.cellgeni.sanger.ac.uk), testis (NCBI: GSE120508), brain (www.gtexportal.org/home/data-sets), retina (NCBI: GSE135922), skin (European Genome-phenome Archive: EGAS00001002927), spleen and esophagus (tissuestabilitycellatlas.org) and fetal tissues (Array Express: E-MTAB-7407 and E-MTAB-8581; developmentcellatlas.ncl.ac.uk). All of the published datasets and relevant data from unpublished sources in this study can be visualized and assessed through a website portal (www.covid19cellatlas.org).

Code availability

Analysis notebooks are available at github.com/Teichlab/covid19_MS1.

References

- Martin, J. C. et al. Single-cell analysis of Crohn's disease lesions identifies a pathogenic cellular module associated with resistance to anti-TNF therapy. *Cell* **178**, 1493–1508 (2019).
- Smillie, C. S. et al. Intra- and inter-cellular rewiring of the human colon during ulcerative colitis. *Cell* **178**, 714–730 (2019).
- Baron, M. et al. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. *Cell Syst.* **3**, 346–360 (2016).
- MacParland, S. A. et al. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nat. Commun.* **9**, 4383 (2018).
- Stewart, B. J. et al. Spatiotemporal immune zonation of the human kidney. *Science* **365**, 1461–1466 (2019).
- Vento-Tormo, R. et al. Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**, 347–353 (2018).

- Guo, J. et al. The adult human testis transcriptional cell atlas. *Cell Res.* **28**, 1141–1157 (2018).
- Henry, G. H. et al. A cellular anatomy of the normal adult human prostate and prostatic urethra. *Cell Rep.* **25**, 3530–3542 (2018).
- Habib, N. et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* **14**, 955–958 (2017).
- Cheng, J. B. et al. Transcriptional programming of normal and inflamed human epidermis at single-cell resolution. *Cell Rep.* **25**, 871–883 (2018).
- Voigt, A. P. et al. Molecular characterization of foveal versus peripheral human retina by single-cell RNA sequencing. *Exp. Eye Res.* **184**, 234–242 (2019).
- Madisson, E. et al. scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. *Genome Biol.* **21**, 1 (2019).
- Popescu, D. M. et al. Decoding human fetal liver haematopoiesis. *Nature* **574**, 365–371 (2019).
- Park, J. E. et al. A cell atlas of human thymic development defines T cell repertoire formation. *Science* **367**, eaay3224 (2020).

Acknowledgments

We are grateful to C. Bargmann, J. Farrar and S. Aldridge for stimulating discussions. We thank J. Eliasova (scientific illustrator) for support with the design of figures, S. Sansum for support in document processing and M. Prete, V. Kiselev and the Wellcome Sanger Cellular Genetics IT Team, as well as P. Bevan, for support with setting up the website portal. The human embryonic and fetal material was provided by the Joint Medical Research Council (MRC)/Wellcome (MR/R006237/1) Human Developmental Biology Resource (www.hdbbr.org). This publication is part of the HCA (www.humancellatlas.org/publications). This work was supported by the Wellcome Sanger Institute core funding (WT206194) and the Wellcome Strategic Scientific Support award 'Pilot projects for the Human Cell Atlas' (WT211276/Z/18/Z), a Seed Network grant from the Chan Zuckerberg Initiative to P.B., T.D., T.E.D., O.E., P.H., N.H., N.K., M.K., K.B.M., A.M., M.C.N., M.N., D.P., J.R., P.R.T., S.Q., A.R., O.R., M.A.S., J.S., J.G.S., C.E.S., H.B.S., D.S., A.M.T., J.W. and K.Z. and by the European Union's H2020 Research and Innovation Program under grant agreement no. 874656 (discovAIR) to P.B., A.B., M.K., S.L., J.L., K.B.M., M.C.N., K.S.P., C.S., H.B.S., J.S., F.J.T. and M.vd.B. W.S. acknowledges funding from the Newton Fund, MRC, The Thailand Research Fund and Thailand's National Science and Technology Development Agency. M.C.N. acknowledges funding from GSK Ltd., Netherlands Lung Foundation project no. 5.1.14.020 and 4.1.18.226. T.D. acknowledges funding from HubMap consortium and Stanford Child Health Research Institute-Woods Family Faculty Scholarship. T.E.D. acknowledges funding from HubMap. P.H. acknowledges funding from LENDULET-BIOMAG grant (2018-342) and the European Regional Development Fund (GINOP-2.3.2-15-2016-00006, GINOP-2.3.2-15-2016-00026, GINOP-2.3.2-15-2016-00037). J.L.B. acknowledges funding from the MRC and the UK Regenerative Medicine Platform (MR/5005579/1). P.B. acknowledges funding from Fondation pour la Recherche Médicale (DEQ20180339158), Agence Nationale de la Recherche (UCAJEDI, ANR-15-IDEX-01; SAHARRA, ANR-19-CE14-0027; France Génomique, ANR-10-INBS-09-03) and Conseil Départemental des Alpes Maritimes (2016-294DGADSH-CV; 2019-390DGADSH-CV). N.E.B. and J.K. acknowledge funding from National Institutes of Health (NIH) grant R01HL145372 and DOD grant W81XWH1910416. I.G. acknowledges funding from NIH (5R24HD000836) and the Eunice Kennedy Shriver National Institute of Child Health and Human. N.H., J.G.S. and C.E.S. acknowledge funding by the Leducq foundation. N.H. is recipient of an ERC Advanced grant. J.K. acknowledges funding from NIH grant K08HL130595 and the Doris Duke Charitable Foundation. N.K. acknowledges funding from NIH grants R01HL127349, U01HL145567 and an unrestricted grant from Three Lakes Foundation. M.K. acknowledges HHMI and Wall Center for Pulmonary Vascular Disease. H.L. acknowledges funding from National Research Foundation of Korea. K.M. acknowledges funding from Wellcome Trust. A.M. acknowledges funding from NIH grants HL135124, AG049665 and AI135964. M.Z.N. acknowledges funding from Rutherford Fund Fellowship allocated by the MRC and the UK Regenerative Medicine Platform (MR/5005579/1 to M.Z.N.). M.Z.N. and M.Y. have been funded by the Rosetrees Trust (M899). M.N. acknowledges funding from a BHF/DZHK grant and British Heart Foundation (PG/16/47/32156). J.O.-M. acknowledges funding from Richard and Susan Smith Family Foundation. D.P. acknowledges funding from Alan and Sandra Gerry Metastasis and Tumor Ecosystems Center. J.P. acknowledges funding from National Health and MRC. P.R.T. acknowledges funding from R01HL146557 from NHLBI/NIH. E.L.R. acknowledges funding from MRC MR/P009581/1 and MR/SO35907/1. A.R. and O.R. acknowledge HHMI, the Klarman Cell Observatory and the Manton Foundation. K.S.-P. acknowledges NIHR Cambridge Biomedical Research Centre. C.S. acknowledges the Swedish Research Council, Swedish Cancer Society and CPI. H.B.S. acknowledges the German Center for Lung Research and the Helmholtz Association. J.S. acknowledges Boehringer Ingelheim, by the German Research Foundation (DFG; EXC2151/1, ImmunoSensation2, the immune sensory system, project no. 390873048), project nos. 329123747, 347286815) and by the HGF grant sparse2big. A.K.S. acknowledges the Beckman Young Investigator Program, a Sloan Fellowship in Chemistry, the NIH (5U24AI118672) and the Bill and Melinda Gates Foundation. F.J.T. acknowledges the German Center for Lung Research. M.vd.B. acknowledges the Ministry of Economic Affairs and Climate Policy by means of the PPP. K.B.W. is funded by the University College London, Birkbeck MRC Doctoral Training Programme. J.W. and Y.Y. acknowledge NIH, U01 HL148856 LungMap phase II. R.X. acknowledges

the NIH (DK043351). H.Z. is supported by the National Key R&D Program (no. 2019YFA0801703) and National Natural Science Foundation of China (no. 31871370). M.A.S. acknowledges funding from NIH grants R01 HL135156, R01 MD010443, R01 HL128439, P01 HL132821, P01 HL107202, R01 HL117004, and DOD Grant W81WH-16-2-0018 and NJH Regenerative Medicine and Genome Editing Program (REGEN).

Author Contributions

P.B., A.M., M.N. and J.R. serve as the coordinators for the HCA Lung Biological Network. W.S., N.H., C.B. and M.B. performed data analyses, initiated by S.A.T. and supervised by S.A.T. and M.C.N. R.Q., M.L., C.T.-L., H.M., D.R., F.S., M.Y., K.B.W. and the HCA Lung Biological Network generated and provided data. W.S., N.H. and the HCA Lung Biological Network interpreted the data. W.S., with significant input from the HCA Lung Biological Network, especially M.C.N., M.vd.B., S.A.T. and K.B.M., wrote the paper. K.B.W., M.Y. and J.L.B. performed experiments for the revision process. All authors read the manuscript, offered feedback and approved it before submission.

Competing interests

N.K. was a consultant to Biogen Idec, Boehringer Ingelheim, Third Rock, Pliant, Samumed, NuMedii, Indaloo, Theravance, LifeMax, Three Lake Partners and Optikira in the last 3 years and received nonfinancial support from MiRagen. J.L. is a scientific consultant for 10X Genomics. J.K. reports advisory board fees from Boehringer Ingelheim, nonfinancial study support from Genentech and grant funding from Boehringer Ingelheim. A.R. is a co-founder and equity holder of Celsius Therapeutics,

an equity holder in Immunitas and an SAB member of Thermo Fisher Scientific, Syros Pharmaceuticals, Asimov and Neogene Therapeutics. O.R. is a co-inventor on patent applications filed by the Broad Institute to inventions relating to single-cell genomics applications, such as in PCT/US2018/060860 and US Provisional Application No. 62/745,259. A.K.S. reports compensation for consulting and/or SAB membership from Merck, Honeycomb Biotechnologies, Cellarity, Cogen Therapeutics, Orche Bio and Dahlia Biosciences. F.J.T. reports receiving consulting fees from Roche Diagnostics GmbH and ownership interest in Cellarity Inc. S.A.T. was a consultant at Genentech, Biogen and Roche in the last 3 years. L.V. is a founder of Definigen and Bilitech, two biotech companies using hPSCs and organoid for disease modeling and cell-based therapy. F.S. is a founder of Bilitech, a biotechnology company using organoids for cell-based therapy.

Additional information

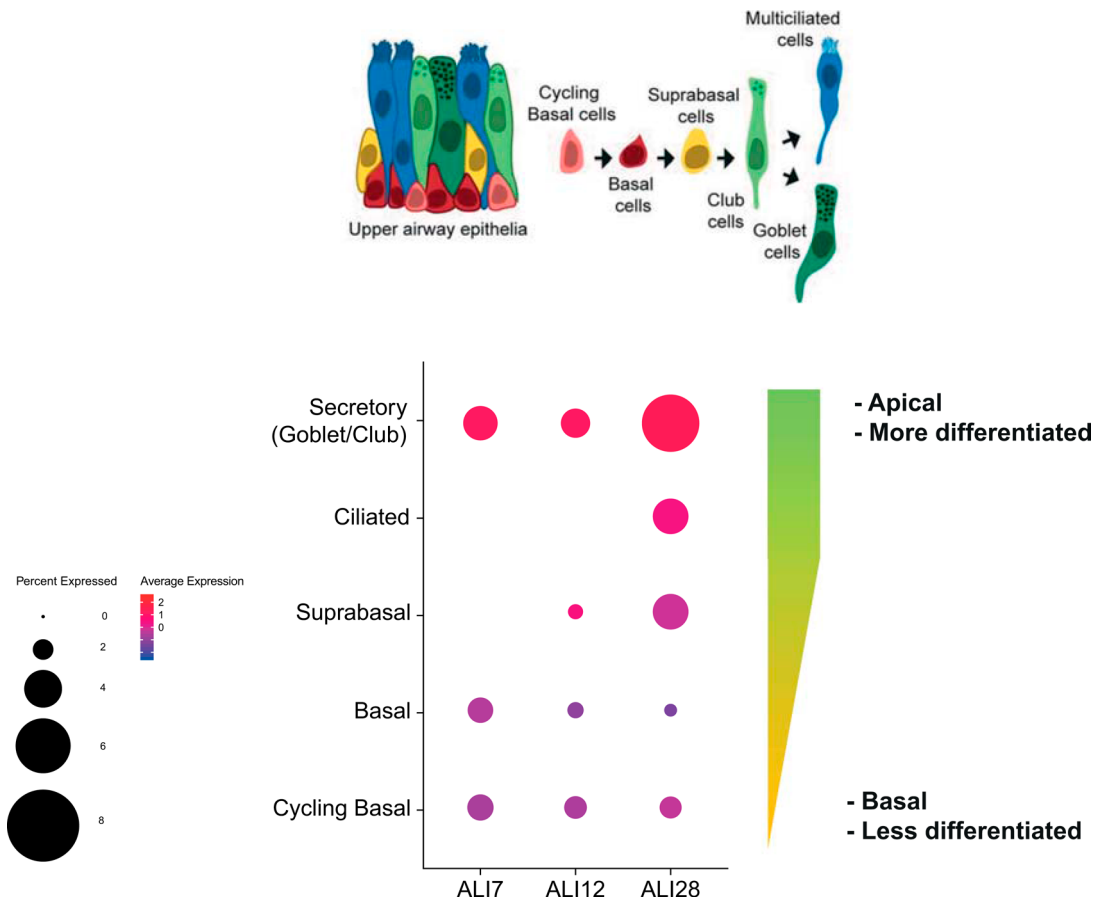
Extended data is available for this paper at <https://doi.org/10.1038/s41591-020-0868-6>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41591-020-0868-6>.

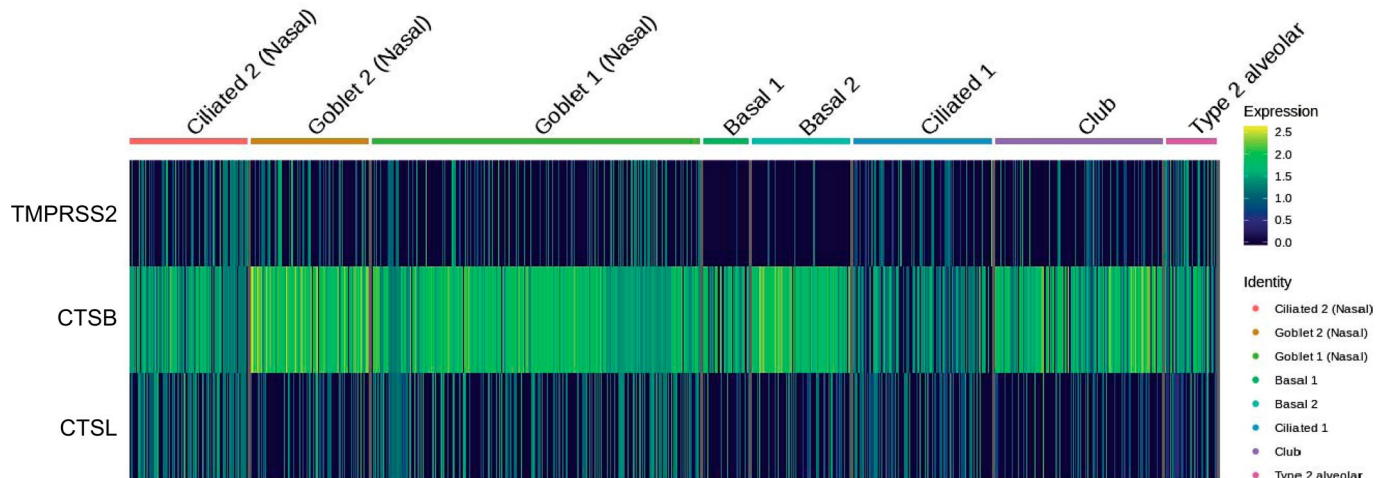
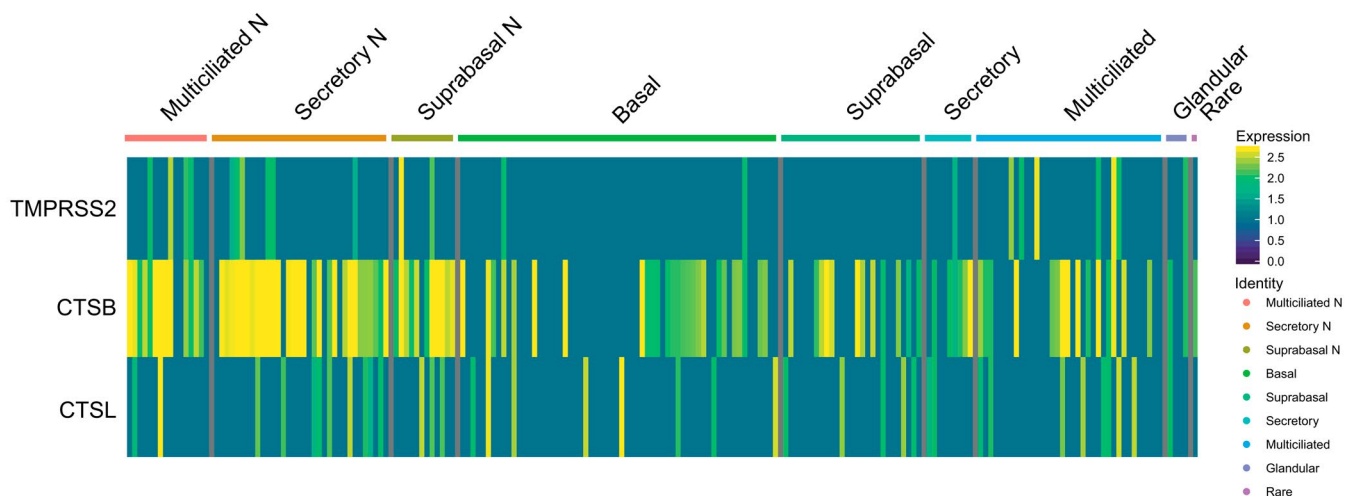
Correspondence and requests for materials should be addressed to W.S.

Peer review information Saheli Sadanand was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.



Extended Data Fig. 1 | Gene expression of ACE2 in an *in vitro* air-liquid interface (ALI) system. Epithelial regeneration system from nasal epithelial cells was used for *in vitro* cultures on successive days (7, 12 and 28), resulting in different epithelial cell types along differentiation trajectory characterized in Ruiz García *et al.* 2019. The cultures were differentiated in Pneumacult media. Schematic illustration depicts the respective cell types in the differentiation trajectory, and the dot plot illustrates the cultured cell types along the differentiation pseudotime, along with their respective location within the epithelial layers. For gene expression results in the dot plot: the dot size represents the proportion of cells within the respective cell type expressing the gene and the dot color represents the average gene expression level within the particular cell type.

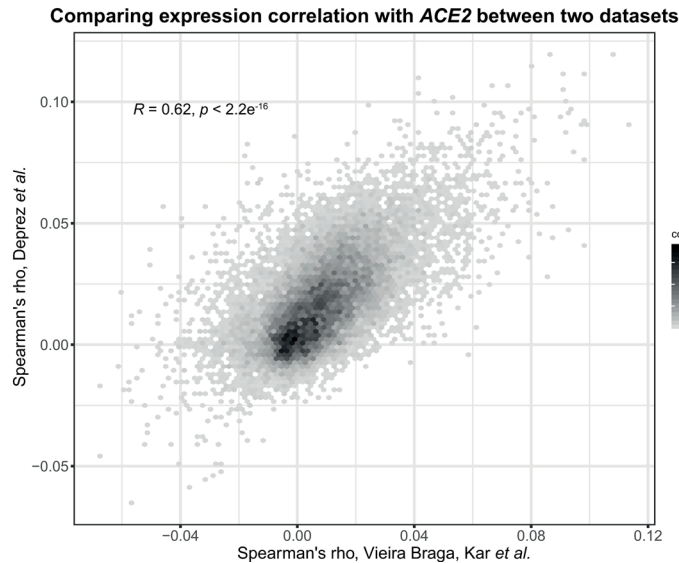
Vieira Braga, Kar *et al.* (*Nature Medicine*, 2019) datasetDeprez *et al.* (*bioRxiv*, 2019) dataset

Extended Data Fig. 2 | Expression and co-expression of SARS-CoV-2 entry-associated proteases in ACE2⁺ airway epithelial cells. The expression of SARS-CoV-2 entry-associated proteases *TMPRSS2*, *CTSB*, and *CTSL* in ACE2⁺ cells from the Vieira Braga, Kar *et al.* (top) and Deprez *et al.* (bottom) airway epithelial datasets is shown. The color represents the expression level at the single-cell resolution and the cells are grouped based on the cell types specified.

a



b



Extended Data Fig. 3 | Spearman's correlation results from the two airway datasets are largely consistent. **a**, Respiratory epithelial expression of the top 50 genes correlated with ACE2 expression based on Spearman's correlation analysis performed on all cells within the Deprez *et al.* dataset. The colored gene names represent genes that are immune-associated (GO:0002376: immune system process). **b**, The Spearman's correlation coefficients of gene expression with ACE2 from the Vieira Braga, Kar *et al.* airway epithelial dataset and the Deprez *et al.* airway dataset are shown in the scatter plot. The number of observations for the genes is counted in each bin, the value on the x-axis represents the Spearman's correlation coefficients from the Vieira Braga, Kar *et al.* dataset, and the value on the y-axis represents the Spearman's correlation coefficients from the Deprez *et al.* dataset.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data were collected based on existing resources and no software was used.

Data analysis

Retrieved data clustering information was retained or reprocessed using scanpy (version 1.4.3) and harmony. Illustration of the results was generated using scanpy (version 1.4.3) and Seurat (version 3.1). The Spearman's correlation with statistical tests were done using the R Hmisc package (version 4.3-1) and the p values were adjusted with Benjamini-Hochberg method with the R stats package (version 3.6.1). Analysis notebooks are available at github.com/Teichlab/covid19_MS1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The published datasets can be found as followed: pulmonary airways (European Genome-phenome Archive: EGAS00001001755, EGAS00001002649; lungcellatlas.org and www.genomique.eu/cellbrowser/HCA), ileum (NCBI: GSE134809), colon (Single Cell Portal: SCP259; portals.broadinstitute.org/single_cell), pancreas (NCBI: GSE84133), liver (NCBI: GSE115469), kidney (www.kidneycellatlas.org), placenta/decidua (EBI Array Express: E-MTAB-6701; maternal-fetal-interface.cellgeni.sanger.ac.uk), testis (NCBI: GSE120508), brain (www.gtexportal.org/home/data-sets), retina (NCBI: GSE135922), skin (European Genome-phenome Archive: EGAS00001002927), spleen and esophagus (tissuestabilitycellatlas.org) and fetal tissues (Array Express: E-MTAB-7407 and E-MTAB-8581; developmentalcellatlas.ncl.ac.uk).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	<input type="text" value="No sample size calculation was performed. Selection of sample size was not relevant to this study."/>
Data exclusions	<input type="text" value="No data were excluded in the study."/>
Replication	<input type="text" value="Each dataset has biological replicates based on particular studies. The findings on airways were confirmed based on 2 independent ex vivo datasets and another in vitro culture dataset."/>
Randomization	<input type="text" value="This is not relevant to the study as we collected datasets from existing resources based on availability."/>
Blinding	<input type="text" value="This is not relevant to the study as we collected datasets from existing resources based on availability."/>

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

Methods

n/a	Involvement
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging